MANAGERIAL ECONOMICS

(DBM05)

(PG DIPLOMA)

ACHARYA NAGARJUNA UNIVERSITY

CENTRE FOR DISTANCE EDUCATION

NAGARJUNA NAGAR,

GUNTUR

ANDHRA PRADESH

CONTENTS

Course Preparation Team

Prof. M. Adhikary
Faculty of Management Studies
University of Delhi
Delhi

Prof. P.R. Banerjee
Indian Institute of Management
Lucknow

Prof. K. Roy Choudhury
St. Stephen's College
University of Delhi
Delhi

Dr. Raghav Gaiha
Faculty of Management Studies
University of Delhi
Delhi

Prof. Rakesh Khurana
Director
School of Management Studies
IGNOU

Prof. K. N. Kabra
Indian Institute of Public Administration
New Delhi

Dr. A.M. Paranjpe
Applied Economic Research Bureau
Bombay

Dr. S. Singh
Punjab National Bank
Bhopal

Dr. Devi Singh
International Management Institute
New Delhi

Dr. M.L. Bhatia
School of Management Studies
IGNOU

Prof. G. Sambasiva Rao
Editor
IGNOU

BLOCK 1 CONCEPTS AND TECHNIQUES

This block, consisting of three units, introduces you to the subject of Managerial Economics.

The first unit concentrates on the scope and methodology of the subject.

In the second unit, teh concepts on scarcity, marginalism, opportunity costs, time perspective, discounting, risk and uncertainty, etc., are briefly explained.

The third unit introduces you to the basic technique of measurement and optimisation, both constrained and unconstrained, using calculus.

Taken as a whole, the Block 1 outlines the conceptual frame work and points towards the technical nature of the subject.

UNIT 1 INTRODUCTION TO MANAGERIAL ECONOMICS

Structure

1.0 INTRODUCTION

In MBDM-103, you have been exposed to the socio-economic environment. Given the environment, the management thinks and takes or makes decision. Almost all such decisions, in one way or the other, relate to economic variables like demand, price, supply, stock, input, output finance, profit and the like. In the name of Managerial Economics, we will now concentrate on these economic decision variables and the economic decision process. Traditionally, the problem of optimal decision by firms and individuals has been studies in microeconomic theory. Managerial Economics, as a separate discipline owes its origin to the growing disenchantment with economic theory in providing solutions to the problem faced by business. This does not mean that Managerial Economics provides ready-made solutions to business problems. What it provides is a tool-box of analysis and a technique thinking which can be helpful in conceptualising the problems faced by management of a business firm. Thus, Managerial Economics is supposed to enrich the **Conceptual and technical skill** of a manager facing business decision problems.

1.1 OBJECTIVES

Managerial Economics is the application of economic analysis to evaluate business decisions. In view of this, in this introductory unit, our objective is to:

- explain the nature and form of economic analysis.
- identify the business decision araes where economic analysis can be applied
- spell out the relationship between Managerial Economics and other disciplines
- outline the methodology of managerial economics.

1.2 ECONOMIC ANALYSIS

The concern of Economics is with the economic problem and its identification, description, explanation and solution, if possible. **An economic problem** is a problem of choice and valuation. The problem of choice arises because limited means (resources) with alternative uses are to be utilised to satisfy ends (wants) which are unlimited and of varying degree of importance. Had resources like men, materials, machines, money, time, energy etc., not been scarce, there would have been no problem of choice; scarcity is at the root of all economic problems of choice. We have to choose between ends, between means, between use of means and satisfaction of ends. Because of scarcity of resources, we have to constantly match ends to means; this is called **'economic activity'**. The optimal economic activity is to maximise the attainment of ends, given the means and their scarcities or to minimise the use of resources, given the ends and their priorities.

Decision-making by management is truly economic in nature because it involves choice from among a set of alternatives- alternative cources of action. A manager, in any function at any level in any organisation, always exercises choice in the name of decision- making. A finance manager chooses the sources and uses of funds. A production manager chooses the product-mix. A personnel manager chooses the staffing pattern. A sales manager chooses the market segments. A purchase manager chooses the quality of materials. Such example canf be multiplied to suggest that each and every manager chooses one thing or the other from among a set of alternatives. A manager's choice is dictated by his objectives and constraints. The optimal decision-making is an act of optimal economic choice, considering objectives and constraints. The optimal decision-making is an act of optimal economic choice, considering objectives and constraints. This justifies an evaluation of managerial decisions through concepts, precepts, tools and techniques of economic analysis.

Economic analysis may be of various types and forms :

i) Micro vs. Macro Analyis : In micro economic analysis the problem of choice is analysed focusing on single individual entitles like a consumer, a producer, a seller, an investor, a commodity, a factor, a market and the like. In macro economic analysis, the problem is approached from the angle of totality or aggregates like national income, national consumption, national investment, general price level etc.

ii) Partial vs. General Equilibrium Analysis : The optimal economic choice defines the state of equilibrium. To attain the state of stable equilibrium, the economic problem may be analysed part by part - one at a time - assuming "other things remaining the same". This is partial equilibrium analysis. For example, given the price, the cost budget, the technology and the input requirement, the purchase manager has to decide on the quantity and quality of the materials he purchases. By contrast, this assumption of "givens" or "other things remaining equal" may be relaxed and interdependence or interactions among variables may be allowed. This is general equilibrium analysis.

iii) Static, Comparative Static vs. Dynamic Analysis : This has reference to the time dimension of the analysis of the problem. A problem may be analysed :

- at a point of time, allowing no change (Static)
- at a point of time, allowing once-for-all change (Comparative Static)
- over a period of time, allowing successive changes (Dynamic)

iv) Positive vs. Normative Analysis : In positive economic analysis, the problem is analysed in objective terms based on principles and theories. In normative economic analysis, the problem is analysed based on value judgement (norms) and, therefore, the policy prescriptions are explicitly stated.

You may note that the distinctions stated above are not watertight. There may be considerable movement from the once to the other, and some overlapping. By and large, Managerial Economics is **Primarily** micro, partial equilibrium type, comparative static, and positivist in approach, though occasionally we may make reference to elements of macro economics, general equilibrium, dynamic and normative policy issues. In fact, a manager has to keep his economic perspective wide open, but for analysis of his decision problem, he must approach his decision environment with a defined set of assumptions and methodologies. The manager himself is a micro economic entity and the firm where he functions is a microeconomic unit; and, therefore microeconomics is relatively more relevant for his tool kit.

Activity 1

a) In what way, is a manger's job an economic one?

....

....

....

....

....

....

....

....

....

b) List a few economic problems :

Micro	Macro
....	....
....	....
....	....
....	....
....	....

Static	Dynamic
....	....
....	....
....	....
....	....
....	....

c) Against each statement listed below, name the type of economic analysis which will be relevant:

i) The procurement price for foodgrains fixed by the Government is designed to protect the farmers ....

ii) There is bearish trend in the stock market ....

iii) The explosive population growth in India is quite alarming ....

iv) The ONGC is planning an expansion programme ....

v) The NTC is making loss ....

vi) The textile industry is facing recession ....

d) Would you accept the view that the same economic problem can be explained through different analyses?

....

....

....

....

....

1.3 BUSINESS DECISIONS

A decision is not a solution by itself; it is only an attempt towards a solution. A decision may solve an existing problem but it may also create a new problem.

Business decisions are often classified into categories depending upon the managerial function to which they relate it. From this standpoint, we may talk about:

i) **Financial Decisions :** Such decisions may relate to costing, budgeting, accounting, auditing, tax-planning, portfolio composition, capital structure, dividend distribution and the like.

ii) **Production Decisions :** Such decisions may relate to quantity and quality of product, choice of technology, product-mix, plant location and lay out, production scheduling, maintenance, pollution control etc.

iii) **Personnel Decisions :** Such decisions may relate to recruitment, selection, induction, training, placement, promotion, transfer, retirement or retrenchment of staff.

iv) **Marketing Decisions :** Such decisions may relate to sales volume, sales force, sales promotion, market research, customer service, packaging, advertisement, new product positioning etc.

v) Miscellaneous Decisions : This category may relate to all residuary items like purchasing, inventory control, information system, data processing, public relations etc.

In the realm of each decision area, one can further distinguish between

a) Scientific and intuitive decisions

b) Strategic and tactical decisions

c) Certain and uncertain decisions

d) Major and minor decisions

e) Hard and soft decisions, and so on.

The point to note is that whatever may be the functional nature or type, all managerial decisions involve some degree of choice and they are, therefore, essential economic in nature. Managerial Economics is, therefore, offered as one of the foundation courses for graduating in any functional area of management.

1.4 MANAGERIAL ECONOMICS

You should now be in a position to form some idea about the intent and content of managerial economics as a discipline. Managerial Economics concentrates on the decision process, decision model and decision variables at the firm level. The firm is viewed as a microeconomic unit located within an industry which exists in the context of a given socio-economic environment of business. The executives and the functional managers are stationed within the firm; these managers, either individually or jointly think, make and execute decisions. Their operation reflects the behaviour and culture of the firm.

Managerial Economics is concerned with the economic behaviour of the firm. To start with, it assumes that the firm maximises profit. Profit is defined as the difference between revenue and costs. The flow of revenue is determined by the demand conditions in the market, whereas the costs are influenced by the supply condition. Demand and supply interact with each other to determine prices — commodity prices in the product market and input prices in the factor market. Much would, of course, depend on the nature of the market structure — perfect or imperfect, free or regulated, buyer's or seller's etc. The firm or the managers on behalf of the firm placed in the contexts of a market environment, decide its economic strategy (in view of its objectives and constraints) and tactics. The tactical decisions are reflected in the course of operational decision variables like price, output, input etc. Such operation through economic decision variables affects the firm's level of profit. The firm can then evaluate its performance in terms of the rate of return on investment — intended and achieved. The firm will thus be in a position to estimate the element of risk and uncertainty it is subject to. In its decision-making process, tl firm's strategy is to minimise such risks and uncertainties through forecasting and forward planning.

The subject matter of Managerial Economics can now be lined up as follows:

Frame work of Concepts and Techniques	Topics of Managerial Economics
• Demand Analysis	• Demand Decisions
• Production and cost analysis	• Input-output Decisions
• Market structure analysis	• Price output Decisions
• Firms behaviour analysis	• Profit maximisation and alternative hypotheses
• Project Appraisal	• Investment Decisions
• Risk and uncertainty analysis	• Economic forecasting and planning

You may appreciate that analytical rigour is very important in our subject of Managerial Economics. At each stage of analysis of any economic decision variable, you should, threfore, be careful about the assumptions, concepts, techniques, logic of reasoning, conclusions and their applicability in the real world business situation.

Activity 2

Run down the list of topics noted above:
Think about the organic relations that exist among them. Can you work out an integrated structure of Managerial Economics? (Kindly note that you do not have to write out an answer. This is just a mental exercise for you).

1.5 RELATED DISCIPLINES

By its nature, Managerial Economics draws heavily on other disciplines like Economics, Accountancy, Mathematics, Statistics, Operations Research, Psychology and Organisational Behaviour. Let us state briefly the contributions made by each of these disciplines to Managerial Economics.

i) Economics and Econometrics: We have already stated that economic theory contributes concepts, precepts, tools and techniques of analysis to Managerial Economics. In the preceding section on Economic Analysis you have been told about the form of analysis which we would like to borrow from economic theory. The additional point which you may note now is that Economics also tells us the art of constructing "models" - a system or relation among variables. In Managerial Economics, we use various types of models: schematic models (diagrams), analog models (flow charts) and mathematical models like Econometric models and stochastic models. Most of these models are rooted in economic logic. Econometrics in particular is a combination of economic logic and statistical techniques. In our course; the empirically estimated functions that we will be using are basically econometric estimates.

ii) Mathematics and Statistics. Most of the deicision models that we have mentioned above are formulated in terms of mathematical symbols and relations. Mathematics gives up precision which is important in deicisions. Decisions have to be precise. To the extent that the economic decisions are measurable and quantifiable, the use of mathematics is welcome. However, as managers, you must remember that symbols and numbers alone do not suffice; what you need additionally is the logic of reasoning behind those numbers and symbols. Scientific decisions arrived through the use of mathematical tools and statistical techniques should have an intuitive appeal and logical support, otherwise the manager will be in trouble. Regression, for example, may help the manager to forecast his sales based on past record, but the manager must make sure that the user of his product has not undergone any change in his tastes and preferences. That way, an econometric model may prove a better aid than pure mathematics and statistics.

iii) Operations Research (OR): OR was developed during the inter-war period and it is inter-disciplinary in character — a combined effort of scientists, economists, mathematicians, statisticians in formulating models to solve specific problems of allocation, transportation, inventory building etc. There are linear programming as well as goal programming models which are very useful for managerial decisions. Managerial Economics, therefore, uses some of these models. However, in our course material here, we will make only passing reference to linear programming techniques, because elsewhere you will have a detailed exposure to those programming techniques. In this course, we will confine ourselves mostly to the calculus (Lagrange Multiplier) techniques of optimisation in specific situations. This is an alternative to linear programming methods of optimisation.

iv) Accountancy: Decision-making requires data and a good deal of it is provided by the accounting system. The information on cost, inventories, receivables, revenue and profits is provided by the accountant. Of course, the economist may modify these data to arrive at optimal economic decisions. There are a number of concepts and techniques which are common to the accountant and the economist. In managerial economics, we make use of both, keeping the distinction between accountancy and economics in the forefront. For example, we will be referring to the 'accounting profit' and 'economic profit' or 'break-even output' of the accountant and the 'shut-down output' of the economist. Accountancy does provide the data support, but in the tradition of economics, managerial economics stresses the logic of reasonig behind those data and uses the data under a set of assumptions.

v) Psychology and Organisation Behaviour (OB): Managerial economics analyses the individual behaviour of microeconomic units like a buyer, seller, an investor, a worker and an employer. Psychology makes a direct contribution towards understanuing the behavioural aspects like attitude and motivation of any individual decisionmaking unit. In fact, there is a new branch of discipline called Psychological Economics which analyses subjects like buyer's behaviour directly

Suppose, for one reason or the other, you are unable to undertake this activity. **Think about** the "constraints" you had as an explanation of your own "economic rationality".

....

....

....

....

....

....

....

....

1.6 ECONOMIC RATIONALITY

You may often hear that as decision-maker, you are asked to act **rationally**. Economics itself assumes rational behaviour on the part of its subjects like consumer, producer and seller. To act rationally is to act objectively, keeping in view the ends and the means, the objectives and constraints. In the decision-making context, a rational behaviour implies the following:

First, all possible courses of action are known to you. **Secondly**, you are able to separate all such courses of action into two categories-feasible and infeasible. **Thirdly**, based on available information, you are in a position to assess the consequences of following from the choice of a particular feasible course of action out of given set of alternatives. **Fourthly**, you can rank the alternatives in terms of priorities. **Lastly**, you choose that alternative (course of action) which occupies the highest position in the order of priority.

Thus economic rationality is a precondition for attaining optimality in a given decision environment. We may take a couple of examples to illustrate it. For example, rationality on the part of a selling firm is to maximise profit or sales revenue. The rationality on the part of an investor is to maximise return. The rationality on the part of a manufacturing unit is to maximise production or to reduce costs. The rationality on the part of an investor is to maximise return. The rationality on the part of a consumer is to maximise satisfaction. You may be tempted to think that rationality thus implies **only** statement of **objectives** in clear terms. That is not right. By way of rationality, you have to take care of **constraints** as well. In fact your judgement about feasibility of a particular course of action is based on your evaluation of objectives and constraints together. To be exact, feasibility, consistency and optimality of a decision — all must be considered to define the dimension of economic rationality.

Activity 5

1 Under what specific conditions (i.e., objectives and constriants) would you consider the following as a piece of **rational economic behaviour** ?

a) A firm is building up its inventory:

....

....

b) A firm is offering heavy discount (in clearance) sales:

....

....

c) The Government announces a wage freeze:

....

....

d) The DTC (Delhi Transport Corporation) workers' union goes on indefinite strike:

....

....

e) A firm employs only "sons of the soil":

....

....

2 Work out an example of "irrational" economic behaviour.

....

....

1.7 METHODOLOGY OF MANAGERIAL ECONOMICS

Managerial Economics is supposed to help us in analysing the rational economic behaviour of the decision-making unit. Economic analysis has its own methods. An economist often starts with a set of **assumptions**. Assumptions are hypothesised premises, some of those assumptions may be too ideal to be real. However, the ideal set of assumptions helps logical construction of analysis. Analysis is followed by deduction of inferences. Such inference, positive or normative, may hold good in reality, provided they have been based on realistic assumptions. Otherwise, the initial assumptions, are to be relaxed and in that light inferences have to be re-worked.

The methodology of economic analysis can be exampled. We may proceed to analyse the purchase behaviour of a household unit, assuming that the household's family size; income level and its tastes and preferences are held constant; and we observe that the household purchases less of vegetables and fruits as and when the price of these items rises. This inference needs to be changed if we relax some of initial assumptions of "other things remaining equal". For example, suppose the household has a new source of income such that its purchasing capacity goes up. Now, this household may continue to buy the same basket of vegetables and fruits even if their prices have gone up.

It is through the relaxation of unrealistic assumptions that a realistic analysis has to be built up. In fact, the criterion of a sound economic analysis is often stated on: **Fewer the assumptions, better it is.** The argument is most assumptions are theoretical constructions, having very negligible empirical relevance.

The statement of assumptions and the construction of a theoretical framework thereby is termed as **model building**. Economic analysis runs in terms of these models. A model can be viewed as an approximation of reality. Sometimes a situation already experienced or what we call case studies may be used to highlight deicision—making context and concepts—complexity of contexts and variety of concepts. Such "case studies" may be regarded as some form of "models" in economic sense; because every case may be unique by itself; it is historic individual; something which has been experienced in the past may not get repeated in exact form in the present or the future; yet the past experience is a valuable reference; it offers lessons or guidelines for the decision maker. In Managerial Economics, we may use both theoretical models and empirical case studies; both may develop the analytical lent of mind of our decision maker. In a subsequent unit, we will have something more to add on these models or cases—their uses and abuses as techniques of analysis.

1.8 SUMMARY AND CONCLUSION

In this unit, we introduced you to the scope and methodology of Managerial Economics. Managerial Economics the application of economic analysis toevaluate business decisions. Economic analysis and related disciplines provide concepts, precepts, tools and techniques to examine the economic decision variables like demand, supply, price, costs and profits. There are various types of economic analysis (theory) and various types of business decisions (practice). In managerial economics, we attempt an integration between economic theory and business practice. This is because, business decision-making itself is an economic problem involving an act of choice.

1.9 ADDITIONAL READINGS

Brigham, E.F. & Pappas, J.L., (2nd edition, 1976), Managerial Economics, The Dryden Press: Illinois, USA Ch. 1.

Adhikary, M., (3rd edition, 1987) Managerial Economics, Khosla Educational Publishers: Delhi, Ch.1.

1.10 SELF-ASSESSMENT TEST

What does economic theory contribute to Managerial Economics?

UNIT 2 FUNDAMENTAL CONCEPTS

Structure

2.0 INTRODUCTION

Managerial Economics, as we have emphasised in the preceding unit, is concept based and technique-oriented. We make use of concepts and techniques basically from Economics and related disciplines. The use of concepts and techniques helps us in developing the analytical rigour of the subject. They aid logical reasoning and precise thinking, typical of a professional manager. In today's business world of complexity, scientific management begin with conceptualisation of the decision problem. This is followed up derivation of decision rules, principles and results through the use of sophisticated tools and techniques. By way of introduction to the subject of managerial economics, we would, therefore, like you to take stock of some of the fundamental **concepts** (in this unit) and basic **techniques** (in the next unit).

In what follows, we have selected at random a few concepts, which are fundamental to any economic analysis, including the analysis of business problems. Though selected at random, these concepts have been ordered in a sequence such that together they may project the conceptual basis and frame work of our discipline.

2.1 OBJECTIVES

The purpose of this unit is to:

- **define** a few concepts, basic to economic analysis
- **exemplify** those concepts with the help of table and diagrams
- **comment** on their applicability in real world business situation
- **point out** the decision principles underlying those concepts.

2.2 SCARCITY

This is one of the fundamental concepts of Economics relevant for Management. The essence of any economic problem, micro or macro, is the scarcity of resources. The managers who decide on behalf of the corporate unit or the national economy always face the economic problem of scarcity of one kind or the other. You may take a couple of examples to illustrate this. As a production manager, you may be facing scarcity of good quality materials or skilled technicians. As a marketing manager, you may be encountering scarcity of sales force at your command. As a finance minister, you may be facing scarcity of funds necessary for expansion or renovation programme. As the finance manager of the country, designated as an important official in the Ministry, your basic problem when you prepare the budget every year, is to find enough revenue resources to finance the necessary expenditure on plans and programmes. Thus, scarcity is a universal phenomenon.

Let us, therefore, attempt a technical definition of the term"scarcity". In economic terms, scarcity may be defined as "excess demand". Any time for any thing, demand (requirement) exceeds supply (availability), that thing is said to be scarce. You may note that scarcity is a relative concept—demand in relation to supply determines the element of scarcity. It is this scarcity which lies at the root of any problem which requires managerial attention. For example unemployment is essentially the scarcity of jobs. Inflation is essentially the scarcity of goods. Unsold stock of inventory is essentially the scarcity of consumers. Underutilised capacity at the plant level may be primarily due to scarcity of power or other support facilities. Had these scarcities not been there, there would have been no managerial economic problem. It is because of this scarcity that a manager has to decide on optimum allocation of scarce resources of men, materials, machines, money, time, energy and what have you.

Activity 1

a) You must have noted that 'excess demand' means that demand exceeds supply. Similarly, 'excess supply' should mean that supply exceeds demand. It follows that 'positive excess demand' means 'negative excess supply'. Could you now redefine your concept of "scarcity"?

....

....

....

b) You must have also observed that excess demand (for labour) pushes the price (wage rate) up. Would you, therefore, argue that "scarcity" is reflected in terms of price escalation? Give some examples.

....

....

....

c) Of the three items X, Y and Z, which one is relatively more scarce than the other two? Why?

....

....

d) Zero excess demand means zero excess supply. True or False?

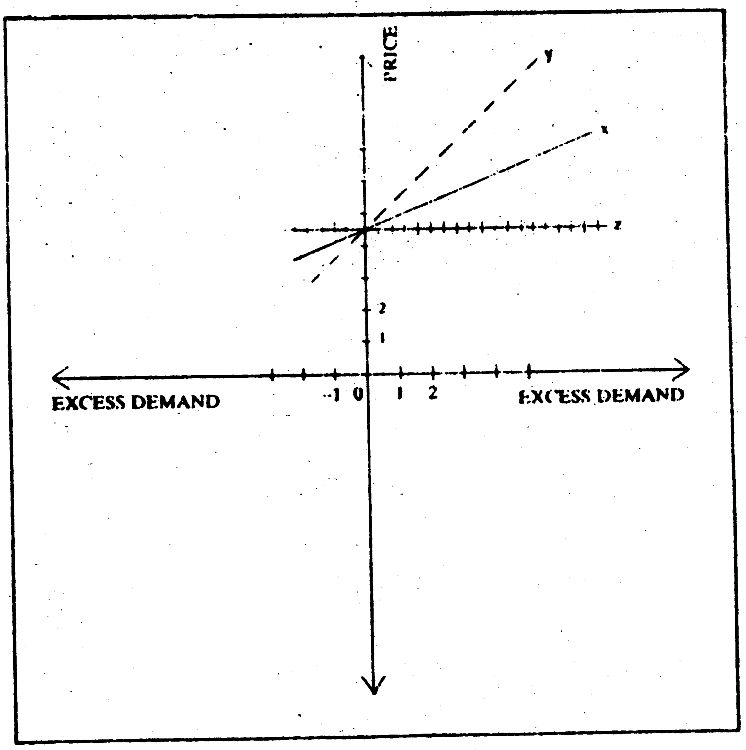
....

....

e) At what price, there is no scarcity of either X or Y or Z?

....

....



2.3 MARGINALISM

If resources at your disposal are scarce, then as managers, you have to be careful about the utilisation of each and every additional unit of resources (inputs). In order to decide whether to use an additional manhour or machine hour or not, you need to know what is the additional output expected therefrom. Similarly, a decision about additional investment has to be taken in view of the additional return from that investment. Economists use the term "marginal" for all such additional magnitudes of output or return. For example, they often use terms like marginal output of labour, marginal output of machine, marginal return on investment, marginal revenue of output sold, marginal costs of production, marginal utility of consumption and so on. Before you attempt to use such concepts of marginalism, be clear about the following points:

a) **The nature of relationship between the variables is to be clearly stated.** If sales depend on advertisement, then you can talk of 'marginal sales of advertisement'; but if advertisement budget depends on the flow of sales revenue, then you should talk of 'marginal advertisement of sales'. In other words, the dependent variable must be clearly distinguished from the independent one.

b) The independent variable is to be changed by just one unit at a time to work out the impact on the dependent variable. The impact of an additional hour of labour is figured out in terms of 'marginal output of labour', just as the return from an additional one crore of investment is stated in terms of 'marginal return on investment'.

c) The marginalism is not be confused with the concept of average. The average product of labour, for example, is the ratio of total product to total labour, whereas the 'marginal product of labour' is the ratio of change in product to one-unit change in labour. The following table may illustrate the relationship between average and marginal concepts.

No. of labourers (L)	Total output (Q)	Average output (Q/L)	Marginal output (Δ Q/ΔL)
1	10	10	—
2	18	9	8
3	24	8	6
4	32	8	8
5	45	9	13
6	60	10	15

Δ is a symbol which denotes change.

You may observe that as the average output falls, the marginal output falls faster than the average output such that the marginal is lower than the average. When the average output remains the same, the marginal equals the average. When the average output increases, the marginal output increases faster than the average output such that the marginal exceeds the average.

At this stage,we may state the marginal principles of the economists:

- Each factor labour may be paid wages according to its margial product $\Rightarrow W = MP_L$
- Each commodity (x) may be priced according to its marginal utility $\Rightarrow P_X = MU_X$

The basic idea is that satisfaction must balance sacrifice.

In case, these principles are not followed, the equilibrium position would be disturbed. In subsequent units, we will examine such propositions.

The marginal concept which measures the rate of change in the dependent variable is very useful concept in Economics. However, in the real world business situation, it is difficult to apply the concept of marginalism. The problem is that the independent variable may be subject to "chunk changes" rather than "unit change". The labourer that the contractor uses on a turnkey project may be changed not by one, but by tens and hundreds. Similarly, the output may change because of a change in process, pattern or a combination of factors which may not always be measured in unit terms. In such cases, the concept of 'marginalism' may be replaced by that of 'incrementalism'. The additional revenue which the firm earns because of say, computerisation of market information, may be termed as "incremental revenue". Similarly the additional cost of installing computer facilities may be termed as "incremental costs". In fact, incrementalism is more general whereas marginalism is more specific. All marginal concepts are incremental concepts; but all incremental concepts need not be confined to marginal concepts alone.

Activity 2

a) Take a graph paper. Plot the table and get the average and marginal output of labour curves. Extend the table with imaginary data and show that the following statements hold true:

i) Total output is maximum, when the marginal output is zero.

ii) Total output falls, when the marginal output is negative. You may note that when labour has zero or negative marginal product, it is said to be "disguisedly unemployed".

iii) Marginal output may be falling when the average output is rising.

b) Review your understanding of:

i) Marginal utility of money spent.

....

....

....

ii) Marginal revenue of advertisement.

....

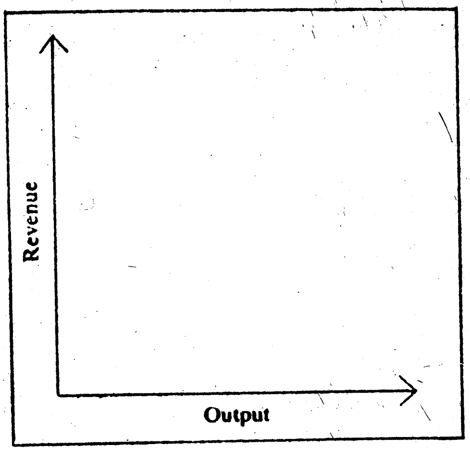
....

iii) Marginal return on investment.

....

....

c) i) Distinguish between 'marginalism' and 'incrementalism' with the help of a diagram.



ii) To apply marginalism, we need a .... function; but if we have got.......................... data, we must use incremental concepts.

Choose the appropriate words from the list below:

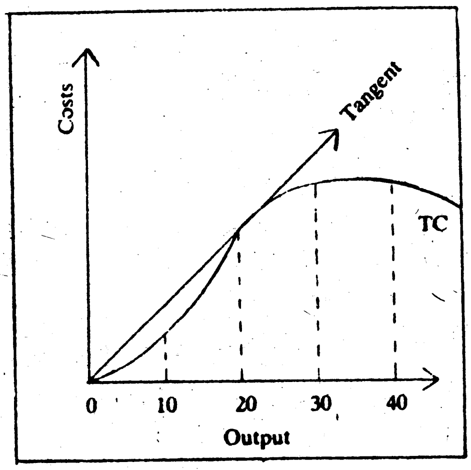
- continuous
- discontinuous
- discrete
- concrete

d) i) For what level of output, the average cost = the marginal cost of production?

....

ii) For what level of output the marginal costs are zero?

....



iii) Can incremental costs be the same as marginal costs?

....

....

2.4 EQUI-MARGINALISM

While developing the concept of marginalism, you must have noticed that we have assumed a single variable function, such as

- Revenue depends on output.
- Costs depend on output.

The implicit decision rule in such cases may be stated as:

- If you want to maximise revenue, then do not sell output beyond the point where marginal revenue is zero.

OR

- If you want to mximise the total product, then do not employ a factor beyond the point where its marginal product is zero.

OR

- If you want to minimise costs, then produce till your average costs equal marginal costs.

Such decision rules are termed as "Absolute Activity Level Principle" (Baumol). At a time, an input has only one use of producing an output—or output has got only one use—either generating revenue or involving costs, but not both.

Now suppose, output has got both uses viz. it involves costs, at the same time, it is capable of generating revenue. In this case, the manager's decision about extra volume or extra batch of production must be based on comparison of both marginal revenue and marginal costs. If marginal revenues exceed marginal costs, additional production should be encouraged. If marginal costs exceed marginal revenue, it is uneconomic to go in for extra production. Thus the optimum level of output should be decided upon the point where marginal revenues just cover the marginal costs. This decision rule is termed as "Relative Activity Level Principle" (Baumol), and this is what the economists call "equi-marginalism". In subsequent units, we will be deriving this principle or concept more rigorously. For the time being, for managerial use, you may find the following as a handy reference:

Nature of the Unit	Equi-Marginal Principle
Multi-market seller	$MR_1 = MR_2 = MR_3 = \ldots = MR_n$
Multi-plant monopolist	$MC_1 = MC_2 = MC_3 = \ldots = MC_n$
Multi-factor employer	$MP_1 = MP_2 = MP_3 = \ldots = MP_n$
Multi-product firm	$M\pi_1 = M\pi_2 = M\pi_3 = \ldots = M\pi_n$
Multi-commodity consumer	$MU_1 = MU_2 = MU_3 = \ldots = MU_n$

You may note that:

MR_i = Marginal revenues ; i=1 ....n markets

MC_i = Marginal costs ; i=1 ....n plants

MP_i = Marginal products ; i=1 .....n factors

$M\pi_i$ = Marginal profits ; i=1 ....n products

MU_i = Marginal utilities ; i=1 ....n commodities

Just take one of these for intuitive explanation. Suppose your firm is a multi-market seller like HMT which sells watches in both home market and foreign market. If marginal revenue from a watch sold in home market exceeds the marginal revenue from a watch sold in the foreign market, as a seller intending to maximise revenue, you are likely to reduce your sale in foreign market in order to push up sales in home market. And this process will continue till the marginal revenues, from both markets are equalised. Of course, whether it always happens this way or not would depend very much on the objectives and constraints of the firm; and economists often state them by way of assumptions.

Let us assume the case of a multi-commodity consumer who is purchasing successive units of X, Y and Z. Each unit costs the same and the consumer is determined to have a combination including all three items. His budget constraint is such that he cannot buy more than six units in all. Lastly but not the least, he is subject to **diminishing marginal utility** i.e. as he has more of an item, he wants to have less of it. The following table illustrates his case:

Units	Marginal Utilities*		
	Item X	Item Y	Item Z
1st	10	9	8
2nd	9	8	7
3rd	8	7	6
4th	7	6	5
5th	6	5	4
6th	5	4	3

* The basic assumption is that marginal utilities are cardinally measurable, say in terms of Rs. and that marginal utility is a monotonically decreasing function of consumption.

You may work out that our utility maximising consumer will end up with a purchase of 3X + 2Y + 1Z because that combination satisfies equi-marginalism:

$MU_x = MU_y = MU_z = 8$

No other combination would give him so much utilities as he gets now (52).

Finally, you may note that in the real world, it may not be always possible to have such a neat simple problem. In particular, you may not have data for each successive units; in that case, you may have to replace "equi-marginalism" by the concept of "equi-incrementalism". But your decision rule or optimising principle will remain the same.

Activity 3

a) Name a couple of firms to suggest real world examples of the following:

Nature of the unit	Real world examples
Multi-market seller	
Multi-plant monopolist	
Multi-product firm	
Multi-factor employer	

b) Distinguish between Absolute and Relative Activity Level Principles.

....

....

....

....

c) Suppose a seller has two markets to serve. The demand schedules in them are reproduced in the Table. Suppose that he has 1400 units to sell and maximise profit threby. What prices will he set in the two markets?

Market A		Market B	
Price	Qty.	Price	Qty
50	400	60	600
40	600	50	800
30	900	40	1100
20	1000	34	1400

Apply equi-incrementalism principle to get your answer. Could you have applied equi-marginalism?

Hint: First get total revenue in each market by multiplying price with quantity.

....

....

....

....

2.5 OPPORTUNITY COSTS

You have learnt the concept of marginal costs and incremental costs. Let us now get another cost concept from Economics, viz. opportunity costs. Here is a very general concept, at the same time a useful one, for decision purpose. The opportunity costs are the "costs of sacrificed alternatives". Whenever the manager takes or makes a decision, he chooses on course of action, sacrificing the other alternative courses. We can, therefore, evaluate the one which is chosen in terms of the other (next best) alternative which is sacrificed. This is the method to compute the opportunity costs of a decision. Let us take a couple of examples to illustrate:

i) You have got an hour at your disposal which you can devote either to reading study material or to watching cricket match on the TV. The opportunity cost of watching the cricket match is the study material which you could have read had you not watched the match. In other words, the opportunity cost of leisure (labour) is the labour (leisure) that one sacrifices.

ii) A machine can produce either X or Y. The opportunity cost of producing a given quantity of X is, therefore, the quantity of Y which it would have produced. If that machine can produce 10 units of X or 20 units of Y, then the opportunity cost of 1X is 2Y.

iii) Suppose we have no information about quantities produced, but have information about their prices. In this case, the opportunity costs can be computed in terms of the ratio of their respective prices, say $\frac{P_x}{P_y}$

iv) The opportunity cost of holding Rs. 1000 as cash in hand for one year is the 10% rate of interest, which could have been earned, had it been kept in the form of fixed deposits in the bank.

You may think of many more similar examples. However, you must note that:

- all decisions which involve choice must involve opportunity costs calculation,
- the opportunity costs may be either real or monetary either implicit or explicit, either non-quantifiable or quantifiable.

The **economists** use many different concepts of "trade-off" (such as Indifference curves, Iso-quants, Phillips) curve etc., which are all based on opportunity cost reasoning. In the same way, the **managers** confront many different decision areas where opportunity cost concept is directly applicable. Make or Buy decision, Breakdown or Preventive Maintenance of machines, Replacement or New Investment decision, Direct recruitment from outside or Departmental promotion from within to man a post and the like. The **accountant** never considers the opportunity costs; he considers only the explicit costs. That is why the "accounting profit" is simply the revenue minus recorded costs, but the "economic profit" is revenue minus explicit and implicit costs; i.e., the "accounting profit" minus opportunity costs.

You may appreciate that for optimal allocation of scarce resources the manager should consider the opportunity costs of using resources, human or non-human, in a given activity; and his decision principle should be minimisation of opportunity costs, given his objectives and constraints.

Activity 4

a) Can opportunity costs, of a given sum of money be ever zero?

....

b) If there is unutilised plant capacity (or if the workers do not have sufficient work to engage them for requisite number of hours), the opportunity cost of unutilised plant capacity (or workers) is zero. True or False? Why?

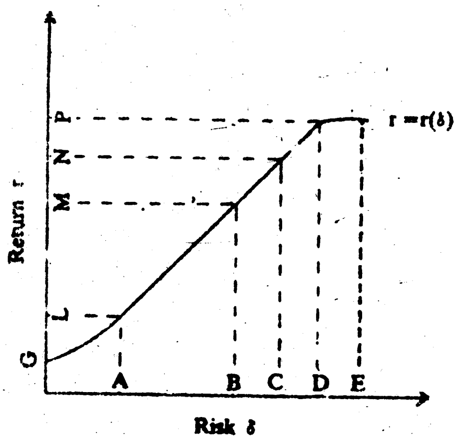
....

....

....

....

c) Here is a risk-return trade off curve. Use the opportunity cost concept to interpret the diagram.



d) Briefly state the "social opportunity costs" of the following project/programme undertaken by our Government.

i) Rs. 2000 crore spent on drought relief measures:

ii) Mathura Refinery:

iii) Clean Ganga Project:

iv) ASIAD 1982 held in Delhi:

v) Would you agree with the statement that the opportunity costs for one is the "opportunity gain" for other?

Activity 5

When you go to your office/factory next time, observe the work/process /pattern/product. Can you find some examples of "opportunity costs"? When you come back home, organise a discussion session with your family members/friends on the subject: **Uses and Abuses of the Concept of opportunity costs.**

2.6 TIME PERSPECTIVE

Decision-making is the task of coordination along the time scale—past, present, and future. Whenever a manager confronts a decision environment he must analyse his present problem with reference to **past** data of facts, figures and observation in order to arrive at a decision, contemplating clearly its **future** implications in terms of actions and reactions likely threupon. Thus the time dimensions of a managerial decision is very important.

The economists often consider the time element in terms of concepts like temporary-run, short-run and long-run. The economists state that in the **temporary-run** the supply of output is totally fixed. By contrast, in the **short-run** the supply can be changed slightly by altering the factor proportion; a fixed factor like plant and machinery can be used more intensively. In the **long-run**, all factors are variable and therefore, the output level can be adjusted freely.

The manager can interpret these economic concepts in a different way. For temporary-run, or short-run, the manager faces a lot of constraints, but for the long-run, there are no constraints, of course, such a long-run is too ideal to be real; in that long-fun, "we are all dead" (Keynes). For a practising manager, the short-run is the present (immediate) period, whereas the long-run is the distant future (remote) period. The manager must calculate the opportunity cost of his decision, if he has to choose between the present and the future. His decision principle is that he must take care of both time periods. A manager cannot afford to have time persepective that is too short. For example, he may set a high price for his product today, but then he should be prepared to face declining sales. Today the advertisement expenditure may inflate the costs, but tomorrow it may increase the revenue flow. Similarly, at present, the management may ignore labour welfare to reduce costs, but in future this may deteriorate industrial relation climate with adverse effect on productivity and profitability. Thus it is important for the manager to take a short and a long view of his decisions.

2.7 DISCOUNTING

This brings us to a very important concept of Managerial Economics. Discounting is both a concept and a technieque, which is borrowed from Accountancy. For its explanation, let us readopt the preceding two concepts together—opportunity costs and time perspective.

Consider the case of a seller. He has to decide between the immediate cash payment of Rs. 1000 by his customer or the future payment of say Rs. 1100 at the end of one year from now. Human nature is such that there is always a time preference in favour of the present. For the seller, it is better to get Rs. 1000 now and put the same in the bank at the 10% rate of interest per annum and then realise Rs. 1100 thereby. Should we say that the present value of the future sum of Rs. 1100 is just Rs. 1000? The question is: how have we arrived at this? It is simple.

An investment of Rs. 1000 (say P, for Principal) yields Rs. 1100 (say A, for Annuity when the prevailing rate of interest (r) is, say, 10%

$$A = P + r.\,P = P(1+r)$$

or $P = \frac{A}{1+r} \Rightarrow$ Rs. 1000 = Rs. $\frac{1000}{(1+10\%)}$

In the same way, you may work out that at the end of the—

second year, $P = \frac{A_2}{(1+r)^2}$

third year, $P = \frac{A_3}{(1+r)^3}$

second year, $P = \frac{A_n}{(1+r)^n}$

Thus if an investment of sum yields a series of return A_i through i period; i = 1 ....n, then in order to calculate its present value, we need to discount $\sum_{i=1}^{n} A_i$ with the help of $(1+r)^i$,

$$P = \sum_{i=1}^{n} A_i/(1+r)^i$$

$$= \frac{A_1}{(1+r)} + \frac{A_2}{(1+r)^2} + \frac{A_3}{(1+r)^3} + \dots + \frac{A_n}{(1+r)^n}$$

You may note that longer the period, larger is the discount factor, $(1+r)^n$ excess $(1+r)^{n-1}$, because nth period is more distant than the (n-1)th period. Heavier discounting for the distant period makes sense because future is uncertain; the distant future involves incalculabe risk. Discounting enables risk-hedging, particularly in the context of investment decisions where the return on investment is spread over a number of years in future. Thus the present value of future return can be estimated by discounting it with the opportunity costs of the safe rate of interest as in our example.

Discounting principle has application in areas other than investment decision. For example, some economists restate their marginal principle of wage determination in the labour market as follows:

"Wages are equal to the discounted value of marginal product of labour". This is because the wage rate as a contractual payment is fixed at the beginning of the period, but the actual productivity of the labour who has been employed on that wage rate is known only at the end of the period. The employer, therefore, must be discounting the future marginal productivity of his worker to settle his present wage rate.

Activity 6

a) Discounting is a technique of computing the opportunity costs of time preference. Explain.

....

....

....

....

b) What does an investor discount? Why does he discount? How does he discount?

....

....

....

....

....

c) Suppose, a firm is going to receive Rs. 10,000 per annum for the coming five years. The firm can earn a safe rate of interest of 10% per annum if it had all the money today. Somebody has calculate that the present value of Rs. 50,000 which the firm is to receive in course of five years is just Rs. 37,907.80. Make your own calculation and show that this estimate is correct. (You may use the PV table is you like: you may certainly use a calculator, if you have one).

....

....

....

....

....

....

....

2.8 RISK AND UNCERTAINTY

Most of the business decisions involve future revenues and costs, and it is seldom possible to predict future with complete accuracy. Future involves changes; there is no guarantee that the present will be repeated in future. The environment changes in course of time.

Changes are not homogeneous. The changes may be **known** or **unknown**. The outcome of known changes may be either **definite** or **indefinite**. The definite outcome assocsiated with known changes define the decision environment of "certainty". The indefinite nature of outcome associated with known changes involves "risk". Such risks can be estimated in terms of acturial probability of events; and accordingly such risks can be insured. But if the changes are unknown, their outcome is indefinite,and therefore, the risk element is incalculable and immesurable. This is the world of "uncertainty" which cannot be insured. Thus a distiniction exists between "risk" and "uncertainty", though these two terms are often used as synonyms in common parlance. You may now concentrate on the schematic presentation below to conceptualise the distinction:

Decision Environment

⇓

Changes

Known | Unknown

Definite ⇓ Certainty

Indefinite ⇓ Risky ($o < \delta > \infty$)

Unknown ⇓ Uncertainty ($\delta \approx \infty$)

δ is the measure of risk element.

The economists go a step further and suggest that interest rate is risk premium, but profit is a reward for uncertainty.

Introduction of risk and uncertainty in the analysis of decision- making is done with the help of statistical concept of probability. Each management, while making policy decision would like to know the likelihood (probability) of its going right or wrong in various estimates about future revenues and costs. An awareness of degree of risk and uncertainty helps the process of decision-making. Of course, management can have different attitudes towards calculable risk and uncertainty (non-calculable risk). Some are risk-avoiders/averters; some are risk-dodgers, some are risk lovers. Some can estimate market risks (due to new product, or new price) and some tend to guesstimate non-market risks (due to new product, or new price) and some tend to guesstimate non-market risks (due to sudden change in Government policy). Risk is often defined as a measure of "standard deviation" between expected returns (costs) and actual returns (costs). We hope to discuss at a later stage the techniques of dealing with risk and uncertainty decision-making in areas like finance and marketing.

Activity 7

a) Based on your day-to-day experience or observation, quote some specific examples of the kind of risks faced by:

i) Production Manager:

....

....

ii) Personnel Manager:

....

....

iii) Finance Manager:

....

....

iv) Marketing Manager:

....

....

v) Purchase Manager:

....

....

vi) Inventory Manager:

....

....

b) Give some examples of "risk diversification" i.e. producing a whole range of products rath[er] than a single product to minimise risk.

....

....

....

....

c) If there is no risk and uncertainty, the manager need not consider the time perspective [of] his decision. Do you agree?

....

....

....

Activity 8

a) Contact an insurance agent (working for any insurance company like LIC, GIC). Try t[o] know how these companies estimate the risk element.

b) Visit a nearby STOCK EXCHANGE. Observe the behaviour of the brokers, and form some idea about their attitude towards risk and uncertainty while dealing in shares.

2.9 PROFITS

Profits of a business enterprise means revenue minus costs. Revenue (TR) depends on physical volume of sales (Q) and the price at which the output is sold (P). On the other hand, the total costs (TC) depend on the volume of factors employed (F) and the average factor price (C). Thus profits (π) can be stated as

$$\begin{aligned} \pi &= TR - TC \\ &= P.Q - C.F \end{aligned}$$

What should be included in costs in order to arrive at profits is, however a controversial subject. In this context, the distinction is made between business (accounting) profit and economic profit

The businessman's concept of profit is produced in the form of a chart below:

Sales Revenue —Direct costs (on labour, materials etc.)
= Contribution — Fixed costs (overheads)
= Operating Margin (E B I T: Earnings before interest & taxes —(Interest + Taxes)
= Net Profit — Dividends
= Retained Earnings

The businessman calculates his return on investment (ROI) by computing profit as a percentage on investment; he uses various concepts of profit such as gross profit, net profit, profit after tax etc. depending upon both accounting convention and accounting convenience.

The concept of economic profit, as we have indicated earlier, is narrower than that of accounting profit, because the opportunity costs have to be deducted from accounting profit to get at economic profit. The costs of employing his "own factors" is the businessman's opportunity costs. Suppose a businessman has employed his 'own' capital, 'own' building and 'own' labour in his own business. The interests, rents and wages which he could have earned, had he not employed his 'own' factors in his "own business" but in somebody else's business, have to be regarded as the opprotunity costs, which when considered, will show that economic profit is less than accounting profit on his books. Suppose a mechanic makes an accounting profit of Rs. 500 per month. He has his own repair shop in the garage of his house. To start this business, he has also invested Rs. 12,000 as read capital. Had he not started this business, he would have earned say Rs. 100 per month as a hired labourer; he would have also earned Rs. 1,200 as interest per annum (i.e Rs. 100 per month) on his capital. Additionally his garage would have fetched a rent of Rs. 200 per month. All total, he would therefore earned Rs. 400 (RS. 100 + Rs. 100 + Rs. 200) per month; this is the opportunity costs of running his "own business". Therefore, his economic profit is Rs. 100 = (Rs. 500 -Rs. 400) only. To the extent, it is difficult to quantify and measure these opportunity costs, it becomes a problem to estimate economic profit. To overcome this conceptual as well as measurement difficulties, most of the economists assume zero opportunity costs such that accounting profits and economic profits do not differ. This is indeed a simplifying assumption.

Lately some behavioural models have been developed by the managerial economists. In these models, the concept of profits has been given a practical orientation. There are now different concepts of profit: maximum profit, critical minimum (target) profit, actual profit, reported profit and so on. We hope to develop these concepts at a later stage.

For the time being, we would like to continue with the pure microeconomic concepts of profit—supernormal profit (TR>TC), normal profit (TR = TC) and subnormal profit (TR<TC). The economists normal profit is no-profit—no-loss situation; and subnormal profits is actually a situation of loss. Our TC, by way of assumption, includes all relevant, measurable opportunity costs. Eventually we have to debate whether the firm really maximises profit or not. At the beginning, we will follow the conventional economic theory of the firm which assumed profit maximisation as the objective.

Activity 9

Take the balance sheet of a Company and make a distinction between "business profit" and "economic profit" by inputting opportunity costs to equity holder's funds. Provide the rationale of choosing a particular opportunity cost.

....

....

....

....

....

....

....

....

Activity 10

a) Construct an example with all imaginary data to show that 'economic profit' is negative even though the "accounting profit" is positive.

....

....

....

....

....

b) Karl Marx regarded profit as a "surplus value". How would you interpret the Marxian view?

....

....

....

....

....

....

....

2.10 SUMMARY AND CONCLUSION

In the present unit, you have been exposed to an overiew of some basic concepts like scarcity, marginalism, incrementalism, equi-marginalism time perspective, opportunity costs, discounting, risk and uncertainty etc. While explaining these fundamental concepts, we have also exposed you to some elementary economic tools like tables and diagrams, function and relations. In the next unit you would be exposed to a few more sophisticated tools and techniques of economic analysis, such techniques are used towards optimisation in a given decision environment.

2.11 ADDITIONAL READINGS

Hailstones, Thomos & Rothwell, John, C, (1985) Managerial Economics, Prentice Hall International: New Delhi, Ch. I.

Davis, J.R & Charg, Simen, (1986) Manageriai Economics, Prentice Hall International: New Delhi, Ch. I.

2.12 SELF-ASSESSMENT TEXT

1 Explain each of the following concepts and state clearly the decision principles associated with each:

a) Incrementalism
b) Equi-marginalism
c) Opportunity costs
d) Discounting

2 Comment on each of the following statements:

a) Managerial decision-making involves calculation of opportunity costs.
b) Equi-incrementation is a practical concept.
c) Despite business profit being positive, the economic profit may be zero or negative.
d) Wages are equal to discounted marginal product of labour.
c) Scarcity is the be-all and end-all of economic problems, micro or macro.

3 "The traditions of mangerial economics are those of specifying technology not in generalised production function but in specified process; of identifying and measuring the goals of an organisation rather than assuring a simple profit maximising goal; of considering the role of information and ingorance instead of certainty and of finding statistical constructs to enable measurement with analysis rather than remaining satisfied with unmeasured concepts' Explain.

It is a difficult question. You may atempt an answer now. It would better. if you attempt another answer when you have completed at least one reading of the entire course material in this subject. You may, then compare the quality of the two answers. That will be an excellent self-assessment of your learningthe subject.

UNIT 3 BASIC TECHNIQUES

Structure

3.0 INTRODUCTION

We now move from 'concepts' to 'techniques'. To begin with, you must note that concepts and techniques are not mutually exclusive. You must have observed in the preceding unit that concepts can be explained better with the help of tools and techniques like tables, diagrams, equations etc. Of course, there are more sophisticated techniques drawn from disciplines like applied mathematics, statistics, econometrics and operations research. The use of such methods and measures makes managerial economics, a highly technical subject. The use of techniques is geared towards measurement and optimisation of economic decision variables.

3.1 OBJECTIVES

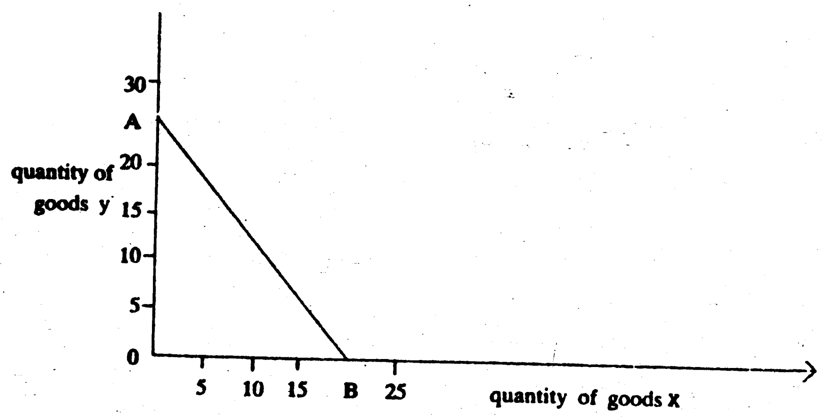
In what follows, our objective is to:

- help you recall some of the basic mathematics, useful, in managerial economics
- introduce you to the formal concept of 'optimisation' in decision-making context
- suggest the use of calculus in measuring economic magnitudes and in optimising economic decision variables.
- make you aware of other tools and techniques like model-building and case-studies.

3.2 BASIC MATHEMATICS

The first concept is that of a set. A set is defined to be a collection of distinct and well defined objects. In fact a set can be defined in two ways: either by enumeration of its members or by specifying a criterion for membership. An example of the first method would be the set of numbers 5,6,7 written as (5,6,7); or the set of alphabets, a,b,c written as (a,b,c). Sometimes it is difficult to define a set by listing its members. For instance, the set of all residents of Delhi or set of all birds in a bird sanctuary. Hence we take recourse to the second method of defining a set i.e. by using some criterion for membership. For instance, the set of all positive integers between 1 and 10; or the set of all points lying on the line x + y = 4.

In managerial economics, we will be concerned with the choice executed by a business firm, often the need arises for specifying the **opportunity set** of the decision maker i.e. the set of alternative actions which are feasible. For instance, the opportunity set of a consumer is the set of all combinations of goods which the consumer can buy with his given income. Given the consumer's money income (budget) and prices of all goods, the opportunity set is well defined, and we can check whether a particular combination is feasible for the consumer, i.e. whether he can afford to buy that combination of goods. If a personconsumesonly two goods (x & y), whose prices are Rs. 5 and Rs 4 per unit, and his money income is Rs. 100, then the consumer's opportunity set is shown in the graph below—all points (combination of x & y) lying on the line AB or below.



If the consumer spends entirely on x, he can buy 20 units of x with no y, and this is shown by point B. On the other hand if he spends only on y, then he buys 25 units of y with no x, and this is indicated by point A. Now all other alternative ways of speding Rs. 100 will lie on the line AB. Hence the area OAB constitutes the consumer's opportunity set for consumption.

In managerial economics you will be dealing with **Variables** like consumption, demand, supply, income, investment, wages, profits, etc. A **variable** is a thing which varies, which can take a set of possible values within a given problem. A **Constant** is a quantity which does not change in a given problem. For instance in the equation Y = a+bX; a and b are constants. Their values remain the same within a given problem, e.g. 'a' can be 2 and 'b' can be 3, the equation then becomes Y=2+3X. A **constant** is also called a **parameter**. Note that the constant can assume other values in different problem, e.g. Y=5+2X is considered to be a different equation from Y=2+3X. Now in the equations Y=2+3X, and Y=5+2X, X and Y are **variables**. Here X can assume different values and this in turn will cause Y to assume different values. In fact X is called the **Independent Variable**, while Y is called the **Dependent Variable**. Sometimes X is called **Exogenous Variable**, while Y is called the **endogenous variable**. The values of X will be given from outside the system, while the values of Y will be determined from within the system (By the given equation).

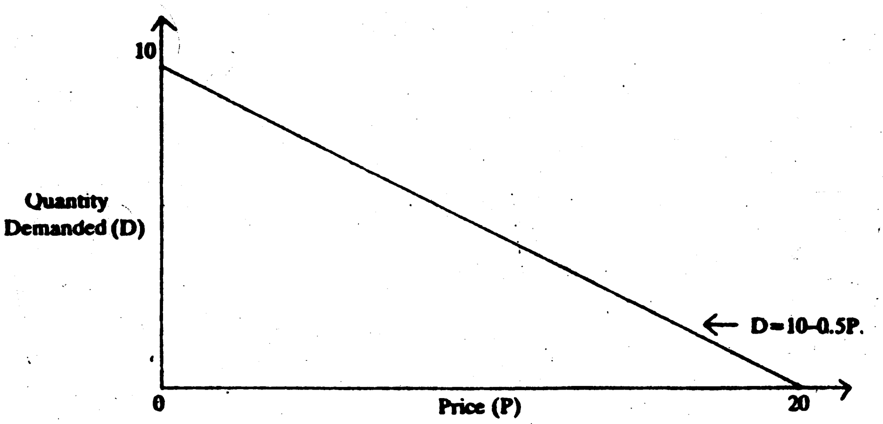
Next you may recall the concept of **function**. A functions such as Q = Q (P) expreses a relationship between two variables Q and P, such that for each value of P there exists one and only one value of Q. Functions are the basic building blocks of formal economic models. The function D = D(P) is called the demand function and its graph on a two-dimensional diagram with price on one axis and quantity demanded an another axis, will be called the **demand curve**. A function maps out a relationship between price and quantity—a mapping of one variable onto another.

Here P is the argument of the function—the independent variable, while D is the dependent variable. A function indicates the cause-effect relationship between variables. In the demand function D=D(P), P is the cause variable while D is the effect variable.

A function can be represented by means of a **table** or by means of a **graph**. A graph is a geometric representation of the relationship embodied in the function. Suppose the specific form of the demand function is D=10-0.5P. In a **table** form, the function can be represented as follows:

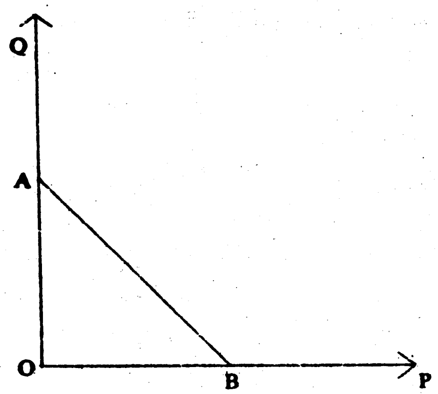
P	1	2	3	4	5	6
D	9.5	9	8.5	8	7.5	7

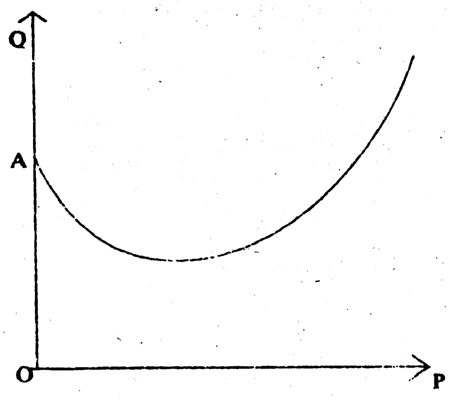
Graphically it would look like the following:



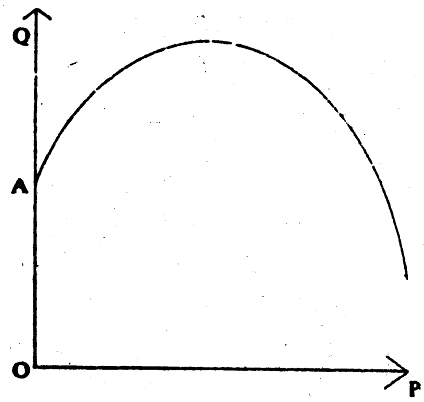
The Quantity intercept is 10, while the slope is-0.5, the price intercept is 20. In general terms the above function can be expressed as D=a-bp where a and b are positive constants, i.e. they do not vary as the independent variable P varies. It should be noted that in managerial economics we normally plot the independent variable on the vertical axis and the dependent variable on the horizontal axis, contrary to what we have done in the above diagram.

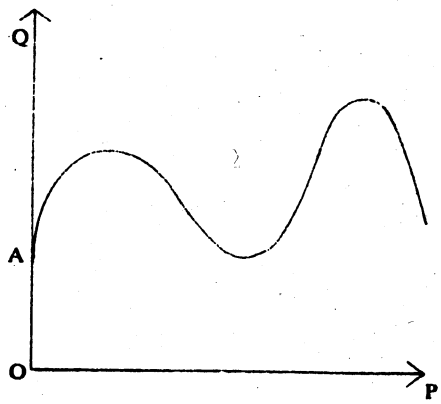
Graphs of functions can take different forms, depending on the form of the functions. **Three functions** frequently encountered in managerial economics involving a single independent variable are given below:





(i) LINEAR $Q = a-bp$ $a=OA$ $b=OA/OB$(iia) QUADRATIC $Q = a-bp+ CP^2$; $a = OA$; $C>O$





(iib) QUADRATIC $Q = a+bp-CP^2$; $a = OA$

(iii) CUBIC $Q = a+bp+ CP^2+dP^3$; $a = OA$

Activity 1

Concentrate on each of the above diagrams, and think about the relationship between Q and P.

a) Draw the graph of the following functions:

i) $Q = 10 - 0.4P$
ii) $Q = 15 - 2p+4P^2$
iii) $C = 100+0.8Y$.

Where C is consumption and Y is income.

b) The following table expresses Savings (S) as a function of income (Y) Graph the function.

Y	0	1000	2000	3000	4000
S	-400	-200	0	200	400

c) A firm's fixed costs are Rs. 6000/- regardless of output (they do not change when output changes); variable costs are Rs. 5/- per unit of output (variable costs are dependent on output). Total cost = fixed costs + variable costs. The selling prices of the goods is Rs. 100/- per unit. Let Q be the output. State the

i) firm's fixed cost function

ii) variable cost function

iii) total cost fucntion

iv) total revenue function

3.3 DERIVATIVES

In your course material on Quantitative Methods, you have been introduced to some basic principles of calculus and their economic applications. The marginal concept and analysis of Economics is easily amenable to the method of calculus. As such, you may begin this section, by recalling some of the standard rules of differentiation in calculus :

i) Basic Rule : $Y = ax^n$, $\frac{dY}{dX} = naX^{n-1}$

ii) Addition Rule : $Y = u(x) + v(x)$, $\frac{dY}{dX} = \frac{du}{dx} + \frac{dv}{dx}$

iii) Product Rule : $Y = u(x) \times v(x)$, $\frac{dY}{dX} = u(x)\frac{dv}{dx} + v(x)\frac{du}{dx}$

iv) Quotient Rule : $Y = \frac{u(x)}{v(x)}$, $\frac{dY}{dX} = \frac{v(x)\frac{du}{dx} - u(x)\frac{dv}{dx}}{v^2}$

v) Chain Rule : $Y = y[u(x)]$, $\frac{dY}{dX} = \frac{dy}{du} \times \frac{du}{dx}$

vi) Logarithm Rule : $Y = log_e x$, $\frac{dY}{dX} = \frac{1}{x}$

vii) Exponential Rule : $Y = e^x$, $\frac{dY}{dX} = e^x$

What is important for you to note now is the *economic interpretation* and use of $\frac{dY}{dX}$

You may recall that $\frac{dY}{dX}$ measures the slope of the curve plotting the function $Y = Y(x)$.

The "Slope" in mathematical sense is the concept of *marginalism* in economic sense.

Thus, if Y = Y(x), the $\frac{dY}{dX}$ stands for change in y as a result of one unit change in x, i.e. marginal y of x. In the same way, we may begin :

$\frac{dD}{dP}$ = Marginal demand of price, when D = d(P)

$\frac{dS}{dA}$ = Marginal sales of advertisement, when S = s(A)

$\frac{dR}{dQ}$ = Marginal revenue of output, when R = r(Q)

$\frac{dC}{dQ}$ = Marginal cost of output, when C = c(Q)

You may also remember that in case of *averagism,*

$\frac{D}{P}$ = Average demand

$\frac{S}{A}$ = Average sales

$\frac{R}{Q}$ = Average revenue

$\frac{C}{Q}$ = Average costs

When we divide the marginal concept by the corresponding average concept, we measure the economic concept of *elasticity*. For example

$\frac{dD}{dP} \cdot \frac{P}{D}$ = Price elasticity of demand

$\frac{dC}{dQ} \cdot \frac{C}{Q}$ = Output elasticity of cost

$\frac{dS}{dA} \cdot \frac{A}{S}$ = Advertisement elasticity of sales revenue

Such 'elasticities' measure the proportion of change. For example, if the percentage change in demand is greater than the percentage change in price, then $[\frac{dD}{dP} \cdot \frac{P}{D}] > 1 \Rightarrow$ elastic demand. On the other hand if $[\frac{dD}{dP} \cdot \frac{P}{D}] < 1 \Rightarrow$ inelastic demand; $[\frac{dD}{dP} \cdot \frac{P}{D}] = 1 \Rightarrow$ unitary elastic demand and so on.

Activity 2

a) What is the difference between 'slope' and 'elasticity'?

....

....

....

....

....

b) How is 'advertisement elasticity of sales' different from 'sales elasticity of advertisement'?

....

....

....

....

....

c) Find the marginal cost (MC) function, given the total cost function, $TC = a+bQ-cQ^2$ where a, b and c are constants and Q stands for output.

....

....

....

d) Consider a firm whose demand function is estimated as P=p(Q), where the total revenue can be defined as TR=PQ, show that the marginal revenue, MR=P(1+1/n) where the n stand for price elasticity of demand.

....

....

....

Do not get upset, if you fail to derive this result. In a subsequent unit, we will solve it for you. You have, therefore, to wait and watch.

You may, therefore, appreciate how is the method of calculus useful for managerial economics – its concepts and measurement.

You may now prepare yourself for an extended use of Differential Calculus in managerial economics. This is where we would like to introduce the mathematical concept of *'Partial derivatives'* and its economic counterpart.

3.4 PARTIAL DERIVATIVES

So far we have considered a function of one independent variable. However, in managerial economics often we have to encounter a function of several independent variables. For instance, a consumer's demand for a product depends on the price of the product, the prices of other related goods, consumer's income, consumer's tastes and so on. When the price of the goods changes, the effect on the quantity demanded of the goods can only be analysed if all other variables are held constant, in the form of 'ceteris paribus' (other things being equal) condition. The funtional relationship that we get between the quantity demanded of a product and its own price is called a Partial Function (a function of one variable when all other variables being used as constant). The process of differentiation can be applied also to the partial functions. The derivative of the partial functions are known as the partial derivative of the original function and is denoted by $\frac{df}{dx_i}$ or $f_i(x)$ or f_x (Instead of using, 'dee f' by 'dee x', we use 'curly dee f' by 'curly dee x').

Note that the partial derivatives are functions of all variables entering into the original function f(x).

Example:

i) If $Y = f(x_1; x_2) = x_1{}^2 x_2{}^2$
Then $\frac{\delta y}{\delta x_1} = 2x_1 x_2{}^2$
And $\frac{\delta y}{\delta x_2} = x^2{}_1 2x^2 = 2x_2 x_1^2$

ii) If $Y = \sqrt{(x_1 x_2)} = (x_1 x_2)^{1/2} = x_1{}^{1/2} x_2{}^{1/2}$
Then $\frac{\delta y}{dx_1} = \frac{1}{2} x_1{}^{-1/2} x_2{}^{1/2} = \frac{1}{2}\frac{x_2^{1/2}}{x_1^{1/2}} = \frac{1}{2}\sqrt{\frac{x_2}{x_1}}$
And $\frac{\delta y}{\delta x_2} = \frac{1}{2} x_1^{1/2} x_2^{-1/2} = \frac{1}{2}\frac{x_1^{1/2}}{x_1^{1/2}} = \frac{1}{2}\sqrt{\frac{x_1}{x_2}}$

iii) If $Z = 4x^2 + 3xy + 5y^2$
Then $\frac{\delta z}{\delta x} = 8x + 3y$
And $\frac{\delta z}{\delta y} = 3x + 10y$

n finding $\frac{\delta z}{\delta x}$ we hold y constant, and we know that the derivative of a constant is zero i.e. $\frac{\delta z}{\delta x}$ of $5y^2$ is zero. Similarly on finding $\frac{\delta z}{\delta y}$ we hold x constant and hence $\frac{\delta z}{\delta x}$ of 4x is zero. This gives $\frac{\delta z}{\delta x}$ to be 8x+3y; and $\frac{\delta z}{\delta y}$ to be 3+10y. Now $\frac{\delta f}{\delta x}, \frac{\delta f}{\delta y}, \frac{\delta y}{\delta x}, \frac{\delta f}{\delta x_2}$ are called first order partial derivatives.

The *second order partial derivatives* indicates that the function has been differentiated partially twice with respect to a given variable, all other variables being held constant. These are denoted in the case of the function z = f(x,y) by $\frac{\delta^2 f}{\delta x^2}$ or f_{xx} And $\frac{\delta^2 f}{\delta y^2}$ or f_{yy}. Thus $\frac{\delta^2 f}{\delta x^2}$ indicates the rate of change of the first order partial derivaties f_x with respect to x with y held constant. Similarly $\frac{\delta^2 f}{\delta y^2}$ is the second order partial derivative of the function with respect to y with x held constant.

The partial derivatives f_{xy} or f_{yx} are the *second order cross partial derivatives* and measure the rate of change of *one* of the first order partial derivatives with respect to the *other* variable. The partial derivative of f_x with respect to y is the second order cross partial derivative f_{xy} or $\frac{\delta y}{\delta y \delta x}$.

Now by *Young's Theorem* $f_{xy} = f_{yx}$: the second order cross partial derivatives are *equal*.

Examples:

i) Z $= 7x^3 + 9xy$

then $\frac{\delta z}{\delta x} = 21x^2 + 9y = f_x$

$\frac{\delta^2 z}{\delta x^2} = 42x = f_{xx}$

$\frac{\delta^2 z}{\delta y \delta z} = 9 = f_{xy}$

ii) Z $= 9xy + 2y^5$

$\frac{\delta z}{\delta y} = 9x + 10y^4 = f_y$

$\frac{\delta^2 z}{\delta y^2} = 40y^3 = f_{yy}$

$\frac{\delta^2 z}{\delta x \delta y} = 9 = f_{yx}$

iii) Z $= 7x^3 + 9xy + 2y$

$\frac{\delta^2 z}{\delta x} = 21x^2 + 9y = f_x$

$\frac{\delta^2 z}{\delta x^2} = 42x = f_{xy}$

$\frac{\delta^2 z}{\delta x \delta x} = 9 = f_{xy}$

$\frac{\delta z}{\delta y} = 9x + 10y^4 = f_y$

$\frac{\delta^2 z}{\delta y^2} = 40y^3 = f_{yy}$

$\frac{\delta^2 z}{\delta x \delta y} = 9 = f_{xy}$

Hence $\frac{\delta^2 z}{\delta y \delta x}$ which is $f_{xy} = 9$

$\frac{\delta^2 z}{\delta x \delta y}$ which is $f_{yx} = 9$

Hence $f_{xy} = f_{yx}$ Young's Theorem is verified.

1.5 OPTIMISATION CONCEPT

Recall the concept of 'economic rationality', we hope you remember that Economic rationality implies optimisation. The idea of optimisation is present in any quantitative decision-making. For instance: a consumer's choice of consumption bundle, a firm's production decision, a planner's choice of resource allocation in the economy and so on. In simple terms, optimisation means the act of choosing the best alternative out of whatever alternatives are available. It describes how decisions (i.e. choice among alternatives) are of should be made (or taken). All optimisation problems consists of three elements.

The Decision Variables: These are variables whose optimal values have to be determined. For example production manager wants to know at what level to set output in order to achieve maximum profit or maximum sales revenue. Here output is the decision variable or choice variable. Similarly a works manager wants to know what amounts of labour, machine time and raw materials to use so as to produce a given output level at a minimum cost. Here the choice variables are labour, machine time, and raw materials. As a consumer you may want to buy that bundle of commodities which you can afford and which makes you feel 'best-off'. Here the choice variables are quantities of commodities. The quantity of any choice variable must be assumed to be measurable (e.g. 20 kg. of sugar, 5 labourers, 10 hours of machine time, Rs. 100 of profit etc.).

The Objective Function: The objective function is a mathematical relationship between the choice variables on the one hand and some variables whose values you wish to maximise or minimise. Thus the objective function could relate

i) profits to the level of output
ii) costs to the amount of labour, machine-time and raw-materials or
iii) an index of consumer's satisfaction to the quantities of commodities he or she may buy. In (i) and (iii) the functions are to be maximised, and in (ii) the function to be minimised with respect to the choice variable.

The Feasible Set: An essential part of any optimisation problem is a specification of exactly what alternatives are available to the decision-maker. The available set of alternatives is called the feasible set.

The feasible set is sometimes described by one or more inequalities which directly define a set of alternative values of the choice variables. It may also be defined by one or more functions or equations which represent a set of alternative values. For instance, in the consumer's choice problem, we can say that a consumer cannot consume negative amount of goods (which is not meaningful in economic). If Q_x and Q_y are the quantities of two goods which the consumer can buy, then we have the following inequalities defining the feasible set $Q_x \geqslant 0$ and $Q_y \geqslant 0$. We also know that in a given consumption period, a consumer's expenditure cannot exceed his income. Hence the feasible set is given by $P_xQ_x + P_yQ_y \leqslant M$ (where M is consumer's money income, P_x and P_y are the prices of the two goods) *plus* the non-negativity requirement, $Q_x \geqslant 0$ and $Q_y \geqslant 0$. These functions and inequalities, which limit or restrict the alternatives which can be considered in defining the feasible set, are known as *'constraints'*.

Solutions: Now a solution to an optimisation problem is that set of values of the choice variables which is in the feasible set and which yields a maximum or minimum of the objective function over the feasible set.

3.6 UNCONSTRAINED OPTIMISATION TECHNIQUE

The calculus technique is extensively used in solving optimisation problem. You may recall from yours 'quantitative methods' course that in the context of decision-making, "Optimisation" may mean either maximisation or minimisation; either of them may be with constraints or without constraints

Unconstrained Optimisation

You may recall that for unconstrained optimisation problem involving single independent variable, you need to satisfy some "conditions":

Order	*Conditions*	*Optimisation*	*(unconstrained)*
		Maximisation	Minimisation
First Order	Necessary conditions	$\frac{\delta y}{\delta x} = 0$	$\frac{\delta y}{\delta x}$
Second Order	Sufficient conditions	$\frac{\delta y}{\delta x} = 0;\ \frac{\delta^2 y}{\delta x^2} < 0$	$\frac{\delta^2 y}{\delta x^2} > 0$

We are assuming that Y = y(x).

You may now note that in economics language, the necessary (first order) conditions is termed "equilibrium conditions" whereas the sufficient (second order) conditions is termed "stability condition". 'Equilibrium' means the balancing of two opposite and equal forces. 'Stability' means continuance of that state of equilibrium. There may be equilibrium, but it may or may not be stable. In other words, the first order condition may be satisfied but it is not a guarantee for the satisfaction of the second order condition. Let us now examine some of the economic uses of these conditions.

Example: The problem

Given a firm's demand function, P=45 − 0.5Q and the average cost function, $AC = Q^2 - 8Q + 57 + Z/Q_1$ find the level of output Q which:

i) maximises total revenue

ii) maximises profits

Solution

i) Since the demand function is p=45 − .5Q. The total revenue will be
$TR=PQ=(45-0.5Q)Q=45-0.5Q^2$
To maximise TR, we find the derivative and set it to zero (the first order or necessary condition).

$$\text{Now } \frac{dR}{dQ} = 45 - 2(.5)Q$$
$$= 45 - Q = 0$$
$$\therefore Q = 45$$

The second order condition (the sufficient condition) requires that $\frac{d^2R}{dy^2}$ be negative.

$$\text{Now since } \frac{dR}{dQ} = 45 - Q$$
$$\therefore \frac{d^2R}{dQ^2} = -1$$

Which is negative because of the minus sign before one. Hence total revenue is maximised when the out put produced is 45 units.

ii) Form the profit function

$$\pi = TR - TC$$

Now TC=(AC) multiplied by Q

$$TC=(Q^2-8Q+57+2/Q)Q \text{ and } TR = (45-0.5Q)Q$$

$=Q^3-8Q^2+57Q+2$ $\qquad = 45Q-0.5Q^2$

After substituting TR and TC in the profit function we get

$\pi = 45Q - 0.5Q^2 - Q^3 + 8Q^2 - 57Q - 2$

$\therefore \frac{d\pi}{dQ} = 45 - Q - 3Q^2 + 16Q - 57$

Now set $\frac{d\pi}{dQ} = 0$

$45 - Q - 3Q^2 + 16Q - 57 = 0$

or $3Q^2 - 15Q + 12 = 0$

or dividing by 3

$Q^2-5Q+4 = 0$

or $(Q-4)(Q-1) = 0$

$\Longrightarrow Q = 4$ or 1

The next step is to test for the sufficient condition (the second order condition)

Thus, $\frac{d^2\pi}{dQ^2}$ must be negative for the profit maximising output.

Now $\frac{d^2\pi}{dQ^2} = -6Q + 15$ (Given $\frac{d\pi}{dQ} = -3Q^2 + 15Q - 12$) when Q = 4 we have

$\frac{d^2\pi}{dQ^2} = -6(4) + 15$

$= -24+15$

$= -9$ which is negative.

When Q = 1 we have $\frac{d^2\pi}{dQ^2} = -6(1) + 15 = -6 + 15$

$= +9$ which is positive.

Hence Q=4 is the profit maximising output and the maximum level of profit = 6 (after substituting Q = 4 in the profit function.)

Maxima and Minima in the two variable case

Now coming back to unconstrained optimisation of a function of several variables, the necessary and sufficient conditions for a funtion z = f(xy) can be stated as follows :

i) **Necessary conditions** (first order conditions) for both maximisation as well as minimisation

$\frac{\delta f}{\delta x} = f_x = 0$

$\frac{\delta f}{\delta y} = f_y = 0$

ii) **Sufficient conditions** (second order conditions). These are *actually two.*

a) For *maximisation*

$\frac{\delta^2 f}{\delta x^2} = f_{xx} < 0$

$\frac{\delta^2 f}{\delta y^2} = f_{yy} < 0$

and $(f_{xx})(f_{yy}) > (f_{xy})^2$

b) For *minimisation*

$\frac{\delta^2 f}{\delta x^2} = f_{xx} > 0$

$\frac{\delta^2 f}{\delta y^2} = f_{yy} > 0$

and $(f_{xx})(f_{yy}) > (f_{xy})^2$

Examples :

Optimise the function

$$Z = 6x^2 - 9x - 3xy - 7y + 5y^2$$

i) Necessary conditions (first order conditions) are

$$f_x = \frac{\delta z}{\delta x} = 0$$

$$f_y = \frac{\delta z}{\delta y} = 0$$

Now $f_x = 12x - 9 - 3y = 0$

i.e. $12x - 3y = 9$

or $4x - y = 3$ (i)

$f_y = 3x - 7 + 10y - 0$

i.e. $10y - 3x - 7$ (ii)

We solve equations (i) and (ii), to get

x=1

y=1

ii) Sufficient conditions

$$\frac{\delta^2 f}{\delta x^2} = f_{xx} = 12$$

$$\frac{\delta^2 f}{\delta y^2} = f_{yy} = 10$$

$$\frac{\delta^2 f}{\delta x \delta x} \quad \frac{\delta^2 f}{\delta_x \delta_y} = f_{xy} = -3$$

$$(f_{xx})(f_{yy}) = 12 \times 10 = 120$$

and $(f_{xy})^2 = (-3)^2 = 9$

$$(f_{xx})(f_{yy}) > (f_{xy})^2$$

120>9.

Hence the function is minimum at x=1 and y=1.

The minimum value of the function is

$$Z = 6 \times 1^2 - 9 \times 1 - 3 \times 1 \times 1 - 7 \times 1 + 5 \times 1^2$$
$$= 6 - 9 - 3 - 7 + 5$$
$$= 11 - 9$$
$$= -8$$

Activity 3

a) A firm is producing two products x and y, and has the following profit function $\pi = 64x - 2x^2 + 4xy - 4y^2 + 32y - 14$

Find the profit maximising levels of output for each of the two products

....

b) A manufacturer in a monopolistically competitive industry produces two different brands of a product for which demand functions are $P_1 = 56 - 4Q_1$ and $P_2 = 48 - 2Q_1$ and the joint cost function is $TC = Q_1^2 + 5Q_1 Q_2 + Q_2^2$.
Find the profit maximising level of output and the price that should be charged for each brand.

3.7 CONSTRAINED OPTIMISATION TECHNIQUE

So far we have been dealing with unconstrained optimisation (either maximisation or minimisation). There are, however, many situations where the objective function has to be maximised or minimised subject to certain constraints being present in the problem. For instance as a producer you may be maximising sales revenue subject to resource constraint or cost constraint; as a consumer you may be maximising utility subject to the income constraint. For a producer the problem can be formally stated in a different form as follows :

Minimise cost subject to the constraint that the firm (the producer) must purchase sufficient inputs to produce the specified output given the firm's production function.

Now the techniques which are used to analyse such a problem are based on the techniques used for unconstrained problems. What we do is to convert the constrained problem into an unconstrained one and solve the latter. This we do with the help of a technique called *Lagrange Multiplier Technique*. In this method we combine the objective funtion and the constraint in one expression, which is called the Lagrange expression. In doing so the constrained maximisation or minimisation problems are reduced to one of unconstrained maximisation or minimisation problem.

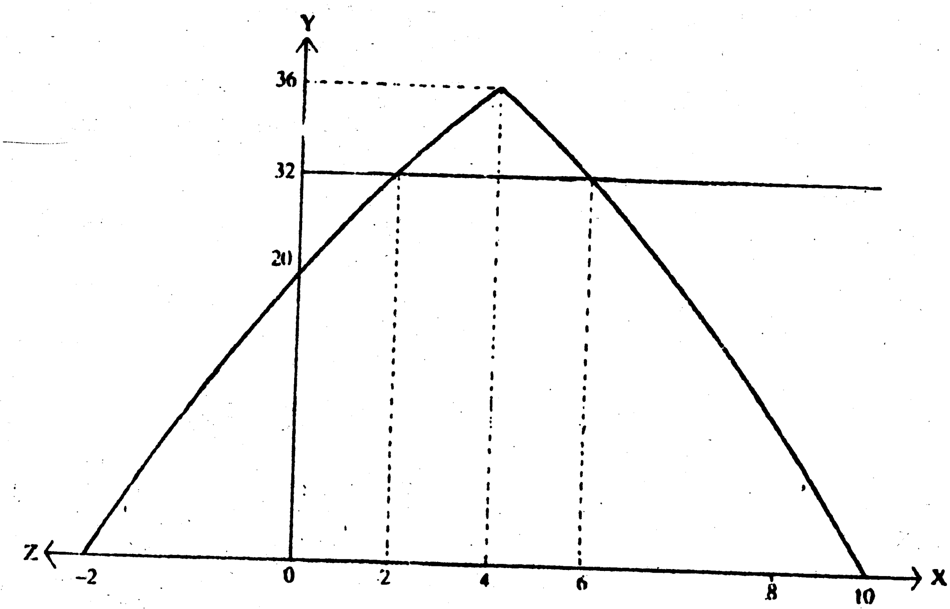
Let us explain the technique with the help of an example.

Maximise $y = -x^2 + 8x + 22$

subject to $x \leqslant 2$.

(We have considered a function of a single independent variable.)

The objective function (the function to be optimised) is $y = -x^2 + 8x + 22$; when plotted it gives the following diagram



The objective function $y = -x^2 + 8x + 20$ can be written as $y = -(x - 4)^2 + 36$. Now $-(x - 4)^2$ has an unconstrained maximum value of zero at x = 4. However, our objective functions is $-(x - 4)^2 + 36$, and hence will have an unconstrained maximum of 36 (at x=4)

This is so for the first derivative test $\frac{dy}{dx} = 0$.

The second derivative test $\frac{dy^2}{dx^2}$ gives -2 (since $\frac{dy}{dx} = -2x + 8$

$$\frac{d^2Y}{dx^2} = -2$$

Thus both the tests would give us x = 4 as the value of the variable at which the objective function attains unconstrained maximum. However, in our problem we have a *constraint* $x \leqslant 2$ (i.e. x has to be less than or equal to 2). This implies that we have to consider the function (as drawn in the graph) only upto the value of x=2, starting from x=−2. The maximum value of the function then is 32 which occurs when x=2. Hence the constrained maximum of the function $y = -(x - 4)^2 + 36$ subject to the constraint $x \leqslant 2$, would occur at x=2 and *not* at x=4.

Let us now see how the Lagrangian expression or function is formed.

We have $y - (x - 4)^2 + 36$ subject to the constraint $x \leqslant 2$ or x−2 = 0. We combine the two to get the Lagrangian expression $L = (-(x - 4)^2 + 36) + \lambda(x - 2)$.

Note that informing the expression L we take the objective function and add the product of λ (which is called the Lagrangian Multiplier, pronounced as *Lamda*) and the constraint function x−2=0. Now L is a function of x and λ. We find out $\frac{dL}{dx}$ and $\frac{dL}{\delta\lambda}$ set them to zero

$\frac{dL}{dx} = -2(x-2)) + \lambda = 0$ (I)

$\frac{dL}{d\lambda} = (x-2) = 0$ (II)

The second equation (II) gives x=2. Hence the constrained maximum occurs at x=2. In this problem λ of course is zero (which is just accidental) but normally it is not zero.

When we apply the Lagrangian technique to solve economic decision problem, λ will have an economic significance. For instance, in consumer's utility maximising problem, λ will be the marginal utility of money income. In producer's cost minimisation profit, λ will be the marginal cost of production. In complex problems, you have to use as many λ, as many constraints. For example if there are two constraints, you may use two λ i.e. λ_1 and λ_2.

Constrained optimisation with two independent variables

Next we consider a problem in which the objective function has two independent variables rather than one considered so far. Let the objective function be

$Z = 4x^2 + 3xy + 6y^2$

subject to x+y=56.

Solution

1st step: Set the constraint equal to zero i.e. x+y-56=0

2nd step: Multiply (x+y-56) by λ and add it to the objective function to form the Lagrange Function.

$L = 4x^2 + 3xy + 6y^2 + \lambda(x + y - 56).$

3rd step: Find the first order partial derivatives and set them equal to zero.

i) $\frac{\delta L}{dx} = [8x + 3y + \lambda] = 0$

ii) $\frac{\delta L}{\delta y} = [3x + 12y + \lambda] = 0$

iii) $\frac{\delta L}{\delta \lambda} = [x + y - 56] = 0$

Solving these three equations we get

x = 36

y = 20

λ = 348

The value of the objective function

L = $4(36)^2+3(36)(20)+6(20)^2+(-348)(36+20-56)$

= 4(1296)+3(720)+6(400)-348(0)

= 9744.

This is the optimised value of the objective function. To find whether it is maximum or minimum, we need the second derivative test. The second order conditions differ from those of unconstrained optimisation. The second order conditions for constrained optimisation also requires the use of second order direct and cross partial derivatives (as is the case with unconstrained optimisation which we discussed earlier). However, it also requires the use of Hessian and bordered Hessian determinant which we will not develop here. Those who are interested can look up: A.C. Chiang, *Mathematical Method for Economics,* (2nd edition).

Activity 4

a) Maximise $z = 10xy - 2y^2$
subject to $x + y = 12$

b) Maximise $R = 737 - 8Q^2 + 14A + QA - 4A^2 + 20Q$
subject to $2Q + A = 2$

....

....

....

....

....

....

....

....

3.8 OTHER TOOLS AND TECHNIQUES

In the preceding sections, we have concentrated on the calculus technique of optimisation because that is what we mostly use in managerial economics. Of course, there are other techniques of optimisation as well.

Linear Programming

You must have already been introduced to this technique in other management courses. This is the reason we do not want to get into the details of this technique here. However, you may appreciate that in linear programming, the statement of the optimisation (maximisation or minimisation, as the case may be) problem runs in a linear form whose variables are restricted to values satisfying a system of linear constraints i.e. a system of linear equations and/or linear inequalities.

In any optimisation problem involving a single inequality constraints, the Lagrangian method discussed in the last section, can still be used and is quite simple. However, when more than one inequality constraint is involved, linear programming will be a better method. In fact, the linear programming technique differs from the classical optimisation technique based on calculus, in that it deals with optimisation problem in which the optimiser faces inequality constraints, and where the constraints as well as the objective function are all linear rather than non-linear. For example, while making his purchase decision, the buyer is required to finance his expenditure out of his budgeted income, B. Thus.

$B = P_x X + P_y . Y$, where x and y are quantities of two items purchased at their respective prices, P_x and P_y. In linear programming, the budget constraint takes the form of an inequality:

$$P_x X + P_y Y \leqslant B$$

Thus, by liberalising the constraint requirement, this new (linear programming) framework of optimisation makes the problem more interesting and more realistic.

Now the *Lagrange Multiplier Technique* (based on Calculus) can be used to deal with inequality constraints, also, but then there will be no way of knowing which constraints are binding and which are not, or whether it is the optimum (unless of course the objective function is strictly concave). Hence the inequality constraint calls for the development of a new method of solution. This is called the *method of algorithm*. You may recall, the feasible area of a Linear Programming problem forms a convex set, while the objective function is concave. When a concave function is optimised over a convex set, the optimum that we get is a global optimum.

You may use a "graphical" method or the "simplex" method to get a solution of linear programming problem. Linear programming as a technique is applicable to a variety of economic problems of choice such as balanced diet, product-mix (allocation) and transportation.

Game Theory

It is a purely mathematical device which has been developed to explain economic behaviour of "players" in a given market environment. The game theory can be treated as an optimisation technique guiding decisions-choices made by individuals in situations in which the consequences of such choices are partly determined by choices of other individuals. In other words, when there is interdependence in decision-making, optimal decision may be arrived at through the use of game theory. For example, in a situation of Duopoly when two sellers confront each other for a given market share, the game theoretic technique may be used to locate a stable equilibrium solution.

There are various types of games—two person zero run, two person constant run etc.; there can also be n-person game, applicable in oligopoly situation. In all such games, the strategy and the tactics are defined. The strategy may be 'minimax' or 'maximia' type; based on this the 'pay-off matrix' is stated. In this case, the interdependence in decision-making is evaluated based on available market information collected by the rational and intelligent players (decision-makers) in the game. The element of risk and uncertainty involved in decision-making can also be taken care of through the game theory approach.

In appropriate context, we will have a more detailed discussion of such technique. Presently you may just note that such techniques are highly mathematical and sophisticated. For example, some background knowledge of matrix algebra may be helpful in using such techniques.

Input-output Model

This is another very popular technique of economic analysis, though it is not an optimisation technique. Input-output model is useful in the context of macro level planning and projection. At the micro level of a corporate unit, input-output model lies at the root of end-use method of demand forecasting.

Again, matrix algebra is basic for understanding the rationale and use of input-output models. Such models essentially states the nature of technological relationship which exists between 'sectors'. Here we consider "intersectoral dependence". For example, coal is used to produce steel from a coal-based steel plant. Steel is used to build railway wagons. Railway is needed to transport coal. In fact, steel may be needed to manufacture coal mining machineries and equipment. Thus 'output' from one sector acts as 'input' for another sector; in this process intersectoral balances assume criticality. It is this critical relationship which is explicitly stated in the form of "technology matrix", "transaction matrix", "Leontief matrix" and so on. For the time being, you do not have to bother about such concepts and techniques. As and when it is immediately required, we will introduce them in appropriate context.

In fact, our idea is to [illegible] such tools and techniques briefly by way of *Appendix* to the relevant unit of study materials which follow subsequently. Presently, we are concentrating on the techniques which are basic but general in nature.

3.9 MODELS AND CASES

The guidelines for decision-making based on analysis can be provided through two alternative methods i.e., models and cases. Both these methods attempt to stimulate real life situations.

Models

The use of models is a very popular technique of economic analysis. In layman's sense, a model is a physical specification—a prototype of an object like the model of a school building, the model of a car etc. In technical sense, a model is a system or relations which help us in understanding the reality. Such relationship can be presented in one or more alternative forms--table, diagram, flow chart, graph, mathematical functions, statistical distribution. Irrespective of the form of representation, a model symbolises the behaviour pattern of a given variable in relation to other. Thus a model, though a theoretical abstraction or just a conceptual construction, can explain the behaviour that is *actually* observed and can predict the behaviour that is *likely* to be observed. In other words, a model has both analytical and predictive value.

From the description above, it follows that all models can be classified into three broad categories:

Iconic Models

These are pictorial or visual representation like drawings, design, prototype etc. which provide information to management.

Analogue Models

Such models present a set of properties of the data in a form which is easily amenable to analysis e.g. a flow chart, funds flow statement, statistical distributions such as bionomial, poisson, normal etc.

Mathematical Models

In these models, relationship are expressed in mathematical symbols and equations. Such models are extensively used in economic analysis. These may be further classified under:

Economic Models: Economists often postulate the basic structural relationship of an aggregate economy in the form of a "macro aggregative model". Sometimes they focus on the structural relationship between various sectors like agriculture and industry; these are called "sectoral models". Sometimes, they shows transactions within an economy between industries dependent on each other. Such "inter-industry models" are usually put in the form of an input-output table—a kind of matrix arrangement. For example, for purpose of forecasting demand, such input-output model is very useful. Similarly, for purpose of planning and projection for the national economy, both aggregative and sectoral models are useful.

Econometric Models: Econometrics is a discipline combining economic theory, statistical method and mathematical precision. We have already stated that a model is a statement of relations. Such relations can be stated among variables, endogenous and exogenous, and constant parameters. Relations are stated in the form of equations. Equations are of two types:

Definitional equations which are identities, definitionally true. All equilibrium conditions are stated in this form e.g. Demand = Supply or Savings + Investment.

- Behavioural equations which explains the behaviour of one variable in terms of other, e.g. Demand depends on price.

  In addition, in a model, you may find some.

- Antonomus terms which are constants influencing a variable.
- Exogenous terms determined outside the system of relations postulated and a few other.
- Parametric information like the slope term.

Given the "structure" of a model defined in above terms, the task is to use the model to derive a "reduced-form solution" — solution of an unknown (endogenous) variable in terms of known constants and other parametric information as well as exogenous terms. Finally, one can use this model to predict the impact of a change in known terms on the unknown variables.

We may take a couple of examples to illustrate this for you.

Example 1: A Competitive Market Model

$D = d_o - d_1 P$ ...... (i)

$S = s_o + s_1 P$ ...... (ii)

$D = S$ ......(iii)

You may note that equation (i) and (ii) are "behavioural" in nature. The equation (i) explains the behaviour of demand (D) in terms of price (P); the equation (ii) explains the behaviour of supply (S) in terms of price (P). The equation (iii) is definitional in nature suggesting the condition for market equilibrium. Also not that d_1 or s_1 are parametric known information: d_1, the slope of demand function is negative and s_1, the slope of supply function is positive. The terms d_o and s_o are antonomus terms: d_o captures the impact of all variables other than P on the demand functions; s_o captures the impact of all non-price variables on the supply function. In all, we have a determinate system of three equations in three unknown (D, S and P); so we can attempt a reduced-form solution:

$$
\begin{aligned}
D &= S \\
d_o - d_1 P &= s_0 + s_1 P \\
s_1 P + d_1 P &= d_o - s_o \\
P(s_1 + d_1) &= d_o - s_o \\
P &= \frac{d_o - s_o}{s_1 + d_1} \quad \text{....... (iv)}
\end{aligned}
$$

The result (iv) is an acceptable solution of an unknown P in terms of known terms d_o, s_o, s_1 and d_1. Also note

$\frac{\delta P}{\delta d_o} > 0 =$ An increase in antonomus demand, other things remaining equal, will raise price and vice versa.

$\frac{\delta P}{\delta s_o} < 0 =$ An increase in antonomus supply, other things remaining equal, will reduce price and vice versa.

In the same way, you may find

$\frac{\delta P}{\delta d_1}$ or $\frac{\delta P}{\delta s_1}$

However, carefully note the assumption in this model

$d_o > s_o$

$d_1 < O$

Example 2: A Macro Model of an economy

$Y^s = Y^d$ .... (i)

$Y^d = C + I + G$ .... (ii)

$C = c_o + c_1 Y^s$ .... (iii)

$I = i_o + i_1 Y^s$ .... (iv)

Here we have a system of four equations or four unknown: Y^d (aggregate demand), Y^s (aggregate supply), C (consumption) or I (Investment). The terms $C_o = i_o$ are antonomus; c_1 or i_1, are slope terms $\bar{G}$ (Government expenditure) is exogenous. Now you may attempt a reduced form solution:

$$
\begin{aligned}
Y^s &= Y^d \\
&= C + I + \bar{G} \\
&= c_o + c_1 Y^s + i_o + i_1 Y^s + \bar{G} \\
&\Rightarrow Y^s - c_1 Y^s - i_1 Y^s = c_o + i_o + \bar{G} \\
&\quad \text{or, } Y^s(1 - c_1 - i_1) = c_o + i_o + \bar{G}
\end{aligned}
$$

$Y^s = \frac{1}{i - c_1 - i_1}[c_o + i_o + \bar{G}]$ .... (v)

Note that in result (v), we have a solution for our unknown Y^s in terms of all known terms : $c_o, i_o, \bar{G}, c_1$ or i_1,

And now, you may operate

$$\frac{\delta Y^s}{\delta c_o} = \frac{\delta Y^s}{\delta i_o} = \frac{\delta Y^s}{\delta \bar{G}} = \frac{1}{1 - c_1 - i_1}$$

You may observe that the behavioural equations in this or earlier model can be estimated through the regression technique. Given the time-series data, you may use least square technique to estimate the intercept term like c_o, i_o, d_o or s_o and the slope terms (regression coefficients) like c_1, i_1, d_1 and s_1.

The econometric models may be of two types:

a) Single equation model as given in the regression analysis, or

b) Simultaneous equations models, as illustrated above.

Regression analysis is useful in those cases where the factors influencing the dependent variables are mutually unrelated. Simultaneous equation models are used where variables are mutually related.

In most of our empirical estimates in this course, you will be exposed to single equation model. For example, you will have reference to statistically estimated (regression) equations for demand function, production function, cost function and profit function.

Quite a few problems of managerial economics can be approached through other types of models as well. Some of these models are listed below:

- Statistical Distribution Models
- Allocation and Transport Models

- Scheduling Models
- Queuing Models
- Simulation Models
- Inventory Models

You may be encountering some of these models in your other courses such as quantitative methods including Operations Research and Linear Programming, production management etc. Therefore, in managerial economics, you are not expected to enter into details of these models.

Activity 5

Recall all the course materials on Management, that you have received till date. Make a list of the models that you are familiar with.

....

....

....

....

....

....

....

Cases

The case method is a pedagogical technique. In managerial economics, the case method is useful to the extent it stimulates a real world business situation. A case represents a depiction of the real life situation; it describes the actual environments, experiences, events or incidents or episodes of the historical past. The factual information comprises objective data and subjective feelings. A case, when thrown for discussion and analysis, the participants think about the problems and come out with probable solutions. Different orientation and perception leads to different points of view. The idea behind the use of case method is to train the participants in the exercise of logical thinking.

There is no definite procedure in analysing a case, but normally the case analyst follows an ordered sequence of steps:

1 Identify the key issue; keep away the trivial issues.

2 Establish the nature of the issue - the problem of choice by examining the available data (facts and figures).

3 Examine the information gap, and make necessary assumptions to bridge that gap.

4 Analyse the facts and assumed information.

5 Work out the range of alternative solutions and the implied consequences.

6 Recommend a particular solution out of the given set of alternatives; this is what can be termed as a "decision".

7 Use the "decision" as a subject for discussion and deliberation by the participants.

The basic idea behind such a method is to raise and resolve the issue. In the analysis of the issue, sophisticated concepts, tools and techniques have to be used.

The use of case method requires some degree of maturity. Particularly for our subject, some prior knowledge of formal concepts and measurement techniques is necessary. We feel that such case method may be appropriate in an advanced course on the basis of introducing courses in economics, statistics, accounts and similar basic discipline. Considering this, we may be using only "caselets" (short cases) rather than fullbodied "cases" in our course materials here. It is, therefore, important for you to grasp the basic concepts and to learn to use of tools and techniques of economics, statistics and econometrics such that you are equipped for case analysis eventually.

3.10 SUMMARY AND CONCLUSION

In this lesson, you have been introduced to a set of mathematical concepts, tools and techniques which have considerable use in managerial economics. These concepts are: set, function, variables, parameters, models, derivatives, maxima, minima, constrained optimisation through lagrange multipliers. There are other relevant tools and techniques like linear programming, game theory, input-output model etc. which we have not considered here; we hope to make passing reference to them in appropriate context by way of Appendices.

At this stage, some of you may find it somewhat difficult to appreciate the managerial uses of such tools and techniques, but eventually you will realise that scientific management is impossible without the use of such tools and techniques This does not mean that such mathematical tools and techniques have no limitations. They have. As practising managers, you have to be equipped with scientific tools and techniques, but you should not over-emphasise their use. Mathematically arrived precise results must have an intuitive appeal before you approve them as "scientific decisions".

3.11 ADDITIONAL READINGS

Allen, R.G.D., *year 3 Mathematical Analysis for Economists,* The English Language Book Society and Macmillan.

Baumol, W.J., 1980, *Economic Theory and Operations Analysis* 4th edition, Prentice-Hall of India.

Birchenhall, C. & Grout, P., 1987, *Mathematics for Modern Economics* 1st edition, Heritage India.

Dowling, E.T., 1980, *Mathematics for Economics* 1st Edition, McGraw-Hill.

Koutsoyiannis, A., 1979, *Modern Micro Economics* 2nd edition, Macmillan International.

3.12 SELF-ASSESSMENT TEST

"Concepts remain abstract, if they cannot be measured; and measurements remain mathematical numbers, if they do not have a conceptual basis". Discuss.

Answers to Activities

3 (a) $x = 40$

$y = 24$

$\pi = 1650$

(b) $Q_1 = 2.75$

$Q_2 = 5.70$
$P_1 = 45$
$P_2 = 36.6$
$\pi = 213.94$

4 (a) $x = 7$
$y = 5$
$\lambda = -50$

(b) $Q = 0.52$
$A = 0.96$
$\lambda = -6.36$

BLOCK II DEMAND DECISION

In this block, we will help you to define, describe and discuss various concepts and techniques of demand analysis.

Demand is a very crucial economic decision variable. As a manager, you have to estimate and analyse sometimes your own demand and sometimes others' demand. As a production manager, you need to know your own demand for factors of production. As marketing manager, you need to know others' demand for your product.

In Unit 4, we introduce you to all such demand concepts: demand for input, demand for output etc.

In Unit 5, we undertake an overview of the theoretical developments with regard to demand analysis.

In Unit 6, we examine the empirical estimates of demand backed by an understanding of demand determinants and elasticities.

In Unit 7, we focus on concepts and techniques of demand forecasting.

All these together should provide you with conceptual and technical background for analysing demand decisions.

UNIT 4 DEMAND CONCEPTS

Structure

4.0 INTRODUCTION

Demand is one of the most critical economic decision variables. Demand reflects the size and the pattern of market. Business activity is always market-determined. The manufacturers' inducement to invest in a given line of production is limited by the size of market. If the market conditions reflect profitability, the firm starts producing for that market. In the process of production, inputs are transformed into flows of output. Output, when sold in the market, yields revenue. Inputs to be obtained for the productive process, involve costs. Profit is the difference between revenues and costs, and it is influenced by the demand-supply conditions for output and input. The demand for output and input; the demand for the firm and the industry; the demand by the consumer and the stockists; and similar other demand concepts become therefore, relevant for managerial decision analysis. Even if the firm pursues objectives alternative to profit-maximisation, demand concepts still remain relevant. For example, suppose the firm is aiming at 'customer service', not profit. How can it ensure quantity and quality of service, without analysing what the customer really wants? Or suppose, the firm is destined to discharge 'social responsibility' of business. Can this be done without evaluating social preferences? Tastes, preferences and choices are all concepts directly built into the economic concepts of 'demand'.

4.1 OBJECTIVES

After studying this unit, you should be able to:

- **appreciate** the significance of demand analysis
- **specify** a demand function, identifying the determinants
- **classify** various types of demand
- **comprehend** a long list of demand concepts
- **locate** the sources of demand
- **relate** demand to the nature of market structure

4.2 SIGNIFICANCE OF DEMAND ANALYSIS

Demand is one of the crucial requirements for the functioning of any business enterprise, its survival and growth. Demand analysis is of profound significance to management information on the size and type of demand helps management in planning its requirements of men, materials, machines, money/and what have you. For example, if the demand for a product is subject to temporary business recession, the firm may plan to pile up the stock of unsold products. If the demand for a product shows a trend towards a substantial and sustained increase in the long run, the firm may plan to install additional plant and equipment to meet the demand on a permanent basis. If the demand for a firm's product is falling, while its rival's sale is increasing, the firm needs to plan its sales strategy and sales tactics; the firm may need to undertake some sales promotion activity like advertisement. If the firm's supply of the product is unable to meet its existing demand, the firm may be required to revise its production plan and schedule; or the firm may have to review its purchase plan for inputs and the suppliers' response to input requirements by the firm. In the same way, larger the demand for a firm's product, the higher is the price the firm can charge. The common theme underlying these examples is that the whole range of planning by the firm—production planning, inventory planning, cost budgeting, purchase plan, market research, pricing decision, advertisement budget, profit planning etc.—call for an analysis of demand. In fact, demand analysis is one area of economics that has been used most extensively by business. The decision which management makes with respect to any functional area, always hinges on an analysis of demand.

Demand analysis seeks to identify and measure the forces that determine sales; it reflects the market conditions for the firm's product. Once the demand analysis is done, the alternative ways of creating, controlling or managing demand can be inferred.

Activity 1

a) Why is a person interested in knowing the *demand for the shares* he has purchased?

....

....

b) Why should the Food Corporation of India be concerned about the *demand for foodgrains* to be released for public distribution system?

....

....

c) Why does the Maruti Udyog Ltd. ask one of its officers to estimate the *demand for Premier* 118 NE in 1990?

....

....

d) Why does a bank analyse the seasonal *demand for credit?*

....

....

You may now think of *any demand* at random, and you are sure to find that there is always some good reason for analysing that demand. The obvious question that may crop up in your mind is : What is demand? It is easy to describe the significance of demand than to define.

4.3 CONCEPT OF DEMAND

As we have indicated earlier, 'demand' is a technical concept from Economics. Demand for a product implies :

a) desires to acquire it,
b) willingness to pay for it, and
c) ability to pay for it.

All three must be checked to identify and establish demand. A poor man's desire to stay in a five-star hotel room and his willingness to pay rent for that room is not 'demand', because he lacks the necessary purchasing power; so it is merely his wishful thinking. Similarly, a miser's desire for and his ability to pay for a car is not 'demand', because he does not have the necessary willingness to pay for a car. One may also come across a well-established person who possesses both the willingness and the ability to pay for higher education. But he has really no desire to have it, he pays the fees for a regular course, and eventually does not attend his classes. Thus, in an economics sense, he does not have a 'demand' for higher education degree/diploma.

It should also be noted that the demand for a product—a commodity or a service—has no meaning unless it is stated with specific reference to the time, its price, price of its related goods, consumers' income and tastes etc. This is because demand, as is used in Economics, varies with fluctuations in these factors. To say that demand for an Atlas cycle in India is 60,000 is not meaningful unless it is stated in terms of the year, say 1983 when an Atlas cycle's price was around Rs. 800, competing cycle's prices were around the same, a scooter's price was around Rs. 5,000. In 1984, the demand for an Atlas cycle could be different if any of the above factors happened to be different. For example, instead of domestic (Indian), market, one may be interested in foreign (abroad) market as well. Naturally the demand estimate will be different. Furthermore, it should be noted that a commodity is defined with reference to its particular quality/brand; if its quality/brand changes, it can be deemed as another commodity.

To sum up, the demand for a product is the desire for that product backed by willingness as well as ability to pay for it. It is always defined with reference to a particular time, place, price and given values of other variables on which it depends.

Activity B

a) Construct some specific examples showing that despite having a 'desire', a person may not have

i) 'willingnes to pay' : ....

....

ii) 'ability to pay': ....

....

b) What do you mean by:

i) A household's demand for a T.V. set : ....

....

....

ii) A firm's demand for labour : ....

....

....

iii) Our Government's demand for defence equipment: ....

....

....

We will now proceed to discuss how the concept of 'demand' differs depending upon factors like:

a) For what reasons does he demand?
b) How do we specify his demand?
c) Who demands?
d) What does he demand?

4.4 SPECIFICATION OF DEMAND

We may start with a restatement of the concept of 'demand'. Demand for a commodity is the necessary step towards consumption of that commodity. In economic terms, a *commodity* is 'a bundle of utilities'; it has usefulness (value-in-use) for the consumer. Unless the commodity has utility, nobody would be willing to pay for it. When one pays the price, one is entitled to acquire it and use it. The process of using the commodity (to be exact, its utility content) is called *consumption*; sometimes this process is called 'destruction of utilities'. The process consumption renders satisfaction. It is this 'satisfaction' which is to be balanced by 'sacrifice'. Sacrifice is involved in the payment for it. The extent of sacrifice in terms of price paid depends not only on the utility (value-in-use) but also on the element of scarcity (value-in-exchange). A person who has the ability to pay and the desire to possess a scarce commodity, would be willing to pay a higher price. What is true of a commodity is also true of a service. Thus, the demand for a product always reflects both value-in-use and value-in-exchange. The intensity of demand itself is reflected in terms of price paid. You must have often heard:

'Demand is always at a price'

However, you should not think that demand is determined only by price. There are other determinants as well. The determinants of demand—its size and intensity—are the subject matter of economic analysis, i.e. an analysis of consumer behaviour. In the next unit, you will be exposed to alternative approaches to that analysis. This unit, is confined to a few basic concepts of demand.

Demand Function and Demand Curve

Demand function is a comprehensive formulation which specifies the factors that influence the demand for the product. For example,

$D_x = D(P_x, P_y, P_z, B, W, A, E, T, U)$

Here D_x stands for demand for item X (say, a car)
P_x, its own price (of the car)
P_y, the price of its substitutes (other brands/models)
P_z, the price of its complements (like petrol)
B, the income (budget) of the purchaser (user/consumer)

W,	the wealth of the purchaser
A,	the advertisement for the product (car)
E,	the price expectation of the user
T,	taste or preferences of user
U,	all other factors.

Briefly we can state the impact of these determinants, as we observe in normal circumstances:

i) Demand for X is inversely related to its own price. As price rises, the demand tends to fall and vice versa.

$\frac{\delta D_x}{\delta P_x} < 0$ This is *Price-demand* relation, showing the *price effect* on demand.

ii) The demand for X is also influenced by its related price—price of goods related to X. For example, if Y is a substitute of X, then as the price of Y goes up, the demand for X also tends to increase, and vice versa. In the same way, if Z is a complement of X, then as the price of Z falls, the demand for Z goes up and, therefore, the demand for X tends to go up.

$\frac{\delta D_x}{\delta P_y} > 0; \frac{\delta D_x}{\delta P_z} < 0.$ This is a **cross-demand** relation, showing the **substitution/complementarity effect.**

iii) The demand for X is also sensitive to price expectation of the consumer; but here, much would depend on the psychology of the consumer; there may not be any definite relation.

$\frac{\delta D_x}{\delta E} \gtrless 0.$

This is *speculative demand.* When the price of a share is expected to go up, some people may buy more of it in their attempt to make future gains; others may buy less of it, rather may dispose it off, to make some immediate gain. Thus the *price expectation effect* on demand is not certain.

iv) The income (budget position) of the consumer is another important influence on demand. As income (real purchasing capacity) goes up, people but more of 'normal goods' and less of 'inferior goods'. Thus *income effect* on demand may be positive as well as negative.

$\frac{\delta D_x}{\delta B} \gtrless 0.$ This is an **income-demand** relation, showing **income effect.**

Sometimes, the demand of a person (or a household) may be influenced not only by the level of his own *absolute* income, but also by *relative* income—his income relative to his neighbour's income and his purchase pattern. Thus a household may demand a new set of furniture, because his neighbour has recently renovated his old set of furniture. This is called *'demonstration effect'.*

v) Past income or accumulated savings out of that income and expected future income, its discounted value along with the present *income-permanent* and *transitory*—all together determine the nominal stock of wealth of a person. To this, you may also add his current stock of assets and other forms of physical capital; finally adjust this to price level. The real wealth of the consumer, thus computed, will have an influence on his demand. A person may pool all his resources to construct the ground floor of his house. If he has access to some additional resources, he may then construct the first floor rather than buying a flat. Similarly one who has a colour TV (rather than a black-and-white one) may demand a V.C.R./V.C.P. This is regarded as the *real wealth effect* on demand.

$\frac{\delta D_x}{\delta W} \gtrless 0.$

vi) Advertisement also affects demand. It is observed that the sales revenue of a firm increases in response to advertisement up to a point. This is *promotional effect* on demand (sales). Thus

$\frac{\delta D_x}{\delta A} \gtrless 0.$

vii) Tastes, preferences, and habits of individuals have a decisive influence on their pattern of demand. Sometimes, even social pressure—customs, traditions and conventions exercise a strong influence on demand. These *socio-psychological determinants* of demand often defy any theoretical construction; these are non-economic and non-market factors—highly indeterminate. In some cases, the individual reveal his choice (demand) preference; in some cases, his choice may be strongly ordered. We will revisit these concepts in the next unit.

You may now note that there are various determinants of demand, which may be explicitly taken care of in the form of a *demand function*. By contrast, a *demand curve* only considers the price-demand relation, other things (factors) remaining the same. This relationship can be illustrated in the form of a table called *demand schedule* and the data from the table may be given a diagrammatic representation in the form of a curve. In other words, a generalised demand function is a multivariate function whereas the demand curve is a single variable demand function.

$D_x = D(P_x)$

In the slope—intercept from, the demand curve which may be stated as

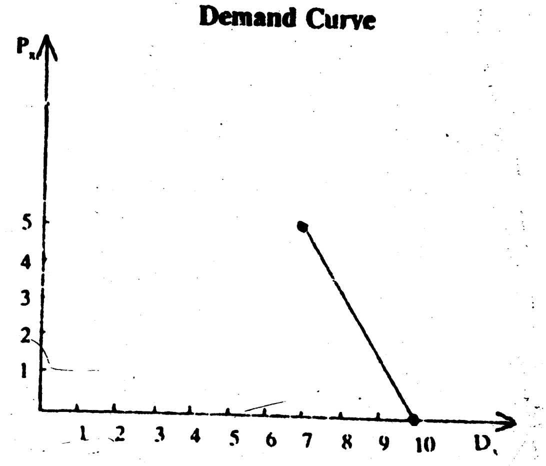
$D_x = \alpha + \beta P_x$, where α is the intercept term and β the slope which is negative because of inverse relationship between D_x and P_x.

Suppose, $\beta = (-)\ 0.5$, and $\alpha = 10$

Demand Schedule

P_x (in Rs.)	D_x (in Units)
1	9.5
2	9.0
3	8.5
4	8.0
5	7.5

Demand Curve



It may be observed from the Demand Schedule and the Demand Curve above that the Price of X and the demand for X move in opposite directions. It may be formally stated :

As price rises, the demand contracts.
As price falls, the demand extends.

This is what is called the LAW OF DEMAND. The statement holds only under *ceteris paribus* assumption i.e., other things remaining equal. In other words, given

$D_x = \alpha + \beta P_x$

$D_x = 10 - 0.5P_x$

$\frac{dD_x}{dP_x} = (-)0.5 \Rightarrow$ Slope of the demand curve.

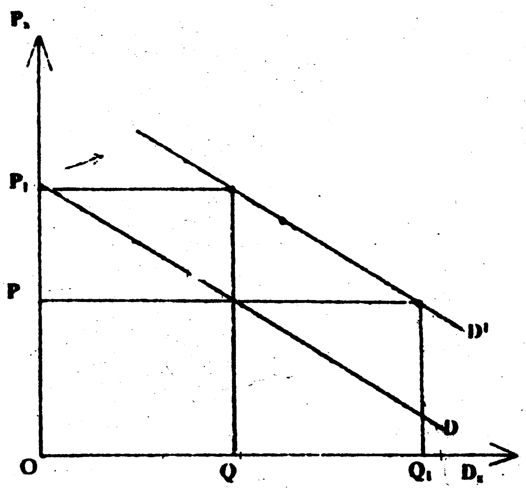
Law of Demand and Changes in Demand

At this stage, it is important to draw a few simple distinctions. The law of demand has reference to 'extension' or 'contraction' of demand, but the changes in demand as a concept has reference to 'increase' or 'decrease' in demand; the former is limited to the movement along the demand curve, but the latter refers to shifts in the demand curve.

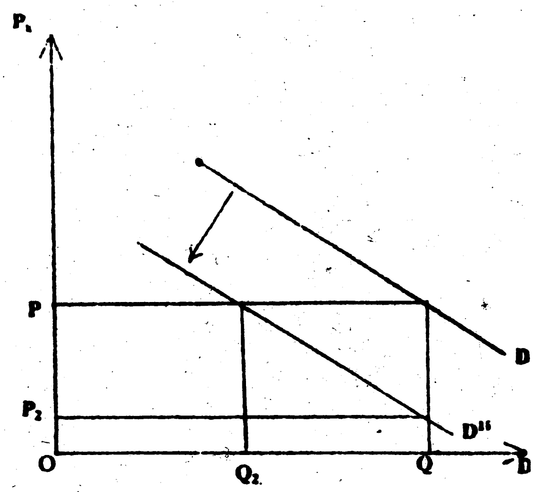
A given demand curve may shift upward to the right or downward to the left—the former will show an 'increase' in demand but the latter will show a 'decrease' in demand.

Shifts in Demand

Increase in Demand

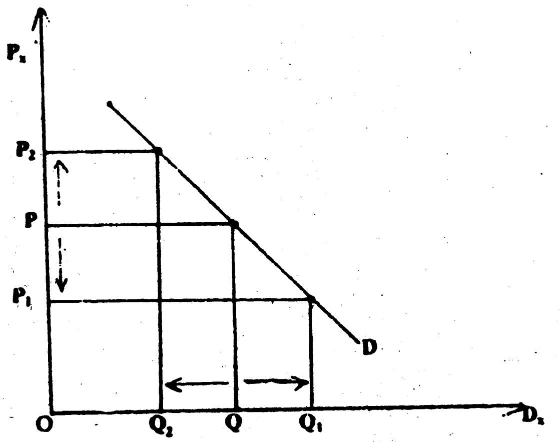


Decrease in Demand



Contraction and Extension of demand

Law of Demand



You may note that if at the same price (OP), more is demanded (OQ$_1$ > OQ) or the same (OQ) is demanded at a higher price ($OP_1 > OP$), you have an **increase in demand.** On the other hand, if at the same price (OP), less is demanded ($OQ_2 < OQ$) or the same (OQ) is demanded at a lower price ($OP_2 < OP$), then you have the case of a **decrease in demand.** In other words, increase or decrease in demand may result from the operation of non-price factors. By contrast, the extension or contraction of demand results from the operation of own—price factor alone. If as result of a rise in price (by $P—P_2$), demand comes down (by $Q—Q_2$), that is **contraction of demand.** Similarly, if as a result of a fall in price (by $P—P_1$), demand goes up (by $Q—Q_1$), that is **extension of demand.** In other words, there is a world of conceptual difference between:

a) Extension and Increase of demand, and

b) Contraction and Decrease of demand.

This distinction is very useful in demand analysis. For example, we may now state more clearly and technically the 'relatedness' of goods/factors that we demand:

When the price of *X goes up* (down) the demand for *X contracts*(extends);and if as a result the demand for *Y increases* (decreases) then X and Y are substitutes.

When the price of *X goes down* (up), the demand for *X extends* (contracts); and if as a result the demand for *Z increases* (decreases), then X and Z are complements., You may note that the demand for related goods like Y and Z has changed not because of a change in its own-price in absolute sense.

Activity 3

a) Make a list of the 'effects' on demand. You need not define/describe them, but you may recall the understanding of each concept. In the subsequent unit, you will be exposed to a more formal treatment of those 'effects'.

....

....

b) Distinguish between a demand curve and a demand function. Can the two be related?

....

....

(Hint : A demand curve plots a single variable demand function).

4.5 TYPES OF DEMAND

In the preceeding unit, you have been told that you may specify demand in the form of a function, schedule or curve. Much of this specification and its form depends on the nature of demand itself—its type and determinants. From this standpoint, we can talk about a few other distinct concepts of demand:

Direct and Derived Demands

Direct demand refers to demand for goods meant for final consumption; it is the demand for ***consumers' goods*** like food items, readymade garments and houses. By contrast, derived demand refers to demand for goods which are needed for further production; it is the demand for ***producers'goods*** like industrial raw materials, machine tools and equipments. Thus the demand for an input or what is called a factor of production, is a derived demand; its demand depends on the demand output where the input enters. In fact, the quantity of demand for the final output as well as the degree of substitutability/complementarity between inputs would determine the derived demand for a given input. For example, the demand for gas in a fertiliser plant depends on the amount of fertiliser to be produced and substitutability between gas and coal as the basis for fertiliser production. However, the direct demand for a product is not contingent upon the demand for other products.

Of course, it must be noted that the distinction between consumers' goods and producers' goods is some what arbitrary, for whether a product is meant for final consumption or further production depends upon its use. For example, a refrigerator located in a house is a consumer product; while a refrigerator located in a chemist's shop is a producer item. However, the distinction is useful because, among other factors, the direct demand for consumers' goods depends on consumers' income, while the derived demand for producers' goods depends on the final output and its technology specification.

ii) Domestic and Industrial Demands

The example of the refrigerator can be restated to distinguish between the demand for domestic consumption and the demand for industrial use. In case of certain industrial raw materials which are also used for domestic purpose, this distinction is very meaningful. For example, coal has both domestic and industrial demand, and the distinction is important from the standpoint of pricing and distribution of coal.

iii) Autonomous and Induced Demands

When the demand for a product is tied to the purchase of some parent product, its demand is called induced or derived. For example, the demand for cement is induced by (derived from) the demand for housing. As stated above, the demand for all producers' goods is derived or induced. In addition, even in the realm of consumers' goods, we may think of induced demand. Consider the complementary items like tea and sugar, bread and butter etc. The demand for butter (sugar) may be induced by the purchase of bread (tea). Autonomous demand, on the other hand, is not derived or induced. Unless a product is totally independent of the use of other products, it is difficult to talk about autonomous demand. In the present world of dependence, there is hardly any autonomous demand. Nobody today consumes just a single commodity; everybody consumes a bundle of commmodities. Even then, all direct demands may be loosely called autonomous.

In the context of econometric estimates of demand, this distinction is used to identify the determinants of demand. For example, in

$$D_x = \alpha + \beta P_x$$

'α' represents the autonomous part which captures the influence of all non-price factors on demand, whereas βP_x represents the induced part—D_x is induced by P_x, given the size of β.

iv) Perishable and Durable Goods' Demands

Both consumers' goods and producers' goods are further classified into perishable/non-durable/single-use goods and durable/non-perishable/repeated-use goods. The former refers to final output like bread or raw material like cement which can be used only once. The latter refers to items like shirt, car or a machine which can be used repeatedly. In other words, we can classify goods into several categories : single-use consumer goods, single-use producer goods, durable-use consumer goods and durable-use producers goods.

This distinction is useful because durable products present more complicated problems of demand analysis than perishable products. Non-durable items are meant for meeting immediate (current) demand, but durable items are designed to meet current as well as future demand as they are used over a period of time. So, when durable items are purchased, they are considered to be an addition to stock of assets or wealth. Because of continuous use, such assets like furniture or washing machine, suffer depreciation and thus call for replacement. Thus durable goods demand has two varieties—replacement of old products and expansion of total stock. Such demands fluctuate with business conditions, speculation and price expectations. Real wealth effect influences demand for consumer durables.

v) New and Replacement Demands

This distinction follows readily from the previous one. If the purchase or acquisition of an items is meant as an addition to stock, it is a new demand. If the purchase of an item is meant for maintaining the old stock of capital/asset intact, it is replacement demand. Such replacement expenditure is to overcome depreciation in the existing stock.

You may observe that this distinction has reference mostly to the durable-use producers' goods like machines. The demand for spare parts of a machine is replacement demand, but the demand for the latest model of a particular machine (say, the latest generation computer) is a new demand. In course of preventive maintenance and breakdown maintenance, the engineer and his crew often express their replacement demand, but when a new process or a new technique or a new product is to be introduced, there is always a new demand.

You may now argue that replacement demand is induced by the quantity and quality of the existing stock, whereas the new demand is of an autonomous type. However, such a distinction is more of degree than of kind. For example, when demonstration effect operates, a new demand may also be an induced demand. You may buy a new VCR, because your neighbour has recently bought one. Yours is a new purchase, yet it is induced by your neighbour's demonstration.

vi) Final and Intermediate Demands

This distinction is again based on the type of goods—final or intermediate. The demand for semi-finished products, industrial raw materials and similar intermediate goods are all derived demands, i.e., induced by the demand for final goods. In the context of input-output models, such distinction is often employed.

vii) Individual and market Demands

This distinction is often employed by the economist to study the size of the buyers' demand, individual as well as collective. A market is visited by different consumers, consumer differences depending on factors like income, age, sex etc. They all react differently to the prevailing market price of a commodity. For example, when the price is very high, a low-income buyer may not buy anything, though a high income buyer may buy something. In such a case, we may distinguish between the demand of an individual buyer and that of the market which is the aggregate of individuals. The following table illustrates this situation :

Price of X (Rs.)	Units of X purchased by			Market (Total)
	A	B	C	
10	2	0	0	2
8	4	2	0	6
6	5	4	1	10
5	7	6	3	16

You may note that both individual and market demand schedules (and hence curves, when plotted) obey the law of demand. But the purchasing capacity varies between indiyiduals. For example, A is a high income consumer, B is middle-income consumer and C is in the low-income group. This information is useful for personalised service or target-group-planning as a part of sales strategy formulation.

viii) Total Market and Segmented Market Demands

This distinction is made mostly on the same lines as above. Different individual buyers together may represent a given market segement; and several market segments together may represent the total market. For example, the Hindustan Machine Tools may compute the demand for its watches in the home and foreign markets separately; and then aggregate them together to estimate the total market demand for its HMT watches. This distinction takes care of different patterns of buying behaviour and consumers' preferences in different segments of the market. Such market segments may be defined in terms of criteria like location, age, sex, income, nationality, and so on. Different segments of the market may be represented by different regions/zones, different uses of the product, different distribution channels, different customer sizes and different sub-products. Each of these segments may differ with respect to delivery prices, net profit margins, element of competition, seasonal pattern and cyclical sensitivity. When these differences are great, demand analysis must be done segmentwise, and accordingly, different market strategies may have to be followed for different segments. For example, the printing press often charges different rates for 'rush order' and 'regular order' jobs; the drycleaner also charges different rates for 'ordinary' and 'urgent' services; the airlines charges different fares for different passengers—'economy class' and 'executive class'.

ix) Short-run and Long-run Demands

This distinction is also drawn on similar lines as above, with specific reference to the time segments in demand analysis. Short-run demand may be taken to mean immediate, existing demand which is based on available taste and technology. Long-run demand, on the other hand, refers to the size and pattern of demand, which is likely to prevail in future, as a result of changes in technology and tastes, product improvement and promotional efforts, and such other factors where adjustments take place over a period of time. Price-income fluctuations are more relevant as determinants of short-run demand, while changes in food habits, urbanisation, work-culture etc. must be considered for long-run demand analysis. Traditional economic theory neglects this distinction. You may recall, the economists normally distinguish between short-run and long-run with reference to production and costs, and not demand. However, from the standpoint of control and management of demand, this distinction is helpful.

x) Company and Industry Demands

An industry is the aggregate of firms (companies). Thus the Company's demand is similar to an individual demand, whereas the industry's demand is similar to aggregated total demand. You may examine this distinction from the standpoint of both output and input.

For example, you may think of the demand for cement produced by the Cement Corporation of India (i.e., a company's demand), or the demand for cement produced by all cement manufacturing units including the CCI (i.e., an industry's demand). Similarly, there may be demand for engineers by a single firm or demand for engineers by the industry as a whole, which is an example of demand for an input. You can appreciate that the determinants of a company's demand may not always be the same as those of an industry's. The inter-firm differences with regard to technology, product quality, financial position, market (demand) share, market leadership and competitiveness—all these are possible explanatory factors. In fact, a clear understanding of the relation between company and industry demands necessitates an understanding of different market structures.

Activity 4

a) Write against each statement, TRUE or FALSE.

   i) In the statement of a demand function in the slope-intercept from. $D_x = \alpha + \beta P_x$, the intercept term α represents the shift parameter ....

   ii) The demand for management trainees is a derived demand ....

   iii) Demand recession in a market may mean either a contraction in demand or a decrease in demand or both ....

   iv) The share market is always subject to speculative demand ....

b) Quote appropriate examples of:

   i) Close-substitutes : ....

   ii) Perfect complements : ....

   iii) New demand : ....

   iv) Intermediate demand : ....

   v) Consumer durables : ....

   vi) Derived demand : ....

   vii) Seasonal demand : ....

   viii) Autonomous demand : ....

   ix) Conspicuous consumption : ....

c) How is 'consumption' different from 'demand'?

....

....

4.6 THE DEMAND FOR PRODUCT : FIRM Vs. INDUSTRY

The demand analysis is inadequate and inefficient, if it has no reference to the market structure i.e. the position of the firm within the industry.

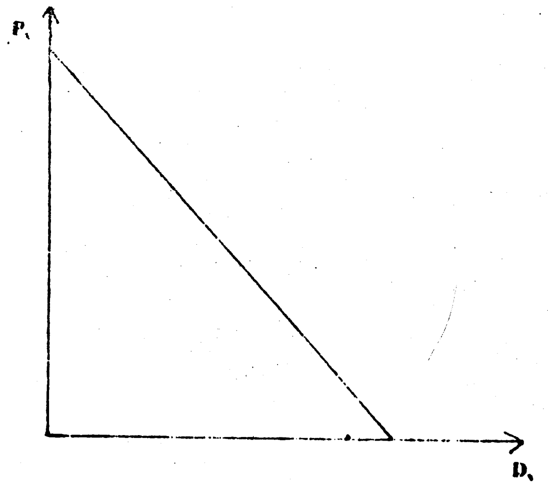
The important market structures are distinguished on the basis of the product differentiation and the number of sellers.

Nature of Products

Number of sellers	*Standardised Products*	*Differentiated Products*
One seller	Monopoly	–
A few sellers	Homogeneous oligopoly	Differentiated oligopoly
Many sellers	Perfect competition	Monopolistic competition

Since a single firm constitutes the whole industry in a monopoly, the company's demand is the same as the industry's demand, while in all other types of market, the two demands are different.

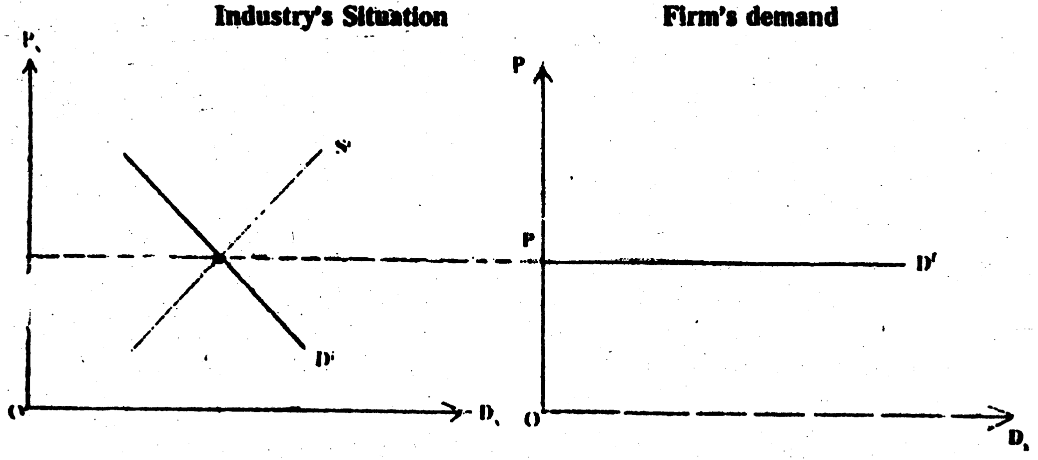
Firm's/Industry's Demand under monopoly



Under ***monopoly***, the demand curve as seen by the seller (firm or industry) is downward falling. A monopolist can sell larger quantity only at a lower price. If he charges a high price, he loses his sales. The demand for postal services and airline services are often quoted as examples of monopoly in India. Of course, with privatisation, Posts and Telegraphs may cease to be a government monopoly. Also, Indian Airlines and Vayudut may have to compete with other modes of transport in certain sectors.

Under ***perfect competition*** the firm's demand is completely divorced from the industry's demand. A company can sell as much as it wishes to at the ruling price, which is determined by the interplay of the forces of industry's demand and industry's supply, as illustrated below:

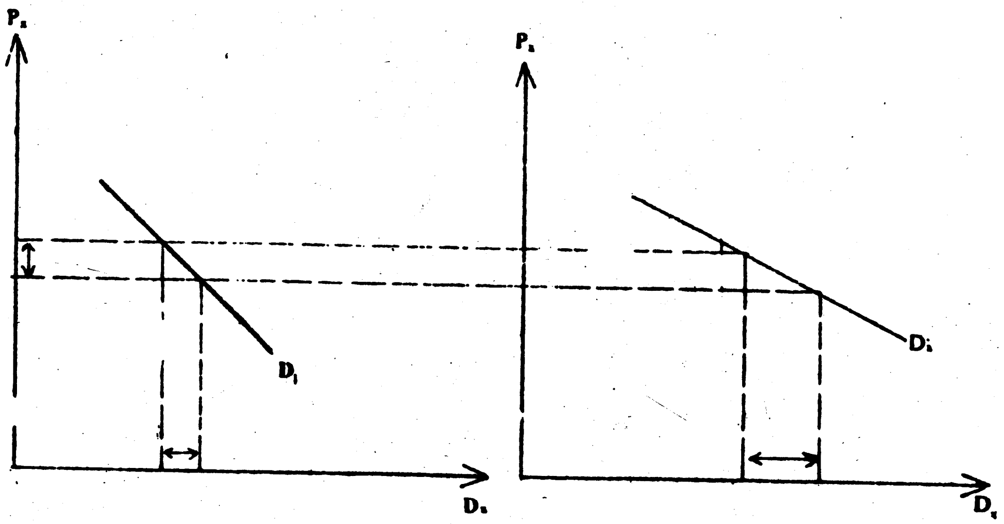
Industry's Situation **Firm's demand**



Thus the industry's demand for the product is downward sloping, but the firm's demand is a horizontal straightline. Such a perfect market situation is very rare, though the market for agricultural goods may approximate such a situation.

In *monopolistic competition*, where there are many sellers with differentiated products (such as toothpastes, tea, textiles etc.) the industry's demand curve has a very little weaning. For, the products of rival firms are advertised like different products and so we have only company demand for each *brand* of a commodity like that in monopoly. The situation is illustrated below:

Industry's demand **Firm's demand**



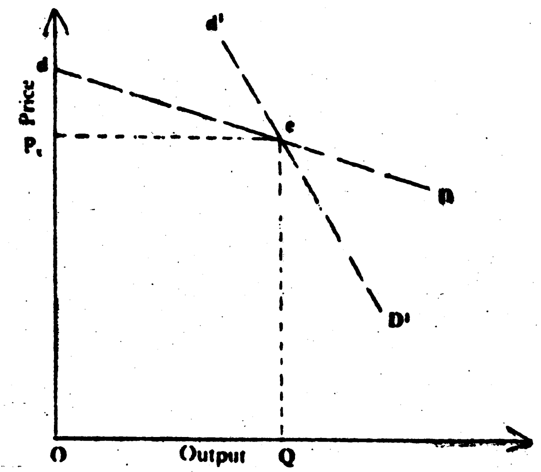
In this case, the industry's demand is relatively stable compared to that of a firm. A firm's demand changes easily and suddenly due to a slight change in rival's strategy and tactics with regard to price, package, colour, discounts etc. which are mostly subject to buyers' illogical preference. Thus for a given change in price, a firm's demand changes proportionately larger than industry's demand.

Under *homogeneous oligopoly,* where sellers are few and products are identical, business is transferable among rivals. The company's own demand is influenced by rivals's action. Cement and steel fall in this category.

In *differentiated oligopoly* also, the demand for an individual company's product is related to industry's demand, but this relationship is less close, compared to the one under homogeneous oligopoly, because some consumers have definite preferences for some particular brands of say, cars, bicycles, refrigerators etc.

A very large number of hypotheses have been suggested to explain such oligopoly situations. One of these hypotheses speaks of Kinked Demand Curve. It is suggested that price rigidity is a characteristic feature of oligopoly. Once the price has been set at OP, there will be resistance to changing it. This is because, if a firm raises its price, its rivals will not raise the price unless

Firm's demand under oligopoly



there is industry-wide increase in costs. If, however, a firm lowers its price, the rivals will follow suit. Thus when we consider two demand cuves simultaneously in the context of anticipated reaction to a rival's action, we find that for a price higher than OP, the relevant demand curve is d-e, whereas for a price lower than OP, the relevant demand curve is e-D^1. Thus the firm's demand curve under oligopoly here, d-e-D^1 contains a kink at point e. Eventually, we will have more to say on such demand curves.

Perfect competition as a situation where a large number of sellers sell identical products is too ideal to be real. Mostly, we encounter imperfect market structure. Based on the analysis here, we may therefore suggest that in reality, the demand for the product of a firm is multi-variate

$d_i = d(P_i, P_o, P, A_i, A_o, S_i, S_o, Y, t_i......)$

where d_1 = demand for the product of the ith firm

P_i = price charged by the ith firm

P_o = price charged by other firms (rivals)

P_r	=	price of related products
A_i	=	advertisement and selling expenses of the ith firm
A_o	=	advertisement and selling expenses of the rivals
S_i	=	Style of product of the ith firm
S_o	=	Style of product of other rivals
Y	=	consumers' income
t	=	consumers' tastes and preferences.

The conclusion from the above discussion is that the demand function of the oligopolist is multivariate. Thus even if prices are sticky, there are other factors that influence the demand for the firm. Based on the historical data on these other factors, one may attempt to estimate statistically a demand function for an oligopoly firm. However, there are many estimation difficulties. Additionally, the market environment of the firm in the real world changes so fast that any statistical demand function becomes inappropriate for future decisions, unless continuously revised.

Finally, in conceptualising the demand for a firm's product's you must consider the *sources of demand*. The traditional economic theory concentrates on final consumers' demand. In reality, the consumers' demand is only a small fraction of the aggregate demand for manufacturing products. The majority of manufacturing products are sold to other businesses, other firms for processing, or to wholesalers and retailers through whom the product reaches the final consumers. In view of this real world business practice, you must try to identify the determinants of demand by the various groups :

1 **Consumers's demand :** It depends on both rational and irrational buyers' preference. The socio-psychological determinants have to be considered along with the traditional price-income factors.

2 **Other firms' demand :** Some firms like ancillary units produce as per the order, size and specification given to them by their parent unit. They enjoy a protected and assured market.

3 **Wholesalers' demand :** Among other things' the volume of expected sales, costs of holding stocks, physical and financial capacity to hold etc. would determine the wholesalers' demand.

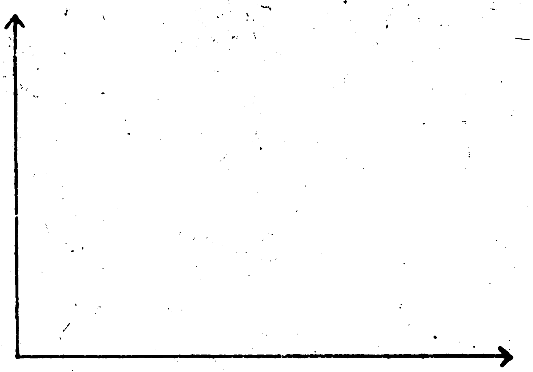
4 **Retailers' demand :** In this case also, the demand will be decided by factors like expected delivery, fixed and variable costs of holding inventory, cyclical or seasonal nature of final consumers' purchase etc.

5 **Suppliers' demand for its own products :** In a modern exchange economy, we mostly produce what we do not consume, and we mostly consume what we do not produce. However, sometimes we may consume a part of what we produce. The farmer himself would like to retain a part of his own produce for self-consumption. This is called 'reservation demand'.

To sum up, when you consider the demand for product, it is necessary to be clear about the structure of market and the sources-cum-uses of the product. After all, there is a conceptual difference between 'demand as seen by the buyer' and 'demand as seen by the seller'. The former is determined by the buyers' preference and purchasing capacity, and the latter is determined by the nature of competition prevailing in the market and its consequent impact on the revenue earning possibility for the seller.

Activity 5

a) Consider the case of a monopolist who is selling the same product at two different prices in two different markets, characterised by two different demand conditions. Use imaginary data to draw the individual demand curves for the two separate markets as well as the aggregated market demand curve.



b) What is 'reservation demand'? Give a couple of examples. How does the reservation demand for foodgrains affect the marketable supply of foodgrains?

....

....

....

....

....

c) Identify the nature of demand for a product like refrigerator from the standpoint of a

i) household purchaser
ii) government purchaser
iii) chemist
iv) distributor
v) manufacturer himself.

....

....

....

....

....

4.7 A FEW MACRO CONCEPTS OF DEMAND

In this unit, you have been exposed so far to a variety of micro-concepts of demand. As a manager, you need to master these concepts. A manager, however, operates within the macro-economic environment; as such a few macro-concepts of demand may also be relevant for him. He may not just stop at the demand for his product within the industry; he may need to understand aggregate demand within the economy as a whole; because national demand may constrain the industry's demand which in its turn may influence a firm's demand.

Some of these macro-concepts of demand are listed and defined below for your ready reference:

- **Effective Demand :** This refers to the aggregate volume of expenditure in an economy, reflecting total demand (size of the market), which induces the manufacturers to match that demand by supply. Thus if demand is 'effective', it should create employment, induce output and generate income in the economy.

- **Consumption Demand :** It is a component of the effective demand; it has reference to the demand for consumer goods i.e., consumption expenditure of a nation which depends on national income.

- **Investment Demand :** It is another component of effective demand; it has reference to the demand for investment goods i.e., investment expenditure in the national economy which is dependent on the net return on investment.

  Both consumption and investment demands may be disaggregated into household, private, corporate and government expenditures.

- **Demand for Money :** This refers to desire to hold money (liquidity) in hand. In this case also, you may think of household demand, private corporate demand and government demand for money—all of them aggregated reflects the total demand for money in the economy as a whole. The demand for money or what is sometimes called 'liquidity preference' originates in either of the three motives—transaction, precaution or speculation. Accordingly, we may speak of *transaction demand for money* to meet day-to-day exchange transactions. The *precautionary demand for money* is to meet contingency requirements. The *speculative demand for money* has got long-term business use; it is mostly influenced by the market rate of interest. In fact, the rate of interest is the opportunity costs of holding money in hand for speculative purposes.

- **Demand for Bonds :** Since money and bonds are substitutes, the demand for bonds is related to the demand for money.

We are not interested in enlarging the list here. The idea is to make you familiar with a few macro-concepts which are also amenable to micro-treatment, if you so desire. In our subsequent units, we will concentrate on the analysis of micro-concepts of demand. And most of this micro-analysis is also relevant for macro-concepts of demand at the aggregate level. That is why we hear a lot about 'Micro-foundations of Macro-economics' in recent literature.

Activity 6

a) The National Thermal Power Corporation (NTPC) has floated bonds to raise capital. State clearly a multivariate demand function for NTPC bonds. Do not forget to define your terms/symbols.

....

....

....

....

....

....

....

b) The export demand for India's turnkey projects is lately showing a downward trend. How do you account for such demand recession? Can you specify an export demand function?

....

....

....

....

....

....

....

....

4.8 SUMMARY

Demand is a critical economic decision variable. Demand analysis is very significant for managerial decisions, particularly decisions related to market strategy and tactics in the area of pricing and advertising. Demand analysis has also a direct bearing on production planning, inventory management, material handling, manpower utilisation, financial evaluation and investment decisions. As a basic concept, demand means desire backed by willingness to pay and ability to pay. Demand is thus always at a price paid. However, there are many determinants of demand—both price and non-price factors. In subsequent units, we will take up a detailed analysis of those factors or 'effects'. In this unit, we have exposed you to a host of concepts of demand. These concepts are very basic to the understanding of demand determinants. Such determinants affect both the direction and the proportion of change in demand. For demand analysis, reference should also be made to the sources of demand, uses of the item demanded, the structure of the market where the firm is located. We will make extensive use of these demand concepts and bring about further sophistication in them in subsequent blocks. The fact remains that the demand concepts introduced in this unit lie at the root of a lot of analysis in managerial economics, and therefore you should aim at *conceptual clarity* while reading the material presented here.

4.9 ADDITIONAL READINGS

1 Mote, V.L., Samuel Paul & G.S. Gupta, 1987. *Managerial Economics—Concepts & Cases.*, Tata Mc-Graw Hill Publishing Company Ltd. : Bombay. (Ch. 2.)

2 Koutsoyiannis, A. 1979. *Modern Microeconomics*, 2nd ed. The Macmillan Press Ltd. : London (Ch. 2, Section IV.)

4.10 SELF-ASSESSMENT TEST

1) Distinguish clearly between (a) Desire and Demand (b) Direct and Derived demand (c) New and Replacement demand (d) Rush and Regular order (e) Autonomous and induced demand (f) Total Market and Market segments (g) Contraction and Decrease in demand.

2) "The demand for the product of a firm within a given industry is influenced by the nature of market structure and sources of demand." Discuss.

3) Why is demand analysis significant for management? Identify various concepts of demand relevant for various functional areas of management.

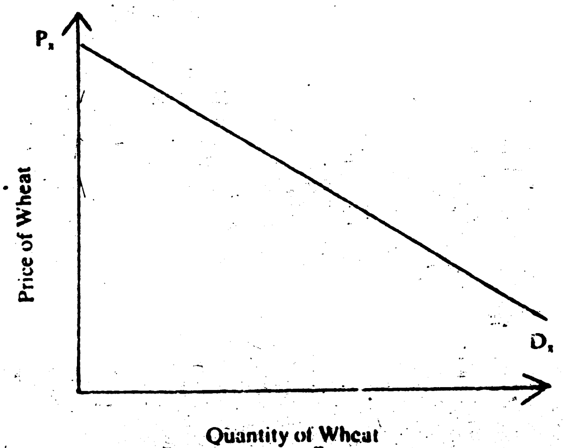
UNIT 5 DEMAND ANALYSIS

Structure

5.0 INTRODUCTION

In the previous unit you have been introduced to the concept of demand and the various influences (factors) on the demand for a commodity. Among the factors mentioned, it is the own-price of the commodity which has the greatest influence on the demand for a particular commodity. You have observed that the demand for a commodity is inversely related to its own price, implying that when own price falls, the quantity demanded extends, and when own price rises, the quantity demanded contracts. If we plot this relationship, we get a downward sloping demand curve.

In this diagram, we have taken the commodity to be wheat, so we call it the demand curve for wheat. Mathematically such a demand curve can be expressed as a demand function of the form $D_x = d(P_x)$ where the form of the function can be linear (straightline) or non-linear (curves). You may recall that when we use demand curve we hold all other influences like income, prices of other goods, consumer's tastes etc. on demand for a commodity constant and only vary the price of own commodity and study how demand responds to such price variations. Normally we find that both individual and market demand curves are downward sloping.



In this unit, we hope to explain why such a demand curve is downward sloping. In other words, we have to explain why the quantity demanded extends when price falls and contracts when price rises, i.e., what has been referred to in the previous unit as the *laws of demand.* This law may be explained in terms of various approaches (theories of consumer behaviours). In particular, we will be considering the following approaches to demand analysis.

i) The Gardinal Utility Approach which was developed by Alfred marshall, the doyen of economists.

ii) The Ordinal Utility Approach (also known as "The Indifference Curve Approach") was developed by two renowned economists of the 20th century—J.R. Hicks and R.G.D. Allen.

iii) The Revealed Preference Approach which has been developed by one of the greatest living economists of our times Paul A. Samuelson who teaches at M.I.T. (Cambridge U.S.A.).

iv) The theory of Consumer's Choice under Risk and Uncertainty which has been developed by authors like Von-Neumann and Margenstern.

v) Finally, the latest developments with regard to some demand models will also be referred.

5.1 OBJECTIVES

On completing this unit, you should be able to :

- *take stock of* various determinants of demand;
- *analyse* some of those determinants in terms of the theories of consumer behaviour;
- *evaluate* the uses of alternatives approaches (theories) to demand analysis i.e., explanation of a downward falling demand curve;
- *understand* the circumstances under which you may get exceptional demand curves; and
- *derive* a demand function from given information about consumer's preference, his objective and constraints imposed by the size of his budget.

5.2 THE CARDINAL UTILITY APPROACH

The question to begin with is : Why does a consumer desire a commodity? Marshall thought, it is because it gives the consumer utility or satisfaction on happiness or subjective sensation from the act of consuming the commodity. According to Bentham, a traditional economist, the utility of an action is the difference between the *pleasure enjoyed* and the *pain suffered* as a consequence of the action. He regarded utility as a magnitude that is in principle measurable in the same way as the body temperature or body weight is. Utility can also be compared among individuals, (i.e. inter-personal comparison of utility can be made). The concept of utility was refined further by the marginalists like Walras, Jevons and Menger. Utility was defined as the power or the ability of a commodity to satisfy human wants. Utility is thus a measure of satisfaction that a consumer receives from the consumption of a commodity. Since utility is something which is experienced by the person, it cannot be measured by someone else. In this sense, it is subjective. However, the marginalist school assumed that utility is measurable (by the person who experiences it). Marshall accepts the Jevonian proposition and builds on it to explain demand.

The Marshallian theory of demand is based on the following set of assumptions.

i) The consumer behaves rationally : i.e. each consumer tries to maximise his own utility; and the choice between various goods by a consumer is ultimately determined by his own evaluation of his own self interest. This implies that each consumer must possess a utility function which reflects the tastes and preference of the consumer. Such a utility function is a relationship between utility, and the quantity and quality of each product or service consumed.

ii) Utility is a cardinal concept : This implies that one can measure it quantitatively on numerical scale. If utility is measurable cardinally, then it implies that there exists a unit of measurement. Marshall assumes that the utility can be measured in terms of money i.e. in terms of price which each consumer is willing to pay for each unit of the commodity rather than do without it. In other words, we ask the consumer. "How much money are you willing to pay for each unit of an apple rather than go without it?". and this gives an indication of the utility derived from its consumption. If utility can be measured, then it can be compared also between persons, between goods and so on.

iii) If money is the measuring rod in terms of which utility is measurable, then we have to assume that the measuring rod itself remains constant. And, therefore, Marshall assumes that *the marginal utility of money remains constant* i.e., the value of Re. 1 to consumer remains constant irrespective of whether his stock of money is increasing or falling. This implies that when the consumer's income rises or falls, the utility of Re. 1 remains constant. This is, however, an unrealistic assumption because utility of money depends upon income.

iv) The proportion of income spent on each product is very small : The budget equation of a consumer is $B = P_xX + P_yY + ...$ This implies that when price is varying, a consumer will not experience a gain or fall in real income, which in turn implies that the utility of rupee always remains constant.

v) Utility is independent : This implies that the utility derived from the consumption of commodity X depends only on the amount of X consumed and not on the quantities of other commodities consumed, i.e. $U_x = f(Q_x); U_y = g(Q_y), U_z = h(Q_z)$. Therefore, the utility function is : $U = U_x + U_y + U_z$ where U = total utility derived from the consumption of various commodities. This shows that the utility function is additive. If X, Y, Z are interrelated, we cannot write the total utility functions in this form. Thus independence of utility implies that there are no related goods.

vi) The consumer is never satiated with a commodity : This is known as the non-satiation assumption, that is, a consumer always gets higher utility from larger quantity consumed than from smaller quantities.

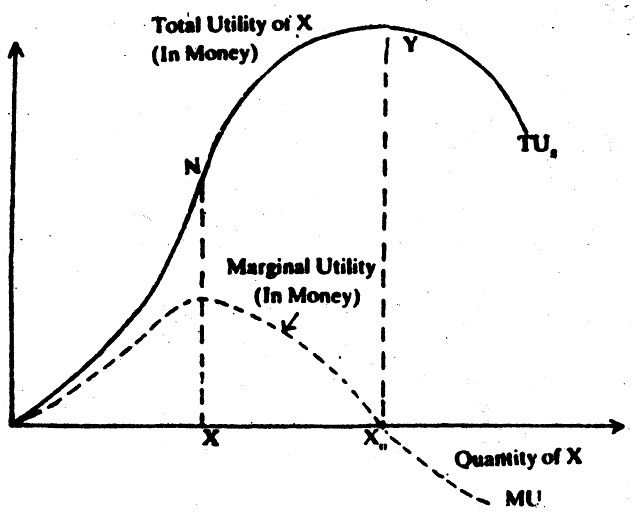
vii) Consumer's tastes remain unchanged : There is no change in his choice pattern and preference at a point of time.

viii) **Consumer's income is constant** and that the money income is all spent in the same period. In other words, in static models savings give no positive utility to the consumer.

ix) Each consumer forms a tiny part of the market and cannot influence the market prices of goods and services through his own actions (i.e. by varying his own demand for the commodity).

x) The consumer behaviour is subject to Law of Diminishing Marginal Utility : The law states that : After sufficient units of a commodity have been consumed, a consumer experiences diminishing marginal utility for the additional units consumed. I.e. additional utility from each additional unit of a commodity goes on diminishing. The total utility will be increasing but at a diminishing rate. The law need not apply to the very first unit consumed, but eventually will, as the consumer goes on consuming more and more units of the same commodity. The extra satisfaction or utility will be falling. It is known as the *Law of Eventual Diminishing Marginal Utility.*

You should be able to recall the concept of marginalism. Presently the concept of Marginal Utility is defined as the addition to total utility caused by consuming additional units of a commodity. This can be illustrated graphically as follows :



Suppose the quantity consumed is measured on the horizontal axis, in physical terms; and total utility or satisfaction is measured on the vertical axis (in units or money).

You may note that the total utility is first increasing at an increasing rate up to point N (which is also called the point of inflexion).

Also note that beyond X_o units consumed total utility starts falling.

$M.U_x = \frac{d(TU_x)}{dx}$ where $d(TU_x)$ is the increase in total utility and dx is the increase in units consumed. Note that TU_x stands for total utility and MU_x for marginal utility of x and Marginal Utility for any unit of the commodity consumed is given by the slope of the tangent line at a point on the total utility curve, corresponding to a given unit consumed. If you are deriving the marginal utility schedule from the total utility schedule you will find that (i) when Total Utility is increasing at an increasing rate, Marginal Utility is increasing, (ii) When TU is increasing at a decreasing rate, MU is falling. You may also recall, in this context, the relation between average and marginal utility.

Therefore, we find that the Law of Diminishing Marginal Utility applies only when consumptio is beyond X. The law actually holds for consumption between X and X_o. If fact at X_o, MU_x 0, i.e., TU_x is maximum. Please note that when TU_x is falling MU_x is negative. Also note th when MU_x is zero the consumer is completely satiated with the goods. The Law of Diminishir MU_x holds when MU_x is positive but declining. The law is also known as the *Law of Satiabl Wants.*

Rationale behind the Law of Diminishing Marginal Utility

When a consumer has fewer units of a commodity then he might use them to satisfy the most important needs. As the number of units available increases, the needs to which these units may be put become less and less important and therefore yield less and less additional utility. For instance, when a consumer has one bucket of water, she might use it for drinking purpose only—the most urgent need. As the quantity of water available increases, she might start using it for cooking, washing and gardening. Some important points about the law must be noted. (i) The units in which the commodity is used must be defined appropriately. For instance, a bucket full of water, one bottle of campa cola, half a litre of mother dairy milk etc. (ii) The commodity must be finely divisible. Oth-rwise we cannot draw a continuous marginal utility curve. (iii) The taste of the consumer must remain unchanged over the period the law is applicable.

Activity 1

Given the following table relating units consumed and total utility, find the incremental/marginal utility for each units consumed, i.e. fill in the blank columns.

a) *Units Consumed (In Units)*	*Total Utility (In Units)*	*Average Utility (In Units)*	*Incremental Utility (In Units)*	*Average Incremental Utility*
10	100			
12	130			
14	170			
16	210			
18	230			
20	240			
22	245			

Units of Utility are sometimes called *utils*, which may be measured in money.

b) *Units consumed (In Units) of commodity*	*Total Utility (In Units)*	*Marginal Utility (In Units)*
1	10	
2	15	
3	27	
4	31	
5	34	
6	36	
7	34	

In the Marshallian approach, the Law of Diminishing MU is the basis of the law of demand (i.e. for the demand curve to be downward sloping). The law of demand can be explained as follows :

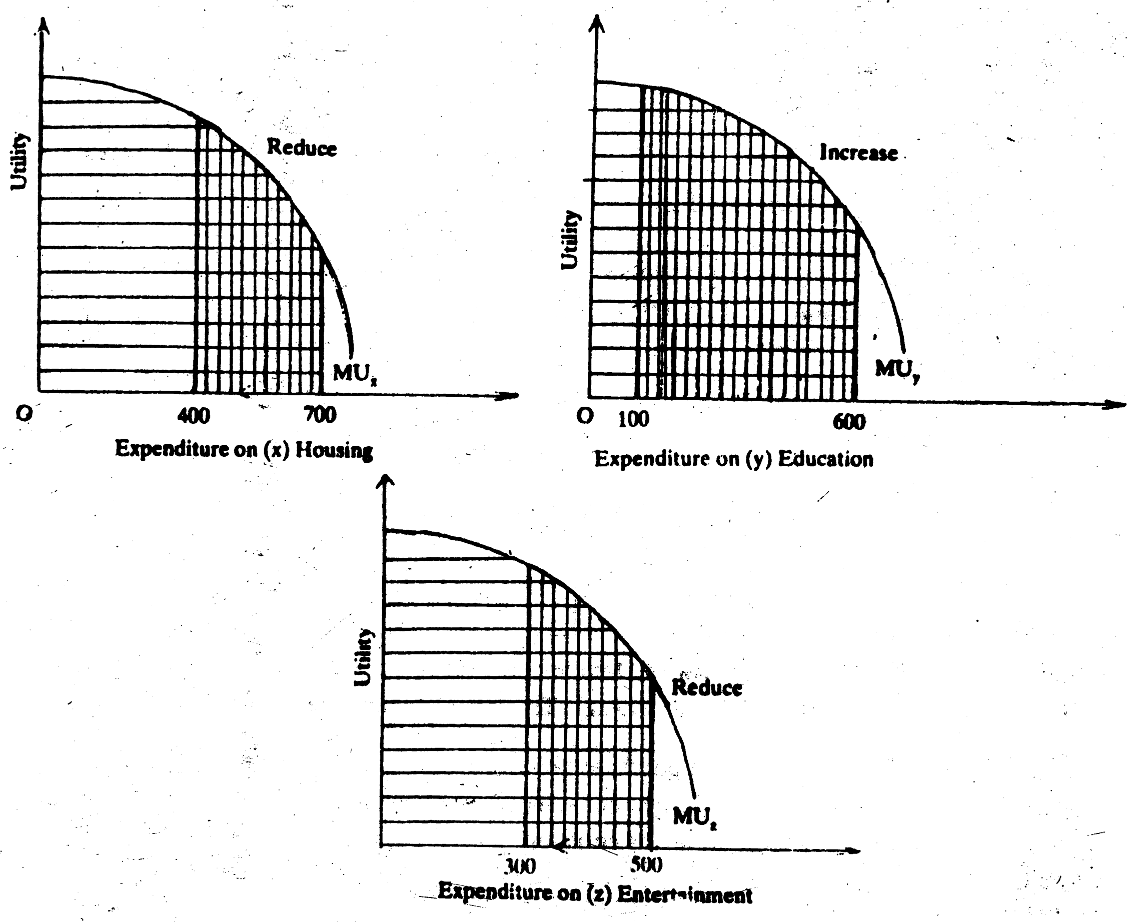
Over the period we are anlysing the purchase behaviour, the consumer's money income is held constant. It is assumed that the consumer consumes only one commodity x say. We also assume that the prices of the commodity remains constant.

It should be noted that anything which yields utility is called *economic good(s)*. Anthing which yields disutility is called *economic bad(s)*. When the consumer maximises utility, we assume that only economic goods are consumed. The consumer's total utility function is given by $U = f(xy)$. This utility function represents the taste of the consumer. You may note that in the utility function the prices of the commodities will never enter. In other words the consumer desired the commodity independent of the price. The price enters when the consumer is making the choice of actual consumption. Thus, the *price enters the demand function for a commodity and not the utility function.* Incidentally the demand function is derived from the consumer's utility function. We shall indicate the derivation subsequently.

The Allocation of consumer's money income between various commodities in order to maximise total utility. This can be explained with the help of the following examples :

Suppose during a given period of time (for instance a month) a consumer has Rs. 1300 as money income and suppose that he spends this income on three goods—goods X (housing), goods Y (education), goods Z (entertainment). Suppose he allocates Rs. 700 to housing, Rs. 100 to education and Rs. 500 to entertainment. The question is whether the consumer is maximising his total utility. The answer is 'No'. In fact by reallocation of Rs. 1300 you can show that the consumer's total utility increases.

As we have stated earlier, in the Marshallian Theory of Demand, the utility functions are independent, separable and additive. In other words, $U_x = f(x)$; $U_y = g(y)$ and $U_z = h(z)$ and total utility $U = U_x + U_y + U_z$. In the intial allocation of Rs. 1300 the consumer's total utility is given by the sum of the areas under the respective marginal utility schedules, corresponding to the amount spent. Graphically it is shown in the figure below :



The alternative way of showing the utility maximising choice is by using the mathematical technique of maximisation.

You will find that if consumer reallocates Rs. 1300 in the following manner the total utility will increase. The expenditure on education ought to be increased from Rs.100 to Rs.600; the expenditure on housing reduce from Rs. 700 to Rs. 400; and expenditure on entertainment reduced from Rs. 500 to Rs. 300. You will find that the increase in total satisfaction in case of education will be greater than the loss of total utility on housing and entertainment. In fact, when the MU of Re. 1 of expenditure is equalised on each and every item, total utility will be maximised.

Suppose the consumers utility function is

$U = u(x)$

and his total income spent (expenditure) on x would be

$\bar{B} = x.P_x$, where his income $\bar{B}$ (budget) is given.

Presumably, the consumer wants to maximise the difference between his Utility (Satisfaction) and Expenditure (Sacrifice). The problem is that of simple maximisation of the function $F = U - \bar{B}$

$$F = U - \bar{B}$$
$$= u(x) - x.P_x$$

The first order condition may be satisfied now :

$$\frac{dF}{dx} = \left[\frac{du}{dx} - P_x\right] = 0$$

The second order condition is

$$\frac{d^2F}{dx^2} < 0$$

Note that x is the decision variable for our single commodity consumer.

The first order condition can be read as an *optimum decision rule* or the condition of consumer's equilibrium. It says that

$$\frac{du}{dx} = P_x$$

or, $MU_x = P_x$ ..... for X consumption

similarly, $MU_y = P_y$ ..... for Y consumption

or, $\frac{MU_x}{P_x} = \frac{MU_y}{P_y} = MU$ of money which is constant by assumption.

This reads that the ratio of MU of x to P_x must be equal to ratio of MU of y to P_y. In other words, the marginal utility per rupee of expenditure must be equialised for both x and y. The condition can also be written $\frac{MU_x}{MU_y} = \frac{P_x}{P_y}$, the marginal utility ratios must equal price ratios, and the budget equation must be satisfied. This is the EQUI-MARGINAL PRINCIPLE in the context of purchase decision.

The sufficient conditions for maximisation are

i) $\frac{\delta^2 F}{\delta x^2} < 0$

ii) $\frac{\delta^2 F}{\delta y^2} < 0$

This would imply that MU of x and y must be declining (i.e. the law of eventual diminishing utility must hold).

Example

A consumer consumes two goods x and y. The MU schedules are given in the following table. Money income of the consumer is Rs. 55. The prices (per unit) are Rs. 5 for x and Rs. 8 for y. The problem is to find the optimal (utility maximising) choice of x and y.

Unit of X	*T.U. of X*	*Unit of Y*	*T.U. of Y*	MU_x	$\frac{MU_x}{P_x}$	MU_y	$\frac{MU_y}{P_y}$
1	35	1	72	35	7	72	9
2	65	2	136	30	6	64	8
3	90	3	192	25	5	56	7
4	110	4	240	20	4	48	6
5	125	5	280	15	3	40	5
6	135	6	312	10	2	32	4
7	140	7	336	5	1	24	3
8	135	8	352	-5	-1	16	2
9	125	9	360	-10	-2	8	1
10	110	10	340	-15	-3	-20	20/8

We find that

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

When the consumer consumes 3 units of x and 5 units of y, the total expenditure on x is 3 × (5) = Rs. 15; while the expenditure on y is 5 × (8) = Rs. 40. Therefore, total expenditure is Rs. 55 which equals total income of the consumer.

The total utility derived (enjoyed by the consumer) is 370 = (90 for 3 units of x plus 280 for 5 units of y).

Suppose, when the consumer is in equilibrium (i.e. consuming three units of x and five units of y), the price of x falls from Rs. 5 to Rs. 3 while the price of y remains unchanged at Rs. 8, and the consumer's money income also remains unchanged at Rs. 55. The question is what will happen to the consumption of x. We recalculate the $\frac{MU_x}{P_x}$ column.

Units consumed	$\frac{MU_x}{P_x}$	$\frac{MU_y}{P_y}$
1	35/3	9
2	10	8
3	25/3	7
4	20/3	6
5	5	5
6	10/3	4
7	5/3	3
8	-5/3	2
9	-10/3	1
10	-5	-20/8

The consumer's new utility maximizing choice is 5 units of x and 5 units of y. [You can verify that the money income of Rs. 55 will all be spent.] Hence the consumer consumes more of x when price of x has fallen. Implying that the quantity demanded of x and the price of x will be negatively correlated. The demand curve for x will be downward sloping.

The demand curve for a commodity is downward sloping because the marginal utility curve is downward sloping. As the price of x falls, the MU of x also falls in the same proportion so that the consumer's equilibrium condition ($MU_x = P_x$) is not disturbed. In order to get that decline in MU of x, the consumption of x must go up as per the law of diminishing utility. Thus with a fall in price of x, the demand for x extends. Hence the law of diminishing marginal utility is the basis of the law of demand.

Activity 2

Study the Total Utility figures given in the following table. A family has a budget of Rs. 34 and has a choice among all three commodities. Explain how the family will allocate its expenditure between the three commodities assuming that it wises to spend all its budget.

Total Utility Schedules

Quantity Consumed (In kg.)	*Cheese (In units)*	*Fish (In units)*	*Meat (In units)*
1	7	8	16
2	13	16	29
3	17	21	41
4	20.5	25	51
5	22	28.5	59
6	25	31.5	65
7	26	33.5	68

The prices are : Rs. 2 per kg. for Cheese Rs. 5 per kg. for Meat
Rs. 4 per kg. for Fish

....

....

....

5.3 THE INDIFFERENCE CURVE APPROACH

In this approach, the basic assumption is that a consumer, when he goes to the market, does not purchase only a single commodity. He actually buys a bundle of commodities, a combination of some selected commodities. For instance, when you visit a vegetable market, you do not buy only one vegetable but other vegetables also; similarly when you visit a grocery shop, you may buy rice, tea, sugar, pulses etc. You always buy a basket of commodities.

The *indifference curve* approach or the preference approach to consumer behaviour assumes that the consumer confronts and exercises the choice. From a given feasible set of consumption bundles, a consumer would choose the one which he prefers most and, therefore, in this approach, one has to start by describing the bundle of commodities and its relation to the preference ordering of the consumer. The consumer is assumed to rank, in terms of preference, all the consumption bundles available in the feasible set. As a *first step,* one has to construct a model of consumer preferences. In the *second step,* the given prices and money income of the consumer have to be considered. As a *third step,* by applying the model of consumer preferences to the feasible set, we arrive at the optimal choice of the consumer. (i.e., a consumer confronts a constrained optimisation problem.)

We assume that a consumer consumes many commodities at the same time. The quantities of various goods is determined by a vector: $X = (x_1; x_2; \ldots x_n)$. The consumption bundle is a collection of n goods. One has to determine the consumer's preference ordering in terms of these consumption bundles. Instead of having x and y, now let us suppose we have two consumption bundles X' and X''.

The preference relation is described as follows :

i) If the consumer prefers or has indifference between X' and X'', then the symbol used is as follows : $X' \geqslant X''$

ii) If the consumer is indifferent betwen X' and X'', it implies $X' = X''$

iii) If $X' \geqslant X''$ and not $X'' \geqslant X'$ it implies $X' > X''$ i.e. X' is strongly preferred to X''.

The symbol [$\geqslant$] refers to either preferred to or indifferent to or sometimes referred to as 'as good as'. This symbol is, therefore, a way of relating pairs of consumption bundles. It describes the consumers' preference-indifference relationship.

Properties of Consumer's Preference Ordering

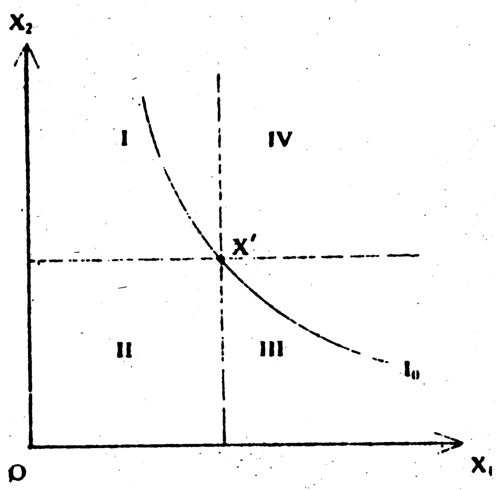
i) Completeness : For any two pairs of consumption bundles, the consumer is able to say whether $X' \geqslant X''$ or $X'' \geqslant X'$ or both. In other words, he must be able to rank all the consumption bundles in the feasible set in terms of preferred to or indifferent to relationship.

ii) Transitivity : Given any three consumption bundles X', X'', X'''. If $X' \geqslant X''$ and $X'' \geqslant X'''$ then $X' \geqslant X'''$

In other words, transitivity implies consistency of consumer behaviour. It implies that if the consumer prefers X' to X'' and also prefers X'' to X''' then it must mean that he must prefer the X' to X'''. Transitivity implies that the consumers' indifference sets are non-intersecting.

iii) Reflexivity : A given bundle is preferred to or indifferent to itself i.e. $X' \geqslant X'$; it implies that each consumption bundle must belong to atleast one indifferent set that is the set containing itself.

iv) Non-satiation : For any two consumption bundles X' and X'', X' is preferred to X'', if X' contains at least more of one commodity and no less of the other. In other words a rational consumer is never satiated with any goods. This implies that a certain relationship exists between the quantities of different goods in different consumption bundles. Take for example the bundle X' There are two goods in the bundle : X_1 and X_2. The geometrical representation of this consumption bundle will be as follows :



Any bundle which lies in the fourth quadrant represents a better bundle than X'. Similarly any bundle which lies in the second quadrant is worse than the given bundle. (Can you guess why? Think in terms of non-satiation assumption.) However we find that the bundles lying in the quadrants I and III cannot be directly compared with a given bundle. As we move from quadrant I to quadrant III we find the consumer is giving up one commodity and acquiring more of the other commodity. Therefore the consumer's indifference set must lie in the first and third quadrants. In other words, all those commodity bundles lying in first and third quadrants must be indifferent to each other.

v) Continuity : This implies that there are no gaps or breaks in the indifference sets. This implies that when you reduce the quantity of a commodity, however small that reduction may be, it is possible to find something of the other commodity which will exactly compensate the consumer and keep him in the same indifference set.

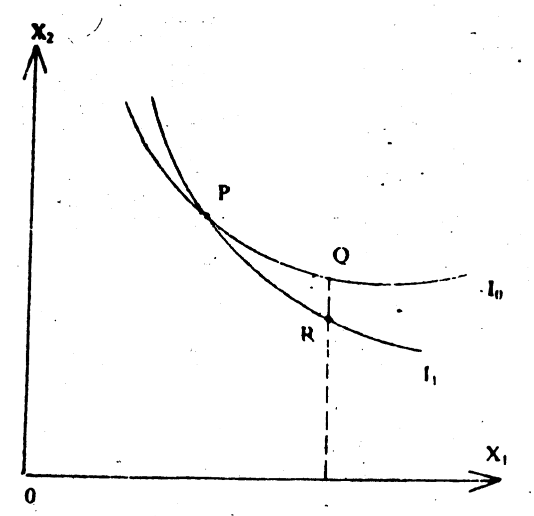
We can therefore define an *indifference curve* as the locus of all those combinations of two goods X_1 and X_2 which are equally preferred by the consumer. The consumer gets the same level of satisfaction or utility from the consumption bundles lying on a given indifference curve. The consumption bundles lying on higher indifference curves will give higher level of satisfaction or utility and, therefore, would be preferred.

An indifference curve has the following properties :

a) It is downward sloping implying that if a consumer has a lesser amount of one commodity he has to be compensated by more and more of the other commodity in order to remain indifferent between the two bundles of commodities. Here it must be pointed out that an indifference curve will never touch the axis. This restriction is due to the fact that a consumer always consumes positive amount of every article.

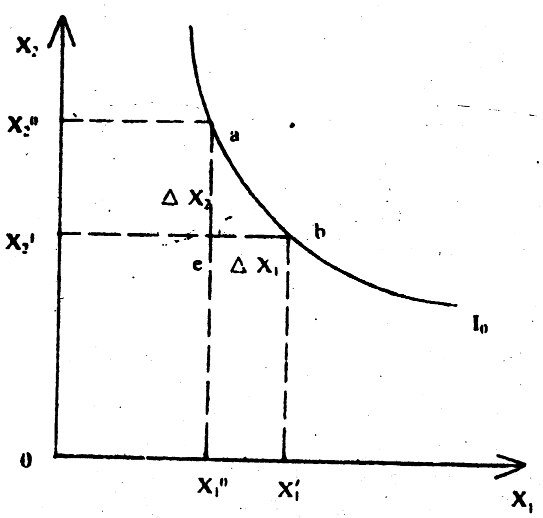
b) The indifference curve reflects the taste of the consumer. Therefore they must not intersect in the positive quadrant, otherwise the transitivity assumption will be violated (implying that the consumer's behaviour is inconsistent). This is explained in the following diagram :

If we compare points P and Q lying on indifference curve I_0 then we find that the consumer will be indifferent between P and Q. Similarly if we compare points P and R lying on indifference curve I_1, we find that the consumer will be indifferent between P and R. Hence transitivity would imply that the consumer must be indifferent between R and Q. But if we compare R and Q we find Q is preferred to R and therefore transitivity assumption is violated. (Since Q has more of X_2 and same amount of X_1, than R).



c) An indifference curve is *strictly convex.* This assumption implies that a small decrease in the amount of X_2 (or to state it differently equal successive decreases in X_2), must be compensated by larger and larger increments of X_1 if the consumer is to remain on the same indifference curve. If the consumer has more and more of X_1 and less and less of X_2, his valuation of the marginal decrease in X_1 relative to the marginal increase in X_2 tends to increase. In other words, the absolute value of slope of the indifference curve tends to increase as we move up the indifference curve or the absolute value decrease as we move down the indifference curve. This assumption also means that any mixture of two commodity bundles between which the consumer is indifferent is preferred to either of them. As the consumer is moving up an indifference curve he has smaller quantities of X_1 and larger quantities of X_2. Therefore, the relative desirability of the two goods would be changing. In the case of successive reduction in the amount of X_1, the consumer needs to be compensated by successive larger increments of X_2, if the consumer is to remain on the same indifference curve.

We can now define an important term called the *Marginal Rate of Substitution* between the two commodities X_1 and X_2. This can be explained with the help of the following diagram :

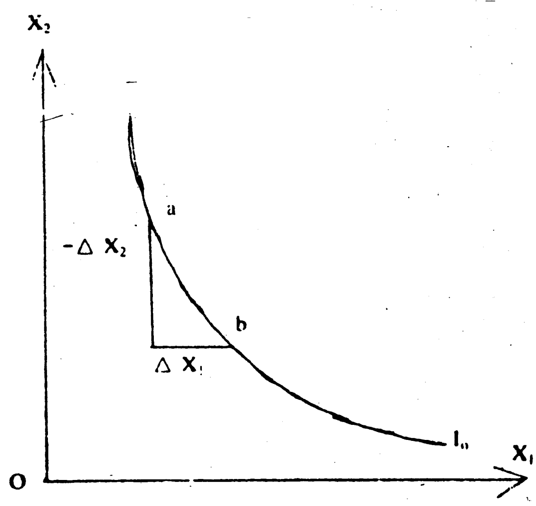


Suppose we consider two consumption bundles a and b. As we move from a to b the consumer is giving up X_2 by $\Delta x_2 = x_2^1 - x_2^0$ and acquiring X_1 by $\Delta x_1 = X_1^1 - X_1^0$.

The consumer is substituting X_1 for X_2. Hence the ratio $\frac{\Delta x_2}{\Delta x_1}$ defines the rate of substitution of the first good for the second good. It shows i.e. the variation in the quantity of the first good necessary to compensate the consumer for a given variation in the quantity of the second good. If we consider infinite variations instead of finite variations we obtain MRS, the marginal rate of substitution between the two goods.

$$MRS\ x_1x_2 = \text{Lim}\ \frac{\Delta X_2}{\Delta X_1} = \frac{dx_2}{dx_1}$$
$$\text{as } \Delta x_1 \rightarrow 0$$

MRS is simply the derivate at a point of the function with a negative slope which expresses the indifference curve. Graphically it is given by the slope of the tangent line at a point on the indifference curve. In fact, MRS is the consumer's psychological rate of substitution between X_1 and X_2 to remain on the same indifference curve. Hence MRS is different at different point on a given indifference curve. As we move from left to right (move down indifference curve), the MRS decreases in absolute value. It must be noted that there are two concepts of MRS, one MRS of X_1 for X_2 and the other MRS of X_2 for X_1. However these two concepts of MRS are symmetrical (one is defined with reference to horizontal axis measuring X_1 and the other is defined with reference to the vertical axis defining X_2). Here it can be pointed out that the concept of MRS is based on the ordinal concept of utility. Even then it can be given a cardinal interpretation as follows :



When the consumer reduces X_2 by ΔX_2 the utility of the consumer will fall by $MU_2 . \Delta Y_2$. On the other hand when he increases consumption of X_1 by ΔX_1 his utility increases by $MU_1 . \Delta X_1$. However, on a given indifference curve total utility is constant (the number that is used to represent total utility is irrelevant)

Therefore

$$MU_2 dX_2 = +MU_1 \Delta X_1$$

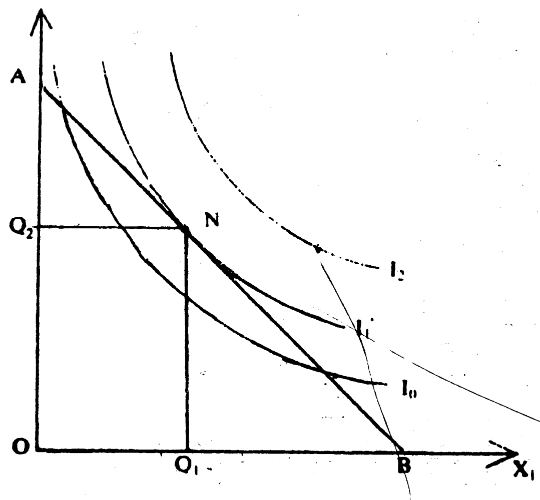
$$\therefore \frac{-dX_2}{dX_1} = \frac{MU_1}{MU_2}$$

$$\therefore MRS_{x_1 x_2} = \frac{MU_1}{MU_2}$$

i.e., MRS is the ratio of marginal utility of the two commodities.

Actual choice of consumption bundle (Consumer's equilibrium)

The consumer's preference scale is described by means of his indifference mapping (i.e. a set of indifference curves). At any given period of time the consumer's money income is given and so are the prices of the goods. The money income and the prices of the goods define the consumer's budget line or price line or real income line.

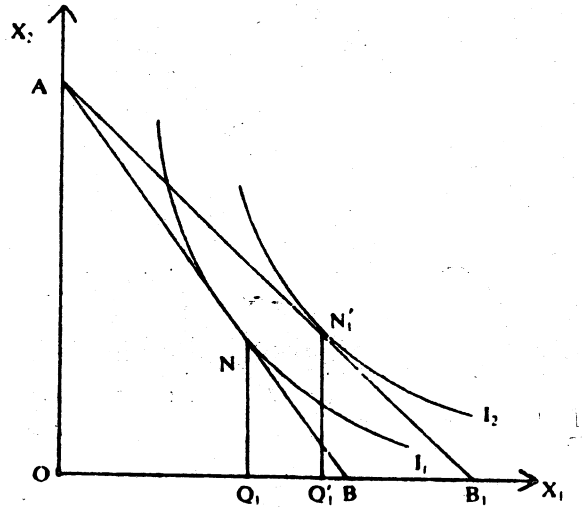


The slope of the budget line is the price ratio of the commodities. The consumer will be maximising utility by operating on the budget line. We super-impose the consumer's preference scale as represented by the indifference mapping on the budget line. The consumer attains equilibrium when he is able to consume the most preferred commodity bundle (which gives him highest utility). In the above diagram this is attained at point N, where the indifference curve I_1 is the highest indifference curve touching tangentially the budget line. The consumer cannot attain the indifference curve I_2 (Why? Think in terms of money income constraint). The utility maximising consumption will be OQ_1 of X_1 and OQ_2 of X_2. Now the characteristics of N is that the indifference curve I_1 is tangential to the budget line implying that their slopes are equal. The slope of the budget line is price ratio $\frac{P_1}{P_2}$, while the slope of the indifference curve is $MRS_{x_1x_2}$. Hence $MRS_{x_1,x_2} = \frac{P_1}{P_2}$.

Since $MRS = \frac{MU_1}{MU_2} \therefore \frac{MU_1}{MU_2} = \frac{P_1}{P_2}$

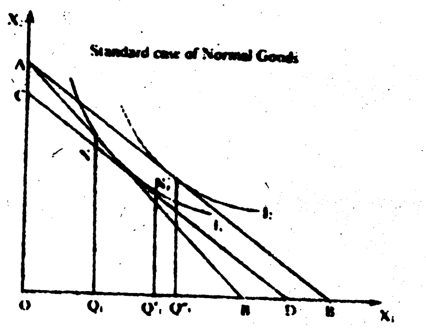
We are back to Marshallian equilibrium condition.

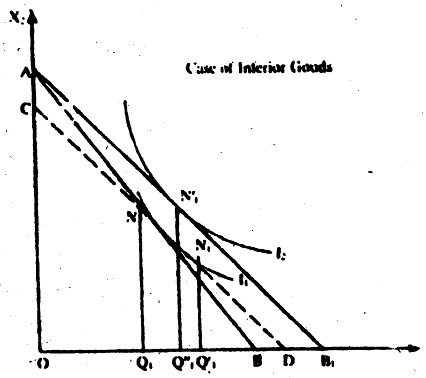
Now we suppose when the consumer is in equilibrium the price of commodity X_1 falls with price of commodity X_2 and money income remains unchanged. The budget line rotates about the point A and becomes flatter. The feasible consumption set expands. The consumer original equilibrium is disturbed. He now moves to a new point of equilibrium N_1 on indifference

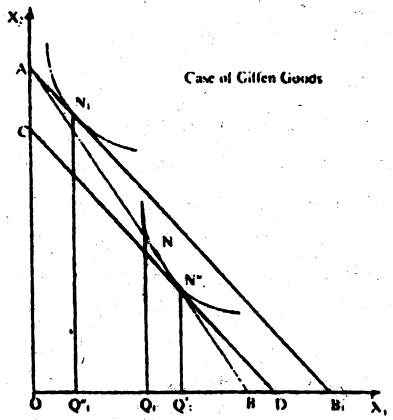


curve I_2. The consumption of commodity X_1 increases from OQ_1 to OQ_1'. Hence we find that consumption of commodity X_1 extends when its price has fallen. The total increase in demand for commodity X_1 can be partitioned into two components real income effect and substitutional effect. In the last unit you have been introduced to these effects. We may now restate them. When the price of commodity X_1 falls, given consumer's money income, consumer's real income will increase. If therefore, we reduce consumer's money income in such a way that the gain in real income is wiped off, then we would be eleminating the real income affect. The increase in consumption of Commodity X_1 will be then due to only substitution effect i.e. due to a change in the relative prices of the commodities. A rational consumer would tend to substitute relatively cheaper goods for the relatively more expensive goods. As a result consumer will always consume more of the goods whose price has fallen. Graphically we show this by drawing an intermediate price line parallel to the new price line and tangential to the original indifference curve. In other words, if the consumer's money income is reduced by AC in terms of commodity X_2 or DB_1 in terms of commodity X_1 then the consumer would be brought back to the original level of utility (which is Hick's way of holding real income constant) but even after the money income is reduced to wipe off the gain in real income the consumer's equilibrium will not be at but N and N_1. Hence the movement from N to N_1 is the substitution affect (by Hick's method). In other words, the increase in consumption of commodity X_1 from OQ_1 to OQ_1' is the *substitution effect.* If the money income deducted is returned to the consumer so that his original money income is restored, the consumer would move from N_1'' to N_1 (from indifference curve I_1 to indifference curve I_2) which according to Hicks is the consumer's real *income effect* of price change. If price ratio is kept constant, the effect of a change in real income can be shown through parallel shift in the budget line. Kindly refer to next page.

It must be pointed out that when consumption of a commodity goes up as a result of increase in real income, we denote such commodity as normal commodity. However, when the consumption of a commodity goes down consequent upon an increase in real income, then such commodities are called *inferior goods*. For *normal goods*, both the substitution effect and the income effect work in the same direction, in other words they reinforce each other. While in case of inferior goods, the two effects move in opposite direction. Hence for normal goods the demand curve will be downward sloping since both substitution effect and income effect would tend to increase the consumption of the goods when price falls. In case of inferior goods, the demand curve may still be downward sloping in case the substitution effect is stronger than the income effect. On the other hand, if the negative income effect dominates the substitution effect then the demand curve will be upward sloping and this happens in case of GIFFEN GOODS (which are inferior goods and for which the negative income effect dominates the substitution effect). These are illustrated in the following diagrams :

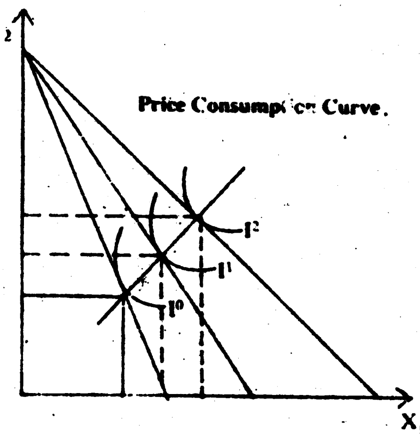


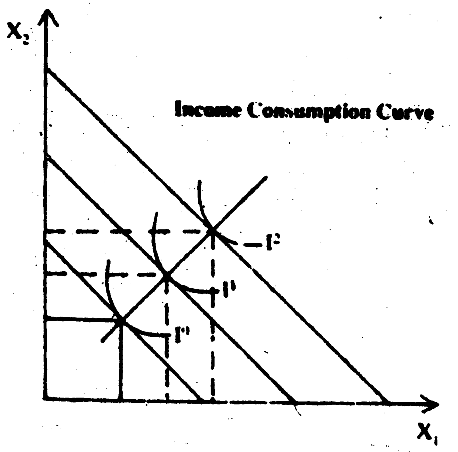


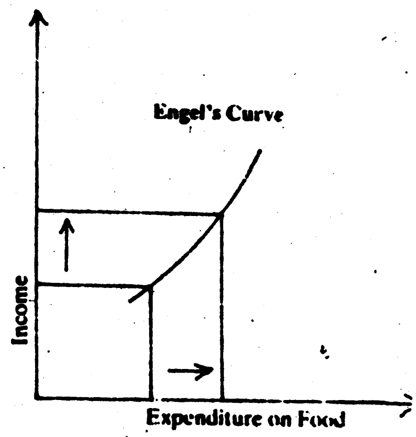


Activity 3

a) Study the [redacted] diagrams and explain each of the curves :







....

....

....

....

b) In what way is the price consumption curve different from the demand curve?

....

....

....

c) Compare and contrast the income consumption curve with the Engle' curve.

....

....

....

We will now move to derive mathematically the optimum choice by a consumer and his demand function therefore, using the Indifference technique.

5.4 DERIVATION OF DEMAND FUNCTION

We start with two functions :
the utility function, $U = u(x,y)$, and
the budget equation, $B = P_x x + P_y Y$

Given $U = u(x,y)$, for indifference curve,

$$du = \left[\frac{\delta u}{\delta x}dx + \frac{\delta u}{\delta y}dy\right] = 0$$

Note that $du = 0$, because an indifference curve is an iso-utility curve, i.e., alongthis curve, the utility level does not change, even if you move from one point to another.

Now, operating du = 0, we have

$$\frac{\delta u}{\delta x}dx = -\frac{\delta u}{\delta y}dy$$

$$\Rightarrow \frac{\frac{\delta u}{\delta x}}{\frac{\delta u}{\delta y}} = -\frac{dy}{dx} = -MRS_{xy} \qquad MRS_{xy} \Rightarrow \text{ slope of the Indifference curve.}$$

Similarly, given $\bar{B} = Px.x + Py.y$, we find

$$x = \frac{\bar{B}}{P_x} - \frac{P_y}{P_x}y$$

or $y = \frac{\bar{B}}{P_y} - \frac{P_x}{P_y}x$

In either case here, we have the equation of the budget line in the slope - intercept for,

$\frac{\bar{B}}{P_x}$ is the intercept on the x - axis.

$\frac{\bar{B}}{P_y}$ is the intercept on the y - axis.

and $\frac{P_x}{P_y}$ or $\frac{P_y}{P_x}$ is the slope of the budget line which is negative.

In the next stage, we may frame the *optimisation problem* for our consumer as either constrained maximisation or constrained minimisation problem, and then derive the optimum decision rule therefrom.

Constrained Maximisation

The consumer intends to maximise utility satisfaction subject to his given budget (income) constraint. Thus the problem is :

Maximise $U = u(x, y)$, subject to $\bar{B} = P_x.x + P_y.y$.

We can state the Lagrangian expression and operate :

$$L = u(x, y) + \lambda[\bar{B} - P_x.x - P_y.y]$$

For first order conditions,

$$\frac{\delta L}{\delta x} = [\frac{\delta u}{\delta x} - \lambda P_x] = 0 \Rightarrow \frac{\delta u}{\delta x} = \lambda P_x \text{ ...... (i)}$$

$$\frac{\delta L}{\delta y} = [\frac{\delta u}{\delta y} - \lambda P_y] = 0 \Rightarrow \frac{\delta u}{\delta y} = \lambda P_y \text{ ...... (ii)}$$

$$\frac{\delta L}{\delta \lambda} = [\bar{B} - P_x.x - P_y.y] = 0 \Rightarrow \bar{B} = P_x.x - P_y.y \text{ .... (iii)}$$

From (i) and (ii) we get

$$MRS_{xy} = \frac{\delta u}{\delta x} / \frac{\delta u}{\delta y} = \frac{P_x}{P_y} \text{ ..... (iv)}$$

This is the consumer's equilibrium condition or his optimum decision rule.

Constrained Minimisation

The consumer intends to minimise the use of budget given his utility (scale of preference) function. Thus, the problem is :

Minimise $\bar{B} = P_x.x + P_y.y$ subject to $\overline{U} = u,(x, y)$

Again, we may operate with another Lagrangian function, L^* with another Lagrangin multiplier λ^*

$$L^* = P_x.x + P_y.y + \lambda^*[\overline{U} - u(x,y)]$$

For first order conditions,

$$\frac{\delta L^*}{\delta x} = [P_x - \lambda^* \frac{\delta u}{\delta x}] = 0 \Rightarrow P_x = \lambda^* \frac{\delta u}{\delta x} \text{ ..... (v)}$$

$$\frac{\delta L^*}{\delta y} = [P_y - \lambda^* \frac{\delta u}{\delta y}] = 0 \Rightarrow P_y = \lambda * \frac{\delta u}{\delta y} \text{ ..... (vi)}$$

$$\frac{\delta L^*}{\delta \lambda} = [\overline{U} - u(x,y)] = 0 \Rightarrow \overline{U} = u(x,y) \text{ ..... (vii)}$$

Operating on (v) and (vi), we get

$$\frac{P_x}{P_y} = \frac{\frac{\delta u}{\delta x}}{\frac{\delta u}{\delta y}} = MRS_{xy} \text{ ......(viii)}$$

As you compare (iv) with (viii), you will observe that you have the same optimum decision rule or equilibrium condition. In either case, the decision variables are x and y, given their prices and income of the consumer. (We have assumed away the second order-stability condition—some of you may like to verify that as well.) We may now attempt an application of this decision principle (equilibrium condition) to locate the optimum choice by a consumer and derive therefrom, his demand function.

Example

Suppose a purchase manager has Rs. 8000 available to be divided between two materials, x and y. The unit price of y is fixed at Rs. 2 by the government. The preference (utility) function of the manager has been estimated based on his firm's technology as u = 3 log x + 9 log y. We have to find the optimum combination of x and y, also the demand function for x.

By applying the rule (iv) or (viii) straight away, we get

$$\left[\frac{\frac{\delta u}{\delta x}}{\frac{\delta u}{\delta y}} = \frac{\frac{3}{x}}{\frac{9}{y}}\right] = [\frac{P_x}{P_y} = \frac{P_x}{2}]$$

$$\text{or, } y = \tfrac{3}{2}x.P_x \text{ ....(ix)}$$

Substituting the result in the budget equation, we have

$$\bar{B} = P_x.x + P_y.y$$

$$\text{Rs. } 8000 = P_x.x + 2[\tfrac{3}{2}x.P_x] = 4x.P_x$$

$$\left.\text{or, } x = \frac{2000}{P_x} \text{ or, } P_x = \frac{2000}{x}\right\} \text{ ....(x)}$$

Either way, we have the ***demand function for x***. Also note, given P_x, we can determine

Optimum choice of x and y combination, because we have in (ix) $\frac{y}{x} = \frac{3}{2}P_x$

Activity 4

1) Derive the demand function for x and y, given

$u = u(x,y) = [xy + x + y]$ and $\bar{B} = P_x.x + P_y.y$.

you should be able to get the following results :-

$$x = \frac{B}{2P_x} + \frac{P_y}{2P_x} - \frac{1}{2} \text{ and } y = \frac{B}{2P_y} + \frac{P_x}{2P_y} - \frac{1}{2}$$

....

....

....

....

5.5 THE REVEALED PREFERENCE APPROACH

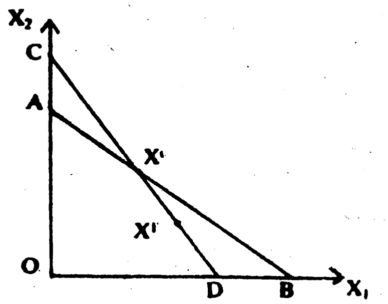
The earlier two approaches are based on the concepts of utility (in one case, it is cardinal and in the other, it is ordinal). However, the concept of utility is non-observable as well as non-measurable. The subjective foundation of the approaches has led to attempts by some economists to develop an alternative approach which is bassed on the behaviour of the consumer in a given market. Since *preference* as well as *utility* are neither observable nor measurable, what we can observe is the behaviour of the consumer i.e. commodity bundles they buy at different price-income situations. This alternative theory is bassed on the assumptions about consumer's behaviour rather than his preference. The basic premise of the theory is that the consumer's choice for a particular commodity or commodity bundle reveals his preference for that commodity or bundle. The assumptions of this new approach are as follows :

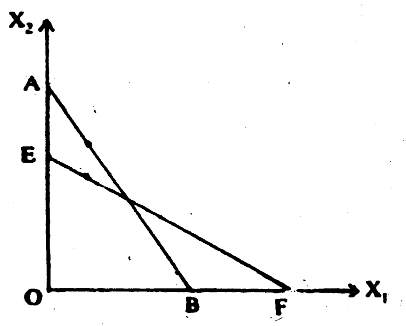
i) The entire money income is spent during the reference period.

ii) Only one commodity bundle is chosen at any price-income situation. If the situation remains unchanged, then the consumer's choice also remains unchanged.

iii) There exists one and only one price-income situation where each commodity bundle is chosen. In other words, X_o cannot be chosen at two different price-income situations.

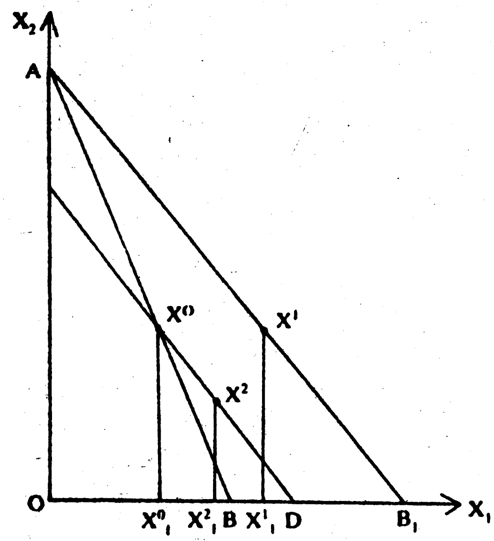
iv) If the consumer choses a bundle X^o when another bundle X' is available and feasible, then X' is chosen, X^o must not be a feasible alternative. This is choice consistency. This is explained graphically as follows :





All commodity bundles on and inside the budget line AB are feasible, but the first assumption necessitates that the choice must be made on the budget line. The choice of X^o reveals the consumer's preference over X'. *Choice reveals preference.*

On the budget line CD, he cannot choose X' because X^o is still feasible. On the budget line EF, X^o is more expensive than X'. Therefore, the choice of X^o in one situation and X' in another situation is not inconsistent. If P^o represents the initial price vector and P' represents the new price vector then the revealed preference axiom would imply (i) $P^oX^o \geqslant P^oX'$ and (ii) $P'X^o > P'X'$. This is known as the *Weak Axiom of Revealed Preference theory* (WARP)



Suppose the consumer's initial situation is as follows:

B_1^o, P_1^o, P_2^o which is represented graphically by means of the budget line AB and consumer's choice is X^o. Suppose next, the price of commodity X_1 has fallen with price of commodity X_2 and money income remaining unchanged. The new budget line AB_1 and on this budget line the consumer's choice is X'. Hence we find that consumption of commodity X_1 extends as its price falls. The substitution effect of the price change will be as follows : We draw the intermediate price line CD through X^o (i.e., the consumer's purchasing power is held constant). That is, at the new prices we reduce his money income till he is allowed sufficient income to buy at least the original basket of goods. The question is where will the consumption point lie on CD? There are three possibilities (i) On CX^o (ii) Coinciding with X^o (iii) On X^oD. In case of (i), the consistency requirement would be violated. It cannot coincide with X^o since the third assumption would be violated. Hence it must lie on X^oD. Hence the consumption point on CD must be X^2. Therefore the movement from X^o to X^2 is the *substitution effect.* While the movement from X^2 to X' is the *income effect.* Thus both the effects reinforce each other and hence the price and quantity demanded are negatively correlated. The demand curve for a commodity is downward sloping. Here it must be pointed out that a positive income elasticity of demand for a commodity is a sufficient condition for the law of demand to be valid in the Revealed Preference theory. For inferior goods, the substitution and income effect will move in opposite directions but in spite of it the law of demand can still remain valid. (Why?)

Activity 5

Show it graphically that it is only in the case of *Giffen Goods* that the law of demand is invalidated.

From the various analyses of demand, we find that the demand for a commodity is determined by the own price of the commodity. In other words, demand is considered to be a function of one variable only, i.e. the price of the same commodity. In reality however, demand for a commodity is a multivariate function. We can restate the other factors or influences on demand as follows :

i) *The price of a related item :* This can be either the price of substitute item or a complementary item. The demand for goods (like wheat) is positively correlated with the price of substitute goods (rice). On the other hand the demand for an item (bread) is negatively correlated with the price of a complimentary item (butter or jam).

ii) *The level of disposable income of the consumer :* This factor has a direct effect on the demand for a goods (except inferior goods which are, of course, very rare in any economy).

iii) *The money supply :* **The relationship between demand for a commodity and money supply is not directly discernible. It works via income creation.**

iv) *The level of advertisement activity undertaken by the industry producing the goods :* A vigorous as well as attractive advertising campaign would tend to increase the demand for a commodity while a bad advertisement may affect adversely. A few other inferences on demand have been examined in the preceding unit.

Since demand is a function of so many explanatory variables it is the job of the econometricians (who are specialists in the empirical measurement of variable) to separate as well as to identify the role of each of the sources of influence on demand. The quality of statistical data or information would play an important role in such cases of estimation.

5.6 SOME RECENT DEVELOPMENTS IN THE THEORY OF DEMAND

The various theories of demand that we have developed in this unit have limited applicability and usefulness in explaining the complexities of the real world. Hence some economists have advocated a pragmatic approach to the theory of demand. Some demand functions have been formulated directly on the basis of market data without reference to the theories of demand that we have studied in this unit. As stated earlier, a demand function is multivariate function and such functions can be estimated statistically. Some recently estimated demand functions refer to *consumers as a group* (not as an individual), and also to *commodity groups,* like demand for food, demand for beverages, demand for consumer durables etc. (rather than individual commodities within a group).

It must be pointed out here that such an empirical estimation of demand function is always beset with difficulties. The foremost problem is the problem of aggregation i.e. aggregation of demand over individuals and over commodities (to arrive at market demand and commodity groups). Index numbers can be used, but there are problems with index numbers as well. There may be estimation problems. One important difficulty arises from the simultaneous change of all the determinants, which would make it difficult to assess the influence of each individual factor separately. Of course, improvements in the econometric techniques are supposed to take care of these problems.

Another recent development is the use of demand function in a dynamic form. Such a dynamic demand function expresses the idea that current purchasing decisions are influenced by past behaviour. For instance it is felt that current purchase behaviour depends on the past levels of income and past levels of demand. In case of durable goods, past purchases constitute a *stock* of this commodity which definitely affects the current and future purchases of such durable goods. In case of non-durables (like food, beverages) past purchases reflect a *habit* which is acquired by buying and consuming the commodity in the past, so that the level of purchases in the previous period influences the current and the future patterns of demand. Another way in which the past behaviour affects the present is that the more recent of past levels of income or demand have a greater influence on present consumption pattern than the more remote ones. For instance, you may be influenced by your income in the last year than by the income earned five years ago. Demand functions which incorporates lagged values of incomes or other variables (prices, demand etc.) are called *distributed lag models.* Such a model may be expressed as follows:

$$Q_t = f(P_t' P_{t-1}; Y_t' Y_{t-1} \text{ ....} Q_{t-1}, Q_{t-2})$$

where Q_t = amount purchased currently i.e. in period t

$\left.\begin{array}{l} Q_{t-1} \\ Q_{t-2} \end{array}\right]$ amount purchased previously in period t-1 and t-2

P_t = Current price i.e. in period t

P_{t-1} = Price in the previous period

Y_t = Current income

Y_{t-1} = Income in the previous period

A widely used model, (both in demand function and in investment function) is the model based on *stock-adjustment principles'* developed by *Nerlove.* The model is applied to the study of demand function of consumer durables. Past purchase (stock) of consumer durables like colour TV determines the likely purchase of other products like VCR. Thus the consumer builds up his stock of durable items.

Houthakker and *Taylor* have extended the stock adjustment principle to non-durable consumer goods, giving it the name, *'habit reaction principle'.* Past purchase reflect the consumer habit, and that influences his current purchases (demand) of non-durables like bread, cigarette etc.

Another recent development have been the *Linear Expenditure System* (LES). Such LES models deal with groups of commodities rather than the individual commodities. Such groups when added yield total consumer expenditure. *R. Stone* was the first to suggest it. The LES are usually formulated on the basis of a utility function from which demand functions are derived in the normal way. In this regard, it is similar to the indifference curve approach. However, the LES differs in that it is applied to groups of commodities between which no substitution is possible. The utility function is additive i.e. total utility is the sum of utilities derived from the various groups of commodities. For instance, suppose that all commodities bought by the consumer are grouped in five categories.

A : Food and Beverages
B : Clothing
C : Consumer durables
D : Household operative expenses
E : Services (Transport, entertainment etc.).

The total utility $U = \sum_{i-1}^{n} U_i$

i.e. $U = U_A + U_B + U_C + U_D + U_E$.

Additivity implies that the utility of the various groups are independent, i.e. there is no possibility of substitution between the goods (Do you note the similarity with the Marshallian theory of demand). Substitution is possible within groups but not between groups. The consumer buys some minimum quantity of each group, irrespective of prices. The minimum quantities are called the 'subsistence quantities' (they are minimum requirements for keeping the consumer alive). The income left is allocated among the various groups on the basis of prices. Thus the income of the consumer is split into two parts — *subsistence income* (spent on the acquisition of minimum quantities of various commodities) and the *supernumerary income* (income left after the minimum quantities are covered). In this approach, the income effect is thus given a new dimension.

Activity 6

a) State clearly but briefly:

i) Stock Adjustment Principle

....

....

....

ii) Habit Reaction Principle

....

....

....

b) What is the basic difference between the above two : (a) and (b) with respect to the analysis of demand determinants?

....

....

....

c) How is 'subsistence income' different from 'supernumerary income'? Why is distinction relevant in demand analysis?

....

....

....

5.7 THEORY OF CONSUMER'S CHOICE UNDER RISK

When you concentrate on a multivariate demand function like the one you observe in a distributed lag model, you will find that price expectations play a very important role as a demand determinant. The price likely to prevail in future is equally important as the prices prevailing in the past and the prices prevailing in the present. When you are considering the role of expected future prices as a determinant of demand, you are considering essentially the risk-and-uncertainty factor associated with the future.

In this context, we would like to make a passing reference to the theory of consumer choice under risk and uncertainty as developed by authors like *Von-Neumann* and *Morgenstren*. According to this theory, if the consumer satisfy certain crucial axioms like *"Complete ordering, Continuity, Independence, Unequal probability* and *Complexity"*, then his utility function can be derived by presenting him with a series of choice between a certain outcome on the one hand and a probabilistic combination of two uncertain outcomes on the other. The utility function thus derived is unique up to a linear transformation and provides a ranking of alternatives in situations that do not involve risk. A consumer maximises expected utility, and these utilities are cardinal in the sense that they can be combined to compute expected utilities and can be used to compare differences in utilities. The expected utility so calculated can be used to determine the consumer's choice and demand decisions in situations involving risk.

1 a) Refer to the book : *Microeconomic theory* by J.M. Henderson and R.E. Quandt (2nd ed.) ch. 2, section 2.9 and define the following axioms :

i) Comple ordering axiom : ....

....

....

ii) Continuity axiom : ....

....

....

iii) Independence axiom : ....

....

....

iv) Unequal probability axiom: ....

....

....

v) Complexity axiom ....

....

....

....

b) Recall the axioms of Revealed Preference. How do they compare with the above set of axioms.

....

....

....

....

....

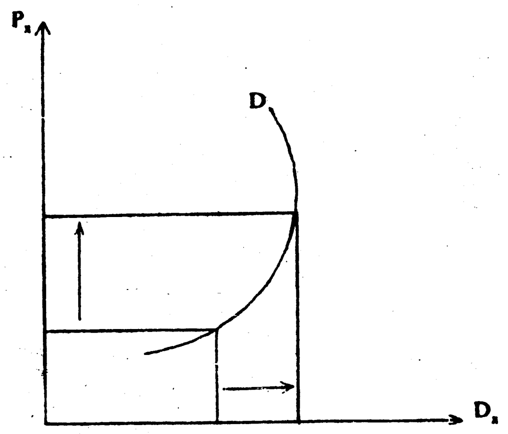
....

5.8 EXCEPTIONAL DEMAND CURVES

You have seen that the normal demand decision of a consumer is to purchase more when the price falls and to purchase less when the price rises. That is how you get downward falling demand curve. However, under certain circumstances, the consumer may purchase less as price falls or may purchase more as price rises. This violates the law of demand and we get exceptional demand curves. We have already indicated such circumstances. Let us now take a stock of these conditions.

1) If the commodity in question is *Giffen type of inferior goods,* with negative income effect outweighing the substitution effect, then you may get a backward bending, demand curve of the

type illustrated here. As price rises, the purchaser's demand extends for commodities like staple food items.



2) In the case of *speculative demand* also, you may encounter such exceptional demand curves. For example, a price-rise may be expected to be a prelude for further rises in future and therefore, people may start buying more. This happens for items of daily necessities subjects to sudden scarcity in the market. It also happens with regard to the demand for shares in the stock market.

3) Another exceptional condition is that of *derived demand.* For example, the demand for labour in an electronic industry is a derived demand. If there is a growing market for electronic products, the demand for labour may go up, even when the wage rate (i.e. the price of labour) goes up.

4) Sometimes, the consumers who are very much concerned about their status, prestige, ego, snob appeal, may not care for price escalation and may continue to buy more. In fact, some of them buy more (of say, diamonds) only because the price is too high and others cannot afford it. Such articles are called items of *ostentatious expenditure* or *conspicuous consumption.*

5) In the case of complementary goods also, you may observe that related prices exercise so much influence that the own-price of the product ceases to be a determinant of demand. When the price of petrol goes up, the demand for driving your own vehicle comes down, even if car prices may be unchanging or falling.

Activity 8

Under each category listed above, think of more examples and put them down.

No.	*Cases*	*Example*
1		
2		
3		
4		
5		

5.9 SUMMARY

In the earlier unit, we introduced you to a set of demand concepts. In this unit, we have attempted a theoretical analysis of various determinants of demand. In the process, we have referred to alternative approaches (theories) like traditional utility approach, modern indifference curve technique and the recent revealed preference approach. We have also made reference to some very recent constructs such as theory of choice involving risk, distributed lag model, stock-adjustment principle and habit reaction principle, and linear expenditure system. An overview of these theoretical developments suggest that for analysing consumers behaviour (demand decision), it is important to distinguish between a single-commodity consumer and multi-commodity consumer, and also between groups of consumers (market demand) and groups of commodities (consumption basket) purchased by a consumer. Additionally, we have also dealt with the distinction between choice and preference, preference intended and actual purchase conditioned by constraints, and between certainty and risk with respect to a consumer's decision environment. The analysis of demand decision is a very complex task; you have to struggle through a lot of theoritical developments in microeconomics. For purposes of managerial economics, you have to go beyond theory and attempt application of some of those hypotheses in real world business situation. From this standpoint, you should try to understand the basic concepts and basic hypotheses before you attempt their application. We shall now take you to concepts of demand forecasting in the next two forthcoming units. Our idea is to take you now from theory to empiricism. The demand functions, statistically estimated and constructed, their introduction and their use as a basis for demand forecasting and forward planning by the firm : all these are now going to be your concern. Surely, such concepts, techniques and measurable magnitude have a lot of relevance for practising managers.

5.10 ADDITIONAL READINGS

Baumol W.J., *Economic Theory and Operations Analysis* (4th ed.)

Kautoyiannis, A., *Modern Microeconomics* (2nd ed.)

Ferguson, C.E., & Gould, J.P., *Microeconomic Theory* (5th ed.)

5.11 SELF-ASSESSMENT TEST

1 Compare and contrast Indifference Curve Approach and Revealed Preference Approach with the traditional utility approach with respect to assumptions, concepts, techniques and decision rules (equilibrium conditions).

2 Review some of the recent developments in demand theory which provide a *pragmatic approach* to the analysis of demand decision variable.

3 Make a list of factors which may determine the demand for,

   a) engineers in a construction company
   b) a consumer durable items like refrigerator
   c) an intermediate goods like cables
   Analyse the common factors.

4 Draw the following :

   a) an Indifference Curve illustrating the trade-off between leisure and labour;

b) an Engel Curve relating income and essential expenditure on food;

c) a normal demand curve derived from the price-consumption curve; and

d) an exceptional demand curve.

5 Why do people purchase more of the following items when their prices go up?

a) Kerosene

b) Car accessories

c) Sweets

d) Shares

[Name the relevant demand hypothesis you have used as an aid to your explanation].

UNIT 6 DEMAND ELASTICITIES AND DEMAND ESTIMATES

Structure

6.0 INTRODUCTION

In the previous two units, you have been introduced to demand analysis, wherein you have been told about various influences on the demand for a commodity as well as the alternative ways of explaining consumer's behaviour. In particular, you have been exposed to the law of demand, and the demand function (curve) for a commodity. The demand schedule and the demand curve indicate the direction of change of demand when price varies. Such a demand curve, individual or market, would in no way indicate how responsive is the quantity demand to a change in the price. The intensity with which demand reacts to price change cannot be captured by the demand curve or demand schedules. Hence we need to introduce a very significant concept. viz *elasticity of demand.* This concept and measure of demand elasticity has a lot of practical use for business manager. Statistically estimated, empirical demand functions give us information on elasticities. Such management information lies at the root of corporate planning and business policy decisions.

6.1 OBJECTIVES

After reading this unit, you should be able to:

- *define* the concepts of elasticity
- *measure* the elasticity coefficients
- *classify* different types of demand elasticity
- *locate* the determinants of demand elasticity
- *interpret* empirical demand functions

6.2 MEANING OF ELASTICITY

The concept of elasticity in economics is actually borrowed from physics. In physics, it is supposed to show the reaction of one variable with respect to a change in the other variable on which it is dependent. *Elasticity is an index of reaction.*

In economics, we define the demand elasticity of a commodity with respect to its price because demand depends on price. We use the concept of (price elasticity of demand) to measure the degree of responsiveness of the quantity demanded with respect to a change in the *own price* of the commodity. It indicates the extent to which demand changes when price of the commodity changes. Formally, it defined as the ratio of the relative variations in the price. In other words, price elasticity of demand is a ratio of two pure numbers; the numerator is the percentage change in quantity demanded and the denominator is the percentage change in the price of the commodity. In fact, instead of percentage change, one can also take proportinate change. Denoting elasticity by e, we have

$$e = \frac{\Delta Q}{Q} \div \frac{\Delta P}{P}$$

where Δ represents any incremental change in Q and P.

If we use infinitesimal (very very small) variations and employ the calculus of derivative, the above formula becomes.

$$e = \frac{dQ}{dp} \cdot \frac{P}{Q},$$

Three points must be noted at this stage :

1) You may observe that price elasticity e thus becomes a ratio of marginal demand $\frac{dQ}{dP}$ to average demand $\frac{Q}{P}$, should we say that elasticity is an extension of our concepts of incrementalism and marginalism.

2) Elasticity is a unitless or dimensionless concept. It is just a pure number. Here it can be pointed out that one can use the slope of the demand curve to express how quantity changes when price changes, since the slope measures the rate of change of one variable in relation to another variable. But slope is not a pure number. It is dependent on the units in which the variables are measured. If we change the unit of measurement of the variable, the slope will change (for instance kg. changed into gms.). Hence one cannot use the slope of demand curves to compare across commodities, when they are measured in different units. We need a measure which is free from the units of measurement; elasticity is such measure. Consider the relative variation of the quantity $\frac{\Delta Q}{Q}$. We cannot alter the ratio simply by changing the unit of measure as both the numerator and the denominator are expressed in the same unit. The same is true of $\frac{\Delta P}{P}$. Hence elasticity is a pure number.

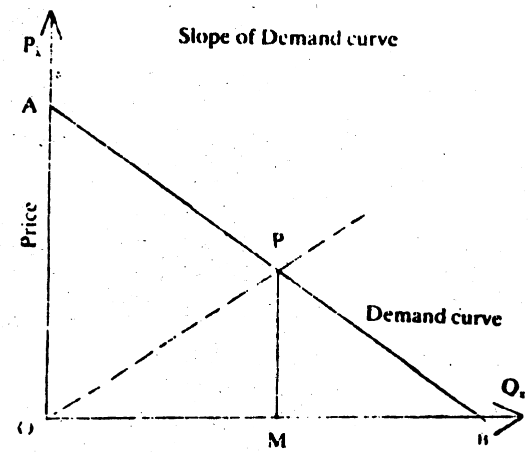
Consider a linear demand curve. The slope of the demand curve is dP/dQ which is in the following diagram PM/MB. Elasticity at point P on the demand curve is

$$e = \frac{dQ}{dP} \cdot \frac{P}{Q}$$

or

$$e = \frac{\frac{P}{Q}}{\frac{dP}{dQ}} =$$ slope of the ray from the origin through point P: slope of demand curve.

Hence $e = \frac{\frac{PM}{OM}}{\frac{PM}{MB}}$



3) The coefficient of elasticity is ordered according to absolute value as opposed to algebraic value. Hence an elasticity of -2 is greater than an elasticity of -1 even though algebraically

the opposite would be true Why does elasticity have a negative sign? If the demand curve slopes downwards from left to right, it follows that if P is greater than zero (i.e. price increases), demand must necessarily be less then zero i.e. the quantity demanded must contract. Hence as long as the demand curve slopes downward, e will always have a negative value. If you keep this in your mind, you can ignore the negative sign.

Activity 1

a) How is the slope of demand curve different from price elasticity of demand?

....

....

....

....

....

....

....

6.3 COMPUTATION OF ELASTICITY COEFFICIENTS

You may use two measures of elasticity :

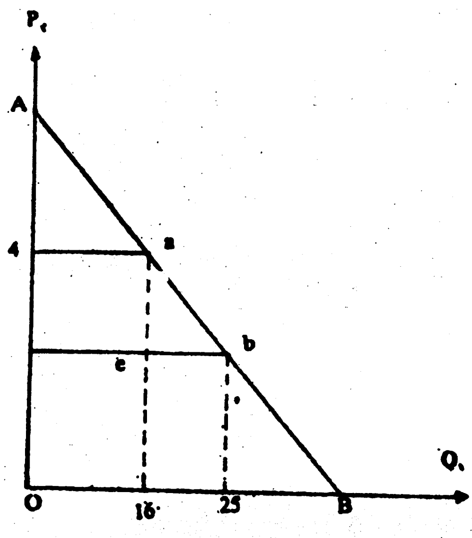
a) ARC ELASTICITY if the data is discrete and therefore incremental changes are measurable.

b) POINT ELASTICITY if the demand function is continuous and therefore only marginal changes are calculable.

Example

Let us see how one can calculate elasticity when the price change is finite (i.e. elasticity measured over a finite stretch of demand curve). The price and quantity situations are given in the following table. We want to calculate easticity when price changes from Rs. 4 to Rs. 3 per unit.

Price of Commodity X (in Rs.)	*Quantity demanded of Commodity X* (in Kg.)
5	10
4	16
3	25
2	30
1	34



When price changes from Rs. 4 to Rs. 3, ΔP = 3−Rs. 4 = −Re. 1.00 (i.e. the price change is negative since it is a price fall). The change in quantity demanded is ΔQ = 25-16 = 9 (Quantity change is positive)

$$e = \frac{\Delta Q/Q}{\Delta P/P} = \frac{9/16}{-1/4} = -9/4 = -2.25$$

Now if we calculate the elasticity when price increases from Rs. 3 to Rs. 4 we find that for the same stretch of the demand curve, elasticity would be different.

$$e = \frac{\Delta Q/Q}{\Delta P/P} = \frac{-9/25}{+1/3} = -9/25 \times 3 = \frac{-27}{25} = -1.08$$

The question is, how is it that we get different demand responses for the same range of price change? The answer is that our initial quantity demanded and price have been different. When we calculate for price fall, they are 16 for initial quantity demanded Rs. 4 for initial price. When we calculate it for price rise they are 25 for initial quantity demanded and Rs. 3 for initial price. Hence elasticity tends to depend on our choice of the initial situation. However, demand response should be the same for the same finite stretch of the demand curve. To get rid of this dilemma created by the choice of the initial situations, we take the arithmatic mean of the two quantities Q and the mean of the two prices P. This gives us a concept of ARC elasticity of demand.

$$\text{ARC elasticity} = \frac{\Delta O}{Q_o+Q_1} \quad \frac{P_o+P_1}{\Delta P}$$

$$\text{or, } e = \frac{\Delta Q}{\Delta P} \quad \frac{P_o+P}{Q_o+Q_1}$$

Where Q_o and Q_1 are the two quantities corresponding to the two points on the demand curve. Similarly P_o and P_1 are the two prices.

Activity 2

You can verify in the above numerical that the elasticity will have the same value whether you consider a price fall from Rs. 4 to Rs. 3 or price rise from Rs. 3 to Rs. 4.

....

....

....

ARC elasticity is based on the notion of average. When we make the ARC small (for non-linear demand curves), the arc elasticity tends to point towards elasticity (the elasticity which we considered to start with). In other words, the limit of arc elasticity as ΔP tends to zero IS point elasticity.

$$\text{i.e. } \lim_{\Delta P \to O} \frac{\Delta Q}{\Delta P} \quad \frac{Q_o+Q_1}{P_o+P_1} \quad = \frac{dQ}{dP} \cdot \frac{P}{Q}$$

For an infinitesimal (very very small) change in price we use point elasticity. However for a finite change in price (however small that change may be), one must always use arc elasticity formula.

Another Example

Consider the following demand function Q=10−0.5P, you are asked to find elasticity at P = 10. The formula to be used is point elasticity i.e. $e = \frac{dQ}{dP} \cdot \frac{P}{Q}$

The first step would be to calculate $\frac{dQ}{dP}$ of the demand function (which is the first derivative of the demand function).

Given Q = 10—0.5P.

$\frac{dQ}{dP} = 0.5$, and $\frac{Q}{P} = \frac{5}{10}$, because when P = 10; Q will be = 10—0.5 (5) = 5

Now $e = \frac{dQ}{dP} \cdot \frac{P}{Q}$

$e = -0.5\ (10/5)$

$e = -0.5\ (2)$

$e = -1.00$

Activity 3

a) Given the demand function, O = 20—0. 4P, find the elasticity when P is Rs. 5.

....

....

....

b) Given the following table find elasticity of demand when price falls from Rs. 8 to Rs. 7. Remember that it is a finite change in price, hence ARC elasticity must be used.

Price of wheat (in Rs.)	*Quantity demanded of wheat (in Kg.)*	*Price elasticity of demand*
10	100	
9	125	
8	140	
7	160	
6	170	

You may wonder what do these calculated coefficients of elasticity really mean. You have to interpret the value coefficients.

Interpretation of Elasticity Coefficients

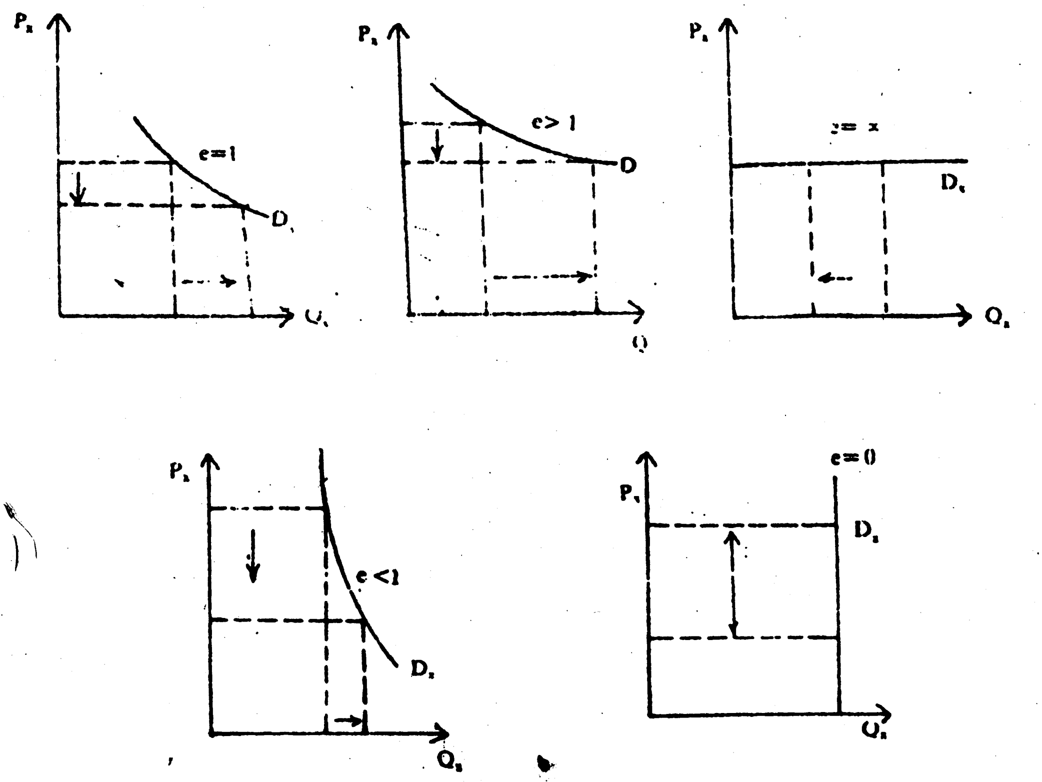
The sign only shows the direction of change. The negative sign shows that price and demand move in opposite direction. Elasticity shows the proportion of change. Suppose the value of elasticity is -1. To start with you may ignore that negative sign. The interpretation of e = /1/ would be that a 1% change in price would lead to a 1% change in quantity demanded. If price falls by 1%, demand would extend by 1%; if price rises by 1% demand would contract by 1%. In this case, demand is said to be of unitary elasticity.

Suppose e = —2. A 1% change in price would lead to 2% change in quantity demanded. In this case, demand is said to be elastic. The consumers respond by buying significantly more when price falls and buy significantly less when price rises. Whenever, demand changes more than price (in opposite direction) demand is said to be responsive and hence demand is said to be relatively elastic (i.e. elasticity is greater than one).

Suppose e = α. It means that even if there is no change in price, demand changes significantly

demand is said to be infinitely elastic.

Suppose e = −0.08. A 1% change in price leads to a less than 1% (in fact 0.8%) change in the quantity demanded. So long as the demand curve is downward sloping, the buyers will buy more when price falls, and buy less when price rises. However in this situation the response is not significant. When price falls, the extension in quantity demanded is not significant. Similarly, the contraction in demand may be insignificant (very small) when price varies. The demand is said to be relatively inelastic or less elastic in this case. In the extreme case, *suppose e is zero* i.e. demand does not change at all, even when price changes. Demand is perfectly inelastic (totally unresponsive) to change in price. In terms of diagram, we may illustrate each of these cases :



Activity 4

Complete the table below :

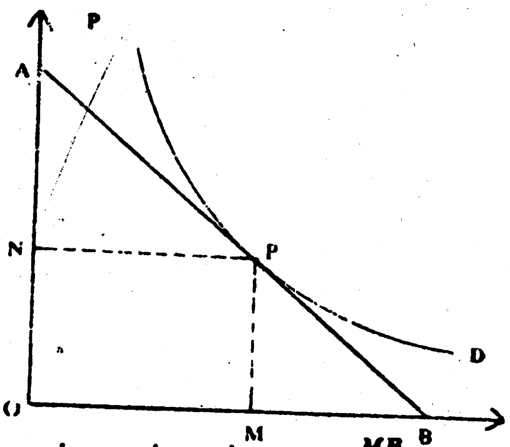
Unitary elastic	$\%\Delta Q = \%\Delta P$	
Relatively elastic		e>1
Perfectly elastic	$\%\Delta P = 0$	
Relatively inelastic		e<1
Perfectly inelastic	$\%\Delta Q = 0$	

6.4 MEASUREMENT OF ELASTICITY

In what follows we will show two methods of measuring elasticity :

Geometrical (diagrammatic) method and arithmetical (tabular) method.

i) A geometrical way of measuring the elasticity at any point on a demand curve is now in order.



Consider point P on the demand curve D_x in the adjacent diagram (we have taken a non-linear demand curve). Draw a tangent line AB at point P on the demand curve. Applying point elasticity formula, it follows the elasticity at point P is

$$e = \frac{dQ}{dP} \quad \frac{P}{Q} \quad = \frac{\frac{dQ}{dP}}{Q/P}$$

$\frac{dQ}{dP}$ is the inverse of the slope of the demand curve, hence is equal to $\frac{MB}{PM}$. Price is equal to PM and quantity is equal to OM.

$$e = \frac{\frac{MB}{PM}}{OM/PM} = \frac{MB}{PM} \quad \frac{PM}{OM}$$

$$= \frac{MB}{OM}$$

In other words, the price elasticity of demand is measured graphically by the ratio of the two segments of the horizontal axis identified by the intersection of the tangent to the point considered, with the horizontal axis and by the perpendicular from that point to the same axis.

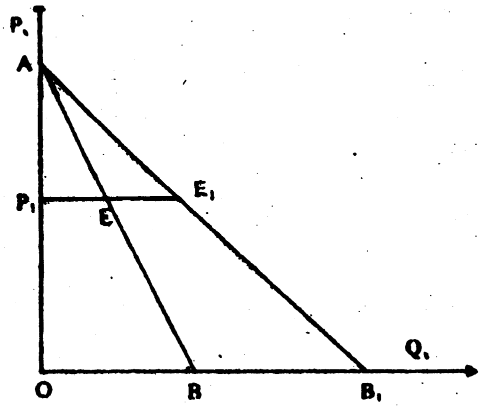
If we now consider the similar triangles APN and PBM, then AP/MB (from properties of similar triangles) or MB/PN = PB/AP. Hence elasticity = MB/OM can be written as equal to PB/AP, i.e. elasticity at P is also equal to PB/AP the ratio of the lower segement of the demand curve to the upper segment.

In the same way, you can show that elasticity is equal to ON/NA (taking again similar triangle and equating the ratio of sides). Obviously, if the demand curve were linear, this method would be further simplified as the tangent to the curve at a point would coincide with the curve itself.

The main advantages of measuring elasticity geometrically is that it allows us to grasp the *law of varying elasticity* straightway, as we move along the demand curve. It can be seen from the previous diagram that the ratio PB/PA falls as you move down the demand curve, since PB decreases and PA length increases. At point A, PA is zero hence in the limit, elasticity is infinite (PB/0). At point B, elasticity is zero since PB is zero. Hence elasticity falls from infinity to zero as you move down a linear demand curve.

Activity 5

It is pointed out that two demand curves originating from the same point on the price axis will have the same elasticity at a given price. They are called *isoelastic demand curve*. Concentrate on the following diagram and explain why price elasticity is the same along the two curves.



Hint

Compare elasticity at points E and E_1 by the ratio method outlined above.

....

....

....

....

....

ii) Another method which can be used to measure the price elasticity of demand is the ***Outlay method*** or the total expenditure method or the total revenue method.

Consider the following table :

Case No.	*Price (Rs.) (P)*	*Quantity (Q)*	*Total revenue spent on Expenditure R = (P × Q)*	*Elasticity (e)*
I	10 5	4 8	40 40	e = 1
II	10 5	4 12	40 40	e > 1
III	10 5	4 6	40 30	e < 1

Activity 6

Instead of a price fall, now assume a price rise which is followed by contraction of demand, as shown in the table below. Fill in the blanks.

Cases	*Price (Rs.)*	*Quantity*	*Expenditure*	*Elasticity*
I	50 100	120 60		
II	50 100		6,000 4,800	
III		120 90	6,000 9,000	

You may now observe that :

i) **If the elasticity is unitary (equal to one), a modest variation in price in either direction will have no effect on the total expenditure of consumers. Total expenditure on the commodity will remain constant whether prices rise or fall.**

ii) If the elasticity is greater than of unity, the total expenditure of consumers will increase following a modest fall in price and total expenditure will fall when price increase.

iii) If the elasticity is less than of unity, a slight fall in price will lead to a fall in total expenditure; and a slight rise in price will lead to a rise in total expenditure.

However, you must note that in outlay method, you can only get some idea about the direction of elasticity (whether equal to one, greater or less than one) but the precise value of elasticity would remain indeterminate.

Activity 7

a) Given the following table on price and quantity, determine whether demand is elastic, unitary elastic, or less elastic when price rises from Rs. 7 to Rs. 8.

Price of commodity gasoline (Rs. per litre)	*Quantity demanded of gasoline (in litres)*
10	100
9	120
8	150
7	175
6	180

b) Why does the Government prefer to levy an excise tax on commodities which are less elastic in demand (like gasoline)?

....

....

....

You may now develop an important relationship between the price of the commodity (average revenue), marginal revenue and price elasticity of demand. You may recall that Marginal Revenue (MR) is the extra revenue from extra units of the commodity sold i.e. dR/dQ. On the other hand, Average Revenue (AR) is the price itself, because

$AR = \frac{TR}{Q} = \frac{PQ}{Q} = P$. Thus you may restate :

$TR = PQ$

$dR/dQ = d/dQ(PQ)$

$dR/dQ = P + QdP/dQ = P\{1 + \frac{Q}{P} \cdot \frac{dP}{dQ}\}$

$\frac{dR}{dQ} = P\{1 + \frac{1}{\frac{dP}{Q}} \cdot \frac{P}{Q}\}$

$\frac{dR}{dQ} = P\{1 + \frac{1}{e}\}$ Since $e = \frac{P}{Q}\frac{dQ}{dP}$,

or, $MR = P\{1 + \frac{1}{e}\}$ or $AR(1 - \frac{1}{e})$ ...

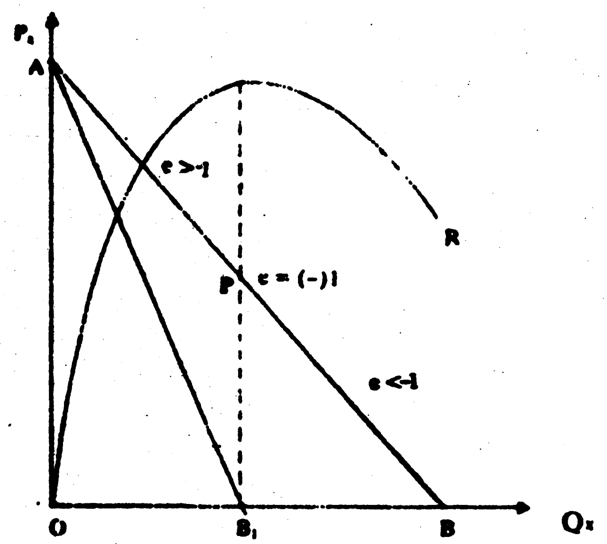
since P is same as Average Revenue (AR) and e is negative by the property of law of demand.

Or, $P = \frac{e}{e-1}$ MR

Or, $e = \frac{P}{P-MR} = \frac{AR}{AR-MR}$

This is a very important measure which has a lot of operational use for decision making.

If you take a linear demand curve, then at the mid point of such a demand curve elasticity will be equal to one (absolute value). When e = 1, implying that MR = 0 and if MR = 0, total revenue will be at and stationary (neither increasing nor decreasing,) is maximum . Abo·: P. the midpoint, e will exceed one. Hence MR will be positive implying that the total revenue is increasing. Similarly below the mid point, e will be less than one. Hence MR is negative implying that total revenue is decreasing. This is shown graphically as follows :



Activity 8

a) Using $e = \frac{AR}{AR-MR}$

explain why at

$A \Rightarrow e = \alpha$ ....

$P \Rightarrow e = 1$ ....

$B' \Rightarrow e = 0$ ....

b) Consider the following data and then fill in the blanks.

Price of X (in rupees)	*Quantity of X (in units)*	*Total revenue (in rupees)*	*Average revenue (in rupees)*	*Marginal revenue (in rupees)*	*Price elasticity coefficient*
10	100	1000			
9	140				
8	190				
7	210				
6	220				

Now consider the demand curve in a hyperbolic form

$Q = aP^{-n}$ or $Q = \frac{a}{P^n}$

where a and n are constants.

You may operate on it to estimate the price elasticity of demand. You know that the measure of point elasticity is

$e = \frac{dQ}{dP} \frac{P}{Q}$

From the above function,

$\frac{dQ}{dP} = n\, a\, P^{n-1}$ and

$\frac{P}{Q} = \frac{P}{aP^n}$

Thus $e = \frac{dQ}{dP} \cdot \frac{P}{Q} = n\, a\, P^{n-1} \cdot \frac{P}{aP^n} = n$

Thus the power term of P is the measure of constant elasticity. It says if P changes by 1%, the quantity demanded of Q will change by n%.

You may also note that if you take a log-linear transformation of the above function,

$Q = aP^n$

You get,

Log Q = log a + n log P

Thus /n/ may appear as an estimated regression coefficient which measures elasticity. This is useful in the context of interpreting statistical demand functions.

6.5 TYPES OF ELASTICITIES

So far, we have considered only *direct price elasticity of demand.*

You may remember that demand depends not only on own price but also on other factors. For example, it may depend on the prices of related goods like coffee and tea (substitutes) or coffee and sugar (complements). In order to evaluate the effect of variations in the price of a product (tea) on the quantity demanded of another product (coffee) we define *cross (price) elasticity of demand* as follows :

The sign of e_{ij} reflects the relation which exists between the two goods. In general we have:

i) $e_{ij}>0$ then i and j are substitute goods. In other words an increase in the price of jth goods (say tea) increases the quantity demanded of the ith goods (say coffee). Hence the sign of cross elasticity will be positive (because of positive correlation between changing price of the jth commodity and units of the ith goods demanded).

ii) $e_{ij} = 0$ if and only if i and j are independent goods (not related).

iii) $e_{ij}<0$ implying i and j are complementary goods. An increase in the price of the J_{th} goods (sugar) reduces the quantity demanded of the ith goods (tea). Hence the relationship is negative correlation, thus cross elasticity is negative.

The greater the absolute value of e_{ij}, the more intense is the relationship (of whatever kind) existing between two goods.

c) Another type of elasticity is *income elasticity of demand.* This measure the degree of responsiveness of quantity demanded of commodity or goods with respect to a change in the level of income of a consumer, other things remaining constant (like prices etc.) It is given by the ratio $\frac{\Delta Q}{Q} / \frac{\Delta R}{R}$

Where R stands for total income (expenditure). It is the ratio of percentage (or proportionate) change in quantity demanded to a percentage (proportionate) change in consumer's income. For a finite change in income we use *arc elasticity* formula, and for an infinitesimal change in income *point elasticity* is used.

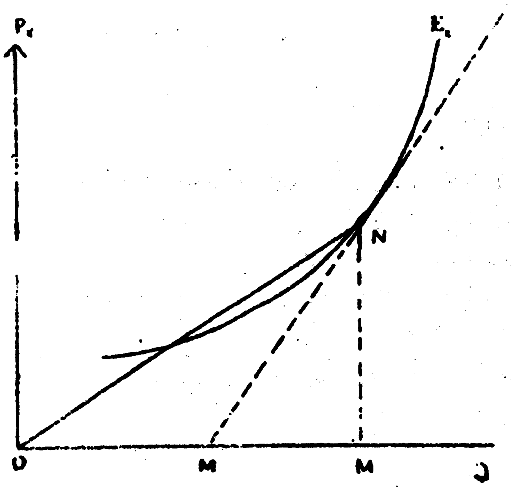
Income elasticity $e_r = \frac{\Delta Q}{Q_0+Q_1} \quad \frac{R_0+R_1}{\Delta R}$

Activity 9

a) At successive stage of income change, calculate income elasticity of demand.

Income of consumer (R) (Rs.)	*Quantity demanded of ready made shirts (in units)*	*Income elasticity (e_r)*
500	100	
1000	150	
2000	185	
3000	225	
4000	300	

b) Determine the income elasticity at point N in the following diagram.



Where E_x is the *Engle's curve for commodity* X (relating income and consumption of X). Use the point method to determine income elasticity at N. Is it equal to one, greater than one or less than one?

....

....

....

....

Few points about income elasticity must be noted :

i) Unlike the price elasticity of demand which is always negative, income elasticity is always positive except for inferior goods.

ii) The value of income elasticity provides us with information regarding the class of goods in question. For necessary goods, income elasticity is less than one; luxury goods will have income elasticity greater than one; for inferior goods, it is negative. Normal goods are goods for which income elasticity is positive but less than one. You should be able to establish it logically if you remember the concept of *income effects* on demand.

iii) Income elasticity varies along a given demand curve. In general, empirical evidence tends to suggest that for a wide range of goods, income elasticity of demand is a decreasing function of income, which means that income elasticity falls as income increases. This is due to the fact that in general, individuals tend to modify their consumption pattern rather than increase their purchases of the same goods, when faced with increasing levels of income.

iv) The weighted sum of income elasticity of demand for various goods must add up to one. Let us suppose that the consumer's income increase by 20%. His consumption of some goods may increase by more than 20%, that of other goods may increase by less than 20%, and that of yet others may actually fall. Since by hypothesis, the entire income must be all spent, there will exist some forms of compensation between groups of goods. For instance for two goods

$$m_1\theta_1 + m_2\theta_2 = 1$$

Where θ_1 and θ_2 are income elasticity of demand for goods one and two, but m_1 and m_2 are the proportion of income spent on goods one and two.

$$m_1 = \frac{P_1X_1}{R} \text{ and } m_2 = \frac{P_2X_2}{R}$$

d) Elasticity of Goods Substitution measures the ease with which one product can be substituted for another in consumption. It determines the rate of substitutability between two goods. It can be measured at a point on the indifference curve as

$$\delta = \frac{\Delta(X_2/X_1)}{X_2/X} \quad \frac{\Delta(dx_2/dx_1)}{dx_2/dx_1}$$

where X_2 and X_1 are, if you remember indifference curve approach, the quantities of two goods consumed $\Delta(\frac{x_2}{x_1})$ is the variation in the ratio between the consumption of two goods. $\frac{dx_2}{dx_1}$ is the MRS (Marginal Rate of Substitution and) $\Delta(\frac{dx_2}{dx_1})$ is the variation in MRS. In other words, δ like all elasticities is the ratio of relative change in the ratio between consumption of the two goods to the relative change in the MRS. It would vary between zero and infinity.

e) Elasticity of Demand with respect to Advertisement is the ratio of percentage change in the quantity demanded of a commodity (Q) to percentage in the advertisement outlay on the commodity (A). It is also called ***promotional elasticity*** of demand or advertisement elasticity. It plays a very important role in the context of marketing management. You may have a point elasticity measure of it as.

$$e_a = \frac{dQ}{dA} \quad \frac{A}{Q}$$

f) ***Elasticity of Price Expectation*** is defined as the percentage change in the level of future prices (P_{t+1}) expected as a result of a change in the level of current prices (P_t). Formally it is

$$E = \frac{dP_{t+1}}{dP_t}\frac{P_t}{P_{t+1}}$$

It measures the ratio of the percentage rise in expected future prices to the percentage rise in its current price. When an increase in current price is expected to result in an equi-proportional increase in future prices then E = 1. If increases in future price is more than proportional to current price rise, then E is greater than one, for less than proportional increase, E is less than one.

Activity 10

a) If you complete the following table, as an exercise, you should be absolutely clear about the concepts and measure of different types of elasticities :

Values / *Concepts*	*Zero*	*One*	*Infinity*	*Greater than one*	*Less than one*
e	$\frac{dQ}{Q} = 0$	$\frac{dQ}{Q} = \frac{dP}{P}$	$\frac{dP}{P} = 0$	$\frac{dQ}{Q} > \frac{dP}{P}$	$\frac{dQ}{Q} < \frac{dP}{P}$
e_{ij}					
e_r					
e_a					
E					

b) In the above table, the symbols stand for :

Elasticity	*Assumption*
e = own price or direct price elasticity of demand	$Q_x = a(P_x)$
e_{ij} =	$Q_i = f(P_j)$
e_r =	$Q_x = r(R)$
e_a =	$Q_x = a(A)$
δ =	$X_1/X_2 = x(MRS_{x1-x2})$
E =	$P_{t+1} = p(P_t)$
e_d = Discretionary income elasticity of demand for consumer durables.	$E^d c = c(I^d)$

Note : The assumption is a restatement from the generalised demand function you have seen earlier except that D_x and a symbol has been replaced by Q_x, and that $E^d{}_c$ stands for expenditure on consumer durable, I^d for discretionary income, and c_d is accordingly termed. You should now be in a position to coin any term of elasticity or use a symbol which has a typical appeal to you.

6.6 DETERMINANTS OF ELASTICITY

So far we have been mainly concerned with the definition and properties of various concepts of elasticity and their measurement. We now discuss the factors which determine the value of a given elasticity. We will confine our reference to price elasticity only.

About the determinants of this elasticity, the following factors are relevant.

i) **The extent of substitutability between goods :** Larger the number of substitutes available to a product, the more will be the elasticity of demand; the smaller the number, the less elastic the demand. For example, consider T.V. set for the first type and salt for the second type.

ii) **The nature of the goods :** A product which accounts for a high percentage of consumer's total expenditure is characterised by high elasticity. You may now examine why salt is inelastic.

iii) **The importance of the goods :** A product which accounts for a high percentage of consumer's total expenditure is characterised by high elasticity. You may now examine why salt is inelastic.

iv) **The price of the product itself :** Highly priced goods tend to have elastic demand, while lower-priced goods have less elastic demand. The expression 'highly priced' is normally taken to mean a price at which the quantity that the consumer plans to buy is close to zero. For example, consider a product like refrigerator.

v) **Whether demand can be postponed or not :** When price has risen and demand cannot be postponed, that will make demand less responsive to price rise than when demand can be postponed.

vi) **Price expectation of buyers :** When the price of the goods has fallen and the buyers expect it to fall further, then they will postpone buying the goods and this will make demand less responsive. On the other hand, if they expect price to go up then they will speed up purchase, which will increase elasticity.

vii) **Time allowed for making adjustment in consumption pattern :** In the short-run, it is very difficult to change habits. Hence the short-term demand is less responsive to price change. The longer the time allowed for making adjustment in consumption pattern, the greater will be the elasticity. The consumers in the long-run would look for better substitutes. Hence the elasticity increases in the long-run.

Activity 11

Think about the determinants of other elasticities and jot down the additional points if any.

....

....

....

....

....

6.7 MANAGERIAL USES OF ELASTICITY CONCEPTS

Regarding the *importance* of the concept of elasticity of demand, it must be pointed out that the concept is useful to the business managers as well as government managers. Elasticity measures help the sales manager in fixing the price of his product. The concept is also important to the economic planners of the country. In trying to fix the production target for various goods in a plan, a planner must estimate the likely demand for goods at the end of the plan. This requires the use of income elasticity concepts.

The price elasticity of demand as well as cross elasticity would determine the substitution between goods and hence useful in fixing the output mix in a production period. The concept is also useful to the policy makers of the government, in particular in determining taxation policy, minimum wages policy, stabilisation programmes for agriculture, and price policies for various other goods (where administered prices are used).

Activity 12

1) Why are market researchers interested in the price elasticity of demand?

....

....

2) If you are the production manager, which concepts of elasticity do you need and why?

....

....

3) To raise revenue, which type of commodities should the Government tax?

....

....

6.8 EMPIRICAL DEMAND ESTIMATES

All managers are concerned with empirical demand estimates because they provide summary information about the direction and proportion of change in demand, as a result of a given change in its explanatory variables. From the standpoint of control and management of external factors, such empirical estimates and their interpretations are therefore, very relevant.

Empirical studies of demand are of three broad types -- consumer surveys, statistical studies and experimental surveys.

Consumer surveys are concerned with the purchaser's intentions. They help in sales forecasting rather than providing information for price policy making. Such surveys thus reveal the 'likely' demand which may not coincide with 'actual' demand.

Statistical studies of a rudimentary type such as trend analysis, postulate time series of the single independent variable. Such studies normally ignore the important determinants of demand such as prices, which are controlled by the management. More sophisticated statistical analysis using multiple correlation and regression techniques, is capable of isolating and measuring the fluctuations in demand which occur in response to principle demand determinants like price and disposable income.

Experimental surveys or controlled experiments can estimate the influence of important demand determinants under management's control. However, care must be taken to reduce the effect of unimportant variables to a minimum. Sometimes, *stimulated* exercises are also undertaken.

The most reliable method of estimating the demand function is to combine the controlled experiments with statistical studies. In fact, in empirical estimation of demand function, there is no escape from statistical techniques.

Some of the statistical demand functions which have been estimated are reproduced below. What is important for you to note is the methodology and findings which may serve as a basis for managerial decisions.

1) **The demand for 8 categories of consumption :** Deaton and Muellbauer (1980) have used a comprehensive approach based on time series data on British consumption, spanning 21 years up to 1974. The reported income and price elasticities are shown in the following table :

Consumption category	*Price Elasticity*	*Income Elasticity*
Food	0 (or negative)	0.21
Clothing	0.92	2.0
Housing	0.31	0.03
Fuel	0.28	1.67
Drinks & Tobacco	0.60	1.22
Transport & Communication	1.21	1.23
Other goods	0.72	1.21
Other services	0.93	1.40

The nature of commodities should explain why price elasticity is generally lower than income elasticity.

2) **The demand for electricity :** Chapman, Tyrrell and Mount (1972) used a combination of time series and cross section data for the years 1949 to 1971 for U.S.A. They estimated that the price elasticity of demand for electricity is about 1.3, income elasticity of demand is 0.3 and the elasticity of demand for electricity with respect to the price of natural gas is 0.15. Positive cross elasticity implies substitutability (through mass) between electricity and gas.

3) **The demand for college education :** Hopkins (1974) considered higher education (4 years bachelors degree programmes at a college or university) in U.S.A. as 'an economy good'. He viewed education at a private institution and education at a public institution as 'different goods' and estimated separate demand functions for each. He hypothesised that the demand for one form of higher education was a function of its own price (or cost), the price of the substitute form of education and the level of income. He did not report elasticity as such, but the coefficient of its linear demand functions were as follows :

[You may remember that if the demand function is in the form of a parabola (semi-log or double-log type), on linear transformation these coefficients do measure constant elasticities.]

Independent Variables → / Dependent Variables ↓	Own price	Price Substitute	Family Income
Demand for Enrolment in Public Institution	−0.022	+0.01	-0.032
Demand for Enrolment in a Private Institution	−0.022	+0.01	+0.036

His result shows that enrolment is 'an economic good' that varies inversely with its price. Enrolment in a public institution is a substitute for enrolment in a private institution. If tution fees at public institutions were to rise, enrolment there would be reduced and those in private institutions would increase. The demand for enrolment in private institution varies directly with family income. But the coefficient on income term for public institution is negative, indicating that enrolment at public institution is 'an inferior good'. Note however, that even though the income effect of a price change for public institution would work against substitution effect, the latter is sufficiently strong to ensure that the price and enrolment varies inversely.

4) **The demand for drinks :** Richard Stone estimated the demand function for beer in U.K. (1920-1930) by the following equation :

$$Q = 1.058\ Y^{0.136}\ R^{0.124}\ G^{0.816}\ P^{-0.727}$$

$$\text{Or } \log Q = \log 1.058 + 0.136 \log Y - 0{,}727 \log P + 0.124 \log R + 0.816 \log G$$

Where Q is the quantity of beer consumed, Y is the aggregate real income, P is the average retail price of beer, R is average retail price of all other commodities, and G is the index of the strength of beer. You may interpret that a price drop of one per cent will increase beer consumption by 0.72%; a rise in aggregate real income of 1% will increase beer consumption by 0.136%; a rise in G will increase it by 0.816%.

5) **The demand for deposits :** G.S. Gupta (1976) estimated that the demand for various types of bank deposits in India as a function of national income (Y), interest rate on 3 month's time deposit (I_n), interest rate on savings deposit (I_s), yield on variable dividend industrial securities (I_i), maximum permissible interest rate on 3 months deposits (I_m) and all scheduled commercial bank offices (B) :

$$\text{Current Deposits, } CD = 1428.85 + 0.033\ 2Y - 572.791\ I_n + 0.4450B$$

$$\text{Savings Deposits, } SD = 2588.12 + 0.387Y + 525.11\ I_s - 176.39\ I_n$$

$$\text{Fixed Deposits, } FD = (-)7441.41 + 0.0873\ Y + 659.94\ I_n - 1082\ I_i + 1791.79\ I_m$$

These equations indicate, among other things,

A Rs. million increase in national income leads to :
a Rs. 0.0332 million increase in demand for CD,
a Rs. 0.387 million increase in demand for SD, and
a Rs. 0.873 million increase in demand for FD.

Activity 13

1) Consider the above functions estimated by G.S. Gupta and find out the impact of

i) 1% increase in I_a

....

....

....

....

ii) 1% increase in I_i

....

....

....

....

iii) 1% increase in I_m

....

....

....

....

iv) An increase in B by 1%

....

....

....

....

6.9 SUMMARY

You have been introduced to the following concepts of elasticity; (i) price, (ii) cross, (iii) income, (iv) substitution, (v) advertisement elasticity, (vi) expectation elasticity. The various forms and measures of price elasticity have been discussed. The difference between point and arc elasticity has also been pointed out. Also, we have derived a formal relationship between price (average revenue), marginal revenue and elasticity. The factors on which the price elasticity of demand depends have been stated and explained. We have mostly confined ourselves to price elasticity, you should have no difficulty in extending same concepts and measures to other types of elasticity which measure the proportion of change in demand. The managerial use of the concept and measure of elasticity makes the subject relevant for practising executives in business and government. However, you must remember that from practical standpoint, our interests lie in both direction and proportion of change in demand; that is what we look for in empirically estimated demand function so that we can use it as a guide to business decision like demand forecasting.

6.10 ADDITIONAL READINGS

A.M. Freeman III, *Introduction to Microeconomic analysis* (1st edition).

A. Koutsoyiannis, *Modem Microeconomics* (IInd edition).

R. Leftwich, *The Price System and the Allocation of Resources* (8th edition).

6.11 SELF-ASSESSMENT TEST

1) An individual spends all his income on three goods. He buys 550 units of X at Re. 1 per unit, 425 units of Y at Rs. 2 per unit, and 200 units of Z at Rs. 3 per unit. The price of X increases by 10% but there is no change in the prices of Y and Z. He now buys 440 units of Y and 190 units of Z. Calculate his price elasticity of demand for X.

2) A consumer demand curve for X is given by the equation P = 100-X. Calculate his point price elasticity when price of X is Rs. 60.

3) The relationship between a consumer's income and the quantity of Q he consumes is given by the equation $R = 1000\ Q^2$. Calculate his point income elasticity of demand for Q when his income is Rs. 64,000.

4) The National Council of Applied Economic Research (NCAER) is the source of the following table about demand coefficients for certain commodities in India, using time series data (1960-61/1975-76). Provide a lucid economic interpretation of the data and reflect on the control or management of demand.

Commodities	*Price Elasticity*	*Income Elasticity*
Rice	0.16	(-)0.19
Wheat	1.25	(-)0.73
Maze cereals	0.46	(-)0.34

5) A statistically estimated demand function for a commodity x is available in the following form :

$$D_x = \frac{1548 P_x^{-0.6} P_y^{0.3} A^{0.4}}{B^{0.5} P_z^{0.2}}$$

where

D_x = Demand for X

P_x = Price of X

P_y = Price of Z

P_z = Price of Z

B = Purchaser's budget (income)

A = Seller's advertisement

a) Is X normal or inferior or Giffen-type?

b) How is X related to Y and Z?

c) Other things remaining equal, if advertisement is doubled, in what direction and how much proportion will the demand for X change?

d) If the seller is interested to increase the sale of X by 50%, how much change should he bring in his price, other things remaining the same?

e) What factors can the seller operate onto control and manage demand? List the factors in order of their importance (strength).

UNIT 7 DEMAND FORECASTING

Structure

7.0 INTRODUCTION

One of the crucial aspects in which managerial economics differs from pure economic theory lies in the treatment of risk and uncertainty. Traditional economic theory assumes a risk-free world of certainty; but the real world business is full of all sorts of risk and uncertainty. A manager cannot, therefore, afford to ignore risk and uncertainty. The element of risk is associated with future which is indefinite and uncertain. To cope with future risk and uncertainty, the manager needs to predict the future event. The likely future event has to be given form and content in terms of projected course of variables, i.e. forecasting. Thus, business forecasting is an essential ingredient of corporate planning. Such forecasting enables the manager to minimise the element of risk and uncertainty. Demand forecasting is a specific type of business forecasting.

7.1 OBJECTIVES

After reading this unit, you should be able to:

- **define** 'forecasting' in contrast to 'projection' and 'prediction'
- **distinguish** between various types of forecasts
- **describe** the techniques of demand forecasting
- **discuss** the uses and abuses of each technique

7.2 CONCEPTS OF FORECASTING

As we have indicated in the introduction, a manager is concerned with problems faced in the immediate present, but he cannot afford to neglect the future. The decision that he takes in the present implies a course of action and reaction in the future.

The manager can conceptualise the future in definite terms. If he is concerned with future *event* — its order, intensity and duration, he can *predict* the future. If he is concerned with the course of future *variables* — like demand, price or profit, he can *project* the future. Thus prediction and projection — both have reference to future; in fact, one supplements the other. Suppose, it predicted that there will be inflation (event). To establish the nature of this event, one needs to consider the projected course of general price index (variable). Exactly in the same way, the predicted event of business recession has to be established with reference to the projected course of variables like sales, inventory etc.

Projection is of two types — forward and backward. It is the forward projection of data variables, which is named **forecasting**. By contrast, the backward projection of data may be named 'backcasting', a tool used by the new economic historians. For practical managers concerned with futurology, what is relevant is forecasting, the forward projection of data, which supports the prediction of an event.

Thus, if a marketing manager fears demand recession, he must establish its basis in terms of trends in sales data; he can estimate such trends through extrapolation of his available sales data. This trend estimation is an exercise in forecasting.

Activity 1

As a manager, you should predict the event, project the course of variables and forecast the data. Keeping this in mind, fill in the blanks in the table below. Note the suggested guideline.

Manager type	Predicted event	Forecast data
Marketing Manager	Demand recession	Sales volume
Finance Manager		Cash flow
Personnel Manager	Employee turnover	
	Machine breakdown	
		Costs of production
Production Manager		Imported material content

7.3 NEED FOR DEMAND FORECASTING

Business managers, depending upon their functional area, need various forecasts. They need to forecast demand, supply, price, profit, costs, investment, and what have you. In this unit, we are concerned with only demand forecasting. The reason is, the concepts and techniques of demand forecasting discussed here can be applied anywhere.

The question may arise : Why have we chosen **demand** forecasting as a model? What is the use of demand forecasting?

The significant of demand or sales forecasting in the context of business policy decisions can hardly be overemphasised. Sales constitute the primary source of revenue for the corporate unit and production for sales gives rise to most of the costs incurred by the firm. Thus sales forecasts are needed for production planning, inventory planning, profit planning and so on. Production itself requires the support of men, materials, machines, money and finance, which will have to be arranged. Thus, manpower, planning, replacement or new investment planning working capital management and financial planning — all depend on sales forecasts. Thus demand forecasting is crucial for corporate planning. The survival and growth of a corporate unit has to be planned, and for this sales forecasting is the most crucial activity. There is no choice between forecasting and no-forecasting. The choice exists only with regard to concepts and techniques of forecasting that we employ. It must be noted that the purpose of forecasting in general is not to provide an exact future data with perfect precision, the purpose is just to bring out the range of possibilities concerning the future under a given set of assumptions. In other words, it is not the '*actual future*' but the '*likely future*' that we build up through forecasts. Such forecasts do not eliminate, but only help you to reduce the degree of risk and uncertainities of the future so that you can undertake operational business policy decisions and actions following therefrom. When you do not forecast, you are simply assuming that the business conditions are going to repeat themselves. This is unrealistic since changes are natural in a dynamic world. It may not be possible to 'estimate' exactly the nature of dynamic changes, but through a set of forecasts you may try to 'guess' them. Forecasting is a step towards that kind of 'guesstimation'; it is some sort of an approximation to reality. It is a type of stimulation exercise to design the probable future state of business. If the likely state comes close to the actual state, it means that the forecast is dependable. If you do not have a dependable forecast about demand, all your business planning will be meaningless. A sales forecast is meant to guide business policy decisions. Without forecasting, forward planning by a corporate unit will be directionless.

Activity 2

a) What kind of demand forecast does a building contractor need? Why does he need such forecasts?

....

....

....

....

....

....

b) State clearly the significance of sales forecast for making :

i. Pricing decision

....

....

....

ii. Advertising decision

....

....

....

iii. Distribution decision

....

....

....

iv. New product decision

....

....

....

7.4 TYPES OF FORECASTS

Demand forecasting is a special kind of economic forecasting. To begin with, we may distinguish between various types of forecasts :

i) **Economic and Non-economic forecasts** : The socialogists often talk about 'social' and 'technological' forecasts. Sometimes they also make 'political' forecasts. For example, one can forecast the crime rate, technological obsolescence, election results and so on. These are all examples of non-economic forecasts; some of them like the technological one may have a lot of economic relevance. New technology may create a demand for new product.

ii) **Micro and Macro-forecasts** : Economic forecasts may be undertaken at various levels — firms, industry or economy. At the firm level, you need a demand or sales forecast. You may attempt similar forecasts at the industry level or the economy level. At the economy level, five-year plan projections are based on trend forecasts concerning national income, investment, employment etc.

iii) **Activity and Passive forecasts** : If the firm extrapolates the demand of previous years to yield the likely/esimated demand for the coming year, it is an example of 'passive' forecast. If the firm, on the other hand, tries to manipulate demand by changing price, product quality, promotional effort etc. then it is an example of 'active' forecast. The expected sales, when planned actions and strategies are undertaken, denote the active forecasts.

iv) **Conditional and Non-conditional forecasts** : By way of 'conditional' forecasting, we estimate the likely impact of certain known or assumed changes in the independent variables on the dependent variable, say demand. 'Non-conditional' forecasting, in contrast, requires the estimation of the changes in the independent variables themselves; such forecasting involves all risks of conditional forecasting.

v) **Short-run and Long-run forecasts** : An important consideration in forecasting is the time span relevant for a particular problem. That is how far ahead the forecast is being made. It can be for a short-run or for a long-run. Duration of the period will vary from problem to problem. In the case of sales, short-run forecasting is generally concerned with projections of established products and are confined up to a year in terms of duration. Introduction of a new product, however, invariably involves a long-term forecast; in this case you have to consider long-term changes in population, tastes and preferences of the buyers, technology, product life-cycle etc. By contrast, short-run forecasting concentrates on a few selected variables; here simple techniques based on analysis of past experience and information can give you fairly accurate forecasts.

7.5 STEPS IN DEMAND FORECASTING

Demand or sales forecasting is a scientific exercise. It has to go through a number of steps. At each step, you have to make critical considerations. Such considerations are categorically listed below :

1) **Nature of Forecast** : To begin with, you should be clear about the use of forecast data — how it is related to forward planning and corporate planning by the firm. Depending up on its use, you have to choose the type of forecasts: Short-run or long-run, active or passive, conditional

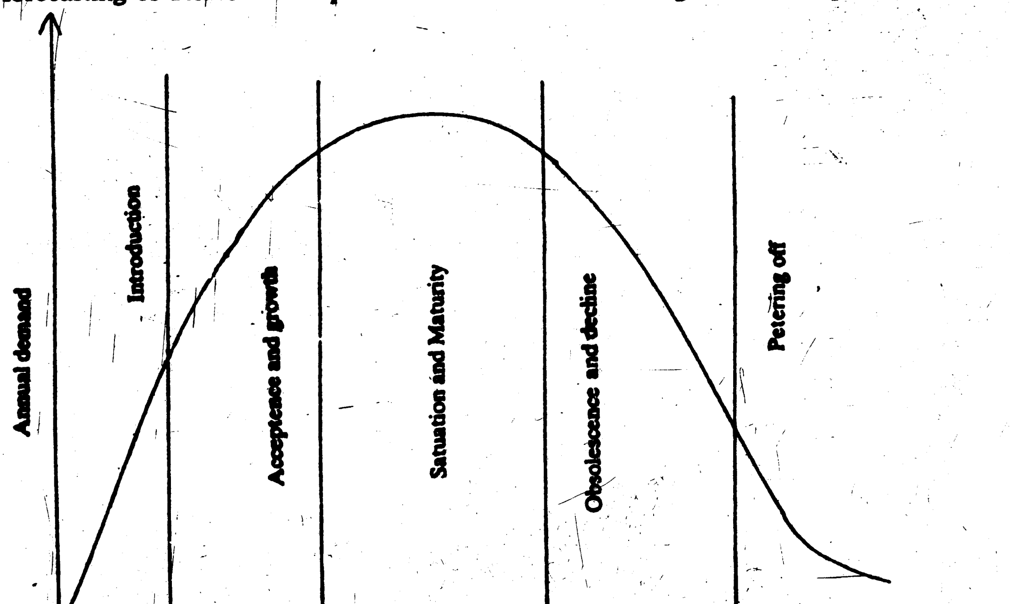
or non-conditional etc.

2) Nature of product : The next important consideration is the nature of product for which you are attempting a demand forecast. You have to examine carefully whether the product is consumer goods or producer goods, perishable or durable, final or intermediate demand, new demand or replacement demand type etc. A couple of examples may illustrate the importance of this factor. The demand for intermediate goods like basic chemicals is derived from the final demand for finished goods like detergents. While forecasting the demand for basic chemicals, it becomes essential to analyse the nature of demand for detergents. Promoting sales through advertising or price competition is much less important in the case of intermediate goods compared to final goods. The elasicity of demand for intermediate goods depends on their relative importance in the price of the final product. If it forms a small component, changes in its price may not be important and demand may be quite inelastic. Similarly, if the final user is in the habit of keeping a stock of intermediate goods, then it should be considered while forecasting the sales of intermediate goods like raw-materials.

The next example may be from consumer durables like radios, bicycles, television sets, cars etc. Such demand has got two components — new demand and replacement demand. As more and more people purchase these goods, market becomes saturated; new demand declines, replacement demand catches up. Buyers can delay replacement demand depending upon their socio-economic conditions, price, availability of consumer credit.

Time factor is a crucial determinant in demand forecasting. Perishable commodities such as fresh vegetables and fruits can be sold over a limited period of time. Here skilful demand forecasting is needed to avoid waste. If there are storage facilities, then buyers can adjust their demand according to availability, price and income. The time taken for such adjustment varies from product to product. Goods of daily necessities that are bought more frequently will lead to quicker adjustments, whereas in case of expensive equipment which is worn out and replaced after a long period of time, adaptation of demand will be spread over a longer duration of time.

Time element enters into demand forecasting in another way. This has reference to the concept of 'product life cycle'. When a new product is introduced, sales will increase slowly as more and more people come to know about it. If it is a success, the market will widen quickly and sales will increase rapidly. As the market begins to get saturated and substitute products also get introduced, sales will begin to taper off. Such a **product life cycle curve** is illustrated below. The forecasting of demand for a product must consider the stage where the product belongs to :



3) Determinants of demand : Once you have identified the nature of product for which you are to build a forecast, your next task is to locate clearly the determinants of demand for the product. Depending on the nature of product and nature of forecasts, different determinants will assume different degree of importance in different demand functions. In the preceding unit, you have been exposed to a number of price-income factors or determinants - own price, related income, advertisement, price expectation etc. In addition, it is important to consider socio-psychological determinants, specially demographic, sociplogical and psychological factors affecting demand. Without considering these factors, long-run demand forecasting is not possible.

Such factors are particularly important for long-run active forecasts. The size of population, the age-composition, the location of household unit, the sex-composition — all these exercise influence on demand in varying degrees. If more babies are born, more will be the demand for toys; if more youngsters marry, more will be the demand for furniture; if more old people survive, more will be the demand for sticks. In the same way, buyers' psychology — his need, social status, ego, demonstration effect etc. — also effect demand. While forecasting, you cannot neglect these factors.

4) Analysis of factors (determinants) : Identifying the determinants alone would not do, their analysis is also important for demand forecasting. In an analysis of statistical demand function, it is customary to classify the explanatory factors into (a) **trend factors**, which affect demand over long-run, (b) **cyclical factors** whose effects on demand are periodic in nature, (c) **seasonal factors**, which are a little more certain compared to cyclical factors, because there is some regularity with regard to their occurrence, and (d) **random factors** which create disturbance because they are erratic in nature; their operation and effects are not very orderly.

An analysis of factors is specially important depending upon whether it is the aggregate demand in the economy or the industry's demand or the company's demand or the consumers' demand which is being predicted. Also, for a long-run demand forecast, trend factors are important; but for a short-run demand forecast, cyclical and seasonal factors are important.

5) Choice of techniques : This is a very important step. You have to choose a particular technique from among various techniques of demand forecasting. Subsequently, you will be exposed to all such techniques, statistical or otherwise. You will find that different techniques may be appropriate for forecasting demand for different products depending upon their nature. In some cases, it may be possible to use more than one technique. However, the choice of technique has to be logical and appropriate; for it is a very critical choice. Much of the accuracy and relevance of the forecast data depends on this choice. The choice itself depends on a number of factors — the degree of accuracy required, reference period of the forecast, complexity of the relationship postulated in the demand function, available time for forecasting exercise, size of cost budget for the forecast etc.

6) Testing accuracy : This is the final step in demand forecasting. There are various methods for testing statistical accuracy in a given forecast. Some of them are simple and inexpensive, others quite complex and difficult. This testing is needed to avoid/reduce the margin of error and thereby improve its validity for practical decision-making purpose. Subsequently you will be exposed briefly to some of these methods and their uses.

Activity 3

a) Recall the steps in demand forecasting. Just name them.

1) ....

method' because of its reliance on the hunches of experts. It is indeed a very simple method, but it replaces analysis by opinions and it can thus turn out to be highly subjective in nature.

2) Reasoned Opinion — Delphi Technique : This is a variant of the opinion poll method. Here is an attempt to arrive at a consensus in an uncertain area by questioning a group of experts repeatedly until the responses appear to converge along a single line. The participants are supplied with responses to previous questions (including reasonings from others in the group by a coordinator or a leader or operator of some sort). Such feedback may result in an expert revising his earlier opinion. This may lead to a narrowing down of the divergent views (of the experts) expressed earlier. The Delphi Techniques followed by the Greeks earlier, thus generates "reasoned opinion" in place of "unstructured opinion"; but this is still a poor proxy for market behaviour of economic variables.

3) Consumers' Survey - Complete Enumeration Method : Under this, the forecaster undertakes a complete survey of all consumers whose demand he intends to forecast. Once this information is collected, the sales forecasts are obtained by simply adding the probable demands of all consumers. For example, if there are N consumers, each demanding D then the total demand forecast is $\sum_{i=1}^{N} D_i$

The principle merit of this method is that the forecaster does not introduce any bias or value judgementof his own. He simply records the data and aggregates. But it is a very tedious and cumbersome process; it is not feasible where a large number of consumers is involved.

Moreover if the data are wrongly recorded, this method will be totally useless.

4) Consumer Survey — Sample Survey Method : Under this method, the forecaster selects a few consuming units out of the relevant population and then collect data on their probable demands for the product during the forecast period. The total demand of sample units is finally blown up to generate the total demand forecast. For example, if there are 1,000 consumers, the forecaster may choose just 50 of them, using random sampling method or stratified sampling method. If the probable demand of the selected consumers is D_i, where i = 1, 2, 3, ..., n; n = 50, then the total demand forecaster for the selected group of commodity will be given by

$$\sum_{1=1}^{50} n_i D_i = n_1 D_1 + n_2 D_2 + \dots\dots\dots\dots + n_{50} D_{50}$$

where n_i is the number of consumers in group i, and $n_1 + n_2 + \dots\dots + n_{50}$ = 1,000.

If all consumers are alike, then the selection may be done on a random basis and the total demand will be given by

$$(D_1 + D_2 + D_3 + \dots\dots + D_{50})\frac{1000}{50}$$

Compared to the former survey, this method is less tedious and less costly, and subject to less data error;.but the choice of sample is very critical. If the sample is properly chosen, then it will yield dependable results; otherwise there may be sampling error. The sampling error can decrease with every increase in sample size.

5) Consumers' Clinics : An improvement over the consumer interview and survey approach is to stimulate the market conditions so as to observe the actual behaviour of the consumers.

2) ....

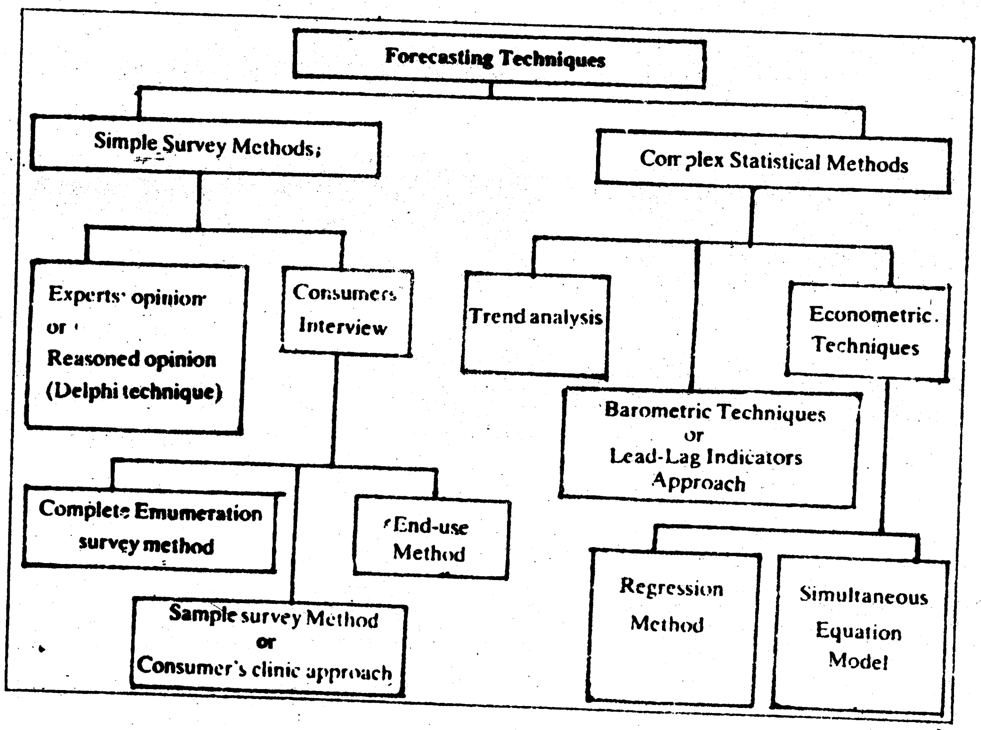
3) ....

4) ....

5) ....

7.6 TECHNIQUES OF DEMAND FORECASTING

Broadly speaking, there are two approaches to demand forecasting – one is to obtain information about the likely purchase behaviour of the buyer through collecting expert's opinion or by conducting interviews with consumers, the other is to use past experience as a guide through a set of statistical techniques. Both these methods rely on varying degrees of judgement. The first method is usually found suitable for short-term forecasting, the latter for long-term forecasting. There are specific techniques which fall under each of these broad methods. We shall begin by presenting a summary view of all these techniques in the form of a schematic model :



In what follows, we shall now take up each one of these techniques under broad category of methods suggested above.

7.7 SIMPLE SURVEY METHODS

For forecasting the demand for existing product, such survey methods are often employed. In this set of methods, we may undertake the following exercise.

1) Experts' Opinion Poll : In this method, the experts on the particular product whose demand is under study are requested to give their 'opinion' or 'feel' about the product. These experts, dealing in the same or similar product, are able to predict the likely sales of a given product in future periods under different conditions based on their experience. If the number of such experts is large and their experience-based reactions are different, then an average – simple or weighted – is found to lead to unique forecasts. Sometimes this method is also called the 'hunch'

In this approach, a number of potential buyers of the product are invited and are given some amounts of money to purchase various products. The price tags are altered and the response of the participants is observed. This method assumes the applicability of Revealed Preference Approach : choice reveals preference. One of the limitations of this method is that the potential buyers may enjoy it as a game and may not behave in a natural manner. Also, it is very expensive; and a large number of buyers can never in involved. The design of the clinic is critical. Despite these limitations, some firms do use this method to find out the brand loyalty of consumers.

6) **End-use Method of Consumers' Survey** : Under this method, the sales of a product are projected through a survey of its end-users. A product is used for **final** consumption or as an **intermediate** product in the production of other goods in the domestic market, or it may be exported as well as imported. the demands for final consumption and exports net of imports are estimated through some other forecasting method, and its demand for intermediate use is estimated through a survey of its user industries.

Given the input-output coefficients of all industries and their production plans, the sum of final consumption demand and export demand net of imports can be obtained with the help of input-output model. (Please see the *Appendix*).

Considering their production plans and input-output model, the demand for steel in India in 1999 can be obtained as

$$[S]_{1999} = [S_c]_{1999} + [S_x]_{1999} - [S_m]_{1999}$$
$$+a_{s1}[Q_1]_{1999} + a_{s2}[Q_2]_{1999} + \dots\dots\dots + a_{sn}[Q_n]_{1999}$$

Where S =aggregate steel demand

S_c = final consumption demand for steel

S_x = export demand for steel

S_m = import of steel

Q_i = output of industry i using steel as an intermediate product

a_{si} = steel requirement of industry i per unit of its output

i = 1, 2,......, n and subscript 1999 stands for the forecast year.

Denoting the sum of final consumption demand and exports net of imports by f, and thus

$Q_{ic} + Q_{ix} - Q_{im} = Q_{if}$, We have

$$Q_1 = a_{11}Q_1 + a_{12}Q_2 + \dots\dots + a_{1n}Q_n + Q_{1f}$$
$$Q_2 = a_{21}Q_1 + a_{22}Q_2 + \dots\dots + a_{2n}Q_n + Q_{2f}$$

$$Q_n = a_{n1}Q_1 + a_{n2}Q_2 + \dots\dots + a_{nn}Q_n + Q_{nf}$$

These n equations can be written in the matrix form as follows :

$$\begin{bmatrix} 1-a_{11} & -a_{12} & -a_{13}\dots\dots\dots\dots -a_{1n} \\ -a_{21} & 1-a_{22} & -a_{23}\dots\dots\dots\dots -a_{2n} \\ \\ -a_{n1} & -a_{n2} & 1-a_{m}\dots\dots\dots\dots -a_{nn} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \\ \\ Q_n \end{bmatrix} \begin{bmatrix} Q_{1f} \\ Q_{2f} \\ \\ Q_{nf} \end{bmatrix}$$

or Q_f = (I-A) Q

Q = Vector of industry outputs, and

Q_f = Vector of final consumption demand and exports net of imports.

I = n × n unit matrix

A = Matrix of input-output coefficients with a_{ij} as its i_{th} row j_{th} column element.

Thus knowing Q_f and A, Q can be easily forecast.

The input-output method of estimating Q_f for any commodity is generally used in the context of national planning and projection. At the industry level, it has limitations. The individual industry has to rely on some other alternative method to estimate the future demand of its products for final consumption, exports and imports. Thus only the intermediate demand or the input demand part of total demand for a commodity can be predicted by the end-use method. This method can be used only when the number of end-users for the product is limited to a few. Also, this method assumes a given set of input-output coefficients; these coefficients may change because of a change in technology. In that case, you have to move from static to dynamic input-output model. At the firm or the industry level, this may turn out be an academic exercise. Otherwise for estimating derived demand for industrial raw-materials, finished and semi-finished producers' goods, this method is quite dependable; it does provide us a use-wise and sector-wise demand forecasts.

Activity 4

a) The National Council of Applied Economic Research (NCAER) has used the end-use method to estimate the demand for steel in India (1968). For a summary of the method and findings, see Chopra, **Managerial Economics**, (pp. 73-75).

b) Explain the 'Delphi' technique. Is it a form of expert's opinion survey ? Or different ?

....

....

....

....

....

c) In developed countries, market experiments are used to assess the demand function for a particular product. By this method, product prices are varied from time to time or between different regions of the country, and then the impact of such price-variation on sales is assessed. Comment on the uses and abuses of such a method of demand forecasting.

....

....

....

....

....

) List the demand forecasting techniques which are based on consumer's interview.

....

....

....

....

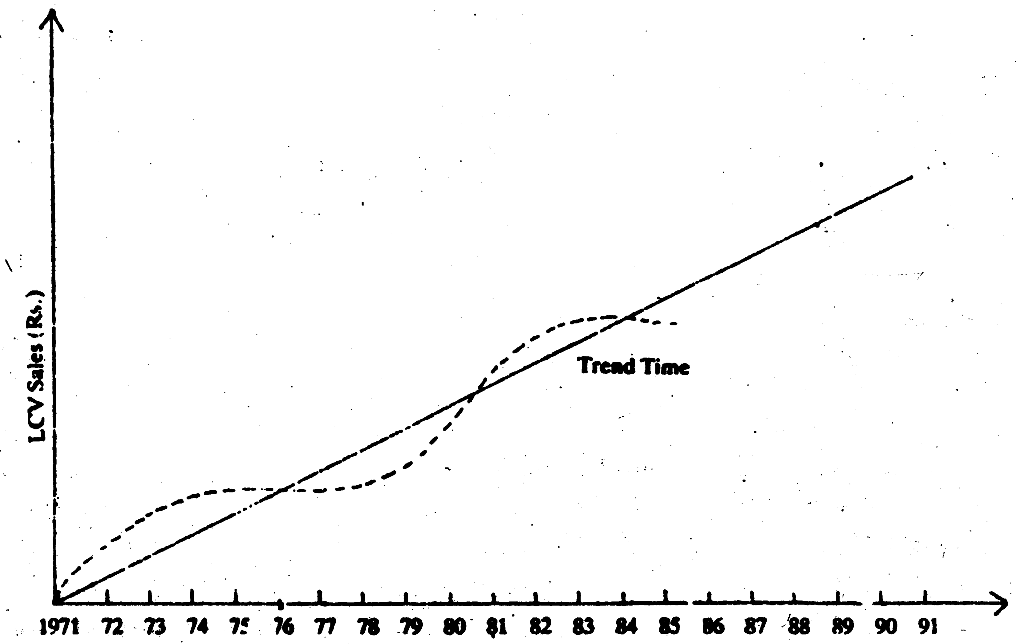
....

7.8 COMPLEX STATISTICAL METHODS

We shall now move from simple to complex set of methods of demand forecasting. Such methods are taken usually from statistics. As such, you may be quite familiar with some of the statistical tools and techniques, as a part of quantitative methods for business decisions.

1) **Time series analysis or trend method** : Under this method, the time series data on the variable under forecast are used to fit a trend line or curve either graphically or through the statisitical method of Least Squares. For example, you may have yearly data on the sale of light commercial vehicles (LCV) in India from 1971-1985. Based on the data plotted on the graph, a line or curve is drawn depicting LCV demand in India; such a line could be drawn up to the present period or the period for which the data is available. It can then be extrapolated to the forecast period. The reading of graph would give you the forecasts. This is a crude version of trend method or what is called a 'naive technique' or the 'graphical method'. This is illustrated in the diagram below :

LCV Sales over Time



You may note in the diagram that the trend line has been fitted around fluctuations. The actua movement of sales of LCV is represented by the dotted line; and the straight line represents the trend line and you will notice that there is an upward trend. Using actual data between A anc B, you may project the line into the future say, any year between 1986-91 i.e. between B and C.

The trend line is worked out by fitting a trend equation to time series data with the aid of an estimation method. The trend equation could take either a linear or any kind of non-linear form. Some of the most suitable trend equations for demand forecasting are the following :

i) **Linear trend :** $D = a + bT$ where D is the demand for the product under forecast, T is the trend variable which may be normalised to take the value of 1 in the first period, 2 in the second period and so on a (intercept) and l (slope) are parametric information which can be estimated. The trend line assumes that there will be a constant absolute amount of change (=b) every period.

ii) **Exponential trend :** $D = ae^{bT}$ or $\log_e D = \log_e a + bT$. This semi-log function assumes a constant growth rate = b each period.

ii) **Second (and higher) degree polynomials trend :** $D = a + bT + cT^2$

In this case, the slope of the parabola is given by the term $\frac{dD}{dT}$ and it changes direction only once, either from positive to negative or vice-versa. The shape and location with respect to axis will vary according to the values of constant a, b and c.

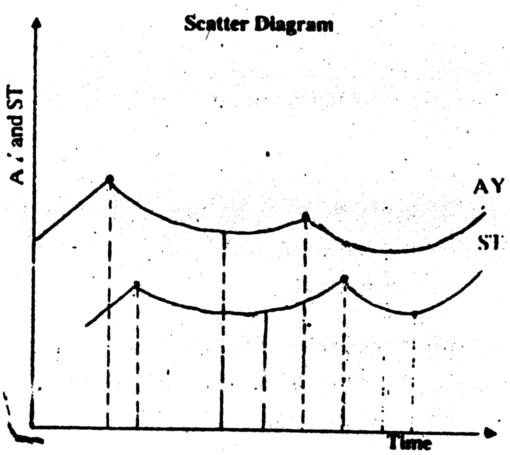
iv) **Double-log (Cobb-Douglas type) trend :** $D = aT^b$

$$\text{Or, } log_e D = log_e a + b \quad log_e T$$

This double-log trend assumes a constant elasticity = b every period.

The trend method outlined above often yields a dependable forecast. The advantage in this method is that it does not require the formal knowledge of economic theory and the market, it only needs the fime series data. The only limitation in this method is that is assumes that the past is repeated in future. Also, it is an appropriate method for long-run forecasts, but inappropriate for short-run forecasts. Sometimes the time series analysis may not reveal a significant trend of any kind. In that case, the moving average method or exponentially weighted moving average method is used to smoothen the series.

2) Barometric Techniques or Lead-Lag indicators method : This consists in discovering a set of series of some variables which exhibit a close association in their movement over a period or time. There are three kinds of time-series : leading series, coincident series and lagging series. The **leading series** are data on the variables which move up or down ahead of some other series; the **coincident series** move along with some other series, and the **lagging series** move up or down behind some other series. For example, the bank rate charged by the Central Bank is a leading indicator of interest rates charged by the commercial bank. If the forecast variable is such that it lags behind the movement of the leading indicator, its value in future could be forecast through a measure of this lead-lag relationship. This relationship can be depicted in the form of a diagram.



It shows the movement of agricultural income (AY series) and the sale of tractors (ST series). The movement of AY is similar to that of ST, but the movement in ST take place after a year's time lag compared to the movement in AY. Thus if one knows the direction of the movement in agriculture income (AY), one can predict the direction of movement of tractors' sale (ST) for the next year. Thus agricultural income (AY) may be used as a barometer (a leading indicator) to help the short-term forecast for the sale of tractors.

We may illustrate the same relationship in an equation form. Suppose.

$$D_t^x = 1.6 + 0.8\, D_t^{y-2}$$

Where D^x and D^y stand for demand (sales) for items X and Y; and there is a two-year lag as per our assumption. In this case, we can forecast the sale of x for the next two periods :

$$D_{t+1}^x = 1.6 + 0.8\, D_{t-1}^y$$

$$DX_{t+2}^x = 1.6 + 0.8\, D_t^y$$

Generally, this barometric method has been used in some of the developed countries for predicting business cycles situation. For this purpose, some countries construct what are known as *'diffusion indices'* by combining the movement of a number of leading series in the economy so that turning points in business activity could be discovered well in advance. Some of the limitations of this method may be noted however. The leading indicator method does not tell you anything about the magnitude of the change that can be expected in the lagging series, but only the direction of change. Also, the lead period itself may change overtime. Through our estimation we may find out the best-fitted lag period on the past data, but the same may not be true for the future. Finally, it may not be always possible to find out the leading, lagging or coincident indicators of the variable for which a demand forecast is being attempted.

3) Correlation and Regression : These involve the use of econometric methods to determine the nature and degree of association between/among a set of variable. Econometrics, you may recall, is the use of economic theory, statistical analysis and mathematical functions to determine the relationship between a dependent variable (say, sales) and one or more independent variables (like price, income, advertisement etc.). The relationship may be expressed in the form of a demand function, as you have seen in Unit 4:

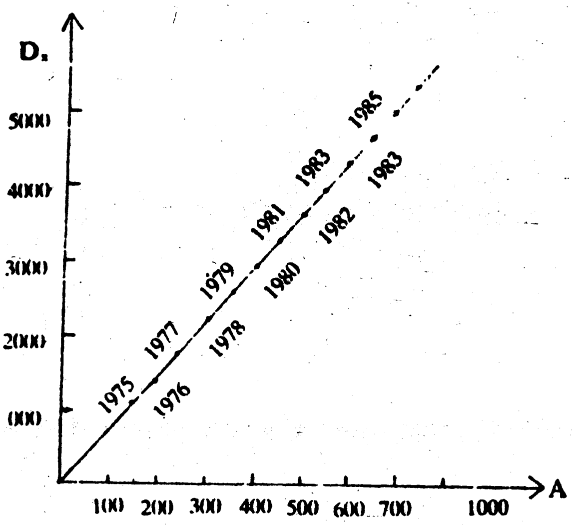
$D_x = d\,(P_x, P_y, P_z\ B, A, \ldots\ldots\ldots)$

Such relationships, based on past data can be used for forecasting. The analysis can be carried with varying degrees of complexity. Here we shall not get into the methods of finding out 'correlation coefficient' or 'regression equation', you must have covered those statistical techniques as a part of quantitative methods. Similarly, we shall not go into the question of economic theory. We shall concentrate simply on the use of these econometric techniques in forecasting.

In the case of simple correlation, the regression equation has one dependent and one independent variable. For example, the sale of X (D_x) is a function of its advertisement (A) . $D_x = D\,(A)$. The analysis involves :

i) determining the nature of association through a regression equation — a straight line or a curve — to describe the average relationship between D_x and A;

ii) determining the extent of association between D_x and A through the estimation of correlation coefficient.



The analysis begins with a look at the scatter diagram. This straight line has been computed through the Least Square method; it shows an upward trend. The regression equation of this line may be in the form :

$D_x = a + bA$

say, $D_x = 400 + 0.5A$

This the next year, if A = Rs. 1,000 then the expected sales will be Rs. 900.

You may note that if l 0, i.e., $\frac{dD_x}{dA}$ 0, implies that A and D_x move in the same direction, similarly if l 0 i.e. $\frac{dD_x}{dA} < 0$, it implies that they are inversely related.

The correlation coefficient, r, is the measure of this degree of relationship between a dependent and independent variable. You may recall, in general form, the formula for its measure is

$$r = \sqrt{\frac{\sum(Y_e-\overline{Y})^2}{\sum(Y-\overline{Y})^2}}$$

where Y is the dependent variable

$\overline{Y}$ is the mean value of Y for all years

Y_e is the value of dependent variable as calculated from the regression equation.

The r may vary between -1 (perfect negative) and + 1 (perfect positive): this reflects the degree of association, it does not measure 'causation', such causation has to be discovered through the logic of reasoning typical of economic theory. The r = 0 suggest absence of any relationship between the dependent and the independent variable.

Let us now move from the case of one to many independent variables. Surely, sales do not depend on advertisement alone; they may depend on a host of price-income factors. If we find a regression equation with more than one explanatory variables, we are on the realm of multiple regression and multiple correlation. The form of the equation may be :

$$D_x = a + b_1A - b_2P_x + l_3B + b_4P_y$$

You may recall from the preceding unit that b_1, b_2, b_3, b_4 .... are regression coefficients which measure elasticities. That is, the l_i reflect the direction as well as proportion of change in demand for x as a result of a change in any of its explanatory variables. For example, $b_2 < 0$ suggest that D_x and P_x are inversely related; $b_4 > 0$ suggest that x and y are substitutes; $b_3 > 0$ suggest that x is a normal commodity with positive income-effect. Given the estimated value of a and b_i, you may forecast the expected sales (D_x), if you know the future values of explanatory variables like own price (P_x), related price (P_y), income (B) and advertisement (A). Lastly, you may also recall that the statistics R^2 (Co-efficient of determination) gives the measure of goodness of fit. The closer it is to unity, the better is the fit, and that way you get a more reliable forecast.

The principal advantage of this method is that it is prescriptive as well as descriptive. That is, besides generating demand forecast, it explains why the demand is what it is. In other words, this techniques has got both explanatory and predictive value. The regression method is neither mechanistic like the trend method nor subjective like the opinion poll method. In this method of forecasting, you may use not only time-series data but also cross-section data. The only precaution you need to take is that data analysis should be based on the logic of economic theory.

Regression analysis is beset with certain limitations. If some of the explanatory variables are not realistically chosen, they tend to be misleading. Likewise, if there is auto-correlation (i.e., when sales of one year are highly correlated with the sales of preceding year), the regression results may again be biased and incorrect. Of course, there are sophisticated econometric techniques to take care of auto-correlation. Thus the choice of explanatory variables is very critical. Lastly, as is true of all statistical methods, the regression method forecasts on the basis of past average relationship, and so, to the extent the future relationship deviates from the past experience (average), the forecast will be wrong.

4) **Simultaneous Equations Method** : Here is a very sophisticated method of forecasting. It is also known as the 'complete system approach' or 'econometric model building'. In your earlier units, we have made reference to such econometric models. Presently we do not intend to get into the details of this method because it is a subject by itself. Moreover, this method is normally used in macro-level forecasting for the economy as a whole; in this course, our focus is limited to micro elements only. Of course, you, as corporate managers, should know the basic elements in such an approach.

You may quickly recall your understanding of terms like model, its structure, its reduced form solution — of the unknown endogenous variables in terms of known variables, parametric information and exogenous terms. Once a complete model is structured, it can explain the behaviour of all the varibales which the decision unit, firm or industry or national economy can control. After the model is theoretically developed, it is estimated through some appropriate methods such as the two-stage least squares method. The model is then solved for each of the endogenous variables in terms of exogenous and lagged endogenous variables, if any. The values of the lagged endogenous variables are obviously known and those of the exogenous variables will have to be predicted. After this is done, the corresponding values of lagged endogenous variables and variables are fed into the equation corresponding to the value variables whose forecast are needed to generate the required forecasts.

The method is indeed very complicated. However, in the days of computer, when package programmes are available, this method can be used easily to derive meaningful forecasts. The principle advantage in this method is that the forecaster needs to estimate the future values of only the exogenous variables unlike the regression method where he has to predict the future values of all, endogenous and exogenous variables affecting the variable under forecast. The values of exogenous variables are easier to predict than those of the endogenous variables. However, such econometric models have limitations, similar to that of regression method. For example, it assumes that the past statistical relationship will hold good in the prediction period. Also, it is a highly complicated, time-consuming and costly method. That is why, despite being theoretically a better method, it has not attracted the corporate level managers. After all, the use of this method largely depends on the degree of modernisation (computerisation etc.) and progressive outlook of the corporate unit itself.

Activity 5

a) Compare and contrast briefly the simultaneous equation and the regression methods of demand forecasting.

....

....

....

....

....

....

....

b) Since statistics can prove or disprove anything, it is better to use non-statistical methods for forecasting. One such non-statistical method, which is free from subjectivity is the **Historical Analogy Method**. Under this method, the demand forecasting is done in two stages :

i) selection of a country A which sometime in the past (period t*), particularly with respect to industry i, the demand of whose product is under forecasting, was in the same stage of development as country B for which forecasts are being made at present (period t),

ii) forecasing of the demand for industry i's product in country B in periods t + 1, t + 2 ...., t + n on the basis of actual demand of that industry's product in country A in periods t* + 1, t* + 2, ..., t* + n.

Thus, the size of colour TV demand in India in 1988 is approximately the same as the one experienced by the USA in 1968. On the assumption of this 20 year lag, as per this method, India's demand in 1988 will roughly be the same as USA's demand in 1978.

State the limitations of this method.

....

....

....

....

....

....

7.9 ACCURACY OF FORECASTS

Since no forecast is going to be correct completely, the forecaster may be interested in measuring the accuracy of his forecasts. This is possible, provided the **actual** values become eventually available i.e., when the forecast period becomes the past or present period.

The forecast accuracy tests compare the forecasts (Y_t) with realised values ($\hat{Y}_t$). There are two kinds of forecasts : ex-ante and ex-post. Ex-ante forecasts relate to the sample periods or cross—sections; but the ex—post forcasts are for future periods. The alternative tests for evaluating the accuracy of ex-ante forecasts are :

i) Coefficient of Determination (R^2) Test

ii) Root Mean Squared Error (RMSE) Test

iii) Percentage Mean Absolute Error (PMAE) Test

iv) Percentage Absolute Error (PAE) Test.

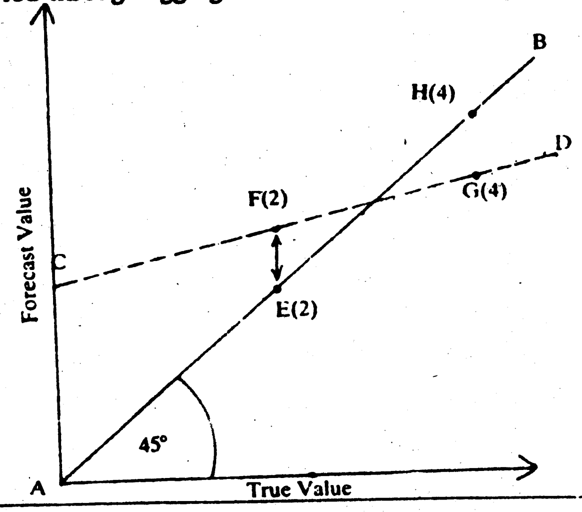
Again, as a part of statistical methods, you should learn these. Suffice it to say here that if the forecast is made for one period only, then the difference between the actual value and the forecast value would measure the forecasting error. If forecasts are made for more than one year, then one can find the 'average error' by simply taking the arithmetic mean of the absolute values of forecasting errors of different periods. You must note that an absolute level of an error does not indicate much; obviously, you should compare the percentage errors. Thus the measure of forecast inaccuracy for one period forecast may be defined by PAE.

$$\text{PAE} = (|\,Y_t - \hat{Y}_t\,|)\frac{100}{Y_t}$$

The same can be generalised for a multiperiod through aggregation of errors over the periods.

Graphically, the forecasts may be plotted against the corresponding time values and then the deviation of this line from the 45° inaccuracy.

In this diagram, AB is the 45° line, CD is the line showing the relationship between the true and forecast value. The vertical distance between these two lines measures the forecasting error. For example, the forecasting error in prediction period 2 is plus EF, while that in period 4 is minus GH.



7.10 SUMMARY

There is no unique method for demand forecasting. The forecaster may choose his techniques depending upon his objective, urgency, data availability, cost-budget, time-budget, and the nature of product whose demand he intends to forecast. For short-term forecast, he may use any of the survey methods or barometric methods; for long-term forecasts, he may use any of the various statistical methods. Similarly, his technique would differ depending upon whether the product is an established (old) one or a new one. For producer goods and intermediate goods, the end-use method may be advisable. Sometimes it is advisable to use a combination of methods. Each of these methods, statistical or non-statistical has its own set of uses and abuses. In general, it may be remarked that no forecast is going to be completely correct; therefore, the forecaster must pay some attention towards minimising forecasing error. Then only, his demand forecasts can provide useful support to his forward planning, corporate planning and business policy decisions. Sometimes forecasting may even be needed for macro level planning and projections. In this unit, we have concentrated on demand forecasting, but the concepts and techniques developed here have application for economic variables other than demand.

7.11 ADDITIONAL READINGS

Adhikary, M. (1987) *Managerial Economics,* Khosla Publishing House : N. Delhi (Ch. VIII).

Gupta, G.S. ... 'Forecasting Techniques', *Management Annual,* Vol. IV, Nov. 1974; pp. 8-21.

Dean, Joel (1976) *Managerial Economics,* Prentice-Hall of India : N. Delhi (Ch. IV).

Chopra, O.P. (1984) *Managerial Economics,* Tata-McGraw Hill.

7.12 SELF-ASSESSMENT TEST

1. Recall your understanding of the concepts of prediction, projection and forecasting. What is the significance of demand forecasting ?

2. Distinguish between the following types of forecasts :

   a) Economic and Non-economic forecasts
   b) Micro and Macro forecasts
   c) Short-term and Long-term forecasts
   d) Conditional and non-conditional forecasts
   e) Ex-ante and Ex-post forecasts.

3. What is the role of time element in the context of demand forecasting? What factors would you normally consider in choosing a forecasiting technique?

4. Discuss the suitability of forecasting technique, if you want to forecast the demand for :

   i) Coal
   ii) Car
   iii) Plastic tea-cups and saucers
   iv) New product

5. Joel Dean has suggested the following set of approaches with regard to techniques of forecasting the demand for new products :

   i) Evolutionary Approach
   ii) Substitute Approach
   iii) Growth Curve Approach
   iv) Opinion Poll Approach
   v) Sales Experience Approach
   vi) Vicarious Approach

Explain each of the approaehes separately.

Note : Some of these terms may be new; but if you get back to the recommended reading No. 3 by Joel Dean and run down those few pages, you should be able to discover all these approaches in this unit itself. And then you can attempt an integration.

Appendix

INPUT-OUTPUT MODEL

The end-use method of demand forecasting is essentially based on the input-output model of the economist. Such models are used towards planning economic activities within a firm or at the macro level i.e. for the country as a whole. In every economy the production of a ***final product*** requires the input of many ***intermediate goods***. For instance, the production of steel requires iron ore, coal, electricity etc. Input-output analysis is essentially a model of production. The problem is essentally technological. What the model seeks to determine is the total demand for a commodity and the quantity of each product which must be used up in the production process, given the quantities of available resources and the state of technology. Though input-output model is not a general equilibrium analysis; it does take into account the interdependence of the production plans and activities of many industries which constitutes an economy. Such an interdependence would exist since each industry employs the output of this industry as a factor of production. For instance, steel is used to make trucks while trucks are used to transport steel, coal and pig iron etc., which are used in the manufacture of steel.

What is the rationale behind the machanics of an input-output model? The analysis starts with an input-output table, (viz. empirical data derived by using some statistical methods) which gives the total output of each and every industry, the distribution of that output among various industries as inputs and the amount allocated for **final use** (for the use of consumers, investors, government and exporters). The left hand margin of the table lists the producing industries, while the top margin that of the purchasing (user) industries and the output of the given industry. The table is a square matrix since each industry is at the same time a producer as well as user. Suppose the economy consists of two industries (sectors like agriculture and manufacturer or consumer goods and investment goods only). Also suppose there is only one primary input i.e. labour (an input which is not produced by the system). Normally the figures in an input-output table is given in money value though they can also represent physical units.

The **input output table** is often put in the form of a **Transaction matrix**. Suppose we have the following transaction matrix.

Sectors of Use → Producing Sectors ↓	Industry I	Industry II	Final Use	Total Output
Industry I	20	9	11	40
Industry II	12	6	12	30
Primary Input	20	10	5	35

The figures are in physical units.

How do we read the table? Let us interpret the first row. Industry I produces 40 units of output, of which 20 units are used by the same industry as an input, 9 units are used by industry II as an input. Hence 11 units are left for final demand (final use). Similar is the interpretation for the second row. The last row shows the amount of labour used by each industry including the final use (labour used as domestic servants, car drivers etc.). The last column is the summation of each row. Since we have used physical units, **column sums have no meaning**. They have meaning only if the figures are given in monetary value.

From the transaction matrix, a technology matrix can be derived. Here we must state one important **assumption** of input-output model. In any production process, all inputs are used in rigidly fixed proportion, as production is characterised by constant returns to scale (which implies that output will expand proportionately with input). Given this assumption, we find that if 20 units of commodity X and 12 units of commodity Y are required to produce 40 units of commodity X, then .5 units of commodity X and .3 units of commodity Y are required to produce each unit of commodity X (20 divided by 40 and 12 divided by 40). Exactly in the same way we find that .3 units of commodity X and .2 units of commodity Y are required to produce commodity Y (divided 9 by 30 and 6 by 30 to get .3 and .2). This information is produced in the following table where the final use and total output columns are omitted

using sectors →	X	Y
Producing sector ↓		
X	.50	.30
Y	.30	.20
Primary Input	.50	.33

Such a table is called a technology matrix. A technology matrix is denoted by A and its elements by a_{ij} where a_{ij} is the quantity of ith commodity required as input to produce one unit of jth commodity. The technology matrix in our example will be

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

a_{ij}'s are called technical coefficient. (In our example there are only two produced commodities and one primary input). Note : The information in transaction matrix and hence technology matrix derived from it are obtained from past record. We assume that such information is available and reliable.

Since business firms as well as government planning agencies are more interested in the future, their problem is one of determining the total output requirements to meet the target of final demand, given existing technology.

The total output for each commodity is the sum of all intermediate demand for the product plus the final demand for the product coming from consumers, investors, government and exporters. Let X_1 and X_2 be the total output of industry 1 and industry 2 (also called total demand) and let F_1 and F_2 be the final demand for the product for Industry 1 and 2. Then we have the following two balance equations.

$$X_1 = a_{11}X_1 + a_{12}X_2 + F_1$$
$$X_2 = a_{21}X_1 + a_{22}X_2 + F_2$$

where the left hand side is the total output to be produced and the right hand side is the total demand for output. The technical coefficient a_{11}, a_{12}, a_{21} and a_{22} are known; F_1 and F_2 are given as targets (the final demand to be achieved in some future year). Hence the unknowns are the total output required to be produced by the two industries to satisfy the final demand F_1 and F_2 and the intermediate demand (($a_{11}x_1 + a_{12}x_2$). There are two equations and two unknowns and hence we can solve with the system of linear simultaneous equations (provided a solution exists). We can use the Cramer's Rule to solve X_1 and X_2. We can rewrite the above two equations in the following forms :

$(X_1 - a_{11}X_1) - a_{12}X_2 = F_1$
$-a_{21}X_1 + (X_2 - a_{12}X_2) = F_2$
or, $(1 - a_{11})X_1 - a_{12}X_2 = F_1$ .... (i)
$-a_{21}X_1 + (1 - a_{22})X_2 = F_2$ .... (ii)
Equation (i) and (ii) can be solved by using **Cramer's Rule**

$$X_1 = \frac{\begin{vmatrix} F_1; & a_{12} \\ F_2; & 1-a_{22} \end{vmatrix}}{\begin{vmatrix} 1-a_{11}; & -a_{12} \\ -a_{21}; & 1-a_{22} \end{vmatrix}}$$

$$X_2 = \frac{\begin{vmatrix} 1-a_{11}; & F_1 \\ -a_{21}; & F_2 \end{vmatrix}}{\begin{vmatrix} 1-a_{11}; & -a_{12} \\ -a_{21}; & 1-a_{22} \end{vmatrix}}$$

But first check whether the determinant of A i.e.

$\begin{vmatrix} 1-a_{11}; & -a_{12} \\ -a_{21}; & 1-a_{22} \end{vmatrix}$ is non-singular and does not vanish.

i.e., $/A/ \neq 0$)
In matrix form we can write the equation system as $X = AX + F$ where X is the column matrix (in fact a vector) $\begin{matrix} x_1 \\ x_2 \end{matrix}$

A is the 2 × 2 technology matrix $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

F is the column matrix (vector) $\begin{bmatrix} F_1 \\ F_2 \end{bmatrix}$

Therefore $X - AX = F$
Therefore $[I - A]X = F$ where I is the identify matrix

Therefore $X = \dfrac{F}{[I - A]}$

Therefore $X = [I - A]^{-1}F$
F is given, A is known and $[I - A]^{-1}$ is to be computed by the method of *inversion of a matrix*. (Those who have a knowledge of inversion of a matrix can use this method instead of the *Cramer's Rule*). The matrix [I - A] is called the *Leontief Matrix*. (In fact, *W. W. Leontief* is the inventor of the input-output model for which he has also ben awarded a Nobel Prize.)
The inverse matrix $[I - A]^{-1}$ has an economic interpretation. While matrix A gives the direct input requirement to produce one unit of output of a given industry required for final use, the matrix $[I - A]^{-1}$ gives the total (direct + indirect) input requirement to produce one unit of output of a given industry required for final use. As we said earlier the production of each and every commodity for final use requires each and every commodity as an input which in turn requires each and every commodity as an input and so on in a never ending succession of continuing cycle. By *direct input* (say cotton textile) we mean the amount of each commodity required as an input for producing one unit of a commodity (say, readymade garments) for final use. By *indirect input* (say, cotton yarn) we mean the amount of each commodity required as an input

for producing commodities (cloth) that are subsequently used as direct input plus the amount of each commodity required as an input for producing the indirect input and so forth. Direct plus indirect input gives rise to total input requirement which is given by the elements of the inverse of the *Leontief Matrix* [I-A]. Remember that the 3rd row gives the amount of primary input (labour) required to produce one unit of output of industry I and industry II to satisfy final demand. (The coefficients a_{31} and a_{32} in the technology matrix). If F_3 is the final demand for primary input and X_3 is the total supply of primary input then given the final demand target F_3 we can compute the amount of labour required for Industry 1 and 2 while producing X_1 and X_2 outputs. The total demand for primary input equals $a_{31} X_1 + a_{32} X_2 + F_3$ and must add up to the total supply if the industry's production plans are not the encounter labour shortages.

You may now appreciate that given the information contained in technology matrix, you may attempt demand forecasts for some output. You may attempt a couple of exercises given below to understand the end-use technique.

Activity 1

a) Given the following technology matrix A, and the final demand vector (F) (i.e. 3 by 1 column matrix), find the total output (demand) which the three industries must produce.

$$A = \begin{bmatrix} 0.3 & 0.4 & 0.1 \\ 0.5 & 0.2 & 0.6 \\ 0.1 & 0.3 & 0.1 \end{bmatrix} \quad F = \begin{bmatrix} 20 \\ 10 \\ 30 \end{bmatrix}$$

(Take X_1, X_2 and X_3 as the output of the three industries.)

Ans. $X_1 = 160.98 \qquad X_2 = 201.99 \qquad X_3 = 118.54$

2) Given the input-coefficient matrix (technology matrix) A and the final demand vector F, find the output levels of the three industries.

$$\begin{bmatrix} 0.0 & 0.3 & 0.3 \\ 0.3 & 0.1 & 0.1 \\ 0.2 & 0.4 & 0.0 \end{bmatrix} \text{ and } F = \begin{bmatrix} 180 \\ 20 \\ 90 \end{bmatrix}$$

Ans. $X_1 = \frac{400}{3}, X_2 = \frac{475}{3}, X_3 = \frac{1325}{3}$

BLOCK III INPUT-OUTPUT DECISIONS

This block introduces you to the basic concepts involved in input-output decisions.

Unit 8 deals with production concepts and techniques useful for analysing production decisions e.g. choice for optimum inputs, optimum product mix, etc.

Unit 9 discusses some basic concepts of costs in relation to production. It explains different types of costs relevant for production, planning and control. It also identifies the ways by which costs advantages could be derived. Finally, it indicates the managerial uses of cost concepts and analysis in decision-making.

Unit 10 deals with various approaches to empirically estimating production and costs. It gives an account of the difficulties in relation to measurement of production and cost functions. Lastly, it identifies managerial uses of empirical estimates about production and costs.

UNIT 8 PRODUCTION CONCEPTS AND ANALYSIS

Structure

8.0 INTRODUCTION

Production in concerned with the supply side of the market. The basic function of a firm is that of readying and presenting a product of sale - presumably at a profit. Production analysis relates physical output to physical units of factors of production. In the production process, various inputs are transformed into some form of output. Inputs are broadly classified as land, labour, capital and entrepreneurship (which embodies the managerial functions of risk taking, organising, planning, controlling and directing resources). In production analysis, we study the least cost combination factor inputs, factor productivities and returns to scale.

In this unit, we shall introduce several new concepts to understand the relationships involved in the production process. We are concerned with economic efficiency of production which refers to minimisation of cost for a given output level. The efficiency of production process is determined by the proportions in which various inputs are used, the absolute level of each input and productivity of each input at various levels. Since inputs have a cost attached, the degree of efficiency in production gets translated into a level of costs per unit of output.

8.1 OBJECTIVES

The objectives underlying this unit are to :

- introduce a set of concepts and techniques relevant for production decision analysis
- state clearly certain decision rules regarding optimum input choice, optimum product mix etc.
- develop a theoretical base for analysing empirical situations
- facilitate comparison between demand analysis and production analysis.

8.2 WHY STUDY PRODUCTION?

When making product decisions, a firm's management must consider both what is to be produced and how to produce it. The discussion in this unit covers decision rules for determining the quantity of various inputs to produce a firm's output under different circumstances.

It also develops a basis upon which a firm's costs can be constructed. After all, a firm incurs costs because it must pay for the productive factors. Thus an understanding of prodution helps provide a foundation for the study of cost. Knowledge about the behaviour of costs can be combined with the revenue concepts (derived from consumer demand) to establish criteria for determining the production price and the optimum level of output.

Business firms produce goods or service as a means to an end. Besides meeting the final consumer needs, the end objective of a firm may be to maximise profits, to gain or maintain market share, to achieve a target return on investment, or any combination thereof. In case of public goods, the objective may be to provide a particular service, such as education and health, within the bounds of a budget constraint. In other words, a firm attempts to combine various inputs in such a way that minimum resources are committed to produce a given product or that maximum production results from a given input. To achieve this, persons in the decision-making position should have a basic understanding of the process of production, and also the time perspective of production. Specially you may recall the concept of **Time Perspective** from introductory units. In production and cost analysis, a distinction is made between the short-run and the long-run. In the short-run, the quantities of some of the inputs are variable while others are in fixed supply. In the long-run, all factors are assumed to vary. Traditionally, the labour has been variable and capital is fixed in the short-run. While using terms like labour and capital, we should think of one unit of labour as including, say, one hour of worker's time plus all the necessary raw materials, fuel, and other variable inputs, and we should think of capital as including all the plant, equipment, land, buildings, and all other expenses that do not vary with the level of output.

8.3 PRODUCTION FUNCTION

A production function expresses the technological or engineering relationship between the output of a product and its inputs. In other words, the relationship between the amount of various inputs used in the production process and the level of output is called production function. For example, a farmer growing peanuts combines soil, fertilisers, pesticides, water, seeds, labour and capital (farm equipment, etc.) to yield a crop of peanuts. Some of these inputs may be fixed in supply: the land available, and the existing amount of capital. The other inputs may be varied and, accordingly the level of output may be expected to vary.

Traditional economic theory talks about land, labour, capital and organisation or management as the four major factors of production. Technology also contributes to output growth as the productivity of the factors of production depends on the state of technology. The point which needs to be emphasised here is that the production function describes only efficient levels of output; that is, the output associated with each combination of inputs is the maximum output possible, given the existing level of technology. Production function changes as the technology changes.

In general, we can represent the production function for a firm as

$Q = f(a,b,c.......z)$

Where a,b,c....z are amounts of various inputs such as land, labour, capital etc., and Q is the level of output for a firm. This is a positive functional relationship implying that the output varies in the same direction as the input quantities. In other words, if all other inputs are held constant, output will go up if the quantity of one input is increased. This means that the partial derivative of Q with respect to each of the inputs is greater than zero. However, for a reasonably good understanding of production decision problems, it is convenient to work with two factors of production. If labour (L) and capital (K) are the only two factors, the production function reduces to:

$Q=f(K,L)$

From the above relationship, it is easy to infer that for a given value of Q, alternative combinations of K and L can be used. It is possible because labour and capital are substitutes to each other to some extent. However, a minimum amount of labour and capital is absolutely essential for the production of a commodity. Thus, for any given level of Q, an entrepreneur will need to hire both labour and capital but he will have the option to use the two factors in any one of the many possible combinations. For example, in an automobile assembly plant, it is possible to substitute, to some extent, the machine hours by man hours to achieve a particular level of output (no. of vehicles). The alternative combinations of factors for a given output level will be such that if the use of one factor is increased, the use of another factor will decrease, and vice versa.

Activity 1

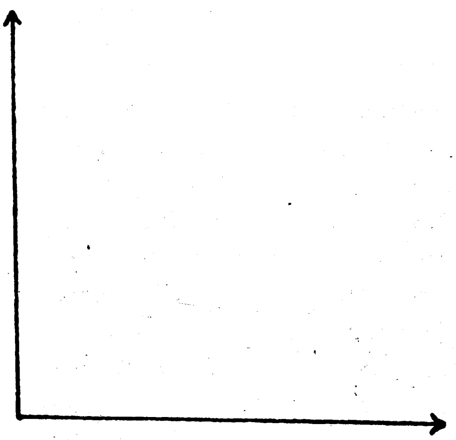
a) Draw a curve relating a single output to a single input. Name the axis. Have you made any assumptions? List them.

....

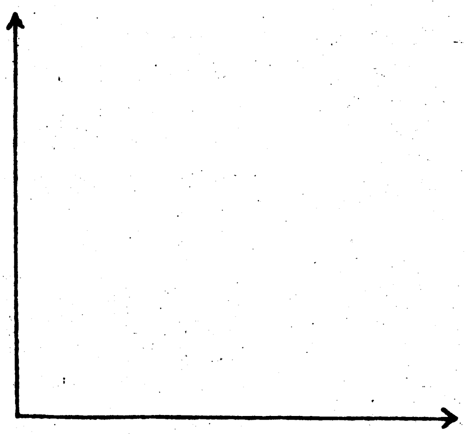
....

....

....



b) Suppose there is a technological innovation in which the same level of output can be produced by lesser input or larger level of output can be produced by the same level of input. Show the impact of this new technology in terms of the curve you have drawn in (a)



c) List the factors which enter into production of medical services by a hospital. Designate the factors in terms of symbols. State them in the form of a production function.

....

....

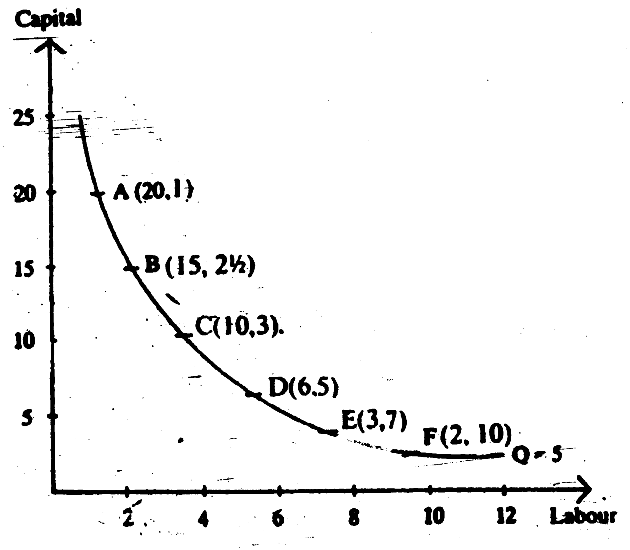
....

....

8.4 ISOQUANTS

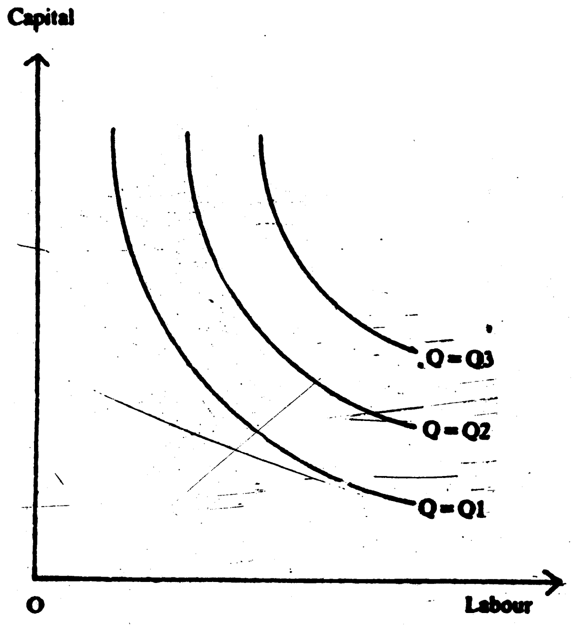
Isoquants are a geometric representation of the production function. As discussed earlier, the same level of output can be produced by various combinations of factor inputs. Assuming continuous variation in the possible combination of labour and capital, we can draw a curve by plotting all these alternative combinations for a given level of output. This curve which is the locus of all possible combinations is called **Isoquant or Iso-product curve.** Each isoquant corresponds to a specific level of output and shows different ways, all technologically efficient, of producing that quantity of output. The Isoquants are downward slopping and convex to the origin. The curvature (slope) of an isoquant is significant becuase it indicates the rate at which factors K&L can be substituted for each other while a constant level of output is maintained. As we proceed north-eastward from the origin, the output level corresponding to each successive isoquant increases, as a higher level of output usually requires greater amounts of the two inputs. Two isoquants don't intersect each other as it is not possible to have two output levels for a particular input combination.

Figure I



ISOQUANT FOR Q=5

Figure II



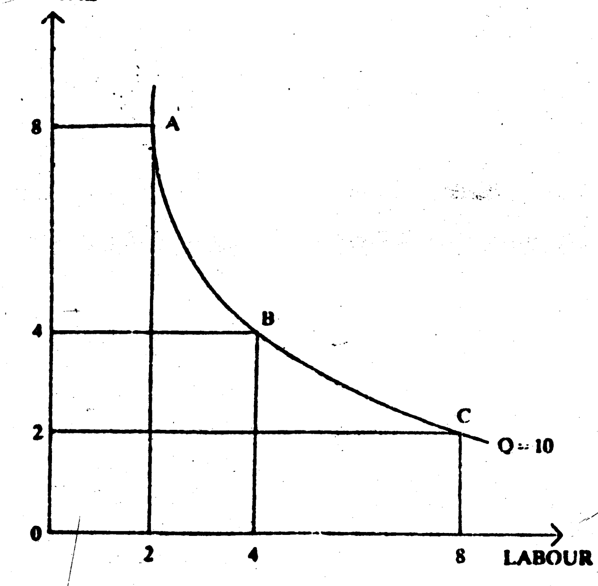
ISOQUANT MAP

Marginal Rate of Technical Substitution: (MRTS) is defined as the rate at wh : two factors are substituted for each other. Assuming that 10 pairs of shoes can be produced i. he following three ways:

Q	K	L
10	8	2
10	4	4
10	2	8

We can derive the MRTS between the two factors by plotting these combinations along a curve (Isoquant).

Figure III DIMINISHING MARGINAL RATE OF TECHNICAL SUBSTITUTION

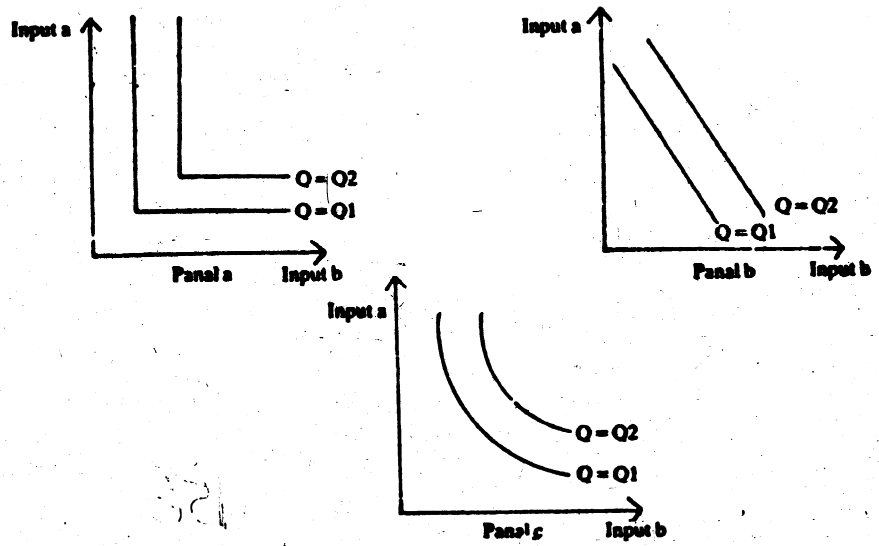


The marginal rate of technical substitution of labour for capital between points A & B is equal to $-\left(\frac{\Delta K}{\Delta L}\right) = \frac{-4}{2} = -2$. Between points B & C, the MRTS is equal to $\frac{-2}{4} = -1/2$. The Marginal rate of technical substitution has decreased because capital and labour are not perfect substitutes for each other. Therefore, as more of labour is added, less of capital can be used (in exchange for another unit of labour) while keeping the output level unchanged.

Substitutability of Inputs

Three general types of shapes that an isoquant might have, are depicted in Figure IV:

Figure IV



In panel a, the isoquants are right angles implying that the two inputs a and b must be used in fixed proportion and they are not substitutable. For instance, there is no substitution possible between the tyres and a battery in an automobile. The MRTS in all such cases would, therefore, be zero.

The other extreme case would be where the inputs a and b are perfect substitutes (panel b). The isoquants in this category will be a straightline with constant slope or MRTS. A good example of this type would be natural gas and fuel oil which are close substitutes in energy production.

The most common situation is presented in panel c. The inputs are imperfect substitutes in this case, and the rate at which input a can be given up in return for one more unit of input b keeping the output constant, diminishes as the amount of input b increases.

Measures of Production

The measure of output represented by Q in the production function is the total product that results from each level of input use. For example, assuming that there is only one factor (L) being used in the production of cigars, total output at each level of labour employed could be:

Labour (L)	Output (Q)	Labour (L)	Output (Q)
1	3	8	220
2	22	9	239
3	50	10	246
4	84	11	238
5	121	12	212
6	158	13	165
7	192	14	94

The total output will be 220 cigars if we employed 8 units of labour.

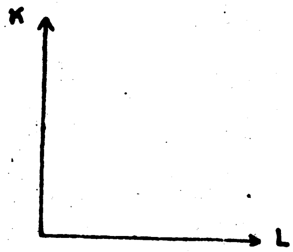
We assume in this example, that the labour input combines with other input factors of fixed supply and that the technology is a constant.

In addition to the measure of total output, two other measures of production i.e. marginal product and average product, are important to understand.

Activity 2

a) Draw an isoquant map using the information available in the following table:

Isoquant I		Isoquant II		Isoquant III	
L	K	L	K	L	K
2	11	4	13	6	15
1	8	3	10	5	12
2	5	4	7	6	9
3	3	5	5	7	7
4	2.3	6	4.2	8	6.2
5	1.8	7	3.5	9	5.5
6	1.6	8	3.2	10	5.3
7	1.8	9	3.5	11	5.5



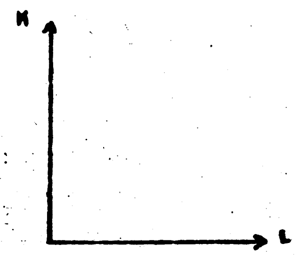
b) Which one of the isoquants provides you with the highest level of output and Why?

c) Take any one of the isoquants and compute $MRTS_{LK}$. What do you observe about computed MRTS? Explain the observed trend.

Isoquant —

L	K	MRTS

d) Illustrate the impact of technological progress in terms of a shift in a given isoquant.



8.5 TOTAL, AVERAGE AND MARGINAL PRODUCTS

Marginal Product

This has reference to the fundamental concepts of marginalism. From the decision making point of view, it is particularly important to know how production changes as a variable input is changed. For example, we want to know if it would be profitable to hire an additional unit of labour for some additional productive activity. For this, we need to have a measure of the rate of change in output as labour is increased by one unit, holding all other factors constant. We call this rate of change the **marginal product of labour**. In general, the marginal product (MP) of a variable factor of production is defined as the rate of change in total product (TP or Q).

Here the output doesn't increase at a constant rate as more of any one input is added to the production process. For example, on a small plot of land, you can improve the yield by increasing the fertiliser use to some extent. However, excessive use of fertiliser beyond the optimum quatity may lead to reduction in the output instead of any increase as per the Law of Diminishing Returns. (For instance, single application of fertilisers may increase the output by 50 per cent, a second application by another 30 per cent and the third by 20 per cent. However, if you were to apply fertiliser five to six times in a year, the output may drop to zero).

It is safe to say that most production functions react in the same general fashion as more of a variable input is added to other fixed factors (land, machinery etc.) In our earlier example of the production function for cigars, first the output increases at an increasing rate (for L less than 5), then increase at a slower rate (for L between 5 and 10), and finally starts declining (for L greater than 10). Also look at Table 2 (Production Function for Shoes) to see what happens to the rate of increase in output (MP) as more of labour is used. The **principle of diminishing marginal productivity** (returns) states that as additional units of a variable input are added to other inputs that are fixed in supply, the increments to output eventually decline (for a constant technology). This phenomenon has been widely observed and there is ~~enough empirical~~ evidence to support it. For business managers, marginal productivity of an input plays an important part in determining how much of that input will be employed.

Average Product

Often, we also want to know the productivity per worker, per kilogram of fertiliser, per machine, and so on. For this, we have to use another measure of prodution: average product. The **average product** (AP) of a variable factor of production is defined as the total output divided by the number of units of the variable factor used in producing that output. Suppose there are n factors (X_1, X_2, ... X_n), and the average product for the ith factor is defined as:

$$AP_i = TP/X_i$$

This represents the mean (average) output per unit of land, labour, or any other factor input.

The concept of average product has several uses. For example, whenever inter industry comparisons of labour productivity are made, they are based on average product of labour. Average productivity of workers is important as it determines, to a great extent, the competitiveness of one's products in the markets. The wage revision are also linked to productivity. If wages are revised upward without much regard to average productivity, a country faces inflationary pressures. The concepts like "quality circles" and other forms of employee participation in management are all aimed at increasing average productivity of worker.

Marginal Average, and Total Product: a Comparison

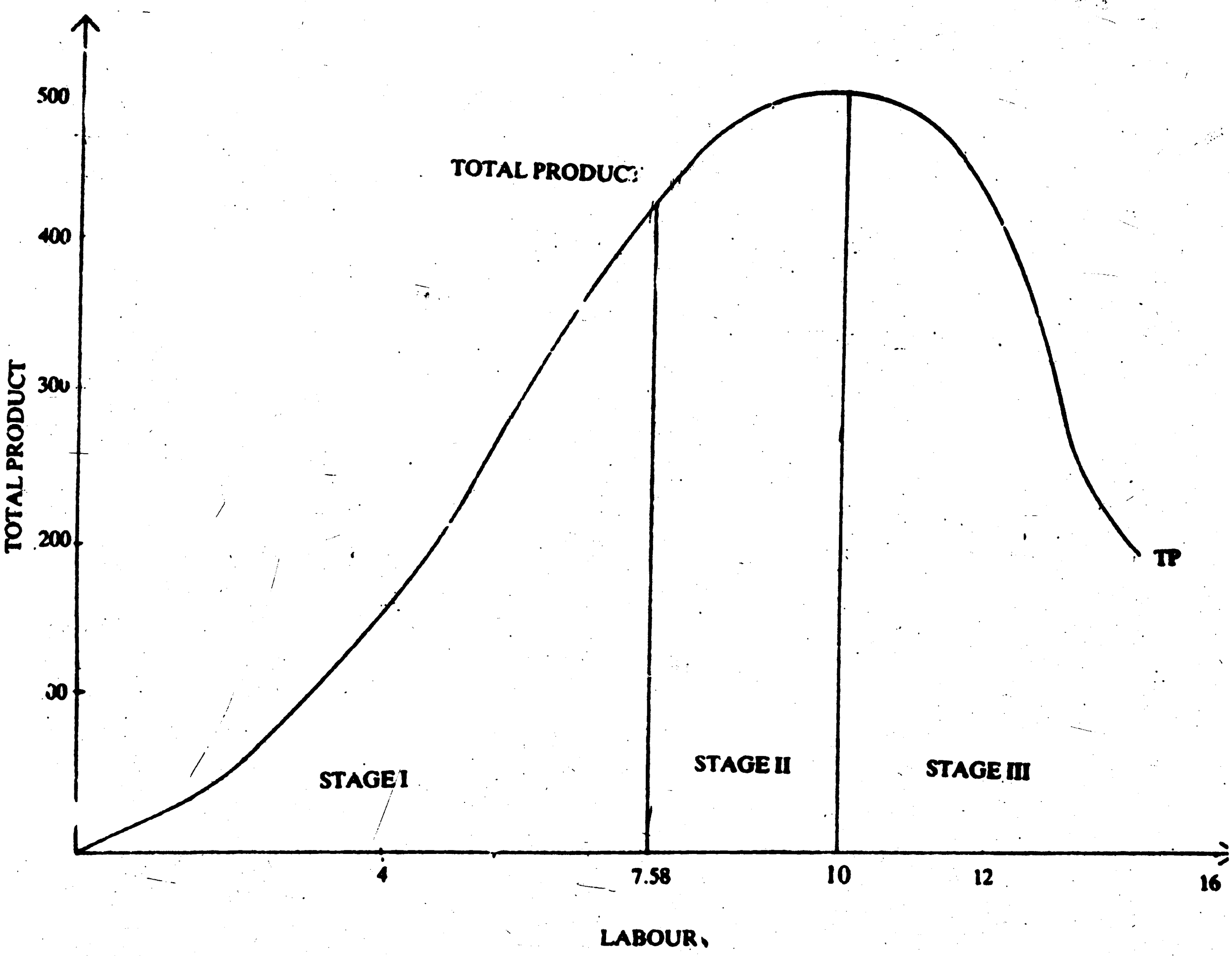
A hypothetical production function for shoes is presented in Table 2 with the total, average and marginal products of the variable factor labour. Needless to say that the amount of other inputs, and the state of technology are fixed in this example.

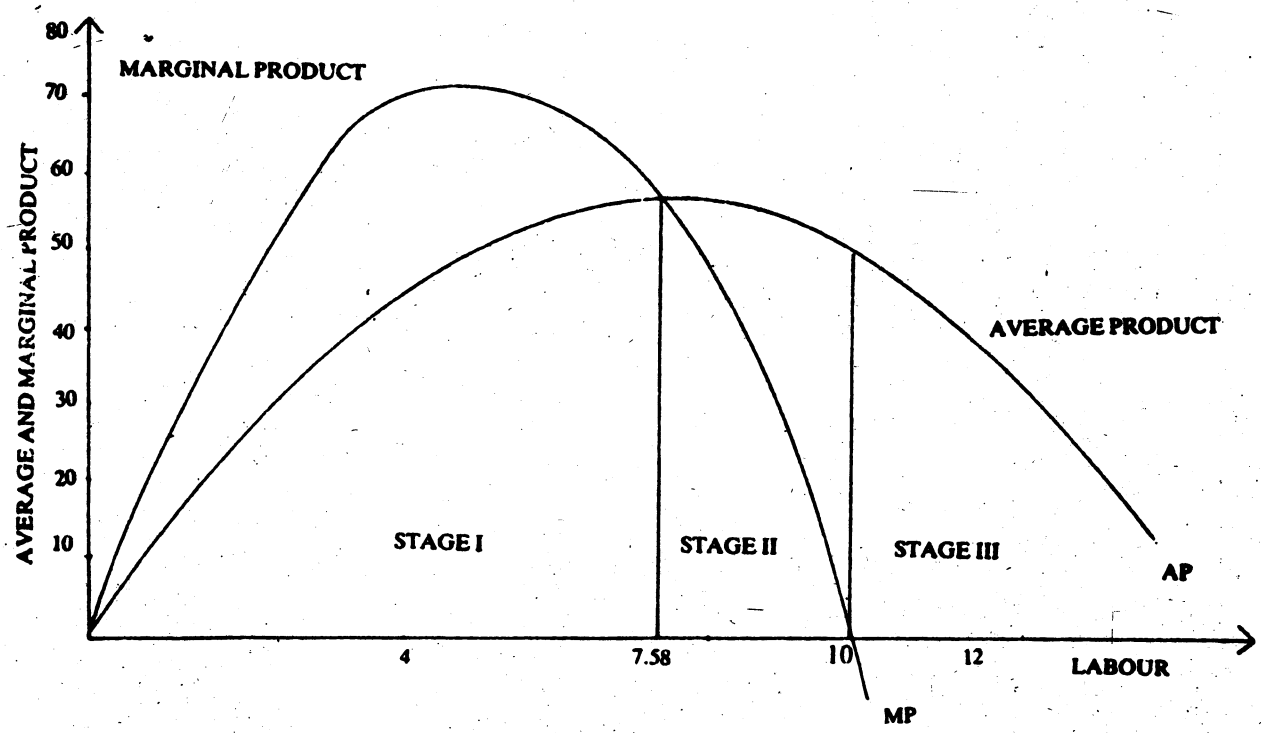
Table 2: Total, Average and Marginal Product

Labour Input (L)	Total Output (TP)	Average Product (AP=TP/L)	Marginal Product $\left(MP = \frac{\Delta TP}{\Delta L}\right)$
0	0	0	0
1	14	14	14
2	52	26	38
3	108	36	56
4	176	44	68
5	250	50	74
6	324	54	74
7	392	56	68
8	448	56	56
9	486	54	38
10	500	50	14
11	484	44	-16
12	432	36	-52
13	338	26	-94
14	196	14	-142

The values for marginal product are written between each increment of labour input because those values represent the marginal productivity over the respective intervals. The data presented in the above table is also graphed in Figure V (See next page)

Figure V. Total, Average and Marginal Product Functions for the Production of Shoes





In both the table and the graphic representation, we see that both average and marginal products first increase, reach the maximum, and eventually decline. Note that MP=AP at the maximum of the average product function. This is always the case. If MP>AP, the average will be pushed up by the incremental unit, and if MP<AP, the average will be pulled down. It follows that the average product will reach it peak where MP = AP.

Activity 3

a) Fill in the blanks in the following table.

Capital	**Labour**	**TP**	**AP_L**	**MP_L**
1	0	0	—	—
1	1	2	2	—
1	2	5	—	3
1	3	—	3	4
1	4	12	3	—
1	5	14	—	—
1	6	—	$2\frac{1}{2}$	1
1	—	15	$2\frac{1}{2}$	—
1	8	14	—	—
1	9	—	$1\frac{1}{3}$	−2

b) State clearly the relation between AP_L and MP_L

c) How many labourers should be employed and Why? State the principle, if any.

(Hint: you may recall the Absolute Activity Level Principle from introductory units)

Our production function for shoes can also be expressed as an algebraic function of the labour input. We can determine the amount of labour to be employed if we want to maximise the production of shoes, using the algebraic function which would be:

$TP = 15L^2 - 1L^3$

It follows that the marginal product is

$MP = \frac{d(TP)}{dL}$

$MP = 30L - 3L^2$.

For TP to be maximum, MP = 0

$$\therefore MP = 30L - 3L^2 = 0$$

$$L(30-3L) = 0$$

which has two solutions: L = 0 and L = 10. We should expect total product to have a maximum at one of these points and minimum at the other. In this case, output will be maximum at L = 10, at which both first order or second order conditions are satisfied. In Figure V, we see that in the production of cigars, both marginal product and average product at first rise, then reach a maximum and then fall. We can find the point at which MP reaches a maximum as follows:

$MP = 30L - 3L^2$

$\frac{d(MP)}{dL} = 30 - 6L = 0$

$$L = 5$$

For amounts of labour input of more than 5, the marginal product has a negative slope

$(i.e. \frac{d(MP)}{dL} < 0 \text{ for } L > 5)$

Similarly, we know that the average product is equal to marginal product when average product is maximum. Thus in this case

$AP = \frac{TP}{L} = 15L - IL^2$

Finding the first derivative of AP, setting it equal to zero, and solving for L, we obtain:

$\frac{d(AP)}{dL} = 15 - 2L = 0$

$L = 7.5$

In order to check that AP=MP when AP is maximum, we must compute the values of AP and MP at L = 7.5. We know that

$MP = 30L - 3L^2$

$\therefore MP = 30(7.5) - 3(7.5)^2$

$= 56.25$

Similarly $AP = 15L - 1L^2$

$\therefore AP = 15(7.5) - 1(7.5)^2$

$= 56.25$

Therefore we find that both are equal at L = 7.5

Note: The production function has been expressed in cubic form as both marginal and average products start falling after a point. If we substitute different values of L in the marginal product function, the MPL will be different from the values presented in Table 2. This is because we have used MPL = $\frac{\Delta TP}{\Delta L}$ in the table which is a discrete function, where as MPL = $\frac{d(TP)}{dL}$ used in algebraic form is a continuous function.

Activity 4

a) Suppose, the production function relating agricultural output to varying amounts of capital input is of the following form:

$$Q = -\frac{K^3}{3} + 2K^2 + 12K$$

Beyond what point do diminishing returns exist?

....

....

....

....

....

....

....

....

Ans: K = 3 (There is another answer which is mathematically correct but economically uninteresting)

b) The law of Diminishing Returns is sometimes called the law of Variable Proportions. How would you justify this?

....

....

....

....

c) Consider the three-stage diagram. Locate the points of

	Total Product	Labour
i) Intensive Margin (at which MP=0):		
ii) Extensive Margin (at which MP=AP):		

Elasticity of Production

This is a concept which is based on the relationship between Average Product (AP) and Marginal Product (MP). The elasticity of production (e_q) is defined as the rate of fractional change in total product, $\frac{\Delta Q}{Q}$ relative to a slight fractional change in a variable factor, say labour, $\frac{\Delta L}{L}$. Thus

$$e_q^1 = \frac{\Delta Q/Q}{\Delta L/L} = \frac{\Delta Q}{\Delta L}\cdot\frac{L}{Q} = \frac{\Delta Q/\Delta L}{Q/L} = \frac{MP_L}{AP_L}$$

Thus labour elasticity of production, e_q^1 is the ratio of marginal productivity of labour to average productivity of labour. In the same way, you may find that **capital elasticity of production** is simply the ratio of marginal productivity to average productivities of capital. Sometimes, such concepts are renamed as **input elasticity of output**. In an estimated production function, the aggregate of input elasticities is termed as the **function coefficient**.

Elasticity of Factor Substitution

This is another concept of elasticity which has a tremendous practical use in the context of production analysis. The **elasticity of factor substitution**, e_s^f, is a measure of ease with which the varying factors can be substituted for others; it is the percentage change in factor proportion $(\frac{K}{L})$ with respect to a given change in marginal rate of technical substitution between factors ($MRTS_{KL}$). Thus,

$$e_s^f = \frac{\Delta(K/L)}{(K/L)} \cdot \frac{(MRTS_{KL})}{\Delta(MRTS_{KL})}$$

$$= \frac{\Delta(K/L)}{\Delta(MRTS_{KL})} \cdot \frac{(MRTS_{KL})}{(K/L)}$$

$$= \frac{\Delta(K/L)}{\Delta(MR_L/MP_L)} \cdot \frac{(MP_K/MP_L)}{(K/L)}$$

The elasticity coefficient of factor substitution, e_s^f, differs depending upon the form of production function. You should be able to see now that factor intensity (factor ratio), factor productivity, factor elasticity and elasticity of factor substitution are all related concepts in the context of production analysis.

8.6 THREE STAGES OF PRODUCTION

The product curves in Figure V have been divided into three stages. The first stage covers the region from the origin to the point at which the average product is maximum at L = 7.5. Throughout Stage I, the average product is rising, even though the marginal product has begun to fall as we near the end of Stage I. During this stage of production, the amount of labour is too small for the existing amount of the fixed resource which results in negative marginal product of that resource.

Stage II starts from the point at which the average product is maximum and ends at the point where the marginal product falls to zero. This is the region between L = 7.5 and L = 10. The marginal and average products are falling but both are positive throughhout Stage II.

Stage III covers the entire region for which the marginal product of labour is negative i.e. the area to the right of L = 10. As in Stage I, this happens because the amount of fixed resources available are not sufficient for the efficient production to take place. What this implies is that of the three stages of production, only Stage II has a positive marginal product for all inputs. The important implication of this relationship for business is that the relevant range of production is restricted to Stage II. No enterprise would knowingly employ a resource for which the marginal product is negative.

8.7 OPTIMAL INPUT CHOICE:LEAST COST COMBINATION OF INPUTS

As already discussed, the production function indicates the alternative combinations of various factors of production which can produce a given level of output. While all these combinations are technically efficient, the final decision to employ a particular input combination is purely an economic decision and rests on cost. An entrepreneur should choose that combination which costs him the least. To aid our thinking in this regard economists have developed the concept of isocost (equal cost) line, which shows all combinations of inputs (a&b) that can be employed for a given cost (in rupees).

In order to determine the least cost combination for a given output we need to have the prices of factors of production. Let us consider, a production function for plastic buckets where the entrepreneur wants to produce 20 buckets. Let the price of L (P_L) be Rs. 10 per unit and the price of capital (P_K) be Rs. 5 per unit. It is assumed that unlimited amounts of labour and capital can be bought at given prices. We can now find the total cost of each of the five possible combinations of labour and capital for Q = 20.

Alternative Combination	Inputs in Physical Units		Cost (Rs.)
	Labour	Capital	
1	4	17	4 × 10 + 17 × 5 = 125
2	5	12	5 × 10 + 12 × 5 = 110
3	6	8	6 × 10 + 8 × 5 = 100
4	7	5	7 × 10 + 5 × 5 = 95
5	8	4	8 × 10 + 5 × 4 = 100

Combination 4 represents the least cost for producing 20 plastic buckets.

Another way to determine the least cost combination is geometrical in natural which uses the isocost and isoquant curves. The isocost line can be defined in the following manner:

$C = P_L \cdot L + P_K \cdot K$, where $C = C$.

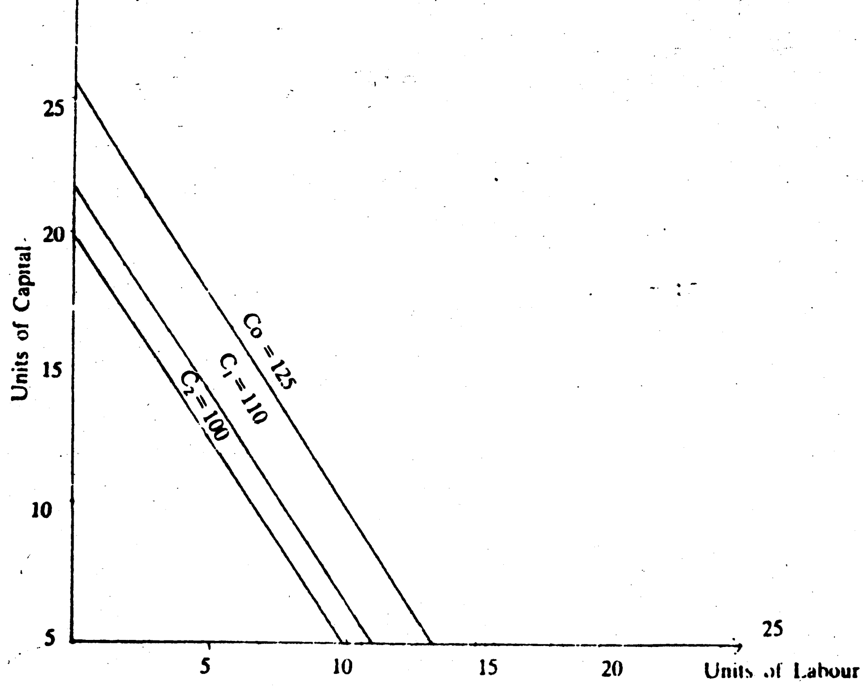
Where C_o is the firm's total cost of inputs for some specified time period, P_L and P_K are the prices of input L and input K, respectively and L and K represent the physical quantities of the inputs. In other words, the isocost equation above states that when the firm's total cost is C_o the price of input L times the amount of input L used plus the price of input K times the amount of input L used must equal C_O. In Figure VI we have drawn isocost lines for C_O = 125, C_1 = 110, C_2 = 100, and C_3 = 95, and C_4 = 100 where PL = 10 and PK = 5. Note that these five isocost lines are parallel.

They must be parallel because the slope of each line is $-\frac{P_L}{P_K}$ or $-\frac{10}{5} = -2$.

Note that the slope of an isocost line must be equal to $-\frac{P_L}{P_K}$, since that represents the rate at which input L can be substituted by input K while maintaining the same level of output. In our example, P_L = 10 and P_K = 5, then we can substitute 1 units of labour for every 2 unitsof capital while maintaining the same cost level. Thus,

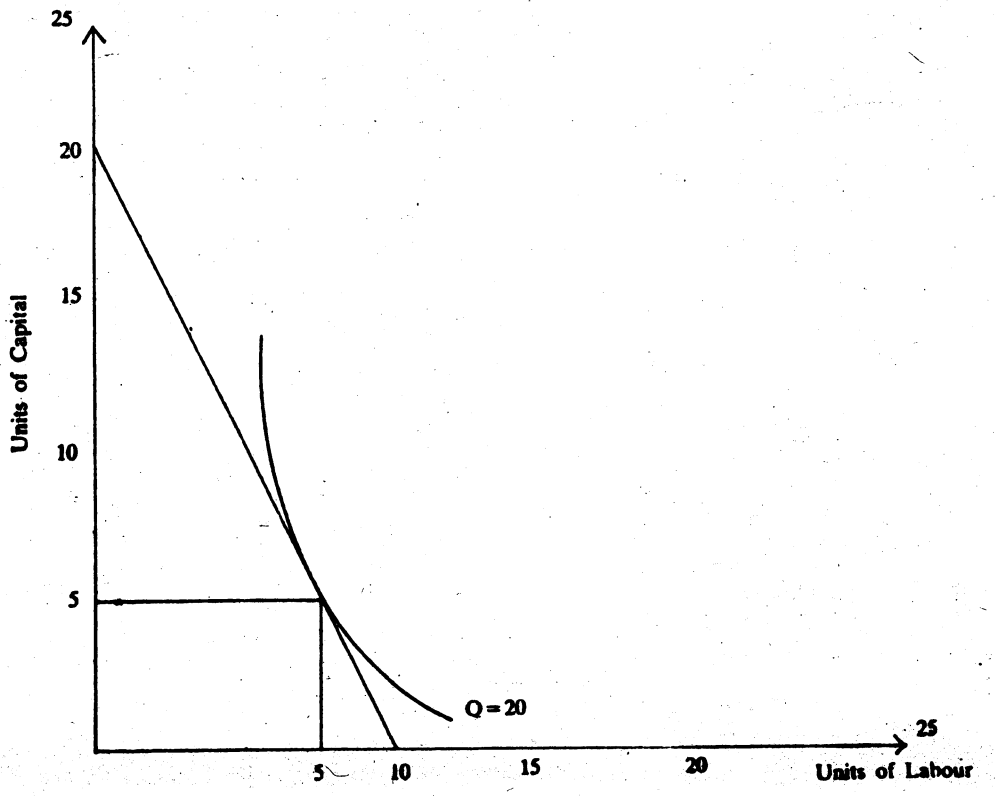
$\frac{\Delta K}{\Delta L} = -\frac{P_L}{P_K} = -\frac{10}{5} = -2$. We should remember that the marginal rate of substitution $-\frac{\Delta K}{\Delta L}$ which is the slope of an isoquant between two points, is also equal to the ratio of margina products of the two inputs $(\frac{MP_L}{MP_K})$. We know that $MP_L = \frac{\Delta Q}{\Delta L}, MP_K = \frac{\Delta Q}{\Delta K}$. Along ar isoquant the increase in output resulting from an additional unit of L must be exactly offset by the decrease in output from a reduction in input K, or $\Delta Q = O = MP_L(\Delta L) + MP_K(\Delta K)$ Thus, $MP_L(\Delta L) = -MP_K(\Delta K)$, or $\frac{MP_L}{MP_K} = \frac{-\Delta K}{\Delta L}$, Q constant.

Figure VI ISOCOST LINES



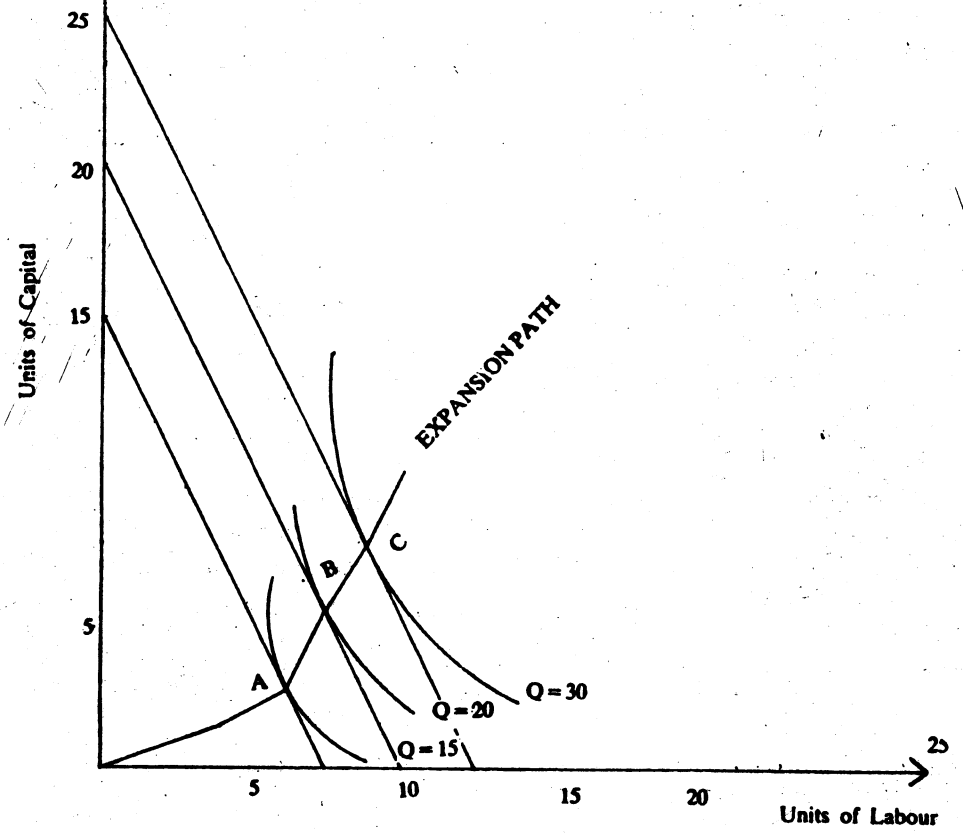
With this relationship clear in mind, we superimpose the isoquant map on the isocost map in order to determine the least cost input combination or the maximum output for a given cost. It can be seen in Figure VII that the maximum output that can be obtained with an outlay of Rs. 95 is 20 buckets where the isocost C = 95 is tangent to the isoquant Q = 20. This is the least cost of producing 20 buckets and the least cost combination of inputs in this case is 7 units of labour and 5 units of capital. Any other combination on the isoquant Q = 20 will have a cost higher than 95.

Figure VII LEAST COST INPUT COMBINATION



The least cost combination for different levels of output i.e. Q = 15, (Q=25) can be found in the same way. The line ABC in Figure VIII thus represents the least cost combination of inputs for different levels of output. This line is called the firm's **expansion path** or **the scale line.**

Figure VIII FIRM'S EXPANSION PATH



Alternatively, you may frame and solve your opimisation problem (of input choice) through the standard calculus technique of Lagrangian multiplier **Constrained Maximisation**

Max. Q(K,L) subject to

$C = P_L.L + P_K.K \Rightarrow$ (cost constraint)

$M = Q(L,K) + \lambda(\bar{C} - P_L.L - P_K.K)$

$\frac{\partial M}{\partial L} = \left\{\frac{\partial Q}{\partial L} - \lambda P_L\right\} = 0 \Rightarrow \left\{\frac{\partial Q}{\partial L} = P_L\right\}$ ..... (i)

$\frac{\partial M}{\partial K} = \left\{\frac{\partial Q}{\partial K} - \lambda P_K\right\} = 0 \Rightarrow \left\{\frac{\partial Q}{\partial K} = P_K\right\}$ ..... (ii)

$\frac{\partial M}{\partial \lambda} = (\bar{C} - P_L L - P_K K) = 0 \Rightarrow (\bar{C} = P_L L + P_K K)$ ..... (iii)

solving (i) and (ii) simultaneously, we get

$\frac{\partial Q/\partial L}{\partial Q/\partial K} = \frac{P_L}{P_K}$

or $MRTS_{LK} = \frac{P_L}{P_K} \Rightarrow$ optimum decision rule

Constrained Minimisation

Min. $C = P_L L + P_K K$ subject to

$\bar{Q} = Q(L, K) \Rightarrow$ (technology constraint)

$M^* = P_L L + P_K K + \lambda^* [\bar{Q} - Q(L,K)]$

$\frac{\partial M^*}{\partial L} = \lambda\left[P_L = \lambda^* \frac{\partial Q}{\partial L}\right] = 0 \Rightarrow \left(P_L = \lambda^* \frac{\partial Q}{\partial L}\right)$ (iv)

$$\frac{\partial M^*}{\partial K} = \Delta\left[P_K = \lambda^* \frac{\partial Q}{\partial K}\right] = 0 \Rightarrow \left(P_K = \lambda^* \frac{\partial Q}{\partial K}\right) \qquad \text{..... (v)}$$

$$\frac{\partial M^*}{\partial \lambda^*} = \left[\bar{Q} - Q(L,K)\right] = 0 \Rightarrow \left[\bar{Q} = Q(L,K)\right] \qquad \text{..... (vi)}$$

Again through similar operations on (iv) and (v) we get

$$MRTS_{LK} = \frac{\frac{\partial Q}{\partial L}}{\frac{\partial Q}{\partial K}} = \frac{P_L}{P_K} \Rightarrow \underline{\text{optimum decision rule}}$$

In other words, the decision rule in the context of optimum input choice remains unchaged irrespective of the nature of constraints.

Activity 5

Find the optumum input combination in each case given below:

a) $Q = \sqrt{L}\sqrt{K}$
   Rs. 80=2L+4K

(Ans: L/K=1/2)

b) **Q = 4 Log L + 8 Log K**
   C = 3L + 9K

(Ans: L/K=3/2)

How Much of an input should be employed?

Would you hire a worker if the cost of doing so, including wages and all the benefits, is more than the revenue you can get by selling the output attributable to that worker? The obvious answer is 'no', because it is not economic to hire him. But, you will go ahead and hire another unit of labour if you can sell the added output for more than the added cost of employing that labour. This is the criterion on which the decision to employ more units of any resource should be based.

Marginal revenue product of labour is defined as the additional revenue that a firm receives by selling the output of an additional worker. This is measured in value terms. It is calculated as:

MRP = (MP) (MR)

Where MP is the marginal product of labour, and MR is the marginal revenue. It is a diminishing function as the MP decreases, employ more of labour. MR can be either diminishing or constant. Marginal Resource Cost (MRC) is the additional cost to the firm for employing one more unit of labour. The optimum number of workers to be employed is determined by the condition:

MRP = MRC.

Activity 6

a) Review your understanding of the concepts of Marginal Product (MRP) and Marginal Resource (Factor) Costs (MRC)

....

....

....

....

b) A profit maximising employer should follow the equi-marginal principle MRP=MRC. Explain.

....

....

....

....

c) Suppose a firm can obtain a valuable input without having to pay for it (e.g. apprentices who receive no pay or free electricity available in a backward area — zero industry area). How does the firm decide how much of the free input to take.

....

....

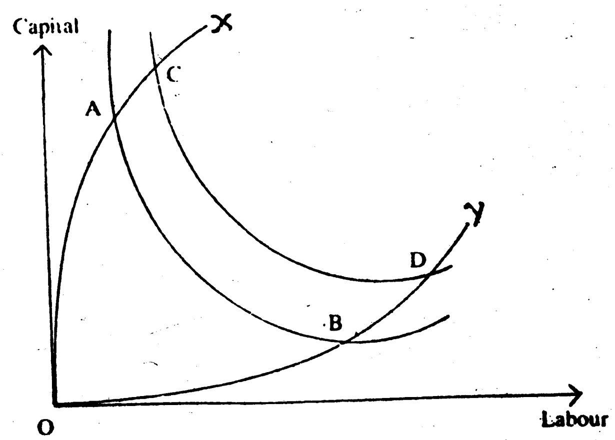
....

....

8.8 ECONOMIC REGION OF PRODUCTION

In the long-run, a firm should use only those combinations of inputs which are economically efficient. A factor should not be used beyond a point, even if it is available free of cost, as it will result in negative marginal product for that factor. These input combinations are represented by the position of an isoquant curve which has a positive slope. A positively slopped isoquant means that the firm will have to use of both the inputs to maintain same level of production. As illustrated in Figure IX the point A on I_0 marks the spot where MP_K = O. Beyond this, the MP_K is negative. Similarly, at point B on I_O, MP_L = O and it is negative there after. The lines OX and OY which are called ridge lines bound the economic region of production beyond which it is economically inefficient to produce.

Figure IX ECONOMIC REGION OF PRODUCTION



8.9 RETURNS TO SCALE

The law of diminishing returns states that as more and more of the variable input is added to the fixed factor base, the increment to total output after some point will decline progressively with each additional unit of the variable factor. The law of diminishing returns is also broadly referred to as the 'law of variable proportions' which implies that as additional units of a variable factor are added to a given quantity of all other factors, the increment to output attributable to each of the additional units of the variable factor will increase at first, decrease later, and eventually become negative. The law of diminishing returns is strictly a short-run phenomenon.

Let us now look at what happens if we change all inputs simultaneously which is possible only in the long-run. What happens to the output level as all factor inputs are increased proportionately? This can be understood with the help of the concept known as **returns to scale.** Under this concept, the behaviour of output is studied when all factors of production are changed in the same direction and in the same proportion. Returns to scale are categorised as follows:

a) **Increasing returns to scale:** If output increases more than proportionate to the increase in all inputs.

b) **Constant returns to scale:** If all inputs are increased by some proportion, output will also increase by the same proportion.

c) **Decreasing returns to scale:** If increase in output is less than proportionate to the increase in all inputs.

For example, if all factors of production are doubled and output increase by more than two times, we have a situation of increasing returns to scale. On the other hand, if output does not double even after 100 percent increase in input factors, we have diminishing returns to scale.

8.10 FORMS OF PRODUCTION FUNCTION

There are five different forms of production function. Understanding the form is important in the context of interpreting statistically estimated production function i.e. empirical situation. However, it may be worthwhile to make a passing reference to the conceptual basis of different types of production function:

a) Cobb-Douglas type

It is a linearly homogeneous production function of degree one i.e. subject to constant returns to scale.

$Q = Q(K,L) = AL^{\alpha} K^{1-\alpha}$

Where A and α are constants; $1>\alpha>0$.

You may note that $[(\alpha) + (1-\alpha) = 1]$.

This means if factors K and L are increased by λ proportion, the output Q will also increase by the same proportion. This means constant returns to scale. Interestingly, you may operate on this form of production function to review all the production concepts you have learnt so far:

i) **Total product**

$$Q = AL^{\alpha}K^{1-\alpha}$$

ii) **Average products**

$$AP_L = A(\tfrac{K}{L})^{1-\alpha}$$

$$AP_K = A(\tfrac{L}{K})^{\alpha}$$

iii) **Marginal products**

$$MP_L = \alpha \,.A(\tfrac{K}{L})^{1-\alpha} = (\alpha \,.\tfrac{Q}{L})$$

$$MP_K = (1-\alpha)A(\tfrac{L}{K})^{\alpha} = \left[(1-\alpha)\tfrac{Q}{K}\right]$$

iv) **Marginal Rate of Technical Substitution**

$$MRTS_{LK} = \frac{MP_L}{MP_K} = \left[\frac{\alpha}{(1-\alpha)}\frac{K}{L}\right]$$

v) **Elasticity of Proudction**

$$e_q^1 = \frac{MP_L}{AP_L} = \alpha$$

$$e_q^1 = \frac{MP_K}{AP_K} = (1-\alpha)$$

vi) **Function Coefficient**

$$f = (\alpha) + (1-\alpha) = 1$$

vii) **Production efficiency**

It is measured by the term 'A'. The firm with higher 'A' produces larger output of 'Q'; high level of 'A' signifies better organisation and managerial efficiency.

viii) **Factor intensity**

This measures the relative factor proportion (or ratio) to produce a given level to output. Along the isoquant, factor intensity may vary or remain constant, depending upon the returns to factor (short-run) and the returns to scale (long-run)

$\frac{K}{L}$ measures capital intensity

$\frac{L}{K}$ measures labour intensity

ix) **Elasticity of factor Substitution**

$$e_s^f = \frac{\%\Delta in(K/L)}{\%\Delta in(MRTS_{KL})}$$

$$\frac{\Delta(K/L)/(K/L)}{\Delta(MRTS_{KL})/(MRTS_{KL})}$$

$$= \frac{\Delta(K/L)/(K/L)}{\left(\frac{\alpha}{1-\alpha}\cdot\frac{K}{L}\right)/\left(\frac{\alpha}{1-\alpha}\right)\left(\frac{K}{L}\right)}$$

$$= \frac{\Delta(K/L)/\left(\frac{\alpha}{1-\alpha}\right)}{\frac{\alpha}{(1-\alpha)}\Delta(K/L)} = 1$$

You may operate on other forms of production to review the same set of concepts. However, here it will suffice to note the following.

b) C E S type

It stands for 'constant' elasticity of substitution production function. As the name implies, here e_s^f remains constant. The Cobb-Douglas function is a special type of C E S function.

c) V E S type

It stands for 'variable elasticity of substitution' production function, because e_s^f is allowed to vary.

d) Leontief type

In this type of production function, e_s^f is assumed to be zero. This means factors are combined in fixed proportion such that there is no scope for factor substitution.

e) Linear type

At the other extreme, e_s^f may be assumed to be infinity, meaning that there is no limit to factor substitution such that a multi-factor production may turn out to be a single factor function like.

$Q = Q(K,L) = Q = 2L$, because K and L are perfect substitutes.

Activity 7

Construct a set of tables with imaginary data to illustrate the following:

a) Diminishing returns to a variable factor

b) Increasing returns to scale

c) Constant returns to scale

d) Diminishing returns to scale

Activity 8

a) Consider the production function

$$Q = 7\frac{L^5}{K^4} + 3\frac{K^2}{L}$$

Establish the form of this function

b) Consider the function

$Q = 86\, L^{0.61}\, K^{0.39}$

Compute the following:

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

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

$MRTS_{LK} =$

$e_q^f =$

$e_q^k =$

$e_s^f =$

f =

A =

c) **Explain each of the following proportions.**

i) **Factor productivity depends on factor-intensity in production.**

....

....

....

....

....

ii) Factor substitutability depends on factor productivity.

....

....

....

....

....

....

iii) Isoquant map for a Leontief type of production has a special shape and slope.

....

....

....

....

....

....

iv) For Cobb-Douglas function, the expansion path is linearly homogeneous of degree one.

....

....

....

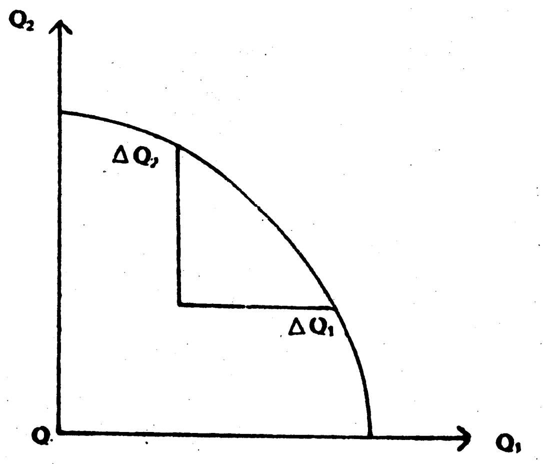
....

8.11 OPTIMAL PRODUCT-MIX

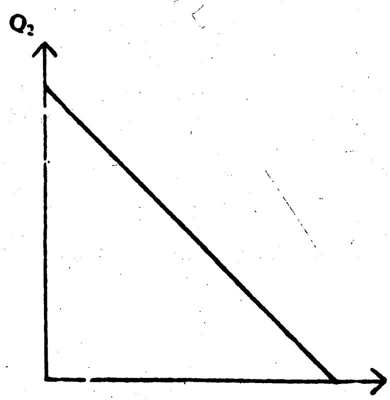
Hitherto we have considered the case of single-product firm. Let us now move on to the case of joint-product or multi-product firm. The joint-product case is a special type of multi-product firm. Most of the firms today produce a combination of products-at least main-product and its by-products, for example, Durgapur Steel Plant produces steel and gas. In other words, you may now assume that there is a firm which produces two products, Q_1 and Q_2 jointly out of a given factor, F. It is like saying both mutton and wool are being produced out of a sheep. The question is what is the optimal product-mix in this case. Before you attempt an answer to this question, you may note:

$F=F(Q_1, Q_2)$ is the transformation or production possibility function, which when plotted gives you the produce transformation curve.

The slope of this curve, $\frac{\Delta Q_1}{\Delta Q_2}$ measures the Marginal Rate of Product Transformation **($MRPT_{Q_1,Q_2}$)**



Now the producer get his revenue from the market by producing and selling both products. Thus, his total revenue, $TR = P_1Q_1 + P_2Q_2$ when P_1 and P_2 are respective prices of Q_1 and Q_2. We now get AB the iso revenue curve, where slope measures the product price ratio $\frac{P_1}{P_2}$

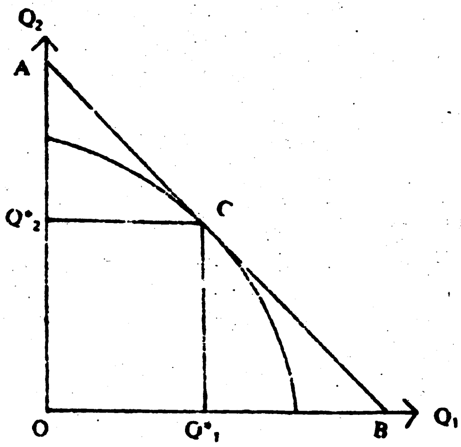


The firm may now try to maximise the revenue subject to the constraint imposed by the production possibility or minimise the resource use to attain a given revenue. Either way, the decision rule for an optimum product-mix works out as:

$$\therefore \frac{\Delta Q_1}{\Delta Q_2} = MRPTQ_1Q_2 = \frac{P_1}{P_2}$$

This is illustrated in the diagram; corresponding to the equilibrium condition the firm produces a combination of $OQ_1^* + OQ_2^*$.

You may note that some degree of substitutability between Q_1 and Q_2 is being assumed. If more of Q_1 is produced out of given resource, F, less of Q_2 can be produced.



The law of diminishing returns is said to operate. Constant returns to scale will be witnessed if the output also exactly doubles in this case.

Activity 9

a) Quote a couple of examples of joint-product or multi-product firms. Name the firm and its products.

....

....

....

....

....

b) Given $F = F(Q_1, Q_2)$.

$\frac{\partial F}{\partial Q_1}$ = marginal resource cost of producing Q_1

$\frac{\partial F}{\partial Q_2}$ = marginal resource cost of producing Q_2

Show the $MRPT_{Q_1Q_2}$ is simply the ratio of these marginal costs (measured in physical resources).

....

....

....

....

c) Given $F = 500\ Q^1, 0.3Q^2\ 0.7$ show that 0.3 or 0.7 measures product elasticity of resource use, $e_f^{q^1}$ or $e_f^{q^2}$.

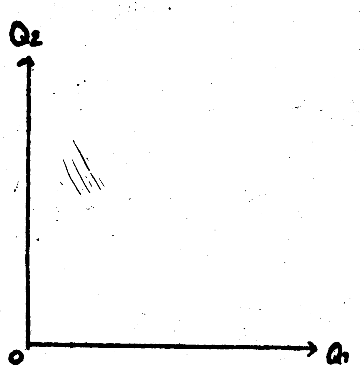
....

d) Also estimate the elasticity of product transformation, e_t^q, in the above estimated function. Provide an economic interpretation of e_t^q. Note that

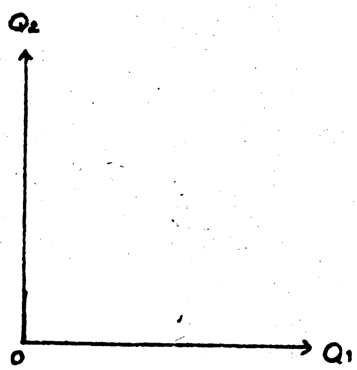
$$e_t^q = \frac{\Delta(Q_1/Q_2)/(Q_1/Q_2)}{\Delta(MRPT_{Q_1Q_2})/(MRPT_{Q_1Q_2})}$$

e) Draw separate diagrams illustrating the case of :

Revenue maximisation subject to given production possibility



Resource minimisation subject to given revenue situation



(You need to draw either production possibility (transformation) map or iso-revenue map as the case may be).

8.10 SUMMARY

Production is the process of providing goods and services that have value to consumers. A profit maximising firm tries to achieve the optimum input-output combination for a given cost. In this unit we have shown that a firm minimises the cost for a given output in order to maximise the profit. Production at the lowest possible cost requires that the additional rupee spent on additional output is equal to the output obtained per additional rupee spent on any other input. We call this the least cost combination of inputs.

We have developed certain concepts like production function, isoquant and isocost curves, and the marginal rate of technical substitution of one input for the other etc. to derive the condition for achieving the least cost combination of inputs. The production function shows the maximum natities of output a firm can produce using various combinations of factor inputs. An **isoquant** curve indicate different combinations of two inputs which can be employed to produce a specific level of output. An **isocost line** expresses different combinations of two inputs which can be utilised for a given rupee cost. The **marginal rate of technical substitution** is the rate at which one input may be substituted for another while the same level of output is maintained and is equal to the negative of the slope of an isoquant curve. The least cost combination of inputs for a particular level of output is reached at the point where an isocost curve is tangent to the isoquant curve corresponding to that level of output. At this point, the marginal rate of substitution, which is also equal to the ratio of marginal product of two factor inputs, is equal to the ratio of inputs prices. In other words,

$MRTS = \frac{MP_L}{MP_K} = \frac{P_L}{P_K}$ assuming there are only two inputs L & K. The line joining all the least cost combination points for different level of output is called the expansion path of the firm.

The **principle of diminsing marginal productivity** affects nearly all types of production functions. Thus, adding additional units of any variable input to fixed inputs resutls in successively smaller increment to output. However, when all inputs can be varied simultaneously, production may be characterised by increasing or constant or decreasing returns to scale.

To maximise profit in the short-run, a firm should employ a variable input (say, input L) upto the point where the additional revenue (that another unit of labour will bring in) is just equal to its cost, or where the marginal revenue product or labour equals its marginal resource cost i.e., $MRP_L = MRC_L$. Subsequently we have moved from single product to joint-product case where we have got comparable concepts like Marginal Rate of Product Transformation ($MRPT_{Q_1Q_2}$) and elasticity of product transformation. However, we have not touched upon the form of **transformation function** as we have done in the case of production function e.g. Cobb-Douglas, Leontief, CES, VES, linear types. In the case of joint-product, the optimum decision rule works out to be one of proportionality between $MRPT_{Q_1Q_2}$ and the product Price ratio $\frac{P_1}{P_2}$, where MRPT is the ratio or resource costs of production.

8.13 ADDITIONAL READINGS

Adhikary, M. 1987, *Managerial Economics,* Khosla Publications: Delhi (Chapter V).

Baumol, William J. 1977, *Economics Theory and Operation Analysis,* (4th edition) Prentice Hall: Englewood-Cliffs. (Chapter II).

Bilas, Richard A, 1971, *Microeconomic Theory,* 2nd edition, McGraw Hill Book Co.: (Chapter 6).

Douglas, Evan J. 1983, *Managerial Economics: Theory Practice and Problems,* Prentice Hall Inc: Englewood-Cliffs.

Mote, V.L., Samuel Paul and G.S. Gupta, 1977, *Managerial Economics: Concepts and Cases,* (Tata McGraw Hill: Bombay)

8.14 SELF-ASSESSMENT TEST

1. Explain clearly the various facets of production decision problems e.g., choice of technique, choice of product etc.

2. Compare the case of single-product firm with that of a joint-product firm with reference to underlying set of concepts, techniques and optimum decision rules.

3. Distinguish between the concepts of:
   a) Production and Productivity
   b) Average and Marginal productivity
   c) Input elasticity and function coefficient
   d) Returns to a factor and Returns to scale
   e) Factor substitution and product transformation
   f) Factor intensity and Factor productivity

4. Take an imaginary production function and show that it is possible for a firm to have "diminishing returns to a factor" along with "constant returns to scale". Bring out clearly the economic implications of such a case.

5. Draw a comprehensive analogy between demand analysis and production analysis.

6. A factory manufactures two types of machines x and y. The transformation function is given by

   $f(x,y) = x^2 + 2y^2 - xy$

   To minimise resource cost, how many machines of each type should be produced if there must be a total of 8 machines?

   (Ans: $x = 5$)
   $y = 3$)

7. Suppose the production function is $16Q = 65 - 2(L-5)^2 - 4(K-4)^2$. The unit price of inputs L and K are 8 and 4 respectively and the unit price of output Q is 32. Determine the maximum profit corresponding to optimum input choice.

   [State profit function, $\pi = R - C = (P_Q Q - (P_L L + P_K K)$
   Also note that the decision variables are L and K]

   (Ans: $L = 4$
   $K = 15/4$
   $\pi = 78\frac{1}{2}$

8. It is suggested that there is a world of difference between **production function and supply function**, though one can be derived from the other. Comment. (Treat it **mathematically**).

   (Hint: See Baumol or Adhikary).

UNIT 9 COST CONCEPTS AND ANALYSIS

Structure

9.0 INTRODUCTION

Cost and revenue are the two major factors that a profit maximising firm needs to monitor continuously. It is the level of cost relative to revenue that determines the firm's overall profitability. In order to maximise profits, a firm tries to increase its revenue and lower its cost. While the market factors determine the level of revenue to a great extent, the cost can be brought down either by producing the optimum level of output using the least cost combination of inputs, or increasing factor productivities, or by improving the organisational efficiency. The firm's output level is determined by its cost.

Product prices are determined by the interaction of the forces of demand and supply. The basic, factor underlying the ability and willingness of firms to supply a product in the market is the cost of production. Thus, cost of production provides the floor to pricing. It is the cost that forms the basis for many managerial decisions like which price to quote, whether to accept a particular order or not, whether to abandon or add a product to the existing product line, whether or not to increase the volume of output, whether to use idle capacity or rent out the faclities, whether to make or buy a product, etc. However, it is essential to underline here that all costs are not relevant for every decision under consideration.

The purpose of this unit is to explore cost and its relevance to decision-making. we begin by developing the important cost concepts, an understanding of which can aid managers in making correct decisions. We shall examine the diffeence between economic and accounting concepts of costs and profits. We shall then consider the concepts of short-run and long-run costs and show that they, in conjunction with the concepts of production studies in the preceding unit, can give us a more complete understanding of the applications of cost theory to decision-making.

9.1 OBJECTIVES

On reading this unit, you should be able to

* relate the concepts of production and costs.
* define different types of costs relevant for production, planning and control.

- **distinguish between economic costs and accounting costs**
- **identify the sources from where cost-advantages follow**
- **indicate application of cost concepts and analysis in managerial decision-making.**

9.2 TYPES OF COSTS

There are many different types of costs that a firm may consider relevant for decision-making under varying situations. The manner in which costs are classified or defined is largely dependent on the purpose for which the cost data are being outlined.

Explicit and Implicit Costs

The opportunity cost (or cost of the foregone alternative) of a resource is a definition cost in most basic form. While this particular definition of cost is the preferred baseline for econo in describing cost, not all costs in decision-making situations are completely obvious; one c skills of a good manager is the ability to uncover hidden costs. For a long time, there has a considerable disagreement among the economists and accountants on how costs shoul treated. The reason for the difference of opinion is that the two groups want to use the co data for dissimilar purposes. Traditionally, the accountants have been primarily connected with collection of historical cost data for use in reporting a firm's financial behaviour and position and in calculating its taxes. The main functions of accountants have been reporting, stewardship and control. They report or record what was happened, present information that will protect the interests of various shareholders in the firm, and provide standards against which performance can be judged. All these have only indirect relationship to decision-making. Business economists, on the other hand, have been primarily concerned with using cost data in decisions making. These purposes call for different types of cost data and classification.

Traditional accounting data is not directly suitable for decision-making. While accountants s rely primarily on historical cost in determining the profit or loss of a firm, economists prefer use the opportunity cost baseline concept for this purpose.

The opportunity cost of a resource can be defined as the value of the resource in its next best use, that is, if it were not being used for the present purpose. The opportunity cost is the benefit of using a resource for the next most attractive alternative. For example, the opportunity cost of a student's doing a full time MBA could be the income that he would have earned if he had employed his labour resources on a job, rather than spending them in studying managerial economics, accounting, and so on. The time cost in money terms can be referred to as implicit cost of doing an MBA.

The out-of-pocket costs on tuition and teaching materials are the explicit costs that a student incurs while attending MBA. Thus, the total cost of doing an MBA to a student is implicit costs (opportunity cost) plus the explicit (out-of-pocket) costs.

Accountants typically use those costs that are recorded in their books as representing an actual transfer of money. These are explicit, or nominal costs and often do not represent full economic costs that should be considered in a given decision. In addition to explicit costs, the business economist uses implicit or imputed cost in evaluating a decision.

Furthermore, in measuring the cost of resource in use, the accountant is only concerned with its acquisition cost. But, for decision-making purposes, we necessarily talk about future costs and revenues, and therefore, past costs have very little relevance. Also, the traditional accounting procedure for valuing assets on the balance sheet is acquisition cost minus depreciation. This is faulty as the true current market value of an asset may differ from its book value.

Direct and Indirect Costs

There are some costs which can be directly attributed to production of a given product. The use of raw material, labour input, and machine time involved in the production of each unit can usually be determined. On the other hand, there are certain costs like stationery and other office and administrative expenses, electricity charges, depreciation of plant and buildings, and other such expenses that cannot easily and accurately be separated and attributed to individual units of production, except on arbitrary basis. When referring to the separable costs of first category accountants call them the direct, or prime costs per unit. The joint costs of the second category are referred to as indirect or overhead costs by the accountants.

Direct and indirect costs are not exactly synonymous to what economists refer to as variable costs and fixed costs. The criterion used by the economist to divide cos into either fixed or variable is whether or not the cost varies with the level of output, where as the accountant divides the cost on the basis of whether or not the cost is separable with respect to the production of individual output units. The accounting statements often divide overhead expenses into 'variable overhead' and 'fixed overhead' categories. If variable overhead expenses per unit are added to the direct cost per unit, we arrive at what economists call as average variable cost.

Private Costs versus Social Costs

A further distinction that is useful to make - especially in the public sector, is between private and social costs. Private costs, are those that accrue directly to the individuals of firms engaged in relevant activity. External costs, on the other hand, are passed on to persons not involved in the activity in any direct way (i.e., they are passed on to society at large). Consider the case of a manufacturer located on the bank of a river who dumps the waste into water rather than disposing it of in some other manner. While the private cost to the firm of dumping is zero, it is definitely positive to the society. It affects adversely the people located downcurrent who are adversely affected and incur higher costs in terms of treating the water for thier use, or having to travel a great deal to fetch potable water. If these external costs were included in the production costs of the producing firm, a true picture of real, or social costs of the output would be obtained. Ignoring external costs may lead to an inefficient and undesirable allocation of resources in society.

Relevant Costs and Irrelevant Costs

The relevant costs for decision-making purposes are those costs which are incurred as a result of the decision under consideration. The relevant costs are also referred to as the incremental costs. Costs that have been incurred already and costs that will be incurred in the future regardless of the present decision are irrelevent costs as far as the current decision problem is concerned.

There are three main categories of relevant or incremental costs. These are the present-period explicit costs, the opportunity costs implicitly involved in the decision, and the future cost implications that flow from the decision. For example, direct labour and material costs, and changes in the variable overhead costs are the natural consequences of a decision to increase the output level. Also, if there is any expenditure on capital equipments incurred as a result of such a decision, it should be included in full, not withstanding that the equipment may have a useful life remaining after the present decision has been carried out. Thus, the incremental costs of a decision to increase output level will include all present-period explicit costs which will be incurred as a consequence of this decision. It will exclude any present-period explicit cost that will be incurred regardless of the present decision.

The opportunity cost of a resource under use, as discussed earlier, becomes a relevant cost while arriving at the economic profit of the firm. This point will be explained in detail in the following section.

Many decisions will have implications for future costs, both explicit and implicit. If a firm expects to incur some costs in future as a consequence of the present analysis, such future costs should be included in the present value terms if known for certain.

Economic Costs and Profits

Our earlier discussion of economic costs suggests that economists and accountants use the term 'profits' differently. Accounting profits are the firm's total revenue less its explicit costs. But economists define profits differently. Economic profits are total revenue less all costs (explicit and implicit, the latter including a normal profit required to retain resources in a given line of production). Therefore, when an economist says that a firm is just covering its costs, it is meant that all explicit and implicit costs are being met, and that, the entrepreneur is receiving a return just large enough to retain his or her talents in the present line of production. If a firm's total receipts exceed all its economic costs, the residual accuring to the entrepreneur is called an economic, or pure profit. In short:

Economic Profit = Total Revenue – Opportunity Cost of all Inputs

This is depicted in the following figure:

	Economic Profits	Accounting Profit
Total Revenue	Economic or Opprotunity Cost (Explicit plus implicit costs, including a normal profit)	Accounting Costs

An economic profit is not a cost, because by definition it is a return in excess of the normal profit required to retain the entrepreneur in a particular line of production.

A Case

Let us consider the case of Mr. Ashis Nathani who is a small store owner. He has invested Rs. 2 lakh as equity in the store and inventory. His annual turnover is Rs. 8 lakh, from which he must deduct the cost of goods sold, salaries of hired staff, and depreciation of equipment and building to arrive at the annual profit of the store. When his accountant Mr. Chhagani, after using the above approach reported the profit to be Rs. 1.5 lakh, Mr. Nathani could not believe it and he deided to go through the books himself. He discovered that his actual profit was only Rs. 75,000 and not Rs. 1.5 lakh. He found out that his accoutant had understimated the costs, by not caring to provide for the resources that he had put into his business, namely his own time and the money. In finding out his economic profit, he added the inputed salary(to himself) and inputed interest cost on his equity. The two estimates are shown below:

Accounting Income Statement
(Mr. Chhagan)

	Rs.	Rs.
Sales		8,00,000
Cost of Goods Sold	6,00,000	
Salaries	40,000	
Depreciation	10,000	6,50,000
Accounting Profit		1,50,000

Economic Statement of profit
(Mr. Nathani)

	Rs.	Rs.
Sales		8,00,000
Cost of Goods Sold	6,00,000	
Salaries	40,000	
Depreciation	10,000	
Inputed Salary to owner Manager	50,000	
Inputed Interest cost on Equity	25,000	7,25,000
Economic Profit		75,000

Separable and Common Costs

Costs can also be classified on the basis of their traceability. The costs that can be easily attributed to a product, a division, or a process are called separable costs, and the rest are called non-separable or common costs. The separable and common costs are also referred to as direct and indirect costs. The distinction between direct and indirect costs is of particular significance in a multi-product firm for setting up economic prices for different products.

Fixed and Variable Costs

Fixed costs are those costs which in total do not vary with changes in output. Fixed costs are associated with the very existence of a firm's plant and therefore must be paid even if the firm's rate of output is zero. Such costs as interest on borrowed capital, rental payments, a portion of depreciation charges on equipment and buildings, and the salaries of top management and key personnel are generally fixed costs.

On the other hand, variable costs are those costs which increase with the level of output. They include payment for raw materials, charges on fuel and electricity, wages and salaries of temporary staff, depreciation charges associated with wear and tear of assets, and sales commission, etc.

This distnction is true only for the short-run. It is similar to the distinction that we made in the previous unit between fixed and variable factors of production under the short-run production analysis. The costs associated with fixed factors are called the fixed costs and the ones associated with variable factors, the variable costs. Thus, if capital is the fixed factor, capital rental is taken as the fixed cost and if labour is the variable factor, wage bill is treated as the variable cost.

However, it is not very easy to classify all costs into fixed and variable. For example, part of the depreciation charges are fixed, and part variable. However. it is very difficult to determine how much of depreciation cost is due to the technical obsolescence of assets and hence fixed cost, and how much is due to the use of equipment and hence variable cost. Nevertheless, it doesn't mean that it is not useful to classify costs into fixed and variable. This distinction is of great value in break-even analysis and pricing decisions. For decision-making purposes, in general, it is the variable cost which is relevant and not the fixed cost.

Activity 1

a) Give specific examples of :

Implicit costs: ....

Social costs: ....

Indirect costs: ....

Sunk costs: ....

Traceable costs: ....

Common costs: ....

b) Comment on the nature of costs involved in **depreciation** from both economic and accounting standpoints.

....

....

....

....

c) Distinguish between historical costs and replacement costs. Why is this distinction useful?

....

....

....

....

d) Give examples to distinguish between 'fixed overheads' and 'variable overheads'.

....

....

....

....

e) Can all 'direct costs' be treated as 'variable costs'?

....

....

....

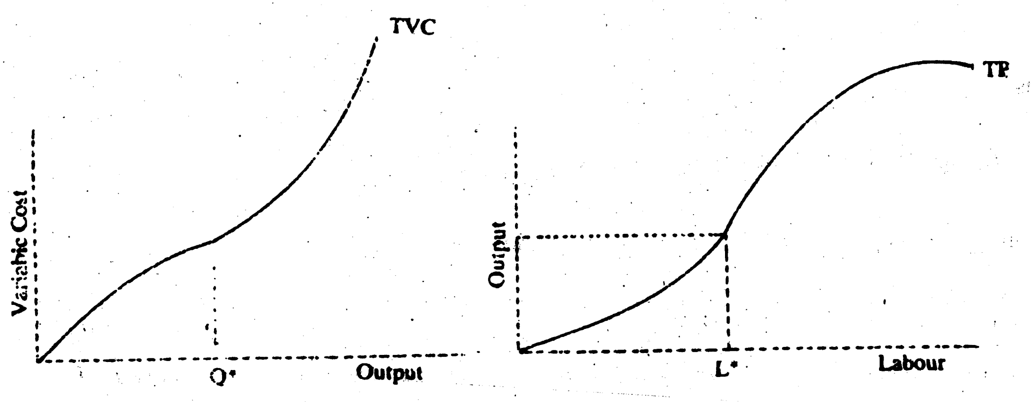
....

9.3 RELATIONSHIP BETWEEN PRODUCTION AND COSTS

The concept of cost is closely related to production theory. A cost function is the relationship between a firm's costs and the firm's output. While the production function specifies the technological maximum quantity of output that can be produced from various combinations of inputs, the cost function combines this information with input price data and gives information on various outputs and their prices. The cost function can thus be thought of as a combination of the two pieces of information i.e. production function and input prices.

Let us consider a short-run production function with only one variable input. As we have already seen in the previous unit, the diminishing returns to the variable input set in after a point. The output grows at an increasing rate in the initial stages implying increasing returns to the variable input, and then diminishing returns to the variable input start. Assuming that the input prices remain constant, the above production function will yield the variable cost function which has a shape that is characteristic of many variable cost function; increasing at a decreasing rate and then increasing at an increasing rate. This relation between the total product curve and the total variable cost is shown in Figure 1.

Figure I



This firm has a production function that exhibits increasing returns to the variable factor input (labour) upto L* and decreasing returns after L*. The input level L* corresponds to the output level Q*. The variable cost function increases at a decreasing rate upto Q* and at an increasing rate beyond Q*

From this, you should be able to derive a relationship between average product and average costs, and marginal product and marginal costs. For example,

TVC = Pr. V

$\therefore AVC = \frac{TVC}{Q} = Pr.\ \frac{V}{Q} = \frac{Pr}{Q/V}$

and $MC = \frac{TVC}{Q} = Pr.\ \frac{V}{Q} = \frac{Pr}{\Delta Q/\Delta V}$

where Pr stands for the price of the variable factor and r stands for amount of variable factor

You may note that Pr being given, AVC is inversely related to the average product of the variable factors.

In the same way, given the wage rate MC is inversely related to the marginal product of labour. We shall explore this relationship in greater detail subsequently.

Activity 2

Given $Q = 100 + 0.L^2 - 0.0005\ L^3$

Where Q is output; L is labour.

Suppose the wage rate is Rs. 10 and that 100 labourers are being employed.

Find the AVC and MC.

....

....

....

....

....

....

....

(Ans : AVC = Rs. 1.66. MC = Rs. 2)

9.4 SHORT-RUN COST FUNCTIONS

Recall the concept of "time perspective" from Block 1.

The short-run is normally defined as a time period over which some factors of production are fixed and others are variable. Needless to emphasize here that these periods are not defined by some specified length of time but, rather, are determined by the variability of factors of production. Thus, what one firm may consider in the long-run may correspond to the short-run for another firm.

In the short-run, a firm incurs some costs that are associated with variable factors and others that result from fixed factors. The former are called variable costs and the latter represent fixed costs. Variable costs (VC) change as the level of output changes and therefore can be expressed as a function of output (Q), that is VC = f(Q). Variable costs typically include such things as raw material, labour, and utilities. In Column 3 of Table 1, we find that the total of variable costs changes directly with output. But note that the increases in variable costs associated with each one-unit increase in output are not constant. As production begins, variable costs will, for a time, increase by a decreasing amount; this is true through the fourth unit of the output. Beyond the fourth unit, however, variable costs rise by increasing amount for each successive unit of output. The explanation of this behaviour of variable costs lies in the law of diminishing returns.

Table 1

Total and Average-Cost Schedules for an Individual Firm in the Short-Run

(Hypothetical Data in Rupees)

Total cost data, per week				Average-cost data, per week			
(1) Total Product	(2) Total fixed cost (TFC)	(3) Total variable cost (TVC)	(4) Total cost (TC) TC = TFC+ TVC	(5) Average fixed cost (AFC) AFC = TFC/Q	(6) Average variable cost (AVC) AVC = TVC/Q	(7) Average total cost (ATC) ATC = TC/Q	(8) Marginal cost (MC) MC = change in TC / change in Q
0	100	0	100				
1	100	90	190	100.00	90.00	190.00	90
2	100	170	270	50.00	85.00	135.00	80
3	100	240	340	33.33	80.00	113.33	70
4	100	300	400	25.00	75.00	100.00	60
5	100	370	470	20.00	74.00	94.00	70
6	100	450	550	16.67	75.00	91.67	80
7	100	540	640	14.29	77.14	91.43	90
8	100	650	750	12.50	81.25	93.75	110
9	100	780	880	11.11	86.67	97.78	130
10	100	930	1030	10.00	93.00	103.00	150

Fixed costs on the other hand, are not a function of the level of output and are constant in the short-run, that is FC = K. Fixed costs may include such things as property taxes, the cost of leases on land, buildings and some types of equipment, interest charges on the long term borrowed funds, and insurance costs. In column 2 of Table 1, we have assumed that the firm's total fixed costs are Rs. 100. Note that, by definition, this fixed cost figure prevails at all levels of output including zero.

The distinction between fixed and variable costs is of great significance to the business manager. Variable costs are those costs which business can control or alter in the short-run by changing levels of production. On the other hand, fixed costs are clearly beyond business executive's control, such costs are incurred in the short-run and must be paid regardless of output level.

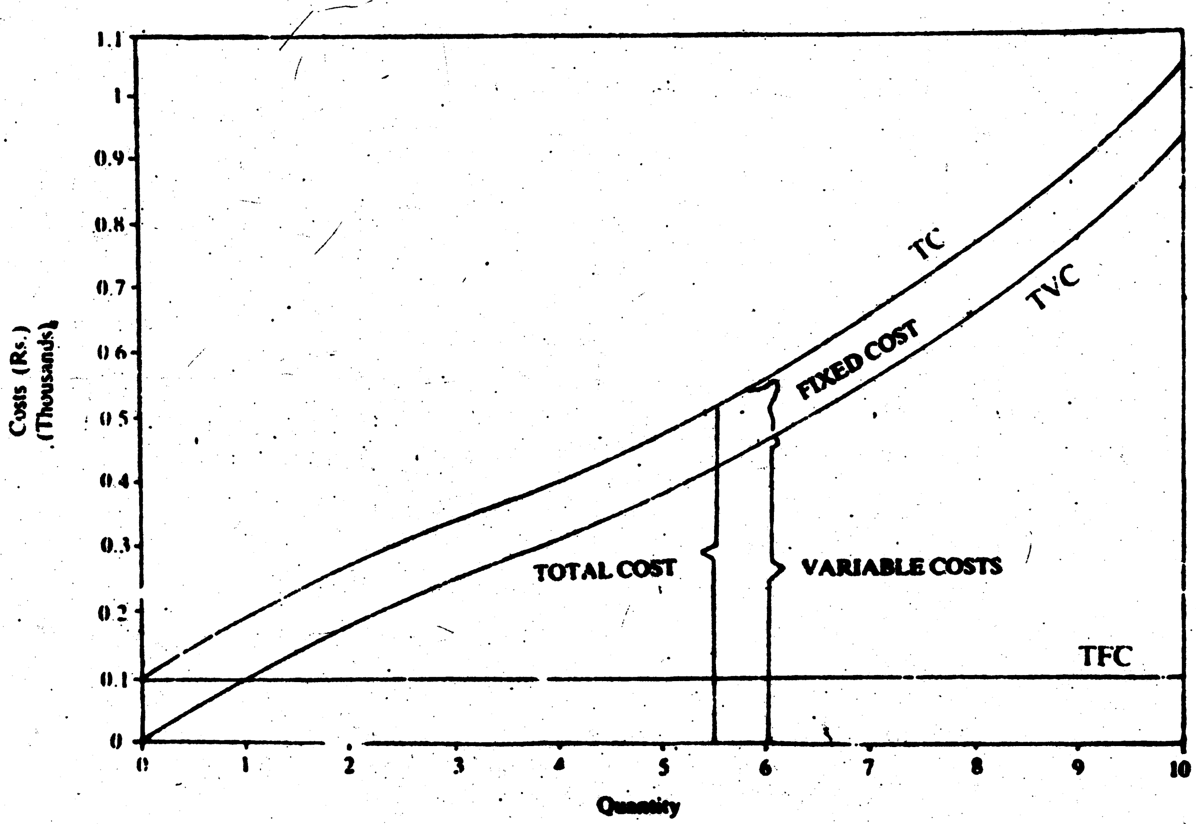
Total Cost

Total cost is the sum of fixed and variable cost at each level of output. It is shown in column 4 of Table 1. At zero unit of output, total cost is equal to the firm's fixed cost. Then for each unit of production (through 1 to 10), total cost varies at the same rate as does variable cost. Figure II shows graphically the fixed, and total cost data of Table 1.

Per Unit, or Average Costs

Besides their total costs, producers are equally concerned with their per unit, or average costs. In particular, average cost data is more relevant for making comparisons with product price, which is always stated on per unit basis. Average fixed cost, average variable cost, and average cost are shown in columns 5 to 7 of Table 1. It is important that we know how these unit-cost figures are derived and how they vary as output changes.

Figure II Fixed, total variable & total costs



TFC+TVC = TC

Average Fixed Costs

Average fixed cost (AFC) is derived by dividing total fixed cost (TFC) by the corresponding output (Q). That is

$$AFC = \frac{TFC}{Q}$$

While total fixed cost is, by definition, independent of output, AFC will decline so long as output increases. As output increases, a given total fixed cost of Rs. 100 is obviously being spread over a larger and larger output. This is what business executives commonly refer to as 'spreading the overheads'. We find in Figure III that the AFC curve is continuously declining as the output is increasing. The shape of this curve is of an asymptotic hyperbola.

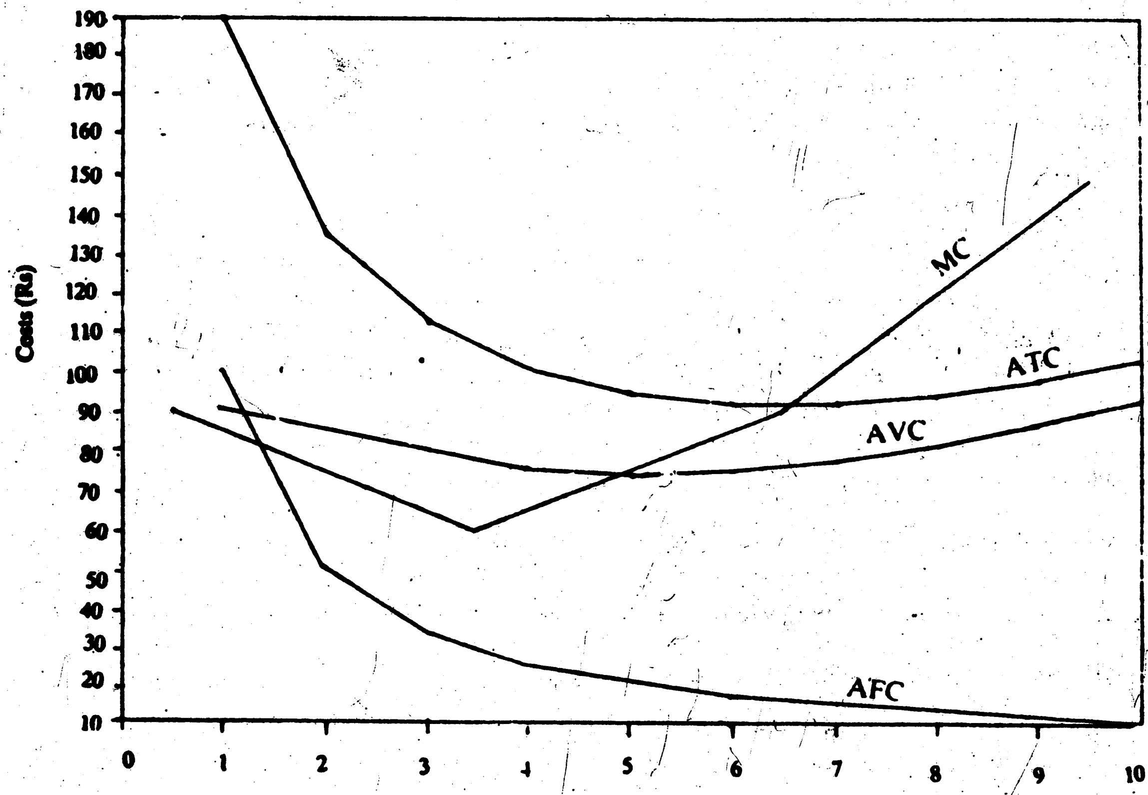
Average Variable Costs

Average variable cost (AVC) is found by dividing total variable cost (TVC) by the corresponding output (Q) :

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

AVC declines initially, reaches a minimum, and then increases again. Graphically, this provides us with a U-shaped or saucer-shaped AVC curve, as shown in Figure III.

Figure III Average fixed, variable & Marginal Costs



AFC+AVC = ATC

$$\frac{\Delta ATC}{\Delta Q} = MC$$

Because total variable cost reflects the law of diminishing returns, so must the AVC figures, which are derived from total variable cost. Because of increasing returns initially, it takes fewer and fewer additional variable resources to produce each of the first four units of output. As a result, variable cost per unit will decline. Thus, AVC hits a minimum with the fifth unit of output, and beyond this point AVC rises as diminishing returns necessitate the use of more and more variable resources to produue each additional unit of output.

Average Total Costs

Average total cost (ATC) can be found by dividing total cost (TC) by total output (Q) or, by adding AFC and AVC for each level of output. That is:

$$ATC = \frac{TC}{Q} = AFC + AVC$$

These data are shown in column 7 of Table 1. Graphically, ATC is found by adding vertically the AFC and AVC curves, as in Figure III. Thus the verticle distance between the ATC and AVC curves reflects AFC at any level of output.

Marginal Cost

Marginal cost (MC) is defined as the extra, or additional, cost of producing one more unit of output. MC can be determined for each additional unit of output simply by noting the change in total cost which that unit's production entails :

$$MC = \frac{Change\ in\ TC}{Change\ in\ Q} = \frac{\Delta TC}{\Delta Q}$$

Our data in Table 1 is so structured that the "change in Q" is always "1", so we have defined MC as the cost of one more unit of output. However, the same concept can be extended for situations where Q is more than 1. Arc marginal cost (over a range of output) can be found by using the above expression.

The marginal cost concept is very crucial from the manager's point of view. Marginal cost is a strategic concept because it designates those costs which the firm has the most direct control. More specifically, MC indicates those costs which are incurred in the production of the last unit of output and therefore, also the cost which can be "saved" by reducing total output by the last unit. Average cost figures do not provide this information. A firm's decisions as to what output level to produce is largely influenced by its marginal cost. When coupled with marginal revenue, which indicates the charge in revenue from one more or one less unit of output, marginal cost allows a firm to determine whether it is profitable to expand or contract its level of production.

Marginal cost is shown graphically in Figure III. Note that marginal cost declines sharply, reaches a minimum and then rises rather sharply. This mirrors the fact that variable cost, and therefore total cost, increases first by decreasing amounts and then by increasing amounts (Figure I).

MC and Marginal Product

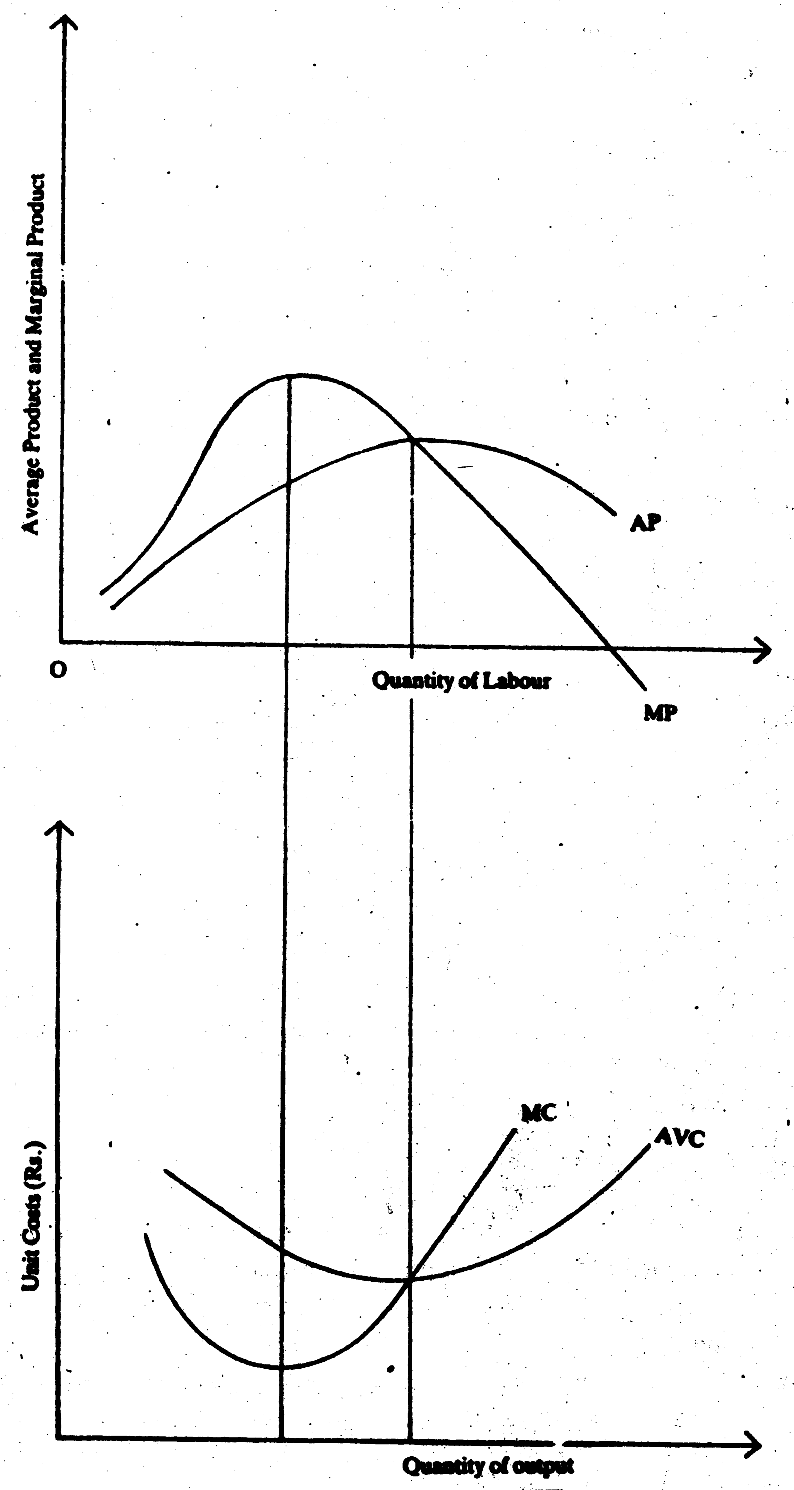
The shape of the marginal cost curve is a reflection of, and the consequence of, the law of diminishing returns. If each successive unit of a variable input say labour, is hired at a constant price, the marginal cost of each extra unit of output will fall so long as the marginal product of each additional worker is rising. This is so because marginal cost is simply the constant wage (cost) of an extra worker divided by his or her marginal product. Thus, given the price (cost) of the variable resource, increasing returns i.e., a rising marginal product, will be reflected in a declining marginal cost and diminishing returns i.e., falling marginal product, in a rising marginal cost. The MC curve is a mirror reflection of the marginal product curve. This relationship is shown in Figure IV. It is clearly shown that when marginal product is rising, marginal cost is necessarily falling. When marginal product is at its maximum, marginal cost is at its minimum. And when marginal product is falling, marginal cost is rising.

Relationship of MC to AVC and ATC

It is also notable that marginal cost cuts both AVC and ATC at their minimum (Figure III). When both the marginal and average variable costs are falling, average will fall at a slower rate. And when MC and AVC are both rising, Mc will rise at a faster rate. As a result, MC will attain its minimum before the AVC. In other words, when MC is less than AVC, the AVC will fall, and when MC exceeds AVC, AVC will rise. This means (Figure III) that so long as MC lies below AVC, the latter will fall and where MC is above AVC, AVC will rise. Therefore, at the point of intersection where MC = AVC, AVC has just ceased to fall and attained its minimum, but has not yet begun to rise. Similarly, the marginal cost curve cuts the average total cost curve at the latter's minimum point. This is because MC can be defined as the addition either to total cost or to total variable cost resulting from one more unit of output. However, no such relationship exists between MC and the average fixed cost, because the two are not related; marginal cost by definition includes only those costs which change with output, and fixed costs by definition are independent of output.

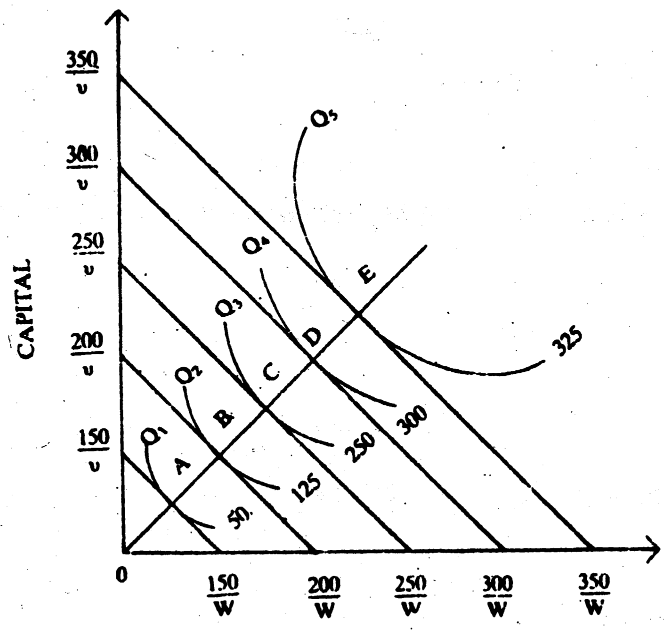
Figure IV

THE RELATIONSHIP BETWEEN PRODUCTIVITY CURVES AND COST CURVES



Taking the values for total cost and output from the expansion path of Figure V (the most efficient points), we can construct the following table for cost and output :

FIGURE V EXPANSION PATH AND RETURNS TO SCALE

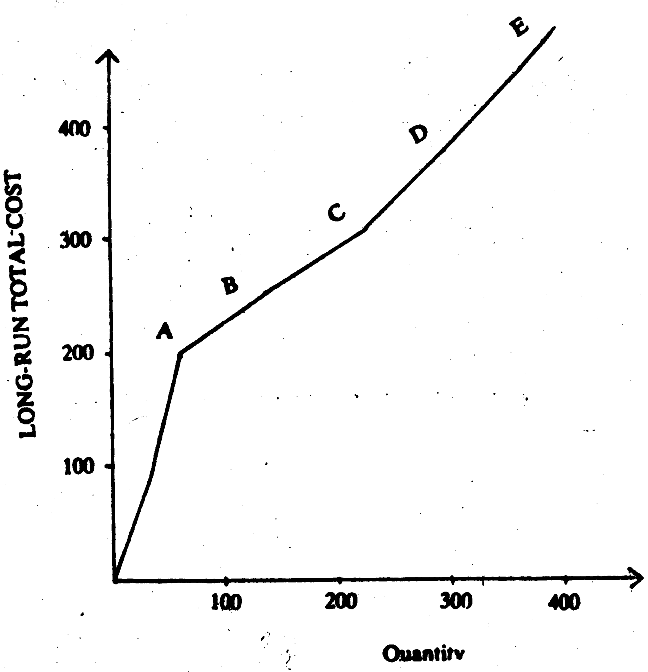


LABOUR

Output (Q)	Long-run Total Cost (LTC)
50	150
125	200
250	250
300	300
325	350

These points are graphed in Figure VI as the long-run total cost (LTC) curve. The points A, B, C, D and E correspond to the equilibrium points in Figure V. Note that the LTC curve at first increases at a decreasing rate, then at a constant rate, and finally at an increasing rate. The LTC curve starts from the origin implying thereby that in the long-run all costs are variable and if nothing is produced, no resources will be used (i.e., the firm will quit the industry altogether). Thus, the LTC curve is analogous to the short-run VC curve. Only difference is, while the shape of VC is due to the law of variable proportions in the short-run, the shape of LTC is due to the existence of increasing, constant, and decreasing returns to scale in the long-run.

Figure VI



Unit Costs in the Long-Run

In the long-run, costs are not divided into fixed and variable components; all costs are variable. Thus, the only long-run unit cost functions of interest are long-run average cost (LAC) and long-run marginal cost(LMC).These are defined as follows :

$$LAC = \frac{LTC}{Q}$$

$$LMC = \frac{\Delta LTC}{\Delta Q}$$

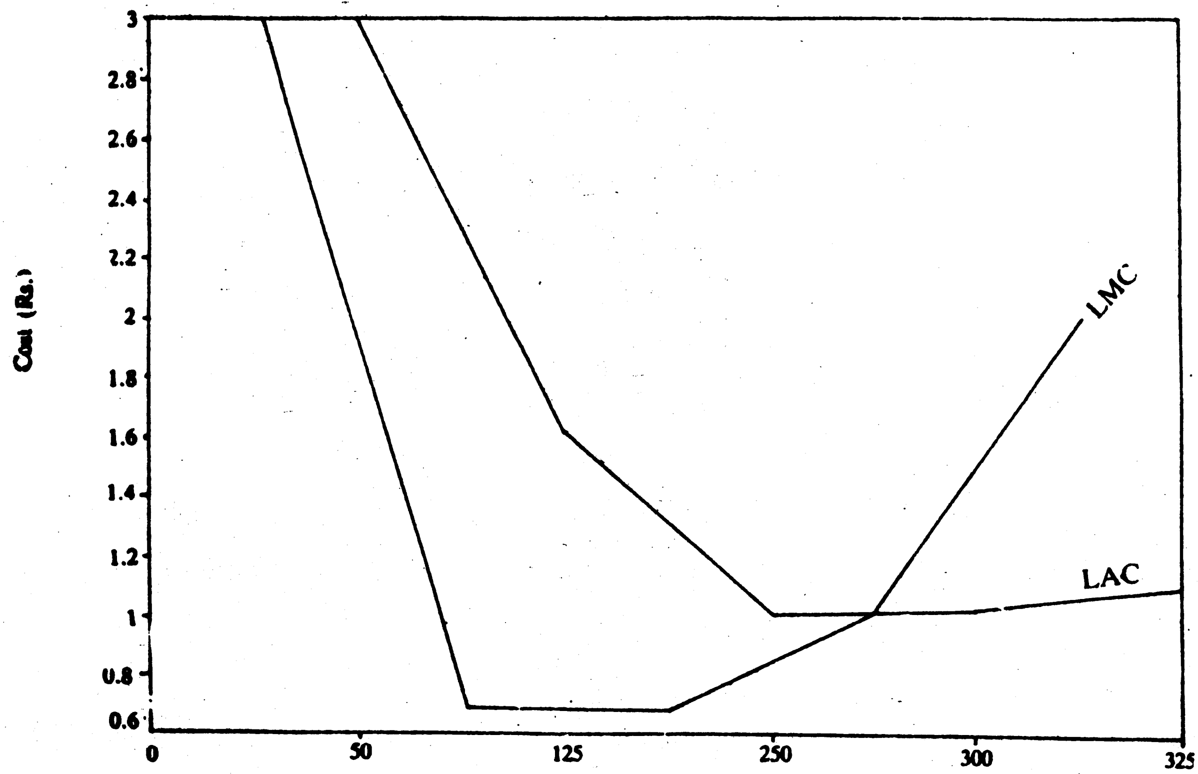
$$LMC = \frac{d(LTC)}{dQ}$$

For the long-run total cost given in Figure VI, these unit costs can be presented in tabular form as follows :

Output Q	Long Run Total Cost (LTC)	Long Run Average Cost (LAC)	Long Run Marginal Cost (LMC)
0	0	-	-
50	150	3.00	3.00
125	200	1.60	0.67
250	250	1.00	0.67
300	300	1.00	1.00
325	350	1.08	2.00

These LAC and LMC and values are graphed in Figure VII. We see, both in the table and in the graph, that LAC and LMC are U-shaped and that they are equal at the minimum of LAC. The values of LMC are graphed at the midpoints of the output intervals they represent.

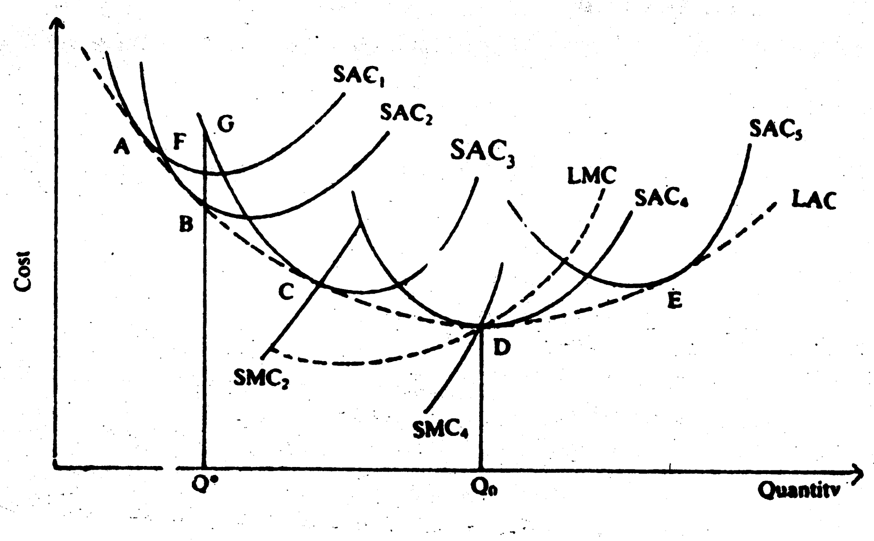
Figure VII : LONG-RUN AVERAGE & MARGINAL COSTS



Long-Run Average Cost as an Envelope Curve

The long-run average cost curve is sometimes shown as the envelope curve of series of all tl. possible short-run average cost curves, as shown in Figure VIII. Five short-run average cost curves, each representing a different sized plant (or set of fixed factors) are illustrated, although many more may exist. The long-run ATC curve shows the least per unit cost at any output can be produced after the firm has had time to make all appropriate adjustments in its plant size. Consider, for example, the production of Q* units in Figure VII. That the level of output could be produced with the plant sizes represented by SAC_1 , SAC_2 or SAC_3. It represents the optimum rate of output for the plant size represented by SAC_1 (i.e., it is at the minimum point of SAC). However, if the firm expects to produce at the rate, the best size of plant is the one related to SAC_2. Q* units could be produced at a cost savings of FB per unit over SAC_2.

FIGURE VIII LONG-RUN AVERAGE COST AS AN ENVELOPE CURVE



The plant size associated with SAC_2 is the optimum size plant because its minimum point is the lowest of all possible unit costs. Given the LAC in Figure VIII, we can say that there are increasing returns to scale (or economies of scale) up to Q_0 and decreasing returns to scale (or diseconomies of scale) beyond Q_0. The firm's LAC curve is often called the 'firm's planning curve'.

Activity 4

a) Why are all costs variable in the long-run?

....

....

b) Why is the LAC called an "envelope curve"?
Why cannot the LMC be an envelope as well?

....

....

....

c) What do you understand by "cost-efficiency"?
Draw a long-run cost diagram and explain.

9.6 ECONOMIES OF SCALE

It has been shown in the preceding section that a larger plant will lead to lower per unit costs in the long-run. However, beyond some point, successively larger plants will mean higher average costs. Exactly, why is the long-run ATC curve U-shaped? This point needs further explanation

It must be emphasised, first of all, that the law of diminishing return is not applicable here; for it presumes that one resource is fixed in supply and, as we have seen, the long-run assumes that all resources are variable. Also, we assume that resource prices are constant. What then, is our explanation? The U-shaped long-run average cost curve is explainable in terms of what economists call "economies and diseconomies" of large scale production.

Economies and diseconomies of scale are concerned with the behaviour of average cost curve as the plant size is increased. Economies of scale explain the downslopping part of the long-run AC curve. As the size of a plant increases, LAC typically declines over some range of output for a number of reasons. The most important is that, as the scale of output is expanded, there is greater potential for specialisation of productive factors. This is most notable with regard to labour but may apply to other factors as well. Other factors contributing to declining LAC include ability to use more advanced technologies and more sophisticated capital equipment; managerial specialisation; opportunity to take advantage of lower costs for some inputs by purchasing larger quantities; effective utilisation of by-products, etc.

But, after sometime, expansion of a firm's output may give rise to diseconomies, and therefore, higher per unit costs. Further expansion of output beyond a reasonable level may lead to problems of over crowding of labour, managerial inefficiencies, etc., pushing up per unit costs.

All these are examples of internal economies and diseconomies of scale arising due to the firm's own expansion. According to Marshall, external economies and diseconomies of scale may arise due to the expansion of industry as a whole. For example, improved infrastructure facilities due to industrial expansion may lead to reduction in per unit costs of production in all the firms in an industry.

9.7 ECONOMIES OF SCOPE

... concept is a recent development. This stands in contrast to the concept of 'economies of scale'. The basic argument is that cost-efficiency in production process has brought about by, variety rather than volume.

According to the concept of economies of scale, cost advantages follow from the increase in volume of production or what is called the scale of output. According to the concept of economies of scope, such cost advantages may follow from variety of output e.g. product diversification within the given scale of plant. If the same plant can produce multiple products, there is the scope for a lot of cost savings because of joint utilisation of inputs. The Government of India has announced the 'broadbanding policy' such that Escorts may now produce in addition to its two-wheeler, four-wheeler using the same plant and technology with slight modifications or adjustments. Broadbanding, therefore, should enable manufacturers to exploit economies of scope through product diversification. In other words, instead of increasing the scale of production of an existing product, the firm can now add new and newer products, if the size of plant and type of technology offer that scope; in this process, the firms will have access to scope-economies in place of scale-economies. In certain processes, the firm can plan wisely to exploit both types of economies simultaneously.

Activity 5

a) Distinguish between internal and external economies of scale. Give examples.

....

....

....

....

....

b) Distinguish between scale-economies and scope-economies. Give examples.

....

....

....

....

....

9.8 COST ELASTICITY AND OUTPUT ELASTICITY

When we take costs determined by output, we can talk about output elasticity of costs, e^c_q. On the other hand, when we find output being determined by costs, we can talk of costs-elasticity of output, e^q_c.

Examining these elasticities, you should be in a position to discover the relationship between average and marginal costs.

$$e^c_q = \frac{dC}{dQ} \cdot \frac{Q}{C} = \frac{MC}{AC}$$

$$e^q_c = \frac{dQ}{dC} \cdot \frac{C}{Q} = \frac{AC}{MC}$$

In other words, given the average costs (AC) and marginal costs (MC); we can derive these elasticities; or given cost elasticity and marginal costs (or average costs) we can derive average costs (or marginal costs).

The economic interpretation of cost elasticity is the usual one i.e., the proportionate change in the level of output resulting from the proportionate change in costs of production.

Activity 6

a) How is production related to costs?

....

....

....

b) Given $C = 45Q^2$, calculate the following for $Q = 2$

AC = MC = e^c_q = e^q_c =

c) In the short-run, the output elasticity of average fixed costs is equal to unity. True or False. Explain.

....

....

....

d) The examination of cost-data of a particular firm shows that in a year, following upward revision of wages the cost per unit of output produced by that very labour has come down. How can that be? Explain with the help of a diagram.

....

....

....

....

....

....

....

....

....

....

Optimum output level

You may recall that optimum has reference to that level (size) of output which minimised the average cost of product alternatively, for which average cost equals marginal costs. Consider the following exam,

$$TC = 128 + 6Q + 2Q^2$$

$$\Rightarrow AC = \frac{128}{Q} + 6 + 2Q$$

$$\frac{d(AC)}{dQ} = \left[-\frac{128}{Q^2} + 2\right] = 0 \Rightarrow Q = 8$$

or, MC = $6 + 4Q$.

Setting AC = MC

$$\left[\frac{128}{Q} + 6 + 2Q\right] = [6 = 4Q] \Rightarrow Q = 8$$

Thus [Q=8] is the optimum level of output (for the short-run). Note that it is a short-run case, because TFC = 128.

Optimum inventory level

All productions are not immediately sold. Sometimes we have to build up inventory or stock of saleable items. In this case, the optimum inventory level may be defined as that for which the average cost of inventory held is at the minimum. In this, there are two types of costs involved ; Carrying costs ana and Reorder costs. The inventory **carrying costs** include storage costs, interest costs on borrowed capital to finance stock etc. On the other hand the **reorder costs** include book keeping costs and telephone charges (i.e., fixed costs); some of the reorder costs also vary with the size of the order placed (i.e., variable costs). Suppose we use the following notations :

S = Expected sale

D = Order quantity to be delivered

S/D = Number of orders delivered

F = Average fixed costs of delivery

V = Coefficient of average variable cost of reorder

K = Average carrying costs

D/2 = Average inventory held beween initial and terminal periods and it is assumed that the demand is spread evenly.

The average costs of inventory held can now be stated explicitly as :

$$AC = \left[K.\frac{D}{2}\right] + [F.V.D.]\frac{S}{D}$$

Carrying costs + Reorder cost.

For deciding on optimum inventory held,

$$\frac{d(AC)}{dD} = \left[\frac{K}{2} - \frac{FS}{D^2}\right] = 0 \Rightarrow D = \sqrt{\frac{2FS}{K}}$$

This is very a very significant derivation. This formula gives us the optimum size of stock or what is called Economic Order Quantity (EOQ).

In other courses, you will be introduced to more sophisticated inventory models. The simple inventory model that we have used here, nevertheless, brings home, some interesting results. For example, inventory should increase in proportion to the square root of expected sales. Also note that variable costs have got no impact on inventory decisions.

Break-even output

This is basically an accounting concept. The break-even chart of the accountant illustrates at what level of output in the short-run, the total revenue just covers the total costs. Let us operate with our familiar algebraic notations :

$$TR = P \cdot Q$$

$$TC = F + VQ$$

At break even point,

$$TR = TC$$

$$\Rightarrow P \cdot Q = F + VQ$$

$$\text{or } Q = \frac{F}{P-V} = \frac{\textit{Total Fixed costs}}{\textit{Price} - \textit{Variable costs per unit}}$$

Here Q stands for **Break-even volume of output**; multiplying this with price (P) we get the **Break-even value of output**. Take an example to illustrate. Suppose

$$TR = 60\,Q$$

$$TC = 1800 + 40\,Q$$

Then the break-even output volume $= 90\left[= \frac{1800}{60-40}\right]$

and the break-even output value = 5400[= 90 × 60]

also note that [P - V] = 20 is called 'contribution margin' per unit of output.

Activity 8

a) Speed-Marine Co. builds motor boat engines. They recently estimated their total costs and total revenue as :

$$TC = 80{,}000 - 600\,Q + 2\,Q^2$$

$$TR = 400\,Q - Q^2$$

(Where Q is the number of engines produced each year)

At what level of production will Speed-Marine co. break-even? How many engines should be produced to maximise profit?

....

....

....

....

....

....

b) Given LTC = $6Q + 2Q^2 - Q^3$, find out the optimum level of output, Q.

c) Write a lucid note on the types of inventory costs, giving suitable examples.

d) Why does not the variable (reorder) cost affect the **economic order quantity?**

e) List the factors which determine the size of optimum buffer stock of foodgrains held by the Food Corporation of India.

....

....

....

9.11 SUMMARY

In this unit, we have emphasised the critical role that costs play in determining the profitability of the firm. The profit-oriented firm manager must consider both opportunity costs and explicit costs in order to use all the resources most economically. Although it is difficult to have totally accurate information on its costs, a firm should have reliable estimates of its fixed costs, how its costs vary with respect to output over the relevant range of production, and of whether or not its costs would be lower with a bigger size plant.

In the short-run, the total cost consists of fixed and variable costs. Variable cost, and thus the total cost, typically increases first at a decreasing rate and then at an increasing rate. The reason for this type of behaviour of variable and total cost is related to the principal of diminishing marginal productivity. If marginal product increases at first, the marginal cost at first declines (that is, the rate of increase in variable and total cost declines). However, once marginal product begins to decline, marginal cost increases. It follows that in the short-run, marginal cost (MC), average variable cost (AVC), or average total cost (AC) curves are generally U-shaped.

The long-run is defined as the period for which all factors of production are variable. Thus, in the long-run total costs are identical to variable costs. The long-run total cost curve may also at first increase at a decreasing rate, and then increase at an increasing rate. If the LTC is of normal shape (inverted S-shaped), the long-run average cost (LTC) will also have a U-shape. The long-run cost function may be derived from the series of most efficient short-run cost curves or from an expansion path. Economies or diseconomies of scale arise either due to the internal factors pertaining to the expansion of output by a firm, or due to the external factors such as industry expansion.

In contrast, economics of scope result from product diversification. Thus the scale-economies have reference to an increase in volume of production, whereas the scope-economies have reference to an improvement in the variety of products from the existing plant and equipment. These cost concepts and analyses have a lot of applications in real world decision-making situation such as optimum output, optimum product-mix, optimum inventory, break-even output etc.

9.12 ADDITIONAL READINGS

Seo, K.K., 1988. *Managerial Economics* (6th ed.) Surjeet Publications: Delhi.

Adhikary, M. 1987. *Managerial Economics* (3rd ed.) Khosla Educational Publishers : Delhi, Ch. V.

Goldhar, J.D. & M. Jelinek, 1983. "Plan for Economics of Scope", in *Harvard Business Review*, December.

9.13 SELF-ASSESSMENT TEST

1. Make a list of relevant Cost-concepts from the standpoint of an (i) accountant, (ii) economist, and (iii) engineer. Take any construction company as a reference to illustrate the cost-concepts.

2. Your Management is interested in cost **reductions and control.** How would you go about it. Recommend some steps which would yield results in

   a) short-run

   b) long-run

3. Explain clearly the managerial use of the concepts like 'economies of scale' and 'economies of scope'. Is it possible to reap both the economies simultaneously? Give examples.

4. Draw and explain the effect of internal and external economies/internal and external diseconomies on the firm's cost curves.

5. The total cost function of a neckwear manufacturing company is estimated as TC = 8 + 12 Q - 6 Q^2 + Q^3 (where Q represents hundreds of ties produced per month)

   Derive equation for AFC, AVC, ATC, MC and plot all these costs in a graph and comment on the relationship between average and marginal costs.

6. A fixed plant is used to manufacture TV sets. If x sets are turned out per week, the total variable costs is Rs. $\left[3x + \frac{1}{25}x^2\right]$. Show that the average variable costs rise steadily with output. What do you expect about marginal costs of production?

7. Given production function.

   $$Q = AL^{\alpha}K^{1-\alpha}$$

   derive the cost function. State your assumptions clearly at the end.

   [Hint : It is a difficult question. You may recall the cost equation, C = $P_L L + P_K K$ and the optimum decision rule.

   $$\frac{P_L}{P_K} = \frac{MP_L}{MP_K}$$

Apply these to get C = C (Q, P_L, P_K, α 1 - a). Have patience; you should be able to do it.]

UNIT 10 EMPIRICAL ESTIMATES OF PRODUCTION AND COSTS

Structure

10.0 INTRODUCTION

For decision-making purposes, a manager should understand clearly the relationship between his inputs and outputs on the one hand, and the output and costs on the other. Estimation of the industry production function may help a manager in taking correct decisions of long-term nature such as capital expenditure. The short-run production estimates at firm level are helpful in arriving at the optimal mix of inputs to achieve a particular output target. This is referred to as the least cost combination of inputs in production analysis. Also, for a given cost, optimum level of output can be found if the production function of a firm is known.

Cost estimation for decision-making is concerned with finding the shape and placement of the firm's cost curves. Both, the short-run cost function and the long-run cost function must be estimated, since both sets of information will be required for some decisions. Knowledge of the short-run cost functions allows the decision makers to judge the optimality of present output levels and to solve decision problems using contribution analysis. We saw in the preceding unit that the concept of incremental cost is fundamental to short-run decision-making on cost issues. Knowledge of long-run cost functions is important when considering the expansion or contraction of plant size, and for confirming that the present plant size is optimal for the output level that is being produced.

In this unit, we shall discuss different approaches to production and cost estimation.

10.1 OBJECTIVES

On reading this unit, you should be able to :

- **relate** theory to empiricism
- **get** a summary-view of various approaches to estimation : economic (econometric), accounting and engineering
- **find** some statistical estimates of production and cost functions
- **understand** both conceptual and statistical difficulties related to measurement of such functions
- **throw** light on empirical determination of production and cost functions; and finally

- Identify managerial use of such empirical estimates.

10.2 SOME ESTIMATES OF PRODUCTION FUNCTION

The production function expresses an engineering relationship between physical inputs and outputs. By definition, the production function denotes an efficient combination of inputs and outputs. As explained earlier, there can be more than one combination of inputs, all of them technically efficient, which can be employed to produce a particular output level. However, for decision-making at the firm level, we use economic, and not technical, criteria to find an optimum input-output mix. It is for this reason that estimation of production function using other than engineering techniques become relevant.

Production function can be estimated by statistical techniques using historical data on inputs and output. One can hypothesise several alternative forms i.e., linear, quadratic, cubic etc., for this function. You may recall the various forms of production function, mentioned in earlier units. However, empirical studies on the subject have found the Cobb-Douglas (named after Cobb and Douglas who first used it for American manufacturing sector) form to be the most appropriate.

The Cobb-Douglas production function, in its general form, is expressed as :

$$Q = A\, K^{a}\, L^{b}$$

where

Q = Total output

A = Constant (or technology factor)

K = Capital Input

L = Labour Input

a,b = exponents of Capital and Labour, respectively

The above production function can also be represented in log-linear form :

$$\text{Log } Q = \text{Log } A + a \text{ Log } K + b \text{ Log } L$$

Various factor intensities, factor productivities and factor elasticities can be easily computed from an estimated production function of Cobb-Douglas type. It can be easily inferred from the above form (log-linear form) that exponents a and b are factor elasticities. Also, by adding up the factor elasticities, a measure of returns to scale can be had.

Cobb and Douglas estimated a production function for the American manufacturing sector using annual time series data for the period 1899 to 1922. Their estimated production function was :

$$X = 1.01\, L^{0.75}\, K^{0.25}$$

Their findings suggested that the American manufacturing sector was facing constant returns to scale during the above mentioned period. The elasticity of output with respect to labour was 0.75, and that with respect to capital was 0.25. This implied that a 10 per cent increase in labour input with no change in capital caused a 7.5 per cent increase in output, and a similar increase in capital input with no change in labour brought about a 2.5 per cent increase in output. The above estimated production function could also be used to generate isoquants for various levels of output, and the least cost input combination for a given output.

Empirical estimate of production functions for industries such as sugar, textiles, cement etc., are available in the Indian context.

Indian Cement Industry

One study estimates the production function facing the Indian Cement industry using time series annual data for the period 1946-1965. [See Gupta, G.S, Production Function and Factor Productivity in the Indian Cement Industry. *Indian Journal of Industrial Relations,* January 1973.] The data in value terms on inputs and output were used for estimation. He hypothesised the following forms of production function :

1) $V = A_1 W^{\alpha 1} K^{\beta 1}$

2) $V = A_2 W^{\alpha 2} K^{\beta 2} M^{\gamma 2}$

Where,

V = Value added at current prices

W = wages, salaries and benefits

K = fixed capital at current prices

M = raw material, fuel, electricity, etc., consumed at current prices

A_1 and A_2 are intercepts, and α, β, γ are value elasticities of labour, capital and raw material, respectively.

The production function estimates are given below :

1) $Q = 1.22\, W^{0.995} \overline{K}^{0.1446}$ ....$R^2 = 0.938$

2) $Q = 1.14\, W^{0.177} K^{0.039} M^{0.075}$....$R^2 = 0.996$

The cement industry depicts diminishing to constant returns to scale $(\alpha + \beta + \gamma \leqslant 1)$ during the period 1946-65. The exponent of labour is positive, as expected. The value elasticity of capital is negative in the first form, and quite low in the second form using raw material as a separate factor. The negative or low, and statistically insignificant elasticity of capital points towards the old technology and dilapidated machinery in the cement industry during the sample period. The high labour elasticity acted as a disincentive to new investment in the cement industry during the above period. The explanatory power (R) of the model is quite high i.e. 0.996.

Indian Sugar Industry

In another study, G.S. Gupta and K. Patel have estimated the production function in Indian Sugar Industry (see *Managerial Economics* by M. Adhikary for a summary of this study). The basic findings of this study are summarised in the following table :

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

Region	Sample period	Regression co-efficients & other statistics					
		Log A	α	β	$\alpha + \beta$	R^2	DW
All-India	1946-66	-1.92	0.91	0.29	1.21	.93	1.95
	1946-58	-1.79	0.92	0.23	1.15	.84	2.27
	1959-66	-5.33	1.41	0.45	1.86	.97	1.32
Uttar Pradesh	1946-66	0.42	0.43	0.28	0.71	.73	2.00
	1946-58	0.72	0.34	0.32	0.66	.53	1.96
	1959-66	-3.15	1.21	0.26	1.47	.83	2.53
Bihar	1946-66	-1.94	0.99	0.12	1.11	.60	1.52
	1946-58	-1.15	0.41	0.50	0.91	.56	0.98
	1959-66	-3.60	1.27	0.30	1.57	.63	2.70

Indian Cotton Textile Industry

The National Productivity Council (NPC) has estimated a Cobb Douglas type of production function for cotton textile output (X) for 1961 and 1966

1) 1961 - $X = 0.0438\ L^{1.054}\ K^{0.018}$

2) 1966 - $X = 0.0039\ L^{1.044}\ K^{0.036}$

The shows that over 1961-66, textile production has become more capital-dependent.

Activity 1

a) Point out your observations on :

   i) Inter-regional variations in factor elasticities.

....

....

....

....

....

   ii) Inter-temporal variations in labour elasticity.

....

....

....

   iii) Inter-temporal variation in capital elasticity.

....

....

....

b) What does $(\alpha + \beta)$ show? Would you agree that our Sugar Industry in general, operated under increasing return to scale ?

....

....

....

....

....

c) Can you infer anything about relative factor productivity in Sugar Industry? Give reasons ror your answer.

....

....

$$Q = Q(K,L) = AL^{\alpha} K^{1-\alpha}$$

Where A and α are constants; $1>\alpha>0$.

You may note that $[(\alpha) + (1-\alpha) = 1]$.

This means if factors K and L are increased by λ proportion, the output Q will also increase by the same proportion. This means constant returns to scale. Interestingly, you may operate on this form of production function to review all the production concepts you have learnt so far:

i) **Total product**

$$Q = AL^{\alpha} K^{1-\alpha}$$

ii) **Average products**

$$AP_L = A(\tfrac{K}{L})^{1-\alpha}$$

$$AP_K = A(\tfrac{L}{K})^{\alpha}$$

iii) **Marginal products**

$$MP_L = \alpha . A(\tfrac{K}{L})^{1-\alpha} = (\alpha . \tfrac{Q}{L})$$

$$MP_K = (1-\alpha)A(\tfrac{L}{K})^{\alpha} = \left[(1-\alpha)\tfrac{Q}{K}\right]$$

iv) **Marginal Rate of Technical Substitution**

$$MRTS_{LK} = \frac{MP_L}{MP_K} = \left[\frac{\alpha}{(1-\alpha)}\frac{K}{L}\right]$$

v) **Elasticity of Proudction**

$$e_q^1 = \frac{MP_L}{AP_L} = \alpha$$

$$e_q^1 = \frac{MP_K}{AP_K} = (1-\alpha)$$

vi) **Function Coefficient**

$$f = (\alpha) + (1-\alpha) = 1$$

vii) **Production efficiency**

It is measured by the term 'A'. The firm with higher 'A' produces larger output of 'Q'; high level of 'A' signifies better organisation and managerial efficiency.

viii) **Factor Intensity**

This measures the relative factor proportion (or ratio) to produce a given level to output. Along the isoquant, factor intensity may vary or remain constant, depending upon the returns to factor (short-run) and the returns to scale (long-run)

$\frac{K}{L}$ measures capital intensity

$\frac{L}{K}$ measures labour intensity

ix) **Elasticity of factor Substitution**

$$e_s^f = \frac{\%\Delta in(K/L)}{\%\Delta in(MRTS_{KL})}$$

$$\frac{\Delta(K/L)/(K/L)}{\Delta(MRTS_{KL})/(MRTS_{KL})}$$

$$= \frac{\Delta(K/L)/(K/L)}{\left(\frac{\alpha}{1-\alpha}\cdot\frac{K}{L}\right)/\left(\frac{\alpha}{1-\alpha}\right)\left(\frac{K}{L}\right)}$$

$$= \frac{\Delta(K/L)/\left(\frac{\alpha}{1-\alpha}\right)}{\frac{\alpha}{(1-\alpha)}\Delta(K/L)} = 1$$

You may operate on other forms of production to review the same set of concepts. However, here it will suffice to note the following.

b) C E S type

It stands for 'constant' elasticity of substitution production function. As the name implies, here e_s^f remains constant. The Cobb-Douglas function is a special type of C E S function.

c) V E S type

It stands for 'variable elasticity of substitution' production function, because e_s^f is allowed to vary.

d) Leontief type

In this type of production function, e_s^f is assumed to be zero. This means factors are combined in fixed proportion such that there is no scope for factor substitution.

e) Linear type

At the other extreme, e_s^f may be assumed to be infinity, meaning that there is no limit to factor substitution such that a multi-factor production may turn out to be a single factor function like.

$Q = Q(K,L) = Q = 2L$, because K and L are perfect substitutes.

Activity 7

Construct a set of tables with imaginary data to illustrate the following:

a) Diminishing returns to a variable factor

b) Increasing returns to scale

c) Constant returns to scale

d) Diminishing returns to scale

....

....

....

....

....

....

....

....

....

....

....

Activity 8

a) Consider the production function

$$Q = 7\frac{L^5}{K^4} + 3\frac{K^2}{L}$$

Establish the form of this function

b) Consider the function

$Q = 86\, L^{0.61}\, K^{0.39}$

Compute the following:

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

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

$MRTS_{LK} =$

$e^{f}_{q} =$

$e^{k}_{q} =$

$e^{f}_{s} =$

$f =$

$A =$

c) **Explain each of the following proportions.**

i) **Factor productivity depends on factor-intensity in production.**

ii) Factor substitutability depends on factor productivity.

iii) Isoquant map for a Leontief type of production has a special shape and slope.

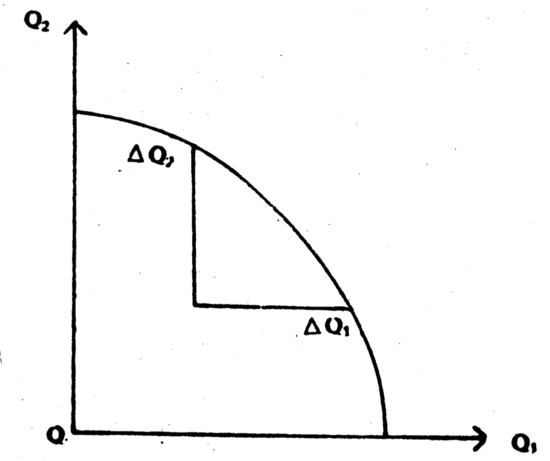
iv) For Cobb-Douglas function, the expansion path is linearly homogeneous of degree one.

8.11 OPTIMAL PRODUCT-MIX

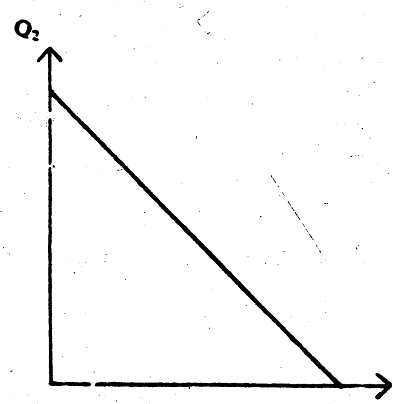
Hitherto we have considered the case of single-product firm. Let us now move on to the case of joint-product or multi-product firm. The joint-product case is a special type of multi-product firm. Most of the firms today produce a combination of products-at least main-product and its by-products, for example, Durgapur Steel Plant produces steel and gas. In other words, you may now assume that there is a firm which produces two products, Q_1 and Q_2 jointly out of a given factor, F. It is like saying both mutton and wool are being produced out of a sheep. The question is what is the optimal product-mix in this case. Before you attempt an answer to this question, you may note:

$F=F(Q_1, Q_2)$ is the transformation or production possibility function, which when plotted gives you the produce transformation curve.

The slope of this curve, $\frac{\Delta Q_1}{\Delta Q_2}$ measures the Marginal Rate of Product Transformation ($MRPT_{Q_1 Q_2}$)



Now the producer get his revenue from the market by producing and selling both products. Thus, his total revenue, $TR = P_1Q_1 + P_2Q_2$ when P_1 and P_2 are respective prices of Q_1 and Q_2. We now get AB the iso-revenue curve, where slope measures the product price ratio $\frac{P_1}{P_2}$

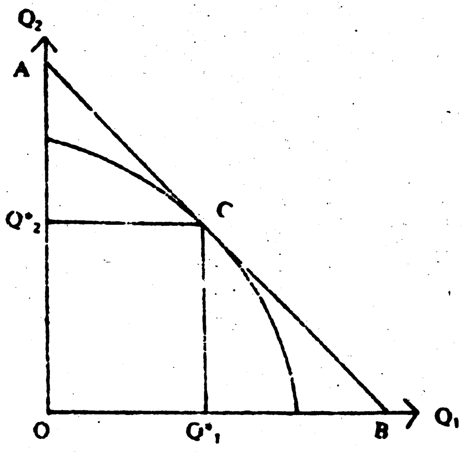


The firm may now try to maximise the revenue subject to the constraint imposed by the production possibility or minimise the resource use to attain a given revenue. Either way, the decision rule for an optimum product-mix works out as:

$$\therefore \frac{\Delta Q_1}{\Delta Q_2} = MRPTQ_1Q_2 = \frac{P_1}{P_2}$$

This is illustrated in the diagram; corresponding to the equilibrium condition the firm produces a combination of $OQ_1^* + OQ_2^*$.

You may note that some degree of substitutability between Q_1 and Q_2 is being assumed. If more of Q_1 is produced out of given resource, F, less of Q_2 can be produced.



The law of diminishing returns is said to operate. Constant returns to scale will be witnessed if the output also exactly doubles in this case.

Activity 9

a) Quote a couple of examples of joint-product or multi-product firms. Name the firm and its products.

....

....

....

....

....

b) Given $F = F(Q_1, Q_2)$.

$\frac{\partial F}{\partial Q_1}$ = marginal resource cost of producing Q_1

$\frac{\partial F}{\partial Q_2}$ = marginal resource cost of producing Q_2

Show the $MRPT_{Q_1Q_2}$ is simply the ratio of these marginal costs (measured in physical resources).

....

....

....

....

c) Given $F = 500\ Q^1, 0.3Q^2\ 0.7$ show that 0.3 or 0.7 measures product elasticity of resource use, $e_f^{q^1}$ or $e_f^{q^2}$.

....

....

....

....

d) Also estimate the elasticity of product transformation, e_t^q, in the above estimated function. Provide an economic interpretation of e_t^q. Note that

$$e_t^q = \frac{\Delta(Q_1/Q_2)/(Q_1/Q_2)}{\Delta(MRPT_{Q_1Q_2})/(MRPT_{Q_1Q_2})}$$

....

....

....

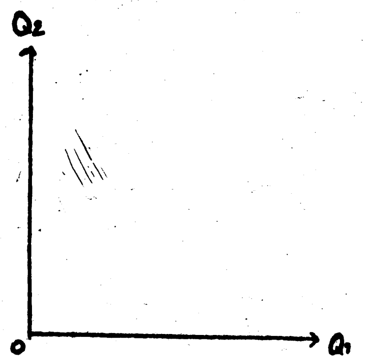
....

....

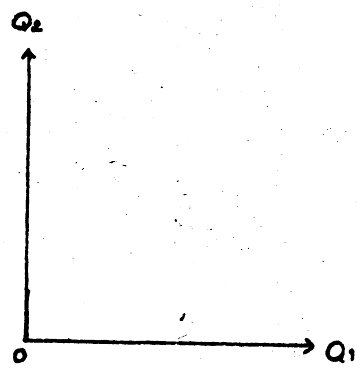
....

e) Draw separate diagrams illustrating the case of :

Revenue maximisation subject to given production possibility



Resource minimisation subject to given revenue situation



(You need to draw either production possibility (transformation) map or iso-revenue map as the case may be).

8.10 SUMMARY

Production is the process of providing goods and services that have value to consumers. A profit maximising firm tries to achieve the optimum input-output combination for a given cost. In this unit we have shown that a firm minimises the cost for a given output in order to maximise the profit. Production at the lowest possible cost requires that the additional rupee spent on additional output is equal to the output obtained per additional rupee spent on any other input. We call this the least cost combination of inputs.

We have developed certain concepts like production function, isoquant and isocost curves, and the marginal rate of technical substitution of one input for the other etc. to derive the condition for achieving the least cost combination of inputs. The production function shows the maximum natities of output a firm can produce using various combinations of factor inputs. An **isoquant** curve indicate different combinations of two inputs which can be employed to produce a specific level of output. An **isocost line** expresses different combinations of two inputs which can be utilised for a given rupee cost. The **marginal rate of technical substitution** is the rate at which one input may be substituted for another while the same level of output is maintained and is equal to the negative of the slope of an isoquant curve. The least cost combination of inputs for a particular level of output is reached at the point where an isocost curve is tangent to the isoquant curve corresponding to that level of output. At this point, the marginal rate of substitution, which is also equal to the ratio of marginal product of two factor inputs, is equal to the ratio of inputs prices. In other words,

$MRTS = \frac{MP_L}{MP_K} = \frac{P_L}{P_K}$ assuming there are only two inputs L & K. The line joining all the least cost combination points for different level of output is called the expansion path of the firm.

The **principle of diminsing marginal productivity** affects nearly all types of production functions. Thus, adding additional units of any variable input to fixed inputs resutls in successively smaller increment to output. However, when all inputs can be varied simultaneously, production may be characterised by increasing or constant or decreasing returns to scale.

To maximise profit in the short-run, a firm should employ a variable input (say, input L) upto the point where the additional revenue (that another unit of labour will bring in) is just equal to its cost, or where the marginal revenue product or labour equals its marginal resource cost i.e., $MRP_L = MRC_L$. Subsequently we have moved from single product to joint-product case where we have got comparable concepts like Marginal Rate of Product Transformation ($MRPT_{Q_1Q_2}$) and elasticity of product transformation. However, we have not touched upon the form of **transformation function** as we have done in the case of production function e.g. Cobb-Douglas, Leontief, CES, VES, linear types. In the case of joint-product, the optimum decision rule works out to be one of proportionality between $MRPT_{Q_1Q_2}$ and the product Price ratio $\frac{P_1}{P_2}$, where MRPT is the ratio or resource costs of production.

8.13 ADDITIONAL READINGS

Adhikary, M. 1987, *Managerial Economics*, Khosla Publications: Delhi (Chapter V).

Baumol, William J. 1977, *Economics Theory and Operation Analysis*, (4th edition) Prentice Hall: Englewood-Cliffs. (Chapter II).

Bilas, Richard A, 1971, *Microeconomic Theory*, 2nd edition, McGraw Hill Book Co.: (Chapter 6).

Douglas, Evan J. 1983, *Managerial Economics: Theory Practice and Problems*, Prentice Hall Inc: Englewood-Cliffs.

Mote, V.L., Samuel Paul and G.S. Gupta, 1977, *Managerial Economics: Concepts and Cases*, (Tata McGraw Hill: Bombay)

8.14 SELF-ASSESSMENT TEST

1. Explain clearly the various facets of production decision problems e.g., choice of technique, choice of product etc.

2. Compare the case of single-product firm with that of a joint-product firm with reference to underlying set of concepts, techniques and optimum decision rules.

3. Distinguish between the concepts of:
   a) Production and Productivity
   b) Average and Marginal productivity
   c) Input elasticity and function coefficient
   d) Returns to a factor and Returns to scale
   e) Factor substitution and product transformation
   f) Factor intensity and Factor productivity

4. Take an imaginary production function and show that it is possible for a firm to have "diminishing returns to a factor" along with "constant returns to scale". Bring out clearly the economic implications of such a case.

5. Draw a comprehensive analogy between demand analysis and production analysis.

6. A factory manufactures two types of machines x and y. The transformation function is given by

   $f(x,y) = x^2 + 2y^2 - xy$

   To minimise resource cost, how many machines of each type should be produced if there must be a total of 8 machines?

   (Ans: $x = 5$)
   $y = 3$)

7. Suppose the production function is $16Q = 65 - 2(L-5)^2 - 4(K-4)^2$. The unit price of inputs L and K are 8 and 4 respectively and the unit price of output Q is 32. Determine the maximum profit corresponding to optimum input choice.

   [State profit function, $\pi = R - C = (P_Q Q - (P_L L + P_K K)$
   Also note that the decision variables are L and K]

   (Ans: $L = 4$
   $K = 15/4$
   $\pi = 78\frac{1}{2}$

8. It is suggested that there is a world of difference between **production function and supply function**, though one can be derived from the other. Comment. (Treat it **mathematically**).

   (Hint: See Baumol or Adhikary).

UNIT 9 COST CONCEPTS AND ANALYSIS

Structure

9.0 INTRODUCTION

Cost and revenue are the two major factors that a profit maximising firm needs to monitor continuously. It is the level of cost relative to revenue that determines the firm's overall profitability. In order to maximise profits, a firm tries to increase its revenue and lower its cost. While the market factors determine the level of revenue to a great extent, the cost can be brought down either by producing the optimum level of output using the least cost combination of inputs, or increasing factor productivities, or by improving the organisational efficiency. The firm's output level is determined by its cost.

Product prices are determined by the interaction of the forces of demand and supply. The basic, factor underlying the ability and willingness of firms to supply a product in the market is the cost of production. Thus, cost of production provides the floor to pricing. It is the cost that forms the basis for many managerial decisions like which price to quote, whether to accept a particular order or not, whether to abandon or add a product to the existing product line, whether or not to increase the volume of output, whether to use idle capacity or rent out the faclities, whether to make or buy a product, etc. However, it is essential to underline here that all costs are not relevant for every decision under consideration.

The purpose of this unit is to explore cost and its relevance to decision-making. we begin by developing the important cost concepts, an understanding of which can aid managers in making correct decisions. We shall examine the diffeence between economic and accounting concepts of costs and profits. We shall then consider the concepts of short-run and long-run costs and show that they, in conjunction with the concepts of production studies in the preceding unit, can give us a more complete understanding of the applications of cost theory to decision-making.

9.1 OBJECTIVES

On reading this unit, you should be able to

- relate the concepts of production and costs.
- define different types of costs relevant for production, planning and control.

- distinguish between economic costs and accounting costs
- identify the sources from where cost-advantages follow
- indicate application of cost concepts and analysis in managerial decision-making.

9.2 TYPES OF COSTS

There are many different types of costs that a firm may consider relevant for decision-makin under varying situations. The manner in which costs are classified or defined is largely depende on the purpose for which the cost data are being outlined.

Explicit and Implicit Costs

The opportunity cost (or cost of the foregone alternative) of a resource is a definition cost i most basic form. While this particular definition of cost is the preferred baseline for econo in describing cost, not all costs in decision-making situations are completely obvious; one c skills of a good manager is the ability to uncover hidden costs. For a long time, there has a considerable disagreement among the economists and accountants on how costs shou treated. The reason for the difference of opinion is that the two groups want to use the co data for dissimilar purposes. Traditionally, the accountants have been primarily connected with collection of historical cost data for use in reporting a firm's financial behaviour and position and in calculating its taxes. The main functions of accountants have been reporting, stewardship and control. They report or record what was happened, present information that will protect the interests of various shareholders in the firm, and provide standards against which performance can be judged. All these have only indirect relationship to decision-making. Business economists, on the other hand, have been primarily concerned with using cost data in decisions making. These purposes call for different types of cost data and classification.

Traditional accounting data is not directly suitable for decision-making. While accountants s rely primarily on historical cost in determining the profit or loss of a firm, economists prefer use the opportunity cost baseline concept for this purpose.

The opportunity cost of a resource can be defined as the value of the resource in its next best use, that is, if it were not being used for the present purpose. The opportunity cost is the benefit of using a resource for the next most attractive alternative. For example, the opportunity cost of a student's doing a full time MBA could be the income that he would have earned if he had employed his labour resources on a job, rather than spending them in studying managerial economics, accounting, and so on. The time cost in money terms can be referred to as implicit cost of doing an MBA.

The out-of-pocket costs on tuition and teaching materials are the explicit costs that a student incurs while attending MBA. Thus, the total cost of doing an MBA to a student is implicit costs (opportunity cost) plus the explicit (out-of-pocket) costs.

Accountants typically use those costs that are recorded in their books as representing an actual transfer of money. These are explicit, or nominal costs and often do not represent full economic costs that should be considered in a given decision. In addition to explicit costs, the business economist uses implicit or imputed cost in evaluating a decision.

Furthermore, in measuring the cost of resource in use, the accountant is only concerned with its acquisition cost. But, for decision-making purposes, we necessarily talk about future costs and revenues, and therefore, past costs have very little relevance. Also, the traditional accounting procedure for valuing assets on the balance sheet is acquisition cost minus depreciation. This is faulty as the true current market value of an asset may differ from its book value.

Direct and Indirect Costs

There are some costs which can be directly attributed to production of a given product. The use of raw material, labour input, and machine time involved in the production of each unit can usually be determined. On the other hand, there are certain costs like stationery and other office and administrative expenses, electricity charges, depreciation of plant and buildings, and other such expenses that cannot easily and accurately be separated and attributed to individual units of production, except on arbitrary basis. When referring to the separable costs of first category accountants call them the direct, or prime costs per unit. The joint costs of the second category are referred to as indirect or overhead costs by the accountants.

Direct and indirect costs are not exactly synonymous to what economists refer to as variable costs and fixed costs. The criterion used by the economist to divide cos into either fixed or variable is whether or not the cost varies with the level of output, where as the accountant divides the cost on the basis of whether or not the cost is separable with respect to the production of individual output units. The accounting statements often divide overhead expenses into 'variable overhead' and 'fixed overhead' categories. If variable overhead expenses per unit are added to the direct cost per unit, we arrive at what economists call as average variable cost.

Private Costs versus Social Costs

A further distinction that is useful to make - especially in the public sector, is between private and social costs. Private costs, are those that accrue directly to the individuals of firms engaged in relevant activity. External costs, on the other hand, are passed on to persons not involved in the activity in any direct way (i.e., they are passed on to society at large). Consider the case of a manufacturer located on the bank of a river who dumps the waste into water rather than disposing it of in some other manner. While the private cost to the firm of dumping is zero, it is definitely positive to the society. It affects adversely the people located downcurrent who are adversely affected and incur higher costs in terms of treating the water for thier use, or having to travel a great deal to fetch potable water. If these external costs were included in the production costs of the producing firm, a true picture of real, or social costs of the output would be obtained. Ignoring external costs may lead to an inefficient and undesirable allocation of resources in society.

Relevant Costs and Irrelevant Costs

The relevant costs for decision-making purposes are those costs which are incurred as a result of the decision under consideration. The relevant costs are also referred to as the incremental costs. Costs that have been incurred already and costs that will be incurred in the future regardless of the present decision are irrelevent costs as far as the current decision problem is concerned.

There are three main categories of relevant or incremental costs. These are the present-period explicit costs, the opportunity costs implicitly involved in the decision, and the future cost implications that flow from the decision. For example, direct labour and material costs, and changes in the variable overhead costs are the natural consequences of a decision to increase the output level. Also, if there is any expenditure on capital equipments incurred as a result of such a decision, it should be included in full, not withstanding that the equipment may have a useful life remaining after the present decision has been carried out. Thus, the incremental costs of a decision to increase output level will include all present-period explicit costs which will be incurred as a consequence of this decision. It will exclude any present-period explicit cost that will be incurred regardless of the present decision.

The opportunity cost of a resource under use, as discussed earlier, becomes a relevant cost while arriving at the economic profit of the firm. This point will be explained in detail in the following section.

Many decisions will have implications for future costs, both explicit and implicit. If a firm expects to incur some costs in future as a consequence of the present analysis, such future costs should be included in the present value terms if known for certain.

Economic Costs and Profits

Our earlier discussion of economic costs suggests that economists and accountants use the term 'profits' differently. Accounting profits are the firm's total revenue less its explicit costs. But economists define profits differently. Economic profits are total revenue less all costs (explicit and implicit, the latter including a normal profit required to retain resources in a given line of production). Therefore, when an economist says that a firm is just covering its costs, it is meant that all explicit and implicit costs are being met, and that, the entrepreneur is receiving a return just large enough to retain his or her talents in the present line of production. If a firm's total receipts exceed all its economic costs, the residual accuring to the entrepreneur is called an economic, or pure profit. In short:

Economic Profit = Total Revenue – Opportunity Cost of all Inputs

This is depicted in the following figure:

	Economic Profits	Accounting Profit
Total Revenue	Economic or Opprotunity Cost (Explicit plus implicit costs, including a normal profit)	Accounting Costs

An economic profit is not a cost, because by definition it is a return in excess of the normal profit required to retain the entrepreneur in a particular line of production.

A Case

Let us consider the case of Mr. Ashis Nathani who is a small store owner. He has invested Rs. 2 lakh as equity in the store and inventory. His annual turnover is Rs. 8 lakh, from which he must deduct the cost of goods sold, salaries of hired staff, and depreciation of equipment and building to arrive at the annual profit of the store. When his accountant Mr. Chhagani, after using the above approach reported the profit to be Rs. 1.5 lakh, Mr. Nathani could not believe it and he deided to go through the books himself. He discovered that his actual profit was only Rs. 75,000 and not Rs. 1.5 lakh. He found out that his accoutant had understimated the costs, by not caring to provide for the resources that he had put into his business, namely his own time and the money. In finding out his economic profit, he added the inputed salary(to himself) and inputed interest cost on his equity. The two estimates are shown below:

Accounting Income Statement (Mr. Chhagan)	Rs.	Rs.
Sales		8,00,000
Cost of Goods Sold	6,00,000	
Salaries	40,000	
Depreciation	10,000	6,50,000
Accounting Profit		1,50,000

Economic Statement of profit (Mr. Nathani)	Rs.	Rs.
Sales		8,00,000
Cost of Goods Sold	6,00,000	
Salaries	40,000	
Depreciation	10,000	
Inputed Salary to owner Manager	50,000	
Inputed Interest cost on Equity	25,000	7,25,000
Economic Profit		75,000

Separable and Common Costs

Costs can also be classified on the basis of their traceability. The costs that can be easily attributed to a product, a division, or a process are called separable costs, and the rest are called non-separable or common costs. The separable and common costs are also referred to as direct and indirect costs. The distinction between direct and indirect costs is of particular significance in a multi-product firm for setting up economic prices for different products.

Fixed and Variable Costs

Fixed costs are those costs which in total do not vary with changes in output. Fixed costs are associated with the very existence of a firm's plant and therefore must be paid even if the firm's rate of output is zero. Such costs as interest on borrowed capital, rental payments, a portion of depreciation charges on equipment and buildings, and the salaries of top management and key personnel are generally fixed costs.

On the other hand, variable costs are those costs which increase with the level of output. They include payment for raw materials, charges on fuel and electricity, wages and salaries of temporary staff, depreciation charges associated with wear and tear of assets, and sales commission, etc.

This distnction is true only for the short-run. It is similar to the distinction that we made in the previous unit between fixed and variable factors of production under the short-run production analysis. The costs associated with fixed factors are called the fixed costs and the ones associated with variable factors, the variable costs. Thus, if capital is the fixed factor, capital rental is taken as the fixed cost and if labour is the variable factor, wage bill is treated as the variable cost.

However, it is not very easy to classify all costs into fixed and variable. For example, part of the depreciation charges are fixed, and part variable. However. it is very difficult to determine how much of depreciation cost is due to the technical obsolescence of assets and hence fixed cost, and how much is due to the use of equipment and hence variable cost. Nevertheless, it doesn't mean that it is not useful to classify costs into fixed and variable. This distinction is of great value in break-even analysis and pricing decisions. For decision-making purposes, in general, it is the variable cost which is relevant and not the fixed cost.

Activity 1

a) Give specific examples of :

Implicit costs: ....

Social costs: ....

Indirect costs: ....

Sunk costs: ....

Traceable costs: ....

Common costs: ....

b) Comment on the nature of costs involved in **depreciation** from both economic and accounting standpoints.

....

....

....

....

c) Distinguish between historical costs and replacement costs. Why is this distinction useful?

....

....

....

....

d) Give examples to distinguish between 'fixed overheads' and 'variable overheads'.

....

....

....

....

e) Can all 'direct costs' be treated as 'variable costs'?

....

....

....

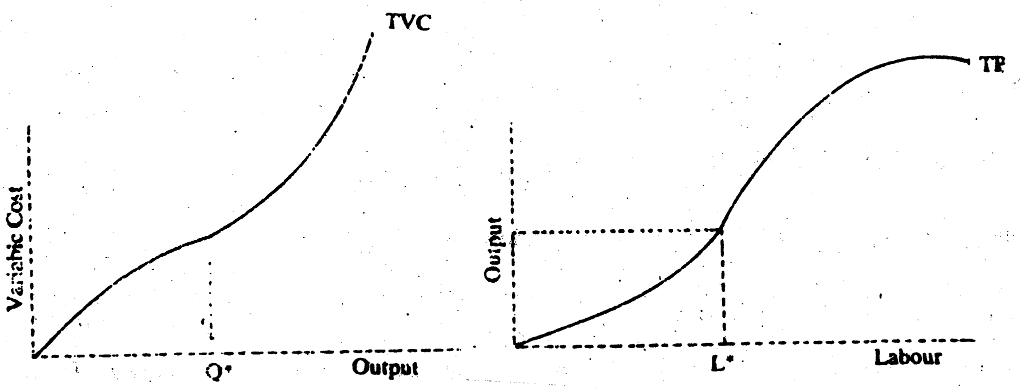
....

9.3 RELATIONSHIP BETWEEN PRODUCTION AND COSTS

The concept of cost is closely related to production theory. A cost function is the relationship between a firm's costs and the firm's output. While the production function specifies the technological maximum quantity of output that can be produced from various combinations of inputs, the cost function combines this information with input price data and gives information on various outputs and their prices. The cost function can thus be thought of as a combination of the two pieces of information i.e. production function and input prices.

Let us consider a short-run production function with only one variable input. As we have already seen in the previous unit, the diminishing returns to the variable input set in after a point. The output grows at an increasing rate in the initial stages implying increasing returns to the variable input, and then diminishing returns to the variable input start. Assuming that the input prices remain constant, the above production function will yield the variable cost function which has a shape that is characteristic of many variable cost function; increasing at a decreasing rate and then increasing at an increasing rate. This relation between the total product curve and the total variable cost is shown in Figure 1.

Figure I



This firm has a production function that exhibits increasing returns to the variable factor input (labour) upto L* and decreasing returns after L*. The input level L* corresponds to the output level Q*. The variable cost function increases at a decreasing rate upto Q* and at an increasing rate beyond Q*

From this, you should be able to derive a relationship between average product and average costs, and marginal product and marginal costs. For example,

TVC = Pr. V

$\therefore AVC = \frac{TVC}{Q} = Pr.\ \frac{V}{Q} = \frac{Pr}{Q/V}$

and $MC = \frac{TVC}{Q} = Pr.\ \frac{V}{Q} = \frac{Pr}{\Delta Q/\Delta V}$

where Pr stands for the price of the variable factor and r stands for amount of variable factor

You may note that Pr being given, AVC is inversely related to the average product of the variable factors.

In the same way, given the wage rate MC is inversely related to the marginal product of labour. We shall explore this relationship in greater detail subsequently.

Activity 2

Given $Q = 100 + 0.L^2 - 0.0005\ L^3$

Where Q is output; L is labour.

Suppose the wage rate is Rs. 10 and that 100 labourers are being employed.

Find the AVC and MC.

....

....

....

....

....

....

....

(Ans : AVC = Rs. 1.66. MC = Rs. 2)

9.4 SHORT-RUN COST FUNCTIONS

Recall the concept of "time perspective" from Block 1.

The short-run is normally defined as a time period over which some factors of production are fixed and others are variable. Needless to emphasize here that these periods are not defined by some specified length of time but, rather, are determined by the variability of factors of production. Thus, what one firm may consider in the long-run may correspond to the short-run for another firm.

In the short-run, a firm incurs some costs that are associated with variable factors and others that result from fixed factors. The former are called variable costs and the latter represent fixed costs. Variable costs (VC) change as the level of output changes and therefore can be expressed as a function of output (Q), that is VC = f(Q). Variable costs typically include such things as raw material, labour, and utilities. In Column 3 of Table 1, we find that the total of variable costs changes directly with output. But note that the increases in variable costs associated with each one-unit increase in output are not constant. As production begins, variable costs will, for a time, increase by a decreasing amount; this is true through the fourth unit of the output. Beyond the fourth unit, however, variable costs rise by increasing amount for each successive unit of output. The explanation of this behaviour of variable costs lies in the law of diminishing returns.

Table 1

Total and Average-Cost Schedules for an Individual Firm in the Short-Run

(Hypothetical Data in Rupees)

Total cost data, per week				Average-cost data, per week			
(1) Total Product	(2) Total fixed cost (TFC)	(3) Total variable cost (TVC)	(4) Total cost (TC) TC = TFC+ TVC	(5) Average fixed cost (AFC) AFC = TFC/Q	(6) Average variable cost (AVC) AVC = TVC/Q	(7) Average total cost (ATC) ATC = TC/Q	(8) Marginal cost (MC) MC = change in TC / change in Q
0	100	0	100				
1	100	90	190	100.00	90.00	190.00	90
2	100	170	270	50.00	85.00	135.00	80
3	100	240	340	33.33	80.00	113.33	70
4	100	300	400	25.00	75.00	100.00	60
5	100	370	470	20.00	74.00	94.00	70
6	100	450	550	16.67	75.00	91.67	80
7	100	540	640	14.29	77.14	91.43	90
8	100	650	750	12.50	81.25	93.75	110
9	100	780	880	11.11	86.67	97.78	130
10	100	930	1030	10.00	93.00	103.00	150

Fixed costs on the other hand, are not a function of the level of output and are constant in the short-run, that is FC = K. Fixed costs may include such things as property taxes, the cost of leases on land, buildings and some types of equipment, interest charges on the long term borrowed funds, and insurance costs. In column 2 of Table 1, we have assumed that the firm's total fixed costs are Rs. 100. Note that, by definition, this fixed cost figure prevails at all levels of output including zero.

The distinction between fixed and variable costs is of great significance to the business manager. Variable costs are those costs which business can control or alter in the short-run by changing levels of production. On the other hand, fixed costs are clearly beyond business executive's control, such costs are incurred in the short-run and must be paid regardless of output level.

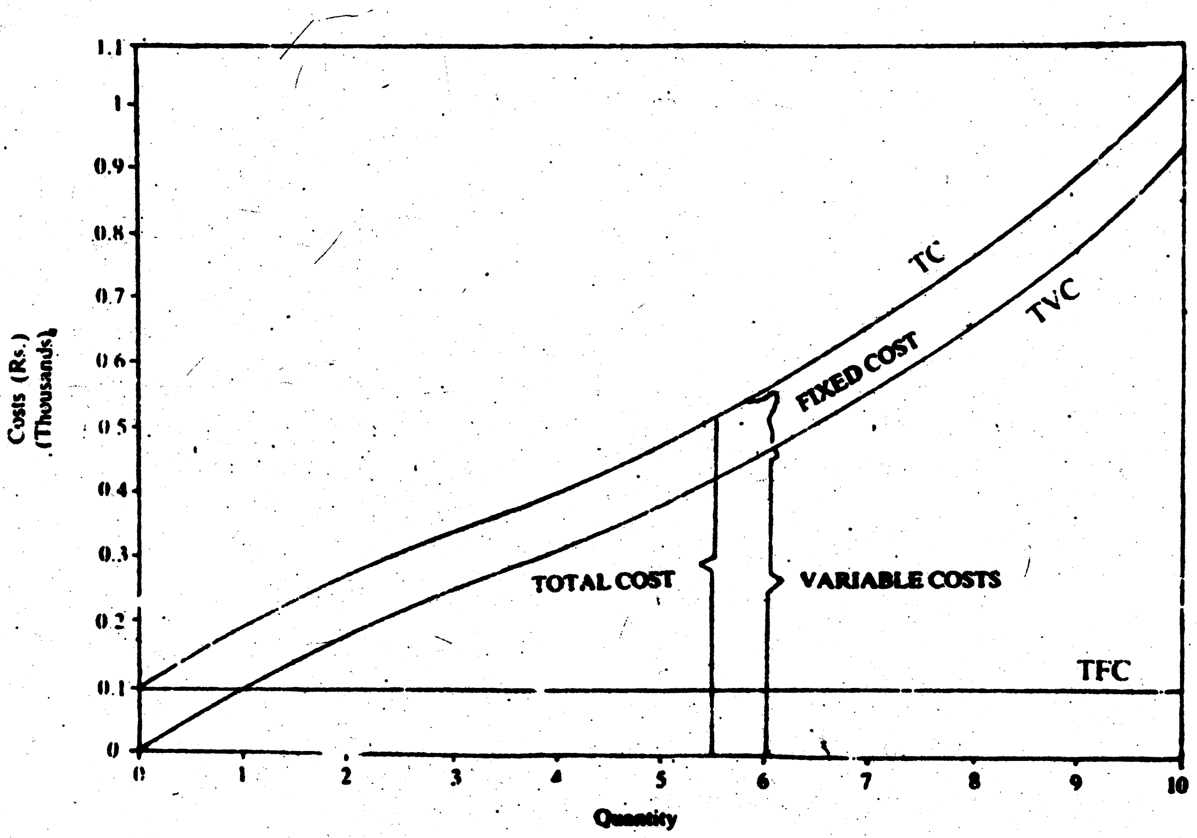
Total Cost

Total cost is the sum of fixed and variable cost at each level of output. It is shown in column 4 of Table 1. At zero unit of output, total cost is equal to the firm's fixed cost. Then for each unit of production (through 1 to 10), total cost varies at the same rate as does variable cost. Figure II shows graphically the fixed, and total cost data of Table 1.

Per Unit, or Average Costs

Besides their total costs, producers are equally concerned with their per unit, or average costs. In particular, average cost data is more relevant for making comparisons with product price, which is always stated on per unit basis. Average fixed cost, average variable cost, and average cost are shown in columns 5 to 7 of Table 1. It is important that we know how these unit-cost figures are derived and how they vary as output changes.

Figure II Fixed, total variable & total costs



TFC+TVC = TC

Average Fixed Costs

Average fixed cost (AFC) is derived by dividing total fixed cost (TFC) by the corresponding output (Q). That is

$$AFC = \frac{TFC}{Q}$$

While total fixed cost is, by definition, independent of output, AFC will decline so long as output increases. As output increases, a given total fixed cost of Rs. 100 is obviously being spread over a larger and larger output. This is what business executives commonly refer to as 'spreading the overheads'. We find in Figure III that the AFC curve is continuously declining as the output is increasing. The shape of this curve is of an asymptotic hyperbola.

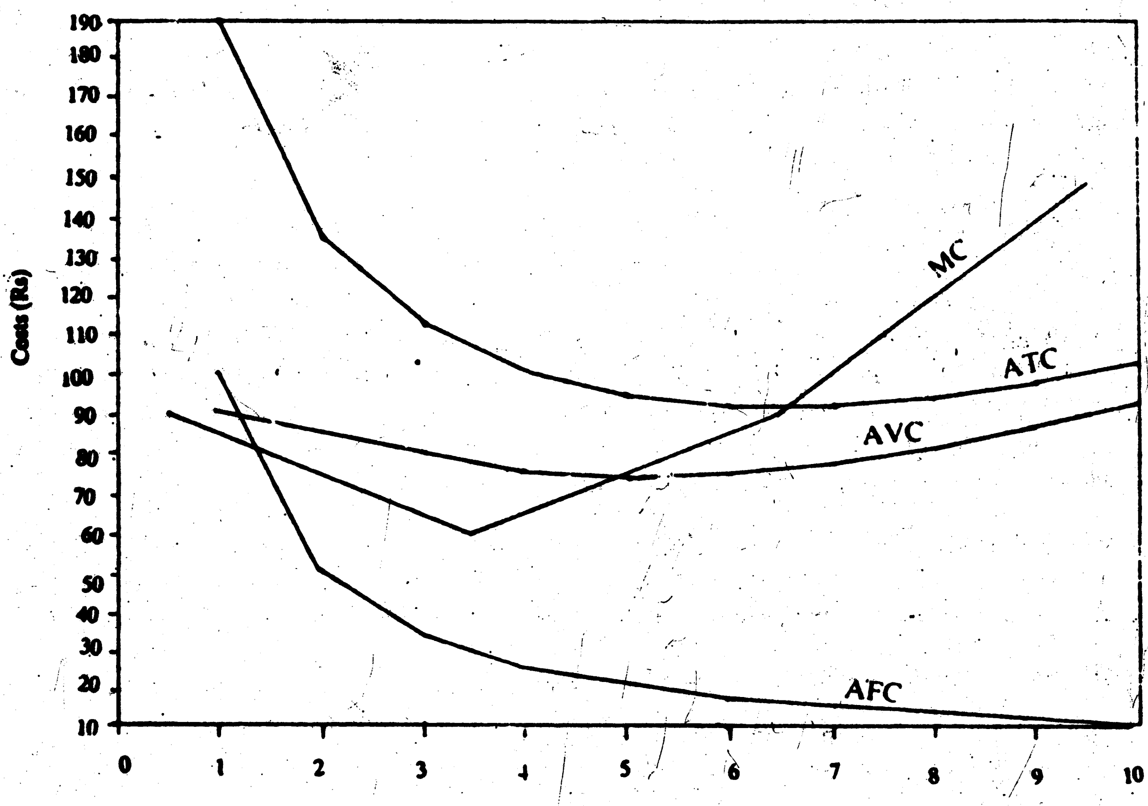
Average Variable Costs

Average variable cost (AVC) is found by dividing total variable cost (TVC) by the corresponding output (Q) :

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

AVC declines initially, reaches a minimum, and then increases again. Graphically, this provides us with a U-shaped or saucer-shaped AVC curve, as shown in Figure III.

Figure III Average fixed, variable & Marginal Costs



AFC+AVC = ATC

$$\frac{\Delta ATC}{\Delta Q} = MC$$

Because total variable cost reflects the law of diminishing returns, so must the AVC figures, which are derived from total variable cost. Because of increasing returns initially, it takes fewer and fewer additional variable resources to produce each of the first four units of output. As a result, variable cost per unit will decline. Thus, AVC hits a minimum with the fifth unit of output, and beyond this point AVC rises as diminishing returns necessitate the use of more and more variable resources to produue each additional unit of output.

Average Total Costs

Average total cost (ATC) can be found by dividing total cost (TC) by total output (Q) or, by adding AFC and AVC for each level of output. That is:

$$ATC = \frac{TC}{Q} = AFC + AVC$$

These data are shown in column 7 of Table 1. Graphically, ATC is found by adding vertically the AFC and AVC curves, as in Figure III. Thus the verticle distance between the ATC and AVC curves reflects AFC at any level of output.

Marginal Cost

Marginal cost (MC) is defined as the extra, or additional, cost of producing one more unit of output. MC can be determined for each additional unit of output simply by noting the change in total cost which that unit's production entails :

$$MC = \frac{Change\ in\ TC}{Change\ in\ Q} = \frac{\Delta TC}{\Delta Q}$$

Our data in Table 1 is so structured that the "change in Q" is always "1", so we have defined MC as the cost of one more unit of output. However, the same concept can be extended for situations where Q is more than 1. Arc marginal cost (over a range of output) can be found by using the above expression.

The marginal cost concept is very crucial from the manager's point of view. Marginal cost is a strategic concept because it designates those costs which the firm has the most direct control. More specifically, MC indicates those costs which are incurred in the production of the last unit of output and therefore, also the cost which can be "saved" by reducing total output by the last unit. Average cost figures do not provide this information. A firm's decisions as to what output level to produce is largely influenced by its marginal cost. When coupled with marginal revenue, which indicates the charge in revenue from one more or one less unit of output, marginal cost allows a firm to determine whether it is profitable to expand or contract its level of production.

Marginal cost is shown graphically in Figure III. Note that marginal cost declines sharply, reaches a minimum and then rises rather sharply. This mirrors the fact that variable cost, and therefore total cost, increases first by decreasing amounts and then by increasing amounts (Figure I).

MC and Marginal Product

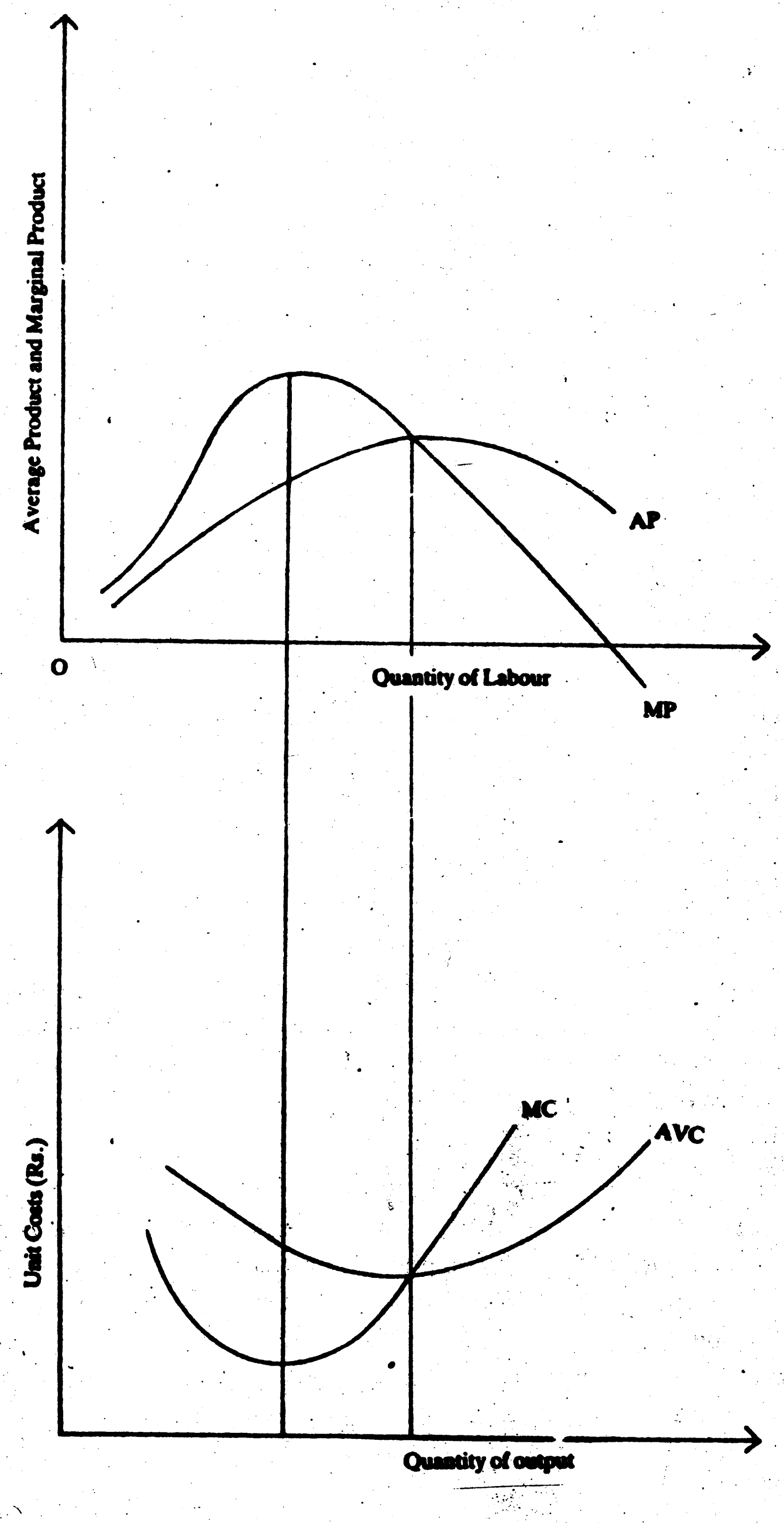
The shape of the marginal cost curve is a reflection of, and the consequence of, the law of diminishing returns. If each successive unit of a variable input say labour, is hired at a constant price, the marginal cost of each extra unit of output will fall so long as the marginal product of each additional worker is rising. This is so because marginal cost is simply the constant wage (cost) of an extra worker divided by his or her marginal product. Thus, given the price (cost) of the variable resource, increasing returns i.e., a rising marginal product, will be reflected in a declining marginal cost and diminishing returns i.e., falling marginal product, in a rising marginal cost. The MC curve is a mirror reflection of the marginal product curve. This relationship is shown in Figure IV. It is clearly shown that when marginal product is rising, marginal cost is necessarily falling. When marginal product is at its maximum, marginal cost is at its minimum. And when marginal product is falling, marginal cost is rising.

Relationship of MC to AVC and ATC

It is also notable that marginal cost cuts both AVC and ATC at their minimum (Figure III). When both the marginal and average variable costs are falling, average will fall at a slower rate. And when MC and AVC are both rising, Mc will rise at a faster rate. As a result, MC will attain its minimum before the AVC. In other words, when MC is less than AVC, the AVC will fall, and when MC exceeds AVC, AVC will rise. This means (Figure III) that so long as MC lies below AVC, the latter will fall and where MC is above AVC, AVC will rise. Therefore, at the point of intersection where MC = AVC, AVC has just ceased to fall and attained its minimum, but has not yet begun to rise. Similarly, the marginal cost curve cuts the average total cost curve at the latter's minimum point. This is because MC can be defined as the addition either to total cost or to total variable cost resulting from one more unit of output. However, no such relationship exists between MC and the average fixed cost, because the two are not related; marginal cost by definition includes only those costs which change with output, and fixed costs by definition are independent of output.

Figure IV

THE RELATIONSHIP BETWEEN PRODUCTIVITY CURVES AND COST CURVES



Managerial Uses of the Short-Run Cost Concepts

As already emphasised the relevant costs to be considered for decision-making will differ from one situation to the other depending on the problem faced by the manager. In general, the total cost concept is quite useful in finding out the break-even quantity of output. The total cost concept is also used to find out whether firm is making profits or not. The average cost concept is important for calculating the per unit profit of a business firm. The marginal and incremental cost concepts are essential to decide whether a firm should expand its production or not

Activity 3

a) Fill in the blanks in the Table below.

b) Take a separate graph paper and draw all the curves.

Short-run Cost-Schedules

Q	TFC	TVC	TC	AFC	AVC	ATC	MC
1.	50		55				
2.	50	8		25			
3.	50		60.5				
4.		13					
5.	50		65				
6.	50	18			3	11.34	3
7.	50		72.5				
8.	50	28					
9.			86				
10.	50	45		5		9.5	9
11.	50	54.5		4.55		9.5	9.5
12.	50	65.2					
13.	50		130				
14.	50	99.1					
15.	50		174.75				
16.	50	162					
17.	50		259.25				
18.		269.5					
19.	50		399				
20.	50	450		2.5	22.5	25	

– **Output Q is measured in '000 units**

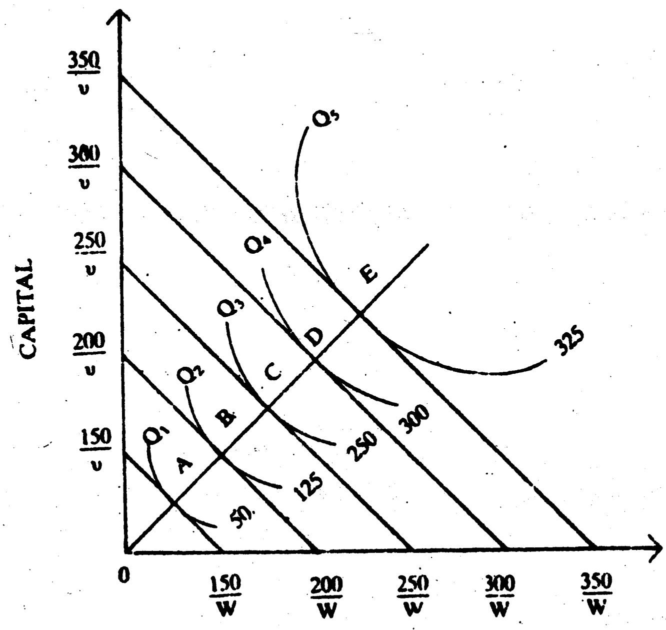
– **All costs are measured in Rs. '000**

9.5 LONG-RUN COST FUNCTIONS 101

Long-run total cost curves are derived from the long-run production functions in which all inputs are variable. Such a production function is represented by the five isoquant curves showing five different levels of output in Figure V. The five isocost curves tangent to these isoquants at the points A, B, C, D and E represent total cost on resources. Since the cost per unit of capital (V) and, labour (W) are assumed to be constant, these five isocost curves are parallel to one another, and the distance between them is constant along the expansion path traced out by A, B, C, D and E.

Taking the values for total cost and output from the expansion path of Figure V (the most efficient points), we can construct the following table for cost and output :

FIGURE V EXPANSION PATH AND RETURNS TO SCALE

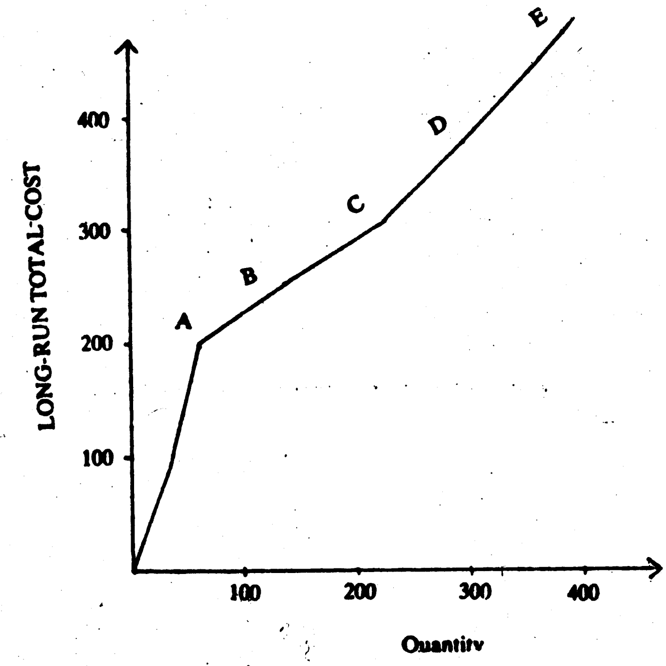


LABOUR

Output (Q)	Long-run Total Cost (LTC)
50	150
125	200
250	250
300	300
325	350

These points are graphed in Figure VI as the long-run total cost (LTC) curve. The points A, B, C, D and E correspond to the equilibrium points in Figure V. Note that the LTC curve at first increases at a decreasing rate, then at a constant rate, and finally at an increasing rate. The LTC curve starts from the origin implying thereby that in the long-run all costs are variable and if nothing is produced, no resources will be used (i.e., the firm will quit the industry altogether). Thus, the LTC curve is analogous to the short-run VC curve. Only difference is, while the shape of VC is due to the law of variable proportions in the short-run, the shape of LTC is due to the existence of increasing, constant, and decreasing returns to scale in the long-run.

Figure VI



Unit Costs in the Long-Run

In the long-run, costs are not divided into fixed and variable components; all costs are variable. Thus, the only long-run unit cost functions of interest are long-run average cost (LAC) and long-run marginal cost (LMC). These are defined as follows :

$$LAC = \frac{LTC}{Q}$$

$$LMC = \frac{\Delta LTC}{\Delta Q}$$

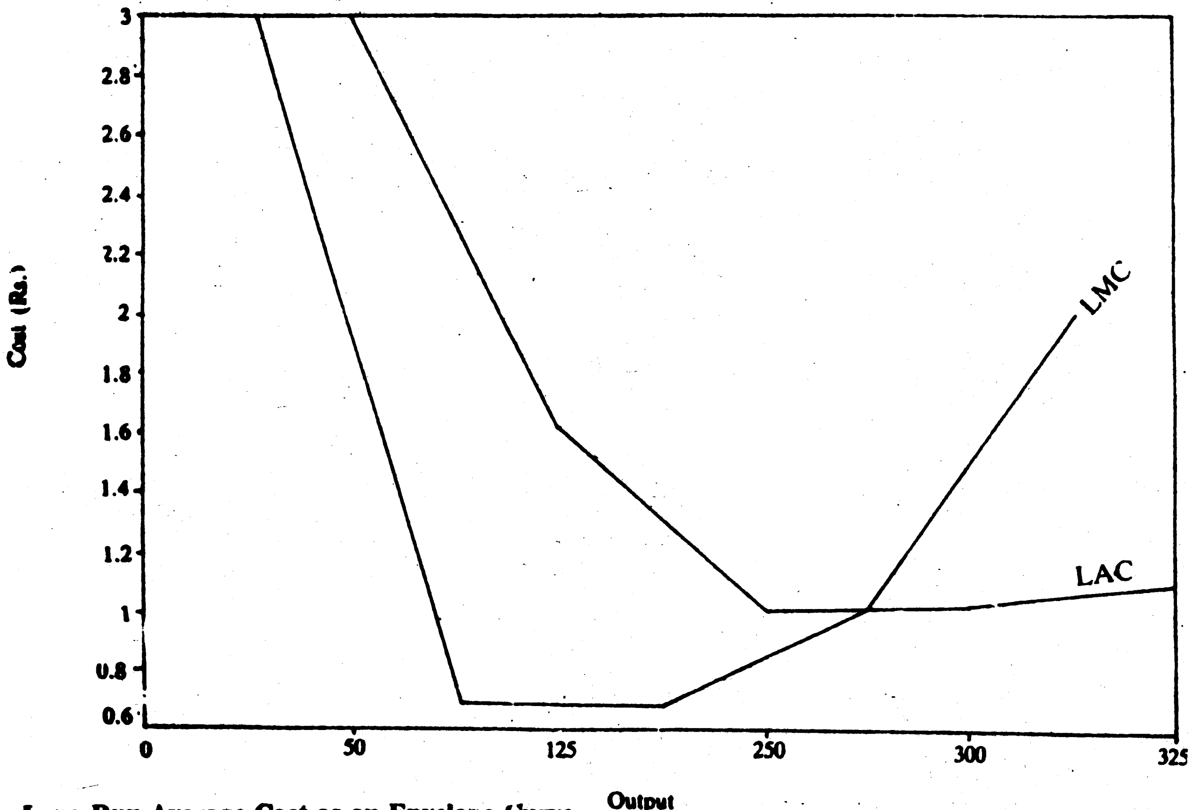
$$LMC = \frac{d(LTC)}{dQ}$$

For the long-run total cost given in Figure VI, these unit costs can be presented in tabular form as follows :

Output Q	Long Run Total Cost (LTC)	Long Run Average Cost (LAC)	Long Run Marginal Cost (LMC)
0	0	-	-
50	150	3.00	3.00
125	200	1.60	0.67
250	250	1.00	0.67
300	300	1.00	1.00
325	350	1.08	2.00

These LAC and LMC and values are graphed in Figure VII. We see, both in the table and in the graph, that LAC and LMC are U-shaped and that they are equal at the minimum of LAC. The values of LMC are graphed at the midpoints of the output intervals they represent.

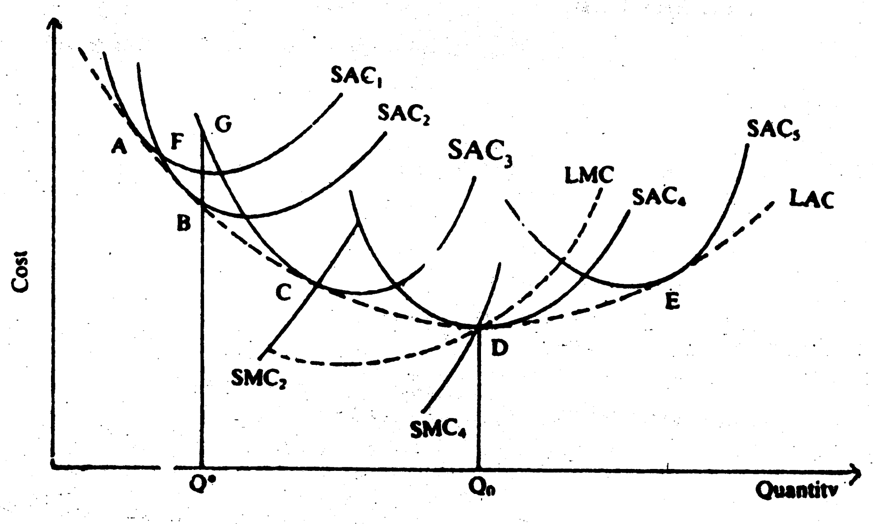
Figure VII : LONG-RUN AVERAGE & MARGINAL COSTS



Long-Run Average Cost as an Envelope Curve

The long-run average cost curve is sometimes shown as the envelope curve of series of all tl. possible short-run average cost curves, as shown in Figure VIII. Five short-run average cost curves, each representing a different sized plant (or set of fixed factors) are illustrated, although many more may exist. The long-run ATC curve shows the least per unit cost at any output can be produced after the firm has had time to make all appropriate adjustments in its plant size. Consider, for example, the production of Q* units in Figure VII. That the level of output could be produced with the plant sizes represented by SAC_1 , SAC_2 or SAC_3. It represents the optimum rate of output for the plant size represented by SAC_1 (i.e., it is at the minimum point of SAC). However, if the firm expects to produce at the rate, the best size of plant is the one related to SAC_2. Q* units could be produced at a cost savings of FB per unit over SAC_2.

FIGURE VIII LONG-RUN AVERAGE COST AS AN ENVELOPE CURVE



The plant size associated with SAC_2 is the optimum size plant because its minimum point is the lowest of all possible unit costs. Given the LAC in Figure VIII, we can say that there are increasing returns to scale (or economies of scale) up to Q_0 and decreasing returns to scale (or diseconomies of scale) beyond Q_0. The firm's LAC curve is often called the 'firm's planning curve'.

Activity 4

a) Why are all costs variable in the long-run?

....

....

b) Why is the LAC called an "envelope curve"?
Why cannot the LMC be an envelope as well?

....

....

....

c) What do you understand by "cost-efficiency"?
Draw a long-run cost diagram and explain.

9.6 ECONOMIES OF SCALE

It has been shown in the preceding section that a larger plant will lead to lower per unit costs in the long-run. However, beyond some point, successively larger plants will mean higher average costs. Exactly, why is the long-run ATC curve U-shaped? This point needs further explanation.

It must be emphasised, first of all, that the law of diminishing return is not applicable here; for it presumes that one resource is fixed in supply and, as we have seen, the long-run assumes that all resources are variable. Also, we assume that resource prices are constant. What then, is our explanation? The U-shaped long-run average cost curve is explainable in terms of what economists call "economies and diseconomies" of large scale production.

Economies and diseconomies of scale are concerned with the behaviour of average cost curve as the plant size is increased. Economies of scale explain the downslopping part of the long-run AC curve. As the size of a plant increases, LAC typically declines over some range of output for a number of reasons. The most important is that, as the scale of output is expanded, there is greater potential for specialisation of productive factors. This is most notable with regard to labour but may apply to other factors as well. Other factors contributing to declining LAC include ability to use more advanced technologies and more sophisticated capital equipment; managerial specialisation; opportunity to take advantage of lower costs for some inputs by purchasing larger quantities; effective utilisation of by-products, etc.

But, after sometime, expansion of a firm's output may give rise to diseconomies, and therefore, higher per unit costs. Further expansion of output beyond a reasonable level may lead to problems of over crowding of labour, managerial inefficiencies, etc., pushing up per unit costs.

All these are examples of internal economies and diseconomies of scale arising due to the firm's own expansion. According to Marshall, external economies and diseconomies of scale may arise due to the expansion of industry as a whole. For example, improved infrastructure facilities due to industrial expansion may lead to reduction in per unit costs of production in all the firms in an industry.

9.7 ECONOMIES OF SCOPE

... concept is a recent development. This stands in contrast to the concept of 'economies of scale'. The basic argument is that cost-efficiency in production process has brought about by, variety rather than volume.

According to the concept of economies of scale, cost advantages follow from the increase in volume of production or what is called the scale of output. According to the concept of economies of scope, such cost advantages may follow from variety of output e.g. product diversification within the given scale of plant. If the same plant can produce multiple products, there is the scope for a lot of cost savings because of joint utilisation of inputs. The Government of India has announced the 'broadbanding policy' such that Escorts may now produce in addition to its two-wheeler, four-wheeler using the same plant and technology with slight modifications or adjustments. Broadbanding, therefore, should enable manufacturers to exploit economies of scope through product diversification. In other words, instead of increasing the scale of production of an existing product, the firm can now add new and newer products, if the size of plant and type of technology offer that scope; in this process, the firms will have access to scope-economies in place of scale-economies. In certain processes, the firm can plan wisely to exploit both types of economies simultaneously.

Activity 5

a) Distinguish between internal and external economies of scale. Give examples.

....

....

....

....

....

b) Distinguish between scale-economies and scope-economies. Give examples.

....

....

....

....

....

9.8 COST ELASTICITY AND OUTPUT ELASTICITY

When we take costs determined by output, we can talk about output elasticity of costs, e_q^c. On the other hand, when we find output being determined by costs, we can talk of costs-elasticity of output, e_c^q.

Examining these elasticities, you should be in a position to discover the relationship between average and marginal costs.

$$e_q^c = \frac{dC}{dQ} \cdot \frac{Q}{C} = \frac{MC}{AC}$$

$$e_c^q = \frac{dQ}{dC} \cdot \frac{C}{Q} = \frac{AC}{MC}$$

In other words, given the average costs (AC) and marginal costs (MC); we can derive these elasticities; or given cost elasticity and marginal costs (or average costs) we can derive average costs (or marginal costs).

The economic interpretation of cost elasticity is the usual one i.e., the proportionate change in the level of output resulting from the proportionate change in costs of production.

Activity 6

a) How is production related to costs?

....

....

....

b) Given $C = 45Q^2$, calculate the following for $Q = 2$

AC = MC = e_q^c = e_c^q =

c) In the short-run, the output elasticity of average fixed costs is equal to unity. True or False. Explain.

....

....

....

d) The examination of cost-data of a particular firm shows that in a year, following upward revision of wages the cost per unit of output produced by that very labour has come down. How can that be? Explain with the help of a diagram.

....

....

....

....

....

....

....

....

....

....

....

....

e) Identify the time perspective relevant for each of the cost function noted below :

Function	Time Perspective
i) C = 100 + 2Q	....
ii) C = 100 Q	....
iii) C = 100 Q² - Q + Q°	....

9.9 FORMS OF THE COST FUNCTION

Costs are related to production. Just as we have talked about the form of production function in the previous unit, we may now talk about the form of cost function. There are four alternative forms of total cost (TC) function which can be estimated by econometric method, and accordingly we can deduce the shape of average costs (AC) and marginal costs (MC) curves under different values of the coefficient of each form of TC function. This is explained categorically in what follows. An understanding of the nature of cost function is basic to the interpretation of empirical cost situation and the designing of managerial decision based thereupon.

i) Linear Total Cost Function

$TC = a_1 + b_1 Q$ .... Where a_1 and b_1, are parametric terms, and Q is the level of output.

$$\Rightarrow AC = \frac{a_1}{Q} + b_1$$

$$\frac{d(AC)}{dQ} = -\frac{a_1}{Q^2}$$

$$\frac{d^2(AC)}{dQ^2} = -\frac{2a_1}{Q^3}$$

The implication is that AC will be a falling curve for $a_1 > 0$ and a rising curve for $a_1 < 0$. The marginal cost function in this case will be

$$MC = \frac{d(TC)}{dQ} = b_1. \qquad \Rightarrow \text{constant MC Curve}$$

ii) Quadratic Total Cost Function

$$TC = a_2 + b_2Q + c_2Q^2$$

$$\Rightarrow AC = \frac{a_2}{Q} + b_2 + c_2Q$$

$$\frac{d(AC)}{dQ} = -\frac{a_2}{Q^2} + c_2$$

$$\frac{d^2(AC)}{dQ^2} = -\frac{2a_2}{Q^3}$$

You may note that AC curve will be falling, if

$$\left[-\frac{a^2}{Q^2} + C_2\right] < 0, \text{ i.e., } Q < \sqrt{\frac{a^2}{c_2}}$$

It will be constant for $Q = \sqrt{\frac{a^2}{c_2}}$

and it will be rising for

$$Q > \sqrt{\frac{a_2}{c_2}}$$

The marginal cost curve in this case will be

$$MC = \frac{d(TC)}{dQ} = b_2 + 2c_2Q$$

$$\frac{d(MC)}{dQ} = \frac{d^2(TC)}{dQ^2} = 2c_2$$

iii) Cubic Total Cost Function

$$TC = a_3 + b_3Q + c_3Q^2 + d_3Q^3$$

$$\Rightarrow AC = \frac{a_3}{Q} + b_3 + c_3Q + d_3Q^2$$

$$MC = b_3 + 2c_3Q + 3d_3Q^2$$

iv) Double-Log Total Cost Function

$$TC = a_4Q^{(b_4)}$$

$$\Rightarrow \log TC = \log a_4 + b_4 \log Q$$

$$AC = a_4Q^{(b_4-1)}$$

$$MC = a_4b_4Q^{(b_4-1)}$$

Activity 7

a) Work out the following in the last two cases (iii) & (iv) :

Case (iii)	Case (iv)
$\frac{d(AC)}{dQ} =$	
$\frac{d^2(AC)}{dQ^2} =$	
$\frac{d(MC)}{dQ^2} =$	
$\frac{d^2(MC)}{dQ^2} =$	

b) Under which case/cases above and why will you get the U-shaped AC curve ?

....

....

....

....

....

9.10 APPLICATION OF COST ANALYSIS

The cost concepts and analyses that you have been exposed to in this unit has a very wide range of application for managerial use. In what follows, we intend to illustrate a few specific uses.

Optimum output level

You may recall that optimum has reference to that level (size) of output which minimised the average cost of product alternatively, for which average cost equals marginal costs. Consider the following exam,

$$TC = 128 + 6Q + 2Q^2$$

$$\Rightarrow \quad AC = \frac{128}{Q} + 6 + 2Q$$

$$\frac{d(AC)}{dQ} = \left[-\frac{128}{Q^2} + 2\right] = 0 \Rightarrow Q = 8$$

or, MC = $6 + 4Q$.

Setting AC = MC

$$\left[\frac{128}{Q} + 6 + 2Q\right] = [6 = 4Q] \Rightarrow Q = 8$$

Thus [Q=8] is the optimum level of output (for the short-run). Note that it is a short-run case, because TFC = 128.

Optimum inventory level

All productions are not immediately sold. Sometimes we have to build up inventory or stock of saleable items. In this case, the optimum inventory level may be defined as that for which the average cost of inventory held is at the minimum. In this, there are two types of costs involved ; Carrying costs ana and Reorder costs. The inventory **carrying costs** include storage costs, interest costs on borrowed capital to finance stock etc. On the other hand the **reorder costs** include book keeping costs and telephone charges (i.e., fixed costs); some of the reorder costs also vary with the size of the order placed (i.e., variable costs). Suppose we use the following notations :

S = Expected sale

D = Order quantity to be delivered

S/D = Number of orders delivered

F = Average fixed costs of delivery

V = Coefficient of average variable cost of reorder

K = Average carrying costs

D/2 = Average inventory held beween initial and terminal periods and it is assumed that the demand is spread evenly.

The average costs of inventory held can now be stated explicitly as :

$$AC = \left[K.\frac{D}{2}\right] + [F.V.D.]\frac{S}{D}$$

Carrying costs + Reorder cost.

For deciding on optimum inventory held,

$$\frac{d(AC)}{dD} = \left[\frac{K}{2} - \frac{FS}{D^2}\right] = 0 \Rightarrow D = \sqrt{\frac{2FS}{K}}$$

This is very a very significant derivation. This formula gives us the optimum size of stock or what is called Economic Order Quantity (EOQ).

In other courses, you will be introduced to more sophisticated inventory models. The simple inventory model that we have used here, nevertheless, brings home, some interesting results. For example, inventory should increase in proportion to the square root of expected sales. Also note that variable costs have got no impact on inventory decisions.

Break-even output

This is basically an accounting concept. The break-even chart of the accountant illustrates at what level of output in the short-run, the total revenue just covers the total costs. Let us operate with our familiar algebraic notations :

$$TR = P \cdot Q$$

$$TC = F + VQ$$

At break even point,

$$TR = TC$$

$$\Rightarrow P \cdot Q = F + VQ$$

$$\text{or } Q = \frac{F}{P-V} = \frac{\textit{Total Fixed costs}}{\textit{Price} - \textit{Variable costs per unit}}$$

Here Q stands for **Break-even volume of output**; multiplying this with price (P) we get the **Break-even value of output.** Take an example to illustrate. Suppose

$$TR = 60\,Q$$

$$TC = 1800 + 40\,Q$$

Then the break-even output volume $= 90\left[= \frac{1800}{60-40}\right]$

and the break-even output value = 5400[= 90 × 60]

also note that [P - V] = 20 is called 'contribution margin' per unit of output.

Activity 8

a) Speed-Marine Co. builds motor boat engines. They recently estimated their total costs and total revenue as :

$$TC = 80{,}000 - 600\,Q + 2\,Q^2$$

$$TR = 400\,Q - Q^2$$

(Where Q is the number of engines produced each year)

At what level of production will Speed-Marine co. break-even? How many engines should be produced to maximise profit?

....

....

....

....

....

....

b) Given LTC = $6Q + 2Q^2 - Q^3$, find out the optimum level of output, Q.

c) Write a lucid note on the types of inventory costs, giving suitable examples.

d) Why does not the variable (reorder) cost affect the economic order quantity?

e) List the factors which determine the size of optimum buffer stock of foodgrains held by the Food Corporation of India.

....

....

....

9.11 SUMMARY

In this unit, we have emphasised the critical role that costs play in determining the profitability of the firm. The profit-oriented firm manager must consider both opportunity costs and explicit costs in order to use all the resources most economically. Although it is difficult to have totally accurate information on its costs, a firm should have reliable estimates of its fixed costs, how its costs vary with respect to output over the relevant range of production, and of whether or not its costs would be lower with a bigger size plant.

In the short-run, the total cost consists of fixed and variable costs. Variable cost, and thus the total cost, typically increases first at a decreasing rate and then at an increasing rate. The reason for this type of behaviour of variable and total cost is related to the principal of diminishing marginal productivity. If marginal product increases at first, the marginal cost at first declines (that is, the rate of increase in variable and total cost declines). However, once marginal product begins to decline, marginal cost increases. It follows that in the short-run, marginal cost (MC), average variable cost (AVC), or average total cost (AC) curves are generally U-shaped.

The long-run is defined as the period for which all factors of production are variable. Thus, in the long-run total costs are identical to variable costs. The long-run total cost curve may also at first increase at a decreasing rate, and then increase at an increasing rate. If the LTC is of normal shape (inverted S-shaped), the long-run average cost (LTC) will also have a U-shape. The long-run cost function may be derived from the series of most efficient short-run cost curves or from an expansion path. Economies or diseconomies of scale arise either due to the internal factors pertaining to the expansion of output by a firm, or due to the external factors such as industry expansion.

In contrast, economies of scope result from product diversification. Thus the scale-economies have reference to an increase in volume of production, whereas the scope-economies have reference to an improvement in the variety of products from the existing plant and equipment. These cost concepts and analyses have a lot of applications in real world decision-making situation such as optimum output, optimum product-mix, optimum inventory, break-even output etc.

9.12 ADDITIONAL READINGS

Seo, K.K., 1988. *Managerial Economics* (6th ed.) Surjeet Publications: Delhi.

Adhikary, M. 1987. *Managerial Economics* (3rd ed.) Khosla Educational Publishers : Delhi, Ch. V.

Goldhar, J.D. & M. Jelinek, 1983. "Plan for Economics of Scope", in *Harvard Business Review*, December.

9.13 SELF-ASSESSMENT TEST

1. Make a list of relevant Cost-concepts from the standpoint of an (i) accountant, (ii) economist, and (iii) engineer. Take any construction company as a reference to illustrate the cost-concepts.

2. Your Management is interested in **cost reductions and control**. How would you go about it. Recommend some steps which would yield results in

   a) short-run
   b) long-run

3. Explain clearly the managerial use of the concepts like 'economies of scale' and 'economies of scope'. Is it possible to reap both the economies simultaneously? Give examples.

4. Draw and explain the effect of internal and external economies/internal and external diseconomies on the firm's cost curves.

5. The total cost function of a neckwear manufacturing company is estimated as $TC = 8 + 12Q - 6Q^2 + Q^3$ (where Q represents hundreds of ties produced per month)

   Derive equation for AFC, AVC, ATC, MC and plot all these costs in a graph and comment on the relationship between average and marginal costs.

6. A fixed plant is used to manufacture TV sets. If x sets are turned out per week, the total variable costs is Rs. $\left[3x + \frac{1}{25}x^2\right]$. Show that the average variable costs rise steadily with output. What do you expect about marginal costs of production?

7. Given production function.

   $$Q = AL^{\alpha}K^{1-\alpha}$$

   derive the cost function. State your assumptions clearly at the end.

   [Hint : It is a difficult question. You may recall the cost equation, $C = P_L L + P_K K$ and the optimum decision rule.

   $$\frac{P_L}{P_K} = \frac{MP_L}{MP_K}$$

Apply these to get $C = C(Q, P_L, P_K, \alpha\ 1 - a)$. Have patience; you should be able to do it.]

UNIT 10 EMPIRICAL ESTIMATES OF PRODUCTION AND COSTS

Structure

10.0 INTRODUCTION

For decision-making purposes, a manager should understand clearly the relationship between his inputs and outputs on the one hand, and the output and costs on the other. Estimation of the industry production function may help a manager in taking correct decisions of long-term nature such as capital expenditure. The short-run production estimates at firm level are helpful in arriving at the optimal mix of inputs to achieve a particular output target. This is referred to as the least cost combination of inputs in production analysis. Also, for a given cost, optimum level of output can be found if the production function of a firm is known.

Cost estimation for decision-making is concerned with finding the shape and placement of the firm's cost curves. Both, the short-run cost function and the long-run cost function must be estimated, since both sets of information will be required for some decisions. Knowledge of the short-run cost functions allows the decision makers to judge the optimality of present output levels and to solve decision problems using contribution analysis. We saw in the preceding unit that the concept of incremental cost is fundamental to short-run decision-making on cost issues. Knowledge of long-run cost functions is important when considering the expansion or contraction of plant size, and for confirming that the present plant size is optimal for the output level that is being produced.

In this unit, we shall discuss different approaches to production and cost estimation.

10.1 OBJECTIVES

On reading this unit, you should be able to :

- relate theory to empiricism
- get a summary-view of various approaches to estimation : economic (econometric), accounting and engineering
- find some statistical estimates of production and cost functions
- understand both conceptual and statistical difficulties related to measurement of such functions
- throw light on empirical determination of production and cost functions; and finally

- Identify managerial use of such empirical estimates.

10.2 SOME ESTIMATES OF PRODUCTION FUNCTION

The production function expresses an engineering relationship between physical inputs and outputs. By definition, the production function denotes an efficient combination of inputs and outputs. As explained earlier, there can be more than one combination of inputs, all of them technically efficient, which can be employed to produce a particular output level. However, for decision-making at the firm level, we use economic, and not technical, criteria to find an optimum input-output mix. It is for this reason that estimation of production function using other than engineering techniques become relevant.

Production function can be estimated by statistical techniques using historical data on inputs and output. One can hypothesise several alternative forms i.e., linear, quadratic, cubic etc., for this function. You may recall the various forms of production function, mentioned in earlier units. However, empirical studies on the subject have found the Cobb-Douglas (named after Cobb and Douglas who first used it for American manufacturing sector) form to be the most appropriate.

The Cobb-Douglas production function, in its general form, is expressed as :

$$Q = A\,K^a\,L^b$$

where

Q = Total output

A = Constant (or technology factor)

K = Capital Input

L = Labour Input

a,b = exponents of Capital and Labour, respectively

The above production function can also be represented in log-linear form :

$$\text{Log } Q = \text{Log } A + a \text{ Log } K + b \text{ Log } L$$

Various factor intensities, factor productivities and factor elasticities can be easily computed from an estimated production function of Cobb-Douglas type. It can be easily inferred from the above form (log-linear form) that exponents a and b are factor elasticities. Also, by adding up the factor elasticities, a measure of returns to scale can be had.

Cobb and Douglas estimated a production function for the American manufacturing sector using annual time series data for the period 1899 to 1922. Their estimated production function was :

$$X = 1.01\,L^{0.75}\,K^{0.25}$$

Their findings suggested that the American manufacturing sector was facing constant returns to scale during the above mentioned period. The elasticity of output with respect to labour was 0.75, and that with respect to capital was 0.25. This implied that a 10 per cent increase in labour input with no change in capital caused a 7.5 per cent increase in output, and a similar increase in capital input with no change in labour brought about a 2.5 per cent increase in output. The above estimated production function could also be used to generate isoquants for various levels of output, and the least cost input combination for a given output.

Empirical estimate of production functions for industries such as sugar, textiles, cement etc., are available in the Indian context.

Indian Cement Industry

One study estimates the production function facing the Indian Cement industry using time series annual data for the period 1946-1965. [See Gupta, G.S, Production Function and Factor Productivity in the Indian Cement Industry. *Indian Journal of Industrial Relations,* January 1973.] The data in value terms on inputs and output were used for estimation. He hypothesised the following forms of production function :

1) $V = A_1 W^{\alpha 1} K^{\beta 1}$

2) $V = A_2 W^{\alpha 2} K^{\beta 2} M^{\gamma 2}$

Where,

V = Value added at current prices

W = wages, salaries and benefits

K = fixed capital at current prices

M = raw material, fuel, electricity, etc., consumed at current prices

A_1 and A_2 are intercepts, and α, β, γ are value elasticities of labour, capital and raw material, respectively.

The production function estimates are given below :

1) $Q = 1.22\, W^{0.995} \overline{K}^{0.1446}$ ....$R^2 = 0.938$

2) $Q = 1.14\, W^{0.177} K^{0.039} M^{0.075}$....$R^2 = 0.996$

The cement industry depicts diminishing to constant returns to scale $(\alpha + \beta + \gamma \leqslant 1)$ during the period 1946-65. The exponent of labour is positive, as expected. The value elasticity of capital is negative in the first form, and quite low in the second form using raw material as a separate factor. The negative or low, and statistically insignificant elasticity of capital points towards the old technology and dilapidated machinery in the cement industry during the sample period. The high labour elasticity acted as a disincentive to new investment in the cement industry during the above period. The explanatory power (R) of the model is quite high i.e. 0.996.

Indian Sugar Industry

In another study, G.S. Gupta and K. Patel have estimated the production function in Indian Sugar Industry (see *Managerial Economics* by M. Adhikary for a summary of this study). The basic findings of this study are summarised in the following table :

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

Region	Sample period	Regression co-efficients & other statistics					
		Log A	α	β	$\alpha + \beta$	R^2	DW
All-India	1946-66	-1.92	0.91	0.29	1.21	.93	1.95
	1946-58	-1.79	0.92	0.23	1.15	.84	2.27
	1959-66	-5.33	1.41	0.45	1.86	.97	1.32
Uttar Pradesh	1946-66	0.42	0.43	0.28	0.71	.73	2.00
	1946-58	0.72	0.34	0.32	0.66	.53	1.96
	1959-66	-3.15	1.21	0.26	1.47	.83	2.53
Bihar	1946-66	-1.94	0.99	0.12	1.11	.60	1.52
	1946-58	-1.15	0.41	0.50	0.91	.56	0.98
	1959-66	-3.60	1.27	0.30	1.57	.63	2.70

Indian Cotton Textile Industry

The National Productivity Council (NPC) has estimated a Cobb Douglas type of production function for cotton textile output (X) for 1961 and 1966

1) 1961 - $X = 0.0438\, L^{1.054}\, K^{0.018}$

2) 1966 - $X = 0.0039\, L^{1.044}\, K^{0.036}$

The shows that over 1961-66, textile production has become more capital-dependent.

Activity 1

a) Point out your observations on :

   i) Inter-regional variations in factor elasticities.

   ii) Inter-temporal variations in labour elasticity.

   iii) Inter-temporal variation in capital elasticity.

b) What does $(\alpha + \beta)$ show? Would you agree that our Sugar Industry in general, operated under increasing return to scale ?

c) Can you infer anything about relative factor productivity in Sugar Industry? Give reasons ror your answer.

....

....

....

d) What do R^2 and DW show?

....

....

....

....

....

Activity 2

a) Get back to various empirical estimates of production function that have been reported earlier. Could you now think of some limitations from which those estimates might be suffereing? Be specific to the estimates pertaining to :

   i) Indian Cement Industry ?

....

....

....

....

   ii) Indian Sugar Industry

....

....

....

....

   iii) Indian Cotton Textile Industry

....

....

....

....

   iv) Look at some latest of Indian Economic Review or other economic journals. Report on some of the recent estimates of production function.

....

....

....

....

....

....

10.3 EMPIRICAL DETERMINATION OF PRODUCTION FUNCTION

Empirical estimates of a production function typically use one of the following statistical approaches :

i) Time series analysis
ii) Cross-sectional analysis
iii) Engineering analysis

Time series analysis is used in identifying the production function of the firm, and data on output and input is obtained relating it to different points of time. Analysis of time series data is appropriate for a single firm that has not undergone significant changes in technology during the time span analysed. To develop a production function for an industry is a different problem. Even if all firms have operated over the same time span, changes in capacity, inputs and outputs may have proceeded at a different pace for each firm. Thus a cross-sectional analysis may be more appropriate.

Cross-sectional analysis deals with data collected at one particular time. The data thus might be linked to a snapshot, frozen in time. Instead of making observations of the variables for each year, the variables are observed for each firm in an industry in a particular year.

Engineering analysis is undertaken when the above two types do not suffice. When good historical data is difficult to obtain or is not available, it may be possible for engineers or agricultural scientists to develop data from controlled experiments or day-to-day working experience. This data is concerned with what the inputs and output ought to be with most efficient combination of raw-materials, labour, capital etc.

Each of the methods outlined above suffers from certain limitations ;

1. Time series and cross-sectional analyses are restricted to a relatively narrow range of observed values. Extrapolation of the production function outside that range may be seriously misleading. For example, in a given case, marginal productivity might decrease rapidly above 85% capacity utilisation; the production function derived for values in the 70% - 85% capacity utilisation range would not show this.

2. Another weakness of time series analysis is the assumption that all observed values of the variables pertain to one and the same production function. In other words, a constant technology is assumed. In reality, most firms or industries, however, find better, faster, or cheaper ways of producing their output. As their technology changes, they are actually creating new production functions. One way of coping with such technological changes is to make it one of the independent variables.

3. Theoretically, the production function includes only efficient (least cost) combinations of inputs. If measurements are to conform to this concept, any year in which the production was less than nominal would have to be excluded from the data. It is very difficult to find a time-series data which satisfy techincal efficiency criteria as a normal case.

4. Engineering data may overcome the limitations of time series data but mostly they concentrate on manufacturing activities. Engineering data does not tell us anything about the firm's marketing or financial activities, even though these activities may directly affect production.

5. In addition, there are both conceptual and statistical in measuring data on inputs and outputs.

It may be possible to measure output directly in physical units such as tons of coal, steel etc. In case more than one product is being produced, one may compute the weighed average of output, the weights being given by the cost of manufacturing these products. In a highly diversified manufacturing unit, there may be no alternative but to use the series of output values, corrected for changes in the price of products. One has also to choose between 'gross value' and 'net value'. It seems better to use 'net value added' concept instead of output concept in estimating production function particularly where raw-material intensity is high.

The data on labour is mostly available in the form of "number of workers employed" or "hours of labour employed". The 'number of workers' data should not be used because, it may not reflect underemployment of labour, they may be occupied, but not productively employed. Even if we use 'manhours' data, it should be adjusted for efficiency factor. It is also advisable that labour should not be measured in money terms as given by expenditure on wages, bonus etc.

The data on capital input has always posed serious problems. Net investment i.e. a change in the value of capital stock, is considered most appropriate. Nevertheless, there are problems of measuring depreciation in fixed capital, changes in quality of fixed capital, changes in inventory valuation, changes in composition and productivity of working capital etc.

Finally, when one attempts an econometric estimate of a production function, one has to overcome the standard problem of multi-collinearity among inputs, autocorrelation, homosedhasticity etc.

10.4 APPROACHES TOWARDS ESTIMATING OF COST FUNCTION

In the preceding unit, we have explained in detail the nature of cost-output relationship. It was shown that the total cost varies directly with the level of output. The knowledge of how the total cost, average cost, and marginal cost behave with the level of output is useful for decision-making. The decisions such as : whether to increase the output, whether plant is having optimum size, would largely depend on the exact amount by which the total cost increases with increases in output from one particular level to another, and the resultant shape of the average cost curve as the plant size is increased.

Thus, a firm should know the cost function facing it. For future planning at firm level also, knowledge of its cost function is absolutely essential. Although the exact future cost-output relationship may r ot be available until the firm really goes for expansion of its output, there are methods through which a firm could get approximate information of its future cost function.

The cost output relationship can be estimated by using the following three methods :

* Engineering Method
* Accounting Method
* Statistical Method

Engineering Method

The engineering method of cost estimation is based directly on the physical relationship of inputs to output, and uses the price of inputs to determine costs. This method of estimating real world cost functions rests clearly on the knowledge that the shape of any cost function is dependent on:

i) The production function, and
ii) The price of inputs

Given the production function and input prices, the optimum input combination for a given output level can be determined (as discussed in the preceding unit). The resultant cost curve can then be formulated by multiplying each input in the least cost combination by its price, to develop the cost function. This method is called engineering method as the estimates of least cost combinations are provided by engineers.

The assumption made while using this a method is that both the technology and factor prices are constant. This method may not always give the correct estimate of costs as the technology and factor prices do change substantially over time. Therefore, this method is more relevant for the short-run.

Accounting Method

The cost-output relationship, under the accounting method, is estimated by dividing the total cost into fixed, variable and semi-variable costs. All these components are then separately estimated. The average variable cost, the semi-variable cost which is fixed over a certain range of output, and fixed costs are determined on the basis of inspection and experience. The total cost and the average and marginal costs for each level of output can then be obtained through a simple arithmetic procedure.

Although the accounting method appears to be quite simple, it is a bit cumbersome as one has to maintain a detailed breakdown of costs over a period to arrive at good estimates of actual cost-output relationship. One must have experience with a wide range of fluctuations in output rate to come up with accurate estimates.

Statistical Method

The most widely adopted method of determining cost function is the statistical estimation of the relationship between cost and output. Under the statistical method, the historical data on cost and output are used to estimate the cost-output relationship. The basic technique of regression is used for this purpose. The data could be a time series data of a firm in the industry or of all firms in the industry or a cross-section data for a particular year from various firms in the industry.

However, depending on the kind of data used, we can estimate short-run or long-run cost functions. For instance, if time series data of a firm whose output capacity has not changed much during the sample period is used, the cost function will be a short-run. On the other hand, if cross-section data of many firms with varying sizes, or the time series data of the industry as a whole is used, the estimated cost function will be the long-run one.

The estimation procedure involves three steps. First, the determinants of cost are identified. Second, the functional form of cost function is specified. Third, the least square method is applied to estimate the chosen form.

The determinants of cost include output, composition of output in the case of multiple products, scale of operation, factor productivities including technology, and factor prices. In general form, the total cost function can be hypothesised as:

$$C = f(Q,X)$$

where, C = Total costs

Q = Level of output

X = 'Other determinants of cost'

f = Unspecified function

You may recall from the earlier unit that the following four functional forms can be considered for estimating the cost function:

i) Linear $C = a_o + bQ + C_iX_i$

ii) Quadratic $C = a_o + b_1Q + b_2Q^2 + C_iX_i$

iii) Cubic $C = a_o + b_1Q + b_2Q^2 + b^2Q^2 + C_iX_i$

iv) Double Log $C = Log \quad a_o + b_1 \quad log \quad Q + C_i \quad log \quad X_i$

With the help of calculus and algebraic manipulations, it can be shown that the linear total cost function would give a constant marginal cost and a monotonically falling average cost curve. The quadratic function would yield a U-shaped average cost curve, and a monotonically rising marginal cost curve. The cubic function would result into a U-shaped AC curve and a U-shaped MC curve. The double log function would imply a falling (or rising) AC and a falling (or rising) MC curve. Thus, to check the validity of theoretical cost-output relationship, one should hypothesis a cubic cost function.

With all the determinants of cost and the functional forms, the data collection is complete. The alternative functional forms can be estimated by using the regression package on the computer. The most appropriate form of the cost function for decision making is then chosen on the basis of the principles of economic theory and statistical inference.

Problem of Measurement

A manager while undertaking the statistical cost analysis does confront some serious problems. First, in collecting the cost and output data, he must be certain that they are properly paired, that is, the cost data is, in fact, applicable to the corresponding data on output. Second, he must also try to obtain data on cost and output during a time period when the output has been produced at relatively even rate. If, for example, a month is chosen as the relevant time period over which the variables are measured, it would not be desirable to have wide weekly fluctuations in the rate of output. The monthly data in such a case would represent an average output rate that could disguise the true cost-output relationship. Not only should the output rate be uniform, but it also should be a rate to which the firm is fully adjusted. Furthermore, there should be no disruptions in the output due to external factors such as power failures, delays in receiving necessary supplies, etc. To generate the data necessary for a meaningful statistical analysis, the observations must include a wide range of rates of output. Observing cost-output data for the last 24 months, when the rate of output was the same each month, would provide little information concerning the appropriate cost function.

Third, the cost data is normally collected and recorded by accountants for their own purposes and in a manner that it makes the information less than perfect from the perspective of economic analysis. While collecting historical data on cost, care must be taken to ensure that all explicit as well as implicit costs have been properly taken into account, and that all the costs are properly identified by time period in which they were incurred.

Finally, for situations in which more than one product is being produced with given productive factors, it may not be possible to separate costs according to output in a meaningful way. One simple approach of allocating costs among various products is based on the relative proportion of each product in the total output. However, this may not always accurately reflect the cost appropriate to each output.

10.5 SOME ESTIMATES OF COST FUNCTION

Long-run average cost functions for 29 Indian manufacturing industries have been estimated by V.K. Gupta using statistical techniques. His estimates have been presented in Table 1. Students may verify the shape of the average and marginal cost curves by using their knowledge of economic theory.

Table 1: Long-Run Average Cost Functions for 29 Manufacturing Industries in India

INDUSTRY	*AVERAGE COST FUNCTION*	*SHAPE*
Sewing Machines	AC=76.6+139/Q+.0003Q	L+
Starch	AC=76.0+413/Q+.02Q	U
Electric Lamps	AC=80.4+224/Q-.005Q	L-
Electric Fans	AC=81.9+302/Q	L
Soap	AC=92.2+130/Q-.014Q	L-
Iron and Steel	AC=92.6-.003Q	S-
Woollen Textiles	AC=89.5+140/Q-.003Q	L-
Bicycles	AC=82.5+65/Q-.003Q	L-
Matches	AC=71.4+876/Q + .004Q	U-
Paints and Varnishes	AC=82.0-.011Q	S-
Paper and Paperboard	AC=73.6+963/Q	L
Fruit and Vegetable Processing	AC=82.0+46/Q+.004Q	L+
Cement	AC=65.6+5622/Q+.005Q	U-
Ceramics	AC=78.4+348/Q	L
Biscuit Making	AC=89.2+143/Q-.031Q	L-
Plywood and Tea Chests	AC=86.1+151/Q-.026Q	L-
Vegetable Oil (edible)	AC=93.0+207/Q-.009Q	L-
Vegetable Oil (not edible)	AC=100-.013Q	S-
Aluminium, Copper, and Brass	AC=88.6-.001Q	S
Distilleries and Breweries	AC=67.8+349/Q+.909Q	U-
Chemicals	AC=73.9-.0002Q	S
Glass and Glassware	AC=89.2+130/Q	L
Tanning	AC=95.7+83/Q-.009Q	L-
Wheat Flour	AC=91.9+54/Q-.003Q	L-
Cotton Textiles	AC=89.7+273/Q	L
Jute Textiles	AC=94.4-.0003Q	S
General and Electrical Engineering	AC=84.4+49/Q+.002Q	L-
Sugar	AC=85.0+124/Q+.001Q	L+
Rice Milling	AC=86.4+151/Q+0.38Q	U

Note : AC = average cost
Q = output
U = U-shaped
U— = U-shaped with flatter rising arm
L+ = L-shaped with slightly rising leg
L = L-shaped with virtually horizontal leg
L— = L-shaped with slightly falling leg
S— = slightly falling straight line
S = virtually horizontal straight line

Source: Adapted from Vinod K Gupta, "Cost Functions, and Barriers to Entry in Twenty-nine Manufacturing Industries of India," *Journal of Industrial Economics*, November 1968, pp. 59-60.

Additionally, you may consider the following two tables summarising a few empirical studies on cost functions.

Table 2: Empirical Studies of Long-run costs

Investigator	*Year*	*Type of Industry*	*Method*	*Findings*
Gupta*	1968	Manufacturing (India)	CS	L-shaped in 18 industries, U-shaped in 5, Linear in 6
Moore	1961	Manufacturing (USA)	E	Economics of scale
Jhonston	1960	Coal Mining (UK)	CS	Wide dispersion of cost per ton
Jhonston	1960	Electricity (UK)	CS	Declining LAC
Holton	1956	Retailing (Puerto Rico)	E	L-shaped LAC but inputs of management undervalued
Barts	1952	Railways (US)	CS	LAC either constant or falling
Lomax	1951	Gas (UK)	CS	Declining LAC
Dean & James	1942	Retail Shoe Store (US)	CS	U-shaped LAC but not because of diseconomies of scale

* Already reported in the preceding table.

Notes: CS = Cross section, E = Engineering Method.

Table 3: Empirical Studies of Short-run costs

Investigator	*Year*	*Type of Industry*	*Findings*
Jhonston	1960	Electricity	Average total cost falls, then flatens, tending towards constant MC up to capacity
Mansfield and Wein	1958	Railways	Constant MC
Eiteman and Guthrie	1952	Manufacturing	MC below AC at all outputs below capacity
Nordin	1947	Light plant	Increasing MC
Ezekiel & Wylie	1941	Steel	Declining MC
Dean	1941	Leather belts	No significant increase in MC
Yntema	1940	Steel	Constant MC

Source: A.A. Watter's article in Econometrics, January 1963

Activity 3

How would you explain the findings of:

a) **Gupta (Manufacturing):**

....

....

....

....

b) **Jhonston (Coal mining):**

....

....

....

....

c) **Barts (Railways):**

....

....

....

....

d) Dean & James (Retail store)

....

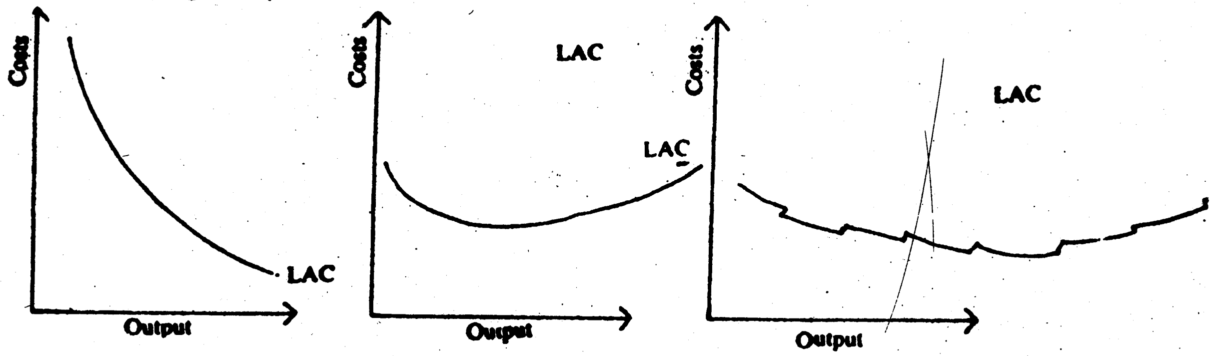
....

....

....

Activity 4

a) Under what specific conditions would you encounter each of the following LAC (long-run average cost) curves?



....

....

....

....

....

....

Hint: Recall the concepts of economies and diseconomies of scale. You may also like to consult some of the additional readings mentioned at the end of this unit.

b) Draw the LMC (long-run marginal cost) curves in each of the above cases.

c) Draw different empirical cost-curves (as indicated in the table) other than the three drawn above. Use the code suggested in the table.

Activity 5

a) How would you reconcile the findings of Yntema with those of Ezekiel and Wylie?

....

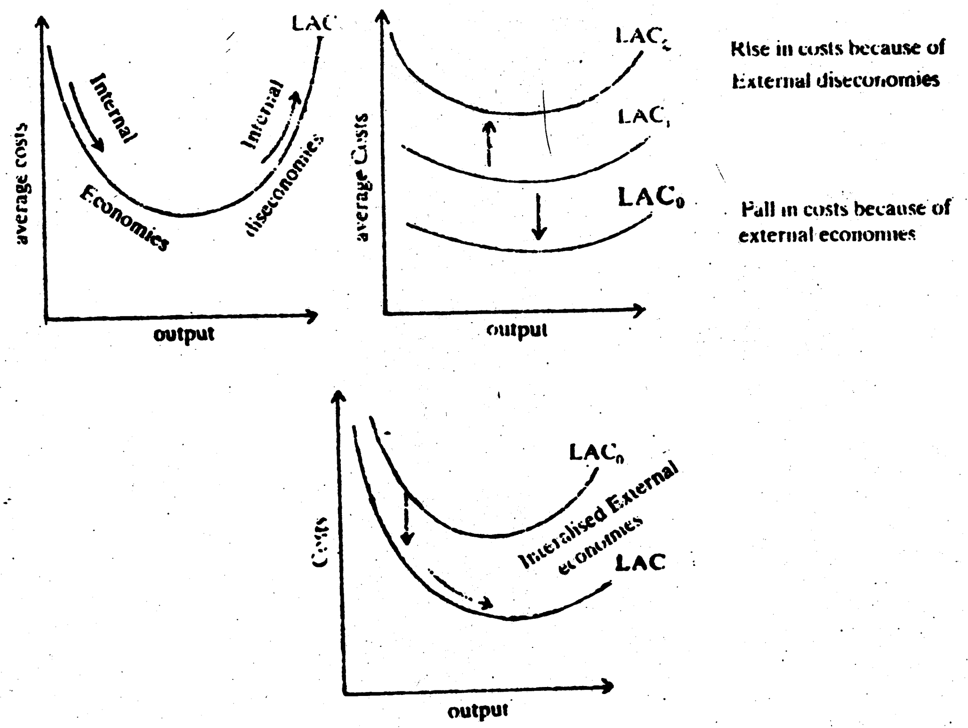
....

....

....

....

i) Explain, quoting appropriate examples, each of the following diagrams:



....

....

....

....

10.6 EMPIRICAL DETERMINATION OF COST FUNCTIONS

There are both statistical as well as economic issues relating to empirical estimation of cost function.

Statistical (metrical) issues

These problems relate to difficulties in methodology and measurement as follows :

i) **Measurement of output of multiple products:** Most attempts to solve this problem have taken the form of weighting the quantity of each product by the proportion of total direct costs that is incurred by the product. In effect, this amounts to determining output as a function of costs, at least to some degree when what is needed is cost as a function of output. Paralleling this problem is the difficulty of measuring the size of the firm in the long-run studies.

ii) **Dealing with technology change:** Whenever technology changes, a new cost function emerges. Thus no cost function can accurately account for technological change. Closely related to this is the problem of variations in the size of the firm.

iii) **Measuring costs of assets:** There are problems relating to measurement of capital as a stock or as a flow, valuation of fixed assets, inventory valuation, unaccounted value changes, valuation of working capital, measurement of depreciation, etc. Some of these problems bring about a debate between accounting approach, economist's approach and engineering approach.

iv) **Choosing a measure of efficiency:** Operating efficiency can be measured in terms of rate of return on investment, net profit, retained earnings after tax and dividends etc. Depending upon the use of a particular measure, a particular cost efficiency may be indicated, but that may not truly reflect 'economic efficiency' (considering opportunity costs) or 'technical efficiency' (in engineering feasibility sense).

Economic (Theoretical) issues

Economic theory often mentions U-shaped nature of average cost curve, but in reality, we come across various other types like the stair shaped one, L-shaped learning curve or a flat bottomed average cost curve. To take care of these empirical situations, the modern theory of costs has been developed. Some explanations are not available to justify the shape and the slope of statistical cost function.

i) **Operating range:** It is possible that the assumptions of economic theory are approximately correct, but total costs tend to be linear or nearly so within the practical operating range.

ii) **Constant returns to scale:** It is possible that the assumptions of economic theory are approximately correct, but constant marginal costs prevail over wide ranges of total costs. If this is true, it means that within relevant range of the data, input or factor proportions are constant and there are no significant economies or diseconomies of scale.

iii) **Dynamic flexibility:** In a dynamic economy, firms have to be flexible to adapt themselves to changing business conditions. Out of experience, the firms 'learn' how to take advantage of continuous technological progress. Thus they prevent the costs from going up i.e. over a long range of output they reap maximum "economies of scale" and "economies of scope". This is what makes their long-run average cost downward falling one or flat bottomed one.

In other words, economic theory, its assumptions and explanation can still aid out understanding of empirical cost situation. You may note that the explanations above are not mutually exclusive. Thus the gap between economic theory and empirical cost situation should not pose any problem.

Activity 6

a) **Production is related to costs. In fact, cost function can be derived from estimated production functions. If so, in view of empirical determination of production functions, can you think of some measurement problems relating to cost function?**

....

....

....

....

b) **Despite the problems of measurement, an estimated cost function is useful to a manager. Can you think of some points to support this contention?**

....

....

....

....

10.7 MANAGERIAL USES OF ESTIMATED COST FUNCTIONS

The estimated cost function can help managers to take meaningful decisions with regard to:

1) determination of optimum plant size
2) determination of optimum output for a given plant, and
3) determination of a firm's supply curve

The optimum plant size, as discussed earlier, is defined in terms of minimum costs per unit of output. In other words, an optimum plant is given by that value of K (plant size) for which average cost is minimum. If the long-run total cost curve is a cubic function, the resultant long-run average cost curve will be a conventional U-shaped curve. The plant level at which the long-run average cost is minimum will be of optimum size. For a given plant, the optimum output level will be achieved at a point where the average cost is the least. This condition can be easily verified from the short-run total cost function.

The level of output that a firm would like to supply to the market will depend on the price that it can charge for its product. In other words, a firm's supply is a positive function of the product price. To get the firm's supply schedule, one needs to know the firm's cost funcction and its objectives.

10.8 SUMMARY

In the preceding sections, we have highlighted the importance of production and cost estimation for decision-making. Various approaches to measure and empirically estimate production and cost functions have also been discussed. A manager has no option but to understand clearly the relationship between his inputs and output, and output and costs to make meaningful decisions of both operational (short-run) and strategic (long-term) nature. In the process, a manager has to blend a number of approaches—economist's approach, accountant's approach and engineer's approach. Though empirical estimates have a lot of use for managerial decision making, yet there are both *conceptual* and *statistical* problems in meeting such functions. Such problems have been discussed in this unit. This is not to discourage managers. The supporting staff of a manager may furnish estimates based on empirical research, but the manager must know the uses and limitations of such estimates before interpreting the findings and basing his decisions on them.

10.9 ADDITIONAL READINGS

Seo, K.K. 1988. *Managerial Economics*, Surjeet Publications: Delhi (Chapter 10 & 12)

Gupta, B.S. January 1973. Article in *Indian Journal of Industrial Relations.*

Gupta, V.K. November 1968. Article in *Journal of Industrial Economics*

Mote, V.L., S. Paul, and G.S. Gupta, 1977. *Managerial Economics*, Tata McGraw-Hill, (Ch. 3).

10.10 SELF-ASSESSMENT TEST

1 Explain clearly the managerial uses of statistically estimated
   a) Product Function
   b) Cost Function

2 State and explain various approaches to estimating a cost function.

3 State and explain the various forms of production function, using the concepts of factor intensity, factor productivity, factor elasticity of output and elasticity of factor substitution.

4 Draw and explain a few empirically encountered cost situations.

5 Review various conceptual and statistical problems of measuring:
   a) Production Function
   b) Cost Function

BLOCK IV PRICE-OUTPUT DECISIONS

This block concentrates on the analysis of price-output decisions taken by the firm, given its market environment. The analysis here is carried on the assumption that the firm maximises profit. In a subsequent block, this assumption will be critically reviewed and relaxed or reworked.

The present block starts with an overview *(unit 11)* of the nature of competition and the form of markets, perfect as well as imperfect. This is followed by an analysis *(unit 12)* of market structure with reference to the large group case — Perfect competition and Monopolistic competition. Next (in *unit 13*), the small group case is analysed with reference to a variety of market structures: monopoly, simple and discriminating, duopoly and oligopoly. We have, of course, refrained from analysing factor market situation like monopsony and oligopsony, though some reference has been made to pricing of services. The discussion on pricing principles, underlying various models, has been followed (in *unit 14*) by an examination of pricing practices, strategies and tactics, as observed in real world business situation. Occasionally we have attempted a brief reference to the regulatory environment whereby market decisions are influenced by Government controls and regulations; this has particular relevance for a country like India. We would like to return to this subject subsequently.

UNIT 11 MARKET ENVIRONMENT OF PRICE-OUTPUT DECISIONS

Structure

11.0 INTRODUCTION

The price-output decision is one of the key managerial decisions. What to produce and how much of it to sell whom, at what price, and what discount, if ~~any~~, constitute a set of very important managerial questions. For most firms, the occasion to ~~decide~~ on what to produce does not come frequently. The product line being given, managers have to continuously bother about the pricing and quantity to be sold. This crucial decision area which occupies so much of the day-to-day managerial attention is referred to as price-volume determination in 'Managerial Economics'.

The natural question to ask is what are the factors that govern this key decision area. Traditional economic theory answered this in terms of demand and supply functions. This traditional analysis assumes that firms strive towards maximising their profits and the consumer utility. The inter-play of demand and supply in the market determines the price which is often referred to as equilibrium price.

There are however many other factors that influence the price-volume relationships of a firm and an industry. The number of firms in the industry, the nature of product, and the possibility of new firms entering the market are just a few of these. In this unit you will understand more about some of the crucial factors that operate in the market place. In the process, you should gain valuable insights into the operations of complex market structures which are more typical of the exciting real world situations.

11.1 OBJECTIVES

The objectives of this unit are to:

- understand the patterns of the market structures that mould competitive behaviour of firm/industry in the real world business situations,
- develop a classification scheme for the different types of market structures
- appreciate the environment of price-output decisions by a typical firm and the industry, and
- Prepare the ground-work for an elaborate and indepth analysis of more complex competitive situations.

11.2 FACTORS INFLUENCING PRICE-VOLUME DECISIONS

There are many factors that influence decisions on market variables. Let us discuss some of these factors:

Competitors

Let us look at a company manufacturing colour television sets (CTV), say Orson Electronics. Obviously, market demand for CTV sets, Orsons's own manufacturing capacities and costs will determine at what price the firm will sell its products. But, Orson will consider several other aspects of the television market. It will study the prices and volumes of other competing brands like Philips, Onida, Uptron, and 10 other firms. The company will also be on guard to prepare for a new reputed firm entering the CTV market. Will the new firm bring out a better product? Will it be priced higher or lower than the Orson CTV? Imagine that Philips and Uptron all of a sudden reduce the prices of their sets. Will Orson follow suit and accept declining profitability or will it keep status quo and accept a possible fall in the market share? What is the goal of the company? Will it always try to maximise profits or can it sometimes accept a temporary beating in the long term interest of the business?

Government Policy

It is quite clear that many and complex factors operate on the company's price-volume decisions. The present price embodies in it a certain proportion of excise duties, customs duties and other taxes. The present volume is limited by the licensed capacity and also by what the market can bear. The government policy can change. Taxes and duties change not only in the annual budget but they are also subjected to variations during the course of the year. Licensing policy is liberalised at one time but can be made restrictive at some other time. It must therefore be recognised that in the Indian context government policy is as much important as market related factors in shaping the price-volume decisions of a firm.

Substitutes and Complements

Continuing with the case of Orson CTV, what about black and white television sets (BWTV)? You might argue that CTV sets do not compete with BWTV sets and therefore what happens in the BWTV market should be of no concern to Orson. To an extent this is right. But then, one must consider prices of related products like VCRs while analysing the CTV market. Items like refrigerators and scooters which satisfy different needs are competing with the limited incomes of the customers. To the extent CTV sets and scooters stake claims on the customer's incomes, a firm making CTV sets cannot ignore the scooters market.

We took the case of a CTV and identified the various factors that may form a part of the decision premise for Orson. CTV is a durable product bought by a typical Indian customer once in a life time. A detergent powder is required daily and even a poor household will use some quantity, say 250 gms, in a month. Customers perceptions of quality and price are different for CTV sets from those of the detergent powder.

In summery, we can now draw a list of the various factors that govern the price-volume decisions of a firm's product. Different factors will carry different weightages in different situations but they will all to be considered by the manager.

- Demand
- Cost of production and capacity
- Objectives of the firm
- Nature of the product
- Nature of competition in the market
- Government policy

Activity 1

Describe the various factors that will affect the price-volume decisions of 'Nirma Detergent' Powder':

Activity 2

We normally seem to talk about a product. What is a firm is selling a service instead? What would be the nature of price and how would one define volume? Try and answer this question for a firm which is in the business of investment counselling for middle-class individuals.

11.3 OBJECTIVES OF THE FIRM

Much of the standard economic theory of a firm assumes that its objective is to maximise profits. Does that mean there are no other goals for the firms to pursue? Is it possible that each and every action of the bussiness managers is directed towards maximising profits? Clearly, business enterprises have other goals and not every action is specially meant for maximising profits. The practical relevance of this goal can best be summerised in the following two statements:

A firm faced with several alternatives having different expected profit outcomes can usually be counted upon to select the alternative with the greatest expected profit" (Arthur A.Thompson, r.)

If a firm is absolutely reckless in calculating costs and revenues, then the Darwinian law of survival of the fittest will probably eliminate it from the economic scene. Therefore, those firms which do manage to survice cannot be completely obvious to the maximisation of profits" (Paul A. Samuelson).

A poser

In the Indian business scene can a firm get eliminated? What is the spectre of sick firms?

In fact many firms, which ought to have been eliminated from the scene, continue as sick units. There are social and political reasons for this phenomenon.

Alternatives to Profit Maximisation

Since the late fifties, several authors like William Baumol, Edith Penrose, F. Machlup and others have questioned the validity of assuming profit maximisation goal. Herbert Simon takes the extreme position of questioning the very ability of human minds working within the firms to calculate profits accurately. He postulates:

"Administrative theory is peculiarly the theory of **intended** and **bounded** rationality - of the behaviour of human beings who **satisfice because they have not the wits to maximise**" (emphasis added).

The whole question of the theoritical aspects of firms' behaviour will be dealt with in a subsequent block but it will be useful to review it.

The issue can be stated thus: while understanding the goals of the business firms, what matters most is **the principal goal** and not **the only goal**. Every one recognises that the firms worry about things like growth, leadership position, technical innovation, market share, public image and what have you. To the extent other goals are subordinate to profits, a firm can be deemed to be operating in a profit maximising manner.

Revenue Maximisation

There are firms where sales revenue comes first and profits next. Very often we read announcements like the following in the newspapers: "In 1987 our firm recorded the maximum sales turnover in the industry and we are proud of this number 1 position". This firm may not be the most profitable one in the industry but its revenues are the highest in the industry and it naturally takes pride in its satus as number one.

It has been observed that several firms try and strive towards revenue maximisation once they are sure of attaining a minimum acceptable level of profitability in the operations. The objective of such firms can be stated thus: Maximise sales revenue subject to attaining say 20% return on net worth. Firms which have operated in an industry for a fairly long period are often prompted to operate by this goal. The record of firms like Asian Paints, Telco, and Bajaj Auto in the Indian context will convince you of the validity of this proposition.

Other Goals

Apart from profits and revenues, firms can and do keep as their principle objectives, parameters like market share, technological excellence, growth rate, customer satisfaction, best employer status etc. Thus, a firm may try to be constantly the top one in the industry in terms of market share. Or it may not care much for sales revenues and market share as long as it can boast of being the first in the industry to introduce product innovation. Lastly, we do come across companies who stress their growth objective. Such a firm would inform the public through an advertisement that 'no other firm in any industry has managed to record a consistent growth of 25% in sales turnover over the last five year period'.

We must bear in mind that all these goals do not imply that profits do not matter. All that these goals signify is that profits are secondary. The only thing that is stressed in the above discussion is the supremacy of one goal over some others. Is it possible for a firm to simultaneously pursue all the goals and succeed?

Activity 3

Rather than trying to find an answer to the above question in theory, make an attempt to identify one firm in the Indian context which meets the following criteria:

- Maximum sales in the given industry
- Highest growth rate in sales during the last decade
- Highest profitability in the industry for a continuous period of five years
- Has the best record of successful new product introductions
- Always keeps all the customers satisfied
- Has no industrial relations disputes
- Is perceived by the people as the best firm

....

Activity 4

What is the principal goal of your company? What other goals the management has set for itself?

If possible arrange these other goals in order of priority.

....

....

....

....

....

....

....

....

....

....

11.4 CLASSIFICATION OF MARKET STRUCTURES

Market and competition

While all of us often use the word 'market', we do not realise that very few markets possess a well defined place in a geographical area or have a postal address. The Bombay Stock Exchange is one such market with a building and an area earmarked for transacting shares. In general, a market is a group of people and firms which are in contact with one another for the purpose of buying and selling some product. It is not necessary that every member of the market be in contact with every other.

The central phenomenon in the functioning of any market is competition. Competitive behaviour is moulded by the market structure of the product under consideration. It is therefore necessary to have a thorough understanding of this concept.

Market Structure

A simple definition of this concept can be found in Pappas and Hirschey (1985). According to them "Market structure refers to the number and size distribution of buyers and sellers in the market for a good or service". (p.312) Implicit in this concept is an idea that the market structure for a product not only includes firms and individuals currently engaged in buying and selling but also the potential entrants.

Markets are traditionally classified into four basic types. These are:

Perfect competition is characterised by a large number of buyers and sellers of an essentially identical product. Each member of the market, whether buyer or seller, is so small in relation to the total industry volume that he is unable to influence the price of the product. Individual buyers and sellers are essentially price takers. At the ruling price a firm can sell any quantity. Since there is free entry and free exit, no firm can earn excessive profits in the long run.

In the monopoly situation there is just one producer of a product. The firm has substantial control over the price. Further, if product is differentiated and if there are no threats of new firms entering the same business, a monopoly firm can manage to earn excessive profits over a long period.

Monopolistic competition a term coined by E.M. Chamberlin implies a market structure with a large number of firms selling differentiated products. The differentiation may be real or is perceived so by the customers. Two brands of soaps may just be identical but perceived by the customers as different on some fancy dimensions like freshness. Firms in such a market structure have some control over price. By and large they are unable to earn excessive profits in the long run. Since the whole structure operates on perceived product differentiation, entry of new firms cannot be prevented. Hence, above normal profits can be earned only in the short run.

Oligopoly is a market structure in which a small number of firms account for the whole industry's output. The product may or may not be differentiated. For example, only 5 or 6 firms in India constitute 100% of the integrated steel industry's output. All of them make almost identical products. On the other hand passenger car industry with only three firms is characterised by marked differentiation in products. The nature of products is such that very often one finds entry of new firms difficult. Oligopoly is characterised by vigorous competition where firms manipulate both prices and volumes in an attempt to outsmart their rivals. No generalisation can be made about profitability scenarios.

Two basic parameters: By now you would have noticed that all the market structures use only two parameters as distinguishing factors - number of firms and degree of product differentiation Other factors like product characteristics and entry of new firms are also important but these determine the level of competition in a given market structure. We shall deal with this aspect in the next section.

It must also be noted that all these market structures can be classified in only two fundamental forms - **Perfect Competition** and **Imperfect Competition**. Under this classification Monopoly, Oligopoly and Monopolistic Competition are treated as special cases of markets which are less than perfect. Thus these forms illustrate the degree of imperfection in a market by using the number of firms and product differentiation as basic criteria.

Classification diagram: A convenient and effective classification scheme depicting types of competitive market structures is shown in the diagram. We may now ask ourselves the question how does the real world look like? A close look at the diagram reveals that most real world markets are neither perfectly competitive nor perfectly monopolistic. Most industries that we come across can be classified in the realm of imperfect competition.

Diagram

Types of Competition

Kind of Competition	**Number of Producers and Degree of Product Differentiation**	**Part of Economy Where Prevalent**	**Degree of Control Over Price**	**Methods of Marketing**
Perfect competition	Many producers; identical products	A few agricultural industries	None	Market exchange or auction
Imperfect competition:				
Many differentiated sellers	Many producers: many real or fancied differences in product	Toothpastes, retail trade, conglomarates	Some	Advertising and quality rivalry: administered prices*
Oligopoly	Few producers: little or no difference in product	Steel,aluminium		
	Few producers:some differentiation of products	Autos, machinery		
Complete monopoly	Single producer: Unique product without close substitutes	A few utilities	Considerable	Promotional and "institutional" public-relations advertising

Source: Samuelson, Economics, p.489.

* The use of the words 'administered prices' should not be confused with the common usage of this term in India. It simply means price administered by the firms. Samuelson defines it as 'posted price' which changes infrequently.

A rider: Markets in the Indian context present an additional dimension. A large number of industries and services are reserved for the exclusive operation by the public sector. We therefore have just one "Indian Railways' catering to the entire country's rail transport needs. There is just one "Bharat Heavy Electricals Ltd.' producing electrical generating sets of 500 MW capacity. In the private sector also, there are some industries, with just one manufacturer. An example is 'multi-cylinder fuel injection pumps' which go into the engines of four wheelers like jeeps, cars and commercial vehicles.

Activitv 5

a) Identify monopoly industries in the public sector and private sector manufacturing and services sectors.

....

....

....

....

....

....

....

....

b) The table below gives information on some Indian industries which exhibit different degrees of imperfection in their market structures. Fill in the last column by specifying the type of competition that is most likely to prevail in these industries.

S.No.	Name of Industry	Number of fimrs	Type of competition
1	Tractor	10	
2	Vanaspati (ghee)	22	
3	Edible oil	3600	
4	Sugar	326	
5	Cigarettes	5	
6	Caustic soda	24	
7	Passenger cars	3	
8	Commercial Vehicles	11	
9	Cement	40	
10	Earthmoving equipment	10	
11	Soaps	20	
12	Synthetic detergents	16	

Sources: Centre for Monitoring Indian Economy, Bombay

c) All the firms in the sugar and cement industries produce almost identical products and there are many firms. Price control prevails in these industries. Each firm in the industry must sell a part of the output at a price determined by the government. In what way the performance of these industries would differ from say the caustic soda industry where no such price control exists?

....

....

....

....

....

....

....

....

....

....

....

....

....

11.5 FACTORS DETERMINING THE NATURE OF COMPETITION

We have already seen that the number of firms and product differentiation are extremely crucial in determining the nature of competition in a market. It has been tacitly assumed that there are a large number of buyers. What would happen if there are several firms producing a standardised product but only one buyer? Obviously, the buyer would control the price, he will dictate how much to buy from whom. The entire price-volume decision takes on a different qualitative dimension. Similarly, product features and characteristics, the nature of production system, the possibility of new entrants in a market have profound impact on the competitive behaviour of firms in a market. Since the 'entry' of new firms has special relevance in business behaviour we reserve it to the next section and deal with other issues in the present one.

Effect of Buyers

We have already referred to the case where there is only one buyer. Such a situation is defined as **monopsony**. For example, thre are just six firms in India manufacturing railway wagons all of which supply to just one buyer, the Railways. Such a situation can also exist in a local labour market where a single large firm is the only provider of jobs for the people in the vicinity. A recent example is the new petro-chemicals complex that is coming up in the rural parts of coastal Maharashtra.

More frequently encountered in the Indian markets is a case of a few large buyers, defined as **oligopsony**. The explosives industry which makes detonators and commercial explosives has three major customers: Coal India Ltd.(CIL), Department of Irrigation and various governmental agencies working on road building activities. Of these, just one customer, CIL takes nearly 60% of the industry's output. There are about 10 firms in the industry which negotiable prices and quantities with CIL to finalise their short term plans.

Most industries manufacturing heavy engineering equipment are typified in India by few manufacturers and few buyers with the Government being the major one. Price and volume determination in such products often takes the form of 'negotiation across table' rather than the operation of any market forces. Since the members in the whole market inclusive of buyers and sellers are not many: very often they know each other. In other situations, like the consumer goods, firms have no direct contact with their customers.

Production Characteristics

Minimum efficient scale of production in relation to the overall industry output and market requirement sometimes play a major role in shaping the market structure. Why there are no more than say, 5 or 10 integrated steel plants even in an advanced country like the U.S.A can be partly explained by this factor. Since the minimum economic size of such a steel plant is a few million tonnes, the entire world steel industry can have no more than 100 efficient and profitable firms. Thus every country has only a handful of steel plants. On the other hand, when one comes to rerolling mills which take the steel billets or bars as input, the minimum eficient size comes down considerably, and given the existing demand, several firms can be seen to operate.

Further, the minimum size does not remain constant but changes drastically with technological advancements. When technical changes push up the economic size of a plant one notices that the number of firms decline over time. This can be noticed in some process industries like synthetic fibre. Conversely, technological innovations may make it possible for smaller sized plants to become economically viable. In such a case of lot of new entrants come and soon the market becomes highly competitive. Notice the personal computer industry in India.

Apart from minimum plant size, factors like availability of the required raw material, skilled labour etc. can also mould the market structures. Presently, only one Indian source (IPCL) provides all the raw material for plastic products. Likewise, enough skilled people are not available work on the sophisticated machines. There factors sometimes restrict output and push up prices even though adequate market potential for expansion exists.

Product Characteristics

Section 11.2 referred to market situations with CTV and detergent powder as product examples. Both these markets have many firms and the products are differentiated. But in case of CTV, there are no close substitutes (BWTV) being a poor one, whereas there are many substitutes to a detergent powder (bar soaps, chips, cakes). Therefore, one notices more violent competition in the detergent market than in the CTV market. In the CTV industry firms are competing with each other's products but in the detergent market the firms are competing with other substitute products as well. Of course you may remind us of the customer income constraint but even with that there should be no difficulty in appreciating the differences in the degree of competitiveness in these two markets. Similarly when two locations are connected by road and rail, firms engaged in passenger bus service are not only competing with themselves but also with an alternate mode of transport.

The physical characteristics of product can also influence the competitive structure of its market. If the distribution cost is a major element in the cost of product, competition would tend to get localised. Within a given region firms would compete and make attempts to set up several plants around all the major markets in a bid to show their presence in all the territories Similarly, for perishable products, the competition is invariably local.

Conflict between physical characteristics and minimum economic size.

An interesting question arises in the case of a product like cement. For reasons of minimising the transport costs on raw materials, most cement plants in the country are located near mine sites. A large efficient plant near a mine site can manufacture cement at the optimum cost, but the local demand is never large enough. If such a plant has to sell in far away markets (from Gujarat to Kerala, for example) the transport costs can be quite high. Customers located in such areas will always buy cement at a much higher price. The government partly offsets this by using the mechanism of levy price which is the same throughout the country.

Activity 6

a) Obtain information about the Freight Equalisation Scheme in steel and analyse its impact on the price of various steel products in different locations.

....

....

....

....

....

....

....

....

....

....

b) Describe the competitive situation in the market for a product with one dominant buyer and one dominant seller.

....

....

....

....

....

11. 6 BARRIERS TO ENTRY

In a classic book J.S. Bain (1956) analysed the character and significance of the condition of entry in manufacturing industries. Till that time, most analyses of how competition works gave little emphasis to the force of the potential or threatened competition of possible new competitors. The attention was simply focused on the competition among firms already established in an industry. Lately, however the meaning of competition is inclusive of 'potential entrants'.

The existence or otherwise of 'entry barriers' in a given industry has profound inpact on its performance and the behaviour of firms in it.

It has been found that the firms in an industry are always worried about the possibility of a new entrant. If the existing number is few then the degree of insecurity will be correpondingly higher. To be sure, the existing firms, especially in an oligopoly, have some advantage over the potential entrant. But because of the threat of new entrants, the existing members cannot exploit these advantages (by raising prices continuously) beyond a point. What that point is and when the new entrants would find it profitable to break the entry barriers are also not known. One thing is clear, that this potential competition always puts a check on the pricing strategies of oligopolists.

What can act as an entry barrier?

Anything that retains the competitive advantages of the existing firms in an industry can act as barrier to those desirous of entering it. Some of the commonly encountered aspects are indicated below.

High Initial Investment

A new passenger can plant with a capacity to assemble say 50,000 automobiles per annum can cost around Rs.100 crores. You know that not many firms have the capacity to mobilise resources of that order. Naturally, there are high entry barriers to the automobile market due to high level of initial investment. For similar reasons, one does not find too many integrated steel plants coming up too often. On the other hand, it takes only a few lakhs of rupees to set up a biscuit making unit. The barrier on account of investment is quite low in such industries.

Economies of Scale in Non-production Activities

Scale economies are not restricted to manufacturing. These extend to distribution, marketing and advertising. Consumer products like soaps, toothpastes display considerable economies of scale in marketing and distribution. A nation-wide presence in these industries presupposes an efficient and penetrating distribution network, high order of brand related marketing skills and ability to service a fairly differentiated product line. Thus, one may find numerous local soap makers but there are substantial entry barriers to a new national brand penetrating the market.

Technology, Patents and Research

The ability to possess and commercially exploit certain specialised technology is one more source of entry barrier. Specially chemicals, drugs, plastics are some of the industries where the difficulty of developing a new product or a process is well understood. These are knowledge related factors. It is very difficulty to penetrate an industry where a few existing firms have a strong research base and a large pool of product related patents. New entrants in such industries are often the employees of the existing firms breaking away to form a new entity.

Switching Costs

Take an industry like earthmoving machinery. For such an industry each firm has a few large customers like contractors, project authorities or coal mines. Consider that a customer has a fleet of say 10 machines of a given brand. When he replaces one machine or augments his fleet, more likely the choice would fall on the same brand. For him it means a familiar machine, known operational details, already trained operators and a host of other things like spare parts stocks. Thus, the cost of switching to a new brand can be fairly high. These costs can act as entry barriers. Along with earthmoving machines the customer also has related equipment like loaders and dump trucks which he had purchased on the ground compatibility with a given brand of the main machine.

Take the case of IBM. Why does every other personal computer (PC) that one comes across claims to be an IBM compatible. It has to be so, because all the software is developed by using IBM standards. The PC cannot work without software. By developing industry level standards, IBM has created 'high switching costs' in an attempt to create entry barriers.

You will have noticed that in oligopoly situations, firms should strive towards creating high entry barriers, if the industry does not possess those necessary characteristics. This is precisely what happens. If there are low entry barriers, new firms entry soon and the profitability of the existing firms drops. Notice the state of the pocket calculator industry. There are virtually no entry barriers and with the existence of cheap smuggled products, it is impossible to create them. As a result, most large firms are almost out of this market leaving it open for the small scale units.

Activity 7

The table below lists some industries. Indicate in column 3 whether the entry barriers are high or low. Give reasons in column 4.

S.No.	Name of the Industry	Entry Barriers	Reasons
1	Computer Software		
2	Mainframe computers		
3	Oil-fed chemicals		
4	CNC machine tools		
5	Breakfast cereals		
6	Aluminium		
7	Ball-point pens		
8	Colour Television Sets		

11.7 THE ROLE OF GOVERNMENT POLICY

All governments, whether in India or abroad, impose taxes and duties. What is special about the Indian governmental policies is their ability to control price, quantity of production, distribution, choice of product, location and almost every business decision of a firm. Some reference to these have been made in the previous sections. Presently we shall see the role of government policy in a synoptic way. Later, a full unit in Block No. 5 will talk about the regulator environment in detail.

Through its industrial licensing policies the Central Government has control over the following business deicisons:

1 Choice of the product
2 Scale of production (capacity)
3 Location of production
4 Choice of technology

The policy on foreign collaborations also regulates the aspects pertaining to choice of technology. Import polices can have significant impact on the types and quantities of raw materials that would become available for production. Choice of machinery is also guided by the import regulations in force.

Through levy of customs and excise duties, the price of the end products as well as the raw materials gets affected. Some industries like sugar, aluminium, steel, edible oils, cement are subject to price controls. These are administered through various Acts and the job of determining prices is often entrusted to the Bureau of Industrial Costs and Prices under the Ministry of Industry. Firms in these industries are thus partly guided by market forces and partly by the Ministry in regard to their pricing deicisions.

Apart from these, several state governments have their own regulations for promoting (or restricting) the growth of certain industries. All things considered, the job of the business manager is made quite difficult in the Indian environment. Ironically, government steps in to correct certain imperfections in the market but in the process adds a few of its own. The existence of many industries with only few firms is mainly attributable to the government policies which have acted as entry barriers for a long period of time. The picture is changing rapidly. There are fair chances that in the future, market related forces would operate more on the price-volume decisions of the firms than the government policy related factors.

11.8 SUMMARY

In this unit, we have made an attept to make you understand the concept of market structure and the impact it has on the competitive behaviour of firms. Various competitive situations were defined and broadly discussed. The number of firms and the level of product differentiation are useful parameters for classifying various market structures. The level of competition also gets influenced by product and production related factors, potential competitors, number of buyers and their behaviour and the governmental policies.

We are now ready to analyse the various market forms in greater detail. That will be attempted in the subsequent units in this block. As a part of the summary, you may now review the following terms that we have used in this unit.

11.9 KEY WORDS

Price-volume decision refers to that activity in a business enterprise where factors pertaining to the prices and quantities of a product or product range are analysed.

Objective of the firm is the main goal towards which part of managerial attention is directed.

Profit maximisation in real life implies that firms behave as if to make all other goals secondary to the attainment of a certain acceptable level of profits.

Market structure refers to the number and size distribution of buyers and sellers in the market for goods or service.

Perfect competition is a market structure where a large number of buyers and sellers deal in nearly identical products. Each is individually so small in relation to the total output that all members are 'price takers'.

Monopoly situation is characterised by just one producer of a product or service.

Monopolistic competition is characterised by many sellers of differentiated product.

Oligopoly situations have fewer sellers with or without the existence of product differentiation.

Product differentiation refers more to the differences in products as perceived by the customers than in real or technical difference in specifications.

Monopsony is a market with only one buyer.

Oligopsony markets have a few buyers.

Bilateral monopoly is a situation where a single seller (monopolist) confronts a single buyer (monopsonist).

Competition is the collective outcome of the forces generated within a given market structure (for a product or service) in combination with product characteristics, number of buyers, potential entrants and government policy.

Barriers to entry refer to the obstacles that impede the entry of new firms in an industry.

11.10 ADDITIONAL READINGS

Samuelson, Paul A. 1973, Economics, McGraw-Hill, Ninth Edition. (Chapters 20 to 26).

Mote, V.L., Samuel Paul and G.S. Gupta, 1977, Managerial Economics- Concepts and Costs, Tata McGraw Hill.

Bain, J.S., 1956, Barriers to New Competition, Harvard University Press.

Dean Joel., 1970, Managerial Economics, Prentice-Hall India: Delhi. (Chapter 2)

Dorfman, Robert., 1965, The Price System, Prentice-Hall India: Delhi.

11.11 SELF-ASSESSMENT TEST

1 List and explain the factors that determine the element competition in a market for either a product or a factor.

2 It is the nature of prevailing competition that decides the classification of a market into perfect and imperfect. Based on your analysis, comment on the degree of perfection in the following markets:

a) Labour market in Dubai
b) Capital market in India
c) Computer market in India
d) Bombay stock market
e) Wholesale vegetable market in a city like Delhi

3 What do you mean by the term "barriers to entry"? State and explain the factors that cause such barriers.

4 "In theory, we talk about 'barriers to entry', in practice (in India), we have 'barriers to exit': our sick units are not allowed to die a natural death". Comment.

5 In what ways, does monopolistic competition differ from perfect competition? Give real world examples to illustrate your answer.

(You may re-do the same question when you have run down the entire reading material of the present Block.)

6 Review your understanding of the following terms:

a) Cut-throat competition
b) Oligopoly/Duopoly
c) Bilateral monopoly
d) Duopsony/Oligopsony
e) Product differentiation
f) Price-volume decision

7 Write a lucid essay on the "Determinants of Price-Output Decisions".

UNIT 12 ANALYSIS OF MARKET STRUCTURE–LARGE GROUP CASE

Structure

12.0 INTRODUCTION

In the preceding unit, you have been exposed to the concepts of market, competition, perfection, monopolistic elements etc. You must have noted that the number and size of firms is an important determinant of the structure of industry and/or market. For example, a large group (of firms) case will differ from a small group case.

In this unit, we shall analyse the behaviour of a firm and the industry when there are many (large number) seller. The terms 'many' and 'few' that we have been using so far have relative meanings. No specific number can be considered as a cut-off point for deciding this issue. If, with respect to the total size of the market no firm is big enough, we can conveniently define such an industry as one with many firms. On the other hand, there are instances when the number of firms can be numerically large, say 50, but one firm accounts for say 15% of the market. The behaviour of such a firm will be markedly different from the rest.

The crucial parameter is the size of the constituent firms in relation to the total industry's output. If, besides the existence of many firms, there is also product differentiation, then monopolistic competition will prevail. If products are almost identical perfect competition will be obtained. While considering product differentiation, the acid test is **customer perception**. Real differences in specifications and performance are relevant, but analytically, two products are differentiated, provided the customers perceive them as different.

Throughout this unit, we shall assume that the firms are guided solely by the motivation of profit maximisation.

12.1 OBJECTIVES

Having seen the various types of market structures, we now propose to:

- **analyse** in greater detail the market structures with many firms, i.e. perfect competittion and monopolistic competition,
- **derive** equilibrium conditions for a firm and the industry in a perfectly competitive situation,
- **explore** some applications of a perfect market equilibrium, and
- **examine** price-output (volume) decisions under monopolistic competition and analyse their relevance to the Indian context.

12.2 PERFECT COMPETITION

A perfectly competitive market is characterised by the following main features.

- Many buyers and sellers exist that no one can influence the price.
- All firms sell identical products or are perceived so by the buyers.
- All resources and inputs like materials, labour and capital are perfectly mobile so that firms can enter the market and fold up shop as and when they wish.
- Members in the market have perfect knowledge; decisions are made as if everything was certain.

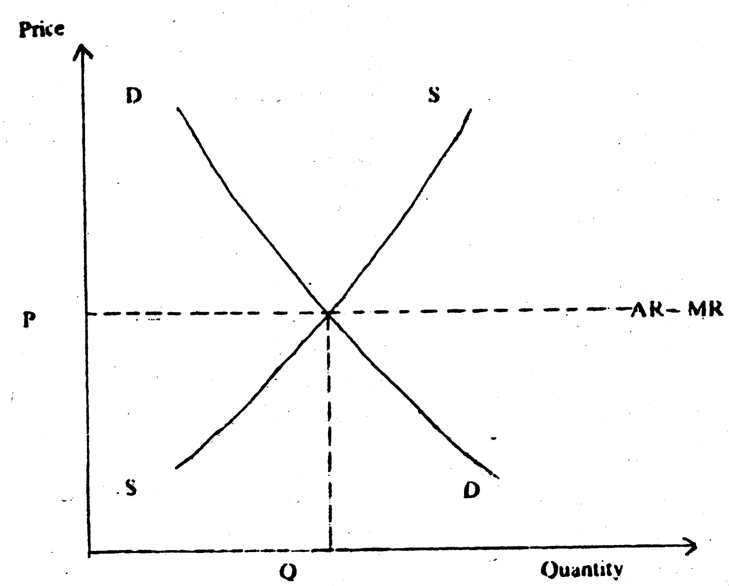
The first two were already referred in Unit 11. The other two features are included in the traditional analysis of perfect competition. With such unrealistic conditions no one expects to see a real world market operating on these lines. Why then such market structures have been theoretically studied? This is so because the analysis of such situations gives insights into the efficiency of resource use. It is used as a yardstick for measuring efficient allocation of resources. To the extent real world markets deviate from this ideal case we get an idea about the inefficiency of resource use prevailing in them. (For a lucid exposition of waste of resources under imperfect competition see Chapter 26 of Samuelson, Economics, Ninth Edition).

Apart from the efficiency aspect, the analysis of perfect competition illuminates several basic principles underlying business behaviour. It is therefore useful to study this market structure in some detail.

Short-run Equilibrium

In the short-run firms cannot increase their production capacities because it takes time to arrange for resources to do so. The industry demand and supply operate in a market where processes similar to an auction are in force. At the intersection of the falling demand curve and the rising supply curve the market price of a commodity for that particular period is settled. Being too small in relation to the total industry's output every individual firm and the buyer have to accept this price. From Figure 1 it can be seen that it price P and quantity Q the industry's equalibrium

Figure 1: Perfect Competition - Industry Short-run Equilibrium



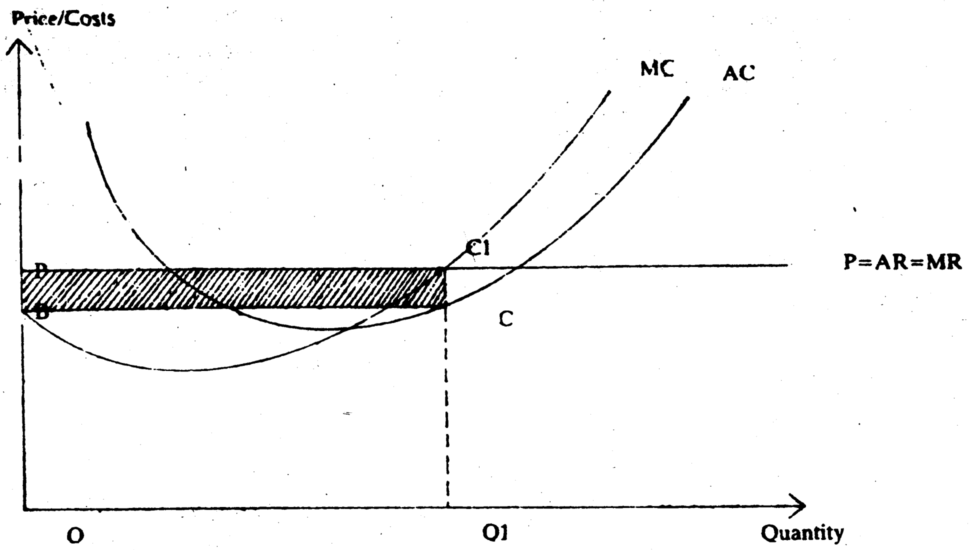
is established. If the price were higher than P, excess supply would come in forcing it downwards. Conversely if it were lower than P excess demand would prevail pushing it up. For an individual firm, the quantity Q_1 that it would offer to the market will depend on its objectives and the cost conditions. Market price being given, the firm is confronted with a horizontal demand curve at the height P. Since all the output can be sold at P, an extra unit of output can also be sold at the same price. Thus, for the firm, the demand curve, the average revenue curve and the marginal revenue curve are identical. We therefore have

P = AR = MR

Maximum profits will be obtained at the output rate where marginal cost MC equals marginal revenue MR. This has to be so because if the cost of producing an additional unit is less than what it can fetch in the market, then profits can be improved by producing and selling it. If, however, it costs more to produce that additional unit than what it earns, the firm would be better off by not producing it. Thus, when MC = MR, the firm is in equilibrium producing an output O_1 as indicated in Figure 2. It has been assumed that the firm is confronted with a U shaped cost curve.

The firm takes the market price P and produces that quantity Q_1 which equates MC and MR so as to fulfil the objective of profit maximisation. The quantum of profit is indicated by the shaded area in Figure 2.

Figure 2: Perfect Competition - Firm Short-term Equilibrium



At the output Q_1 the total cost is $Q_1 \times C$ which is nothing but the area of the rectangle OQ_1CB. Total revenue is $P \times Q_1$ which is equal to the area of rectangle OQ_1C_1P. The difference between total revenue and total cost is the profit as indicated by the rectangle BCC_1P.

Illustration

Imagine a firm operating in a perfectly competitive market. The following data are available.

Price $P = AR = MR =$ Rs.20/- Unit.

Total cost function is $C(Q) = 8+170\ Q-4Q^2+Q^3$

Let us now find out the profit maximising output and the maximum profit by using the concepts developed above.

By definition, marginal cost will be available if the first derivative of the total cost function is obtained. Thus,

$$MC = \frac{dC(Q)}{dQ} = 17 - 8Q + 3Q^2$$

Miximum profit will be earned when MC and MR are equal:

$$20 = 17-8Q+3Q^2$$

Solving this equation gives two values for Q; $-\frac{1}{3}$ and 3. It is obvious that negative output cannot be produced; hence at Q = 3 the firm will maximise profits. Total revenue will be Rs. 60 and total cost Rs. 50(8+17× 3−4×9+27). The maximum profit at the output of 3 units is Rs.10.

Long-run Equilibrium

The shaded area BCC_1P shown in Figure 2 is the profit that a firm in a perfectly competitive industury earns in the short-run. This is defined as 'economic profit' and represents an above-normal profit situation for the firm. A normal profit is defined as a rate of return on capital which is just sufficient to attract the investment necessary to set up and operate a firm. It is customary to include normal profit as a part of 'economic costs'. Thus any profit which is more than whatever is already included as an element of cost becomes above-normal profit.

Over the long-run, any such positive economic profits will attract new firms in the industry or an expansion by the existing firms or both. As this happens, the industry supply gets expanded depressing the price of the product. Long-run equilibrium will be reached when each and every firm operates at a level of output that minimise average economic costs of producing it (which includes normal profit). Under this condition price will equal not only marginal cost but also average cost

$$P = MC = AC = MR$$

This long-run equilibrium situation is depicted in Figure 3. It must be appreciated that as long as the price is above AC there is room for above-normal profit and hence new firms will enter. Conversely, if for some firms the AC is above the price, they will not even earn the minimum incentive to stay in business (normal profit) and will fold up shop in the long-run. When every firm is making just the normal profit, no new firms enter, none of the existing firms quit and equilibrium prevails. The industry as such is in equilibrium when no firm is earning above-normal profits.

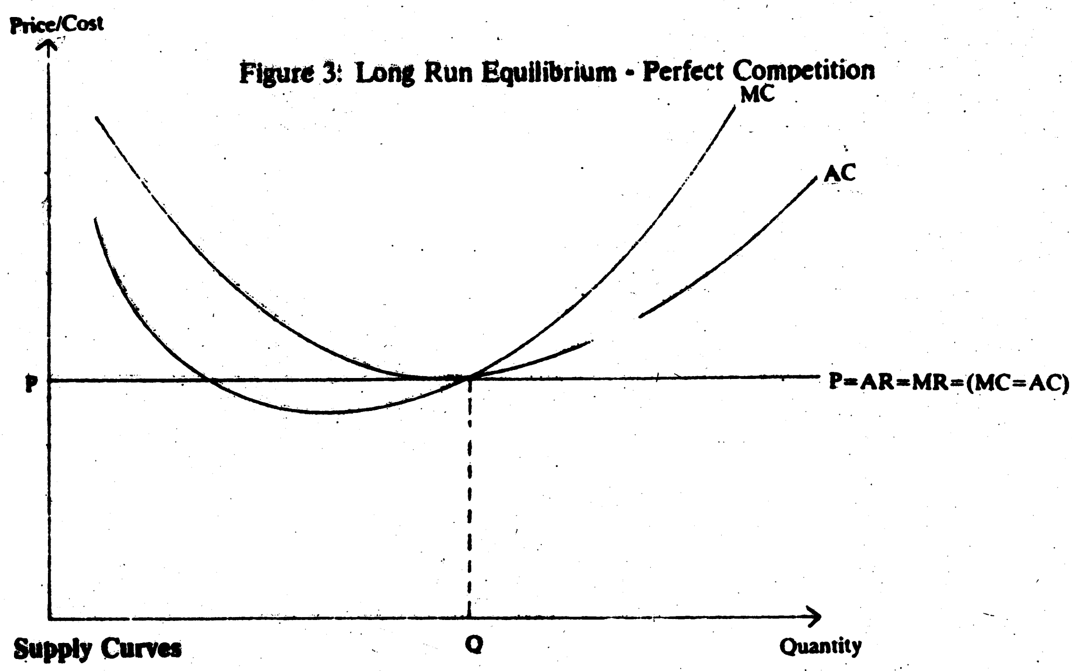


Figure 3: Long Run Equilibrium - Perfect Competition

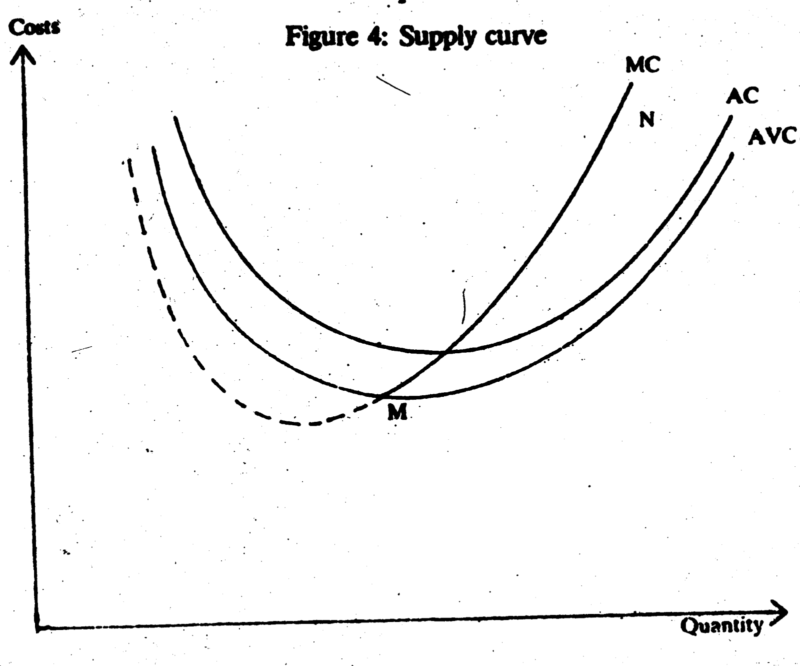
Supply Curves

We have just see that a profit maximising firm will proauce that quantity at which the marginal cost equals the price. Suppose in a particular short period the market price is lower than the profit maximising level. What will the firm do in such a situation? Since the firm is a price-taker, will it just accept the consequences? Its response will depend on how low the price is. If it is low that the firm is unable to recover its variable costs, it will simply stop production and incur a loss equal to its fixed costs since these cannot be escaped. The crucial point for the firm to 'run the show' in the short run is to recover its variable costs. This gives us the following result:

In the short run, the supply curve of firm in a perfectly competitive market will correspond to that portion of the marginal cost curve that lies above the average variable cost curve. As long as the price exceeds average variable cost, every unit of output provides some profit contribution which can be applied to cover fixed costs and earn some profit. The point at which the price just covers the average variable cost is known as the "shut down point" implying if price drifts below that point, the firm will stop production in the short-run. In the long-run where all costs are variable, price does cover AC = AVC = MC at the point of perfectly competitive equilibrium.

Figure 4 represents various situations for such a firm. The supply curve (MN) is indicated by the portion of MC curve above the minimum point of the AVC curve.

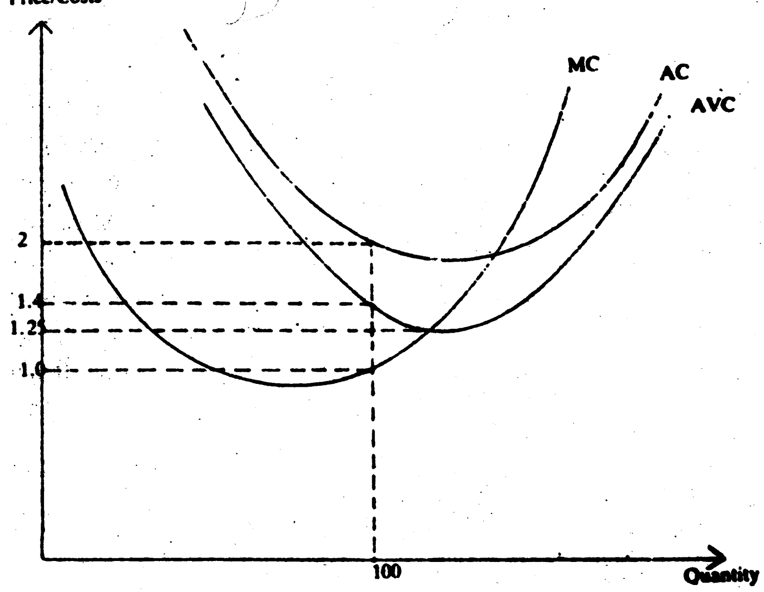
Figure 4: Supply curve



Activity 1

Scrutinise Figure 5 and explain the response of a firm at various prices indicated by the dotted lines.

Figure 5



....

....

....

....

....

....

....

In the long run, all costs are variable costs, machines can be bought or scrapped, new workers can be hired or the old ones fired (provided the union agrees!). Accordingly using the same principles, that portion of the firm's long-run marginal cost curve which lies above its long-run average cost curve (fixed plus variable) represents its long-run supply position.

Activity 2

Ajay enterprises is small firm in the steel office chairs industry which is perfectly competitive. The market price of each chair is Rs. 640 and the company's total cost and marginal costs are given as:

$$C = 240Q - 20Q^2 + Q^3$$

$$MC = 240 - 40Q + 3Q^2$$

A normal profit is included in the cost function.

a) Determine the profit maximising output, average cost at this output and the total profits.

b) If this firm is typical one in the industry, is the industry in equilibrium? Give reasons.

....

....

....

....

....

....

....

....

....

....

....

....

....

....

12.3 TAXATION, SPATIAL DISTRIBUTION AND PERFECT COMPETITION

Spatially Distributed Firms

Consider an industry with 100 firms. Fifty are located in Faridabad and fifty in Jaipur. New Delhi is the only market for the product. It costs Rs. 6 to transport one unit of the output from Faridabad to Delhi whereas the transport cost goes upto Rs. 10 if brought from Jaipur. All the firms in both the locations operate under identical production and cost conditions. The cost function for a typical firm is given as

$$C = 0.5Q^2$$

But, transport costs have to be included if the product is to be sold. If Q_F and Q_J are the quantities for the representative firms in two locations, the cost functions inclusive of transport costs can be represented as

$$C_F = 0.5Q_F^2 + 6Q_F$$
$$C_J = 0.5Q_J^2 + 10Q_J$$

All the firms in both locations are profit maximisers and hence would equate their marignal costs to the price prevailing in Delhi.

If P is the price, the equilibrium condition for a typical Faridabad firm is

$$P = MC = Q_F + 6$$

Similarly, the Jaipur firm must have

$$P = MC = Q_J + 10$$

What happens if the price is less than Rs. 6? Obviously, there will be no supplies at all, since even the transport costs from the nearest location are not covered. If the price is Rs. 6 or more but less than Rs. 10, only the Faridabad firms would supply and the Jaipur firms would abstain from servicing the market.

At a price of Rs. 10 or more, all the 100 firms will compete in the market. The aggregate supply function will be

$$\begin{aligned} S &= 50\,Q_F + 50\,Q_J \\ &= 50(P-6) + 50(P-10) \\ &= 100\,P - 800 \end{aligned}$$

Suppose the aggregate demand function is

$$D = -20P+1600$$

The Market equilibrium will be reached when the aggregate demand equals the aggregate supply. Thus

$$100\,P - 800 = -20P+1600$$

It can be easily seen that at P = 20 the market is cleared and demand equals supply at 1200 units. At this price, a typical firm in Faridabad supplies 14 units and that in Jaipur 10 units. Notice that $14 \times 50 + 10 \times 50 = 1200$. As regards profits for a Faridabad firm, we have

$$\begin{aligned} \text{Profits} &= \text{Total Revenue} - \text{Total Costs} \\ &= P \times Q - (0.5Q^2 + 6Q) \\ &= 20 \times 14 - (0.5 \times 196 + 6 \times 14) \\ &= 280 - 182 \\ &= 98 \end{aligned}$$

Thus each firm earns a profit of Rs. 98. You can check that the firms in Jaipur make only Rs. 50 in this deal.

A fairly simple conclusion emerges from this analysis. If all firms are operative under identical cost conditions, output and profit are inversely related to the level of unit transport costs. Faridabad firms not only supply more quantity but they also earn more profit. In the long-run, you should therefore expect that either more factories get set up in Faridabad or the Jaipur ones transfer their machinery to that favourable location or a little of both.

Suppose now that the firms in Jaipur organise under the banner of JMA (Jaipur Manufacturers' Association) and represent their case before the Delhi Administration (DA). After a lot of persuasion the DA passes a law that so and so goods coming from Faridabad have to pay an entry tax of Rs. 2 per unit and the same goods coming from Jaipur will receive a subsidy of Rs. 2 per unit.

Activity 3

In the above example, after the imposition of entry tax and subsidy, both the locations after identical quantity of supplies and the customer does not suffer. What else happens? Why?

....

....

Taxation

Suppose a perfectly competitive industry consits of 100 firms with identical cost functions.

$$C = 0.1Q^2 + Q + 10$$

Assume usual profit maximising behaviour

$$P = MC = 0.2Q + 1$$

Thus if P is less than one, no supplies can be made and for other prices the supply function can be derived as:

$$P = 0.2\,Q + 1$$
$$\therefore 5P = Q + 5$$

Substituting S for the quantity supplied. We have,

$$S = 5P - 5$$

For the entire industry consisting of 100 firms the aggregate supply function is

$$S = 500P - 500$$

If the demand function is

$$D = -400p + 4{,}000$$

then the equilibrium quantity and price for the industry can be seen as

$$P = 5,\ D = S = 2000$$

Now, consider that a sales tax of Rs. 't' per unit is imposed. This is called 'specific'since it is fixed per unit. A tax which is based on value i.e. so much percentage of output value is called 'ad valorem'. The cost function for a typical firm then becomes

$$C = 0.1A^2 + Q + 10 + tQ$$
$$= 0.1Q^2 + (1+t)\,Q + 10$$

The industry supply function will then be

$$S = 500(p - t) - 500$$

The equilibrium price is given as

$$-400\, P + 4{,}000 = 500\,(p - t) - 500$$
$$\text{or } P = 5 + 5/9t$$

If the tax is 90 paise per unit then

$$P = 5.5,\ D = S = 1800$$

Let us analyse this result. Even though the tax is 90 paise, the price has increased only by 50 paise. What has happened to the balance 40 paise? The firm has to absorb this in its production costs.

Secondly, since the price has gone up, the equilibrium quantity for the industry as such dropped by 200 units. Some customers cannot afford Rs. 5.50 and would rather not buy the product. At the old price without tax, the typical firm supplied 20 units (5P-5, at P = 5) and earned a profit of Rs. 30.

After the imposition of tax, the typical firm would supply 18 units [5(5.5-0.90) -5] and earn a profit of Rs.22.40. The sales tax collection for the government is Rs.16.20 per firm.

Suppose the objective of the local government was to collect Rs. 16.20 from each firm by some means. As an alternative, it could impose a tax on profits at the rate of 54% and collect exactly Rs. 16.20. In this case the firm's after-tax profits will be Rs. 13.80. But then, more customers can be served, since at Rs. 5 market price the total quantity would be 2000.

Can we, therefore, conclude that it is better to impose 54% tax on profits since only 100 firms get reduced profits compared to 200 customers going without the product in case of sales of 90 paise? Can you compare the agony of 100 firms with the misery of 200 customers?

12.4 EFFECT OF PRICE AND QUANTITY CONTROL

Let us now analyse a case typically found in the Indian context. Items of daily consumption like sugar are made available to the customers through fair price shops at a price much below the prevailing market rate. This is accomplished by following a policy of what is known as '**dual price control**'. A part of the output is sold in the free market while a part is released to the ration shops at 'fixed price'. Every manufacturer supplies levy sugar to the government at the levy price and subjects the remaining amout to the market forces.

As an example, imagine that there are 100 firms in the sugar industry each having an identical cost function:

$$C = 0.1Q^2+Q+10$$

The demand for sugar is represented by

$$D = -\,400\,P + 4000$$

For the sake of simplicity, we have assumed the same functions that were encountered in the taxation problem of the previous section. The result is also known, that is, in equilibrium the price of sugar will be Rs. 5 per unit, the industry will supply 2000 units and there will be a profit of Rs. 30 for each sugar mill. Now let us impose price and quantity control as follows:

Each sugar mill must supply 30% of its production to the government at a fixed price of Rs. 4 per unit. The remaining 70% of the output can be sold in the free market.

Assume that by this action, the demand function is not affected and the costs are any way not going to change. What is the outcome of this policy?

Approach this problem in the following manner. Obtain the solution for free market price in identical fashion with the only change that the MR for each firm will now be

$$(.3) \times (4) + (.7)P$$

This is so because 30% of output is sold to government at Rs. 4 and the P that you get in market is available for the balance 70%. Using this MR value it can be seen that each firm will sell 19.2 units in free market at a price Rs. 5.20. Since 19.2 is only 70% of the total output, the part sold at levy price of Rs. 4 is 8.23 units. Thus each firm is compelled to produce 27.43 units for maximising profits. The situation with this price- quantity control is compared with the total free market in the table below:

Items	Free market	Dual Price
Price per unit	Rs. 5	Rs. 20 Rs. 4.00(Levy)
Equilibrium quantity per sugar mill	Rs.20	Rs.27.43
Industry Output	Rs.2000	Rs.2.743
Profit per sugar mill	Rs.30	Rs.20.20

The consequences of this dual-pricing policy can now be seen for each party in operation. The total industry produces more sugar but earns less profits. Customers who want more quantities of sugar than what the ration shops provide have to buy them at a higher price. Some people who could not afford any quantity at the old equilibrium of Rs. 5 are now in a position to buy some at Rs.-4.

Reflection: Sections 12.3 and 12.4 have helped us realise that although perfect competition rarely exists in real life, the analytical methods developed within the ambit of perfectly competitive market can throw light on important issues of public policy. Do not treat these examples and illustrations as what actually happens. However, the direction of decision-making can certainly be understood if such economic analysis is employed.

12.5 MONOPOLISTIC COMPETITION

We now know that many firms selling differentiated products provide the essence of monopolistic competition. Each firm in the industry strives hard to differentiate its products from the competitors be it soap or toothpaste or toy. Products of no two firms will be perceived as identical. This perceived differentiation gives each firm an element of control over the price it can charge. At the same time, the firm cannot expect to reap the benefits of a differentiated product too long since others can always duplicate the effort albeit with a time lag. Similarly, the price variation between two competing brands of a given product is also not too large.

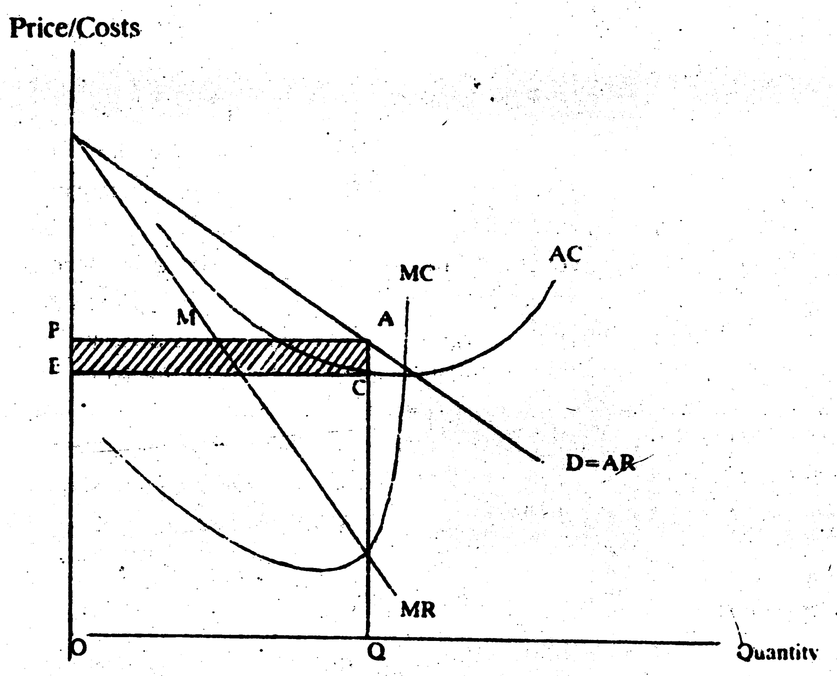
Monopolistic competiton therefore has several interesting aspects. The nature of competition is not restricted to variations in price and volume but extends to promotion, distribution, research and development also. **Cross elasticities** of demand for various products are fairly high. This means that a small upward variation in the price of say 'Colgate' may tempt many of its customers to switch to 'Promise'. If this happens, 'Colgate' may not revise its price downwards but instead would change its advertising or innovate on the product features so as to convince the customers that they are getting 'more value for their money'. As a result, you will notice that products in such market are close substitutes to each other.

All these elements reflect in a downward sloping demand curve for each firm in a monopolistic competitive situation. This is in sharp contrast to perfect competition where each firm faces horizontal demand curve.

Equilibrium in the Short-run

Figure 6 depicts the profit-maximising price-volume relation of a firm in a situation of monopolistic competiton.

Figure 6: Short-run Equilibrium of a firm in Monopolistic Competition



For the sake of simplicity, the demand (which is nothing but the average revenue AR) curve is assumed to be linear. The MR curve is below AR so that the distance of MR line is half that of the AR line from the price axis. The cost curves depict the usual U shaped structure of costs. The equality between MR and MC provides the equilibrium price-quantity combination. The maximum profit is indicated by the shaded area.

Suppose the firm's demand function is

$$P = 11{,}100 - 30Q$$

and the total cost function is

$$C = 4{,}00{,}000 + 300\,Q - 30\,Q^2 + Q^3$$

For this firm, let us equate MR and MC

$$\begin{aligned} TR &= \text{Total Revenue} = \text{Price} \times \text{Quantity} \\ &= P \times Q = (11{,}100 - 30Q) \times Q \\ &= 11{,}100\,Q - 30\,Q^2 \end{aligned}$$

$$\therefore MR = d(TR)/dQ = 11{,}100 - 60Q$$

$$MC = 300 - 60Q + 3Q^2$$

MR = MC implies

$$11{,}100 - 60Q = 300 - 60Q + 3Q^2$$

$$\text{or } 3Q^2 = 10{,}800$$

$$\text{or } Q = +60, -60$$

Since output cannot be negative, the profit maximising volume is naturally 60 units. At this output, the price will be (11,100-30 × 60) or 9,300. Please notice that at a volume of 60 units with a price per unit of Rs. 9,300 the profit for the firm is Rs. 32,000 during that period.

Conditions of Supply

Intense competition, freedom of entry (or the existence of low entry- barriers) often make life difficult for firms in such an industry. What happens if a given period offers weak demand? Whether the firm supplies any volumes at all or decides to close operations will depend on whether it can get price enough to cover its variable cost. The MR = MC rule yields maximum pro but the maximum that a firm can obtain may as well be negative. Generally, firms would not close down so long as variable costs can be covered by the equilibrium price obtained through the equality of MR and MC. When even that becomes difficult a product line may be dropped temporarily. But, there is a danger of the customer forgetting about that brand once it is out of the market for a while. The advertisement expenditure required to restore the brand may have to be reckoned along with the decision to reopen the product line. Alternatively some more differentiation could be added to a product, in which case it would face a new demand function. Notice 'Tomco' launching a detergent powder for exclusive use in washing machines to supplement the sales of the regular product. A new market segment was created, thereby putting the company at a slight advantage. You will have also noticed that just after a few months one more company (Levers) came out with similar product dampening the advantage.

These are normal happenings in a typical monopolistically competitive market. Since each firm faces a unique pric volume-promotion mix, the concept of an 'industry supply' function becomes ambiguous. Indeed non-price competition is very prominent in such markets. In the short-run, the total industry demand for all types of, say shampoos, will not change. Yet, market shares can undergo some changes. Rather than fighting the market battle on price and end up with low profits, firms choose to compete on non-price related factors-mainly promotion and product variation. We will have something to say on this later.

Long-run Adjustments

Long-run adjustments in a monopolistically competitive market can be compared to those observed under perfect competition. Many producers are in the market, often selling closely substitutable products. Any above-normal profit would invite new entrants and the firms will have to innovate almost continuously to maintain that position. Theoretically normal profits should previal in the long-run. Since the very concept of an industry is rather loose in such markets, defining normal profitability criteria for a typical firm becomes a difficult task. One thing can be surely observed. The nature of competition does not allow too many firms to earn above normal profits over a long time span. The margins are always under pressure and therefore the positive deviation from normal profits is always small. Firms constantly try and reduce costs, introduce minor product differentiation, change advertising strategies and try to maintain that small margin. Although entry-barriers are not too high some firms can manage to keep a lead over the others in patenting some unique product features, thereby acquiring a competitive advantage even over a long-term horizon.

In summary, no clear-cut rules of behaviour can be formulated for a market which is characterised by monopolistic competition.

12.6 ADVERTISING, PRODUCT VARIATIONS AND MONOPOLISTIC COMPETITION

Advertising

We have already seen that non-price competition is more likely to be a dominant feature in monopolistic competition. Advertising is an important element in the competitive strategy of such a firm. But it does not come free. Any expenditure on promotion pushes up the cost curves. Whether or not that has any impact on the demand curve is uncertain at the time of launching a promotional campaign. One only hopes that the expenditure so incurred would result in higher sales at given prices or hopefully even at higher prices. Again, rival firms do not keep quiet; they also have their own promotion strategies. In some cases, a great portion of sales promotion efforts by firms are self-cancelling; one firm's actions are matched by rivals resulting in only minor gains in sales and output. In such instances, unit costs go up all the same. Ideally, firms would like to have increased sales so as to more than cover the promotion expense leading to a net gain in profitability. In worst cases, promotion gets you nowhere—indeed there is a drop in profitability. The additional costs of promotion are more than the incremental sales revenue. Which outcome finally materialises cannot be forecast with perfect certainty.

Product Variation

Like advertising and sales promotion, this is also an activity with uncertain outcome. Firms introduce variations in order to inject an element of differentiation into their products. Whether or not market perceives it that way is the real test of a successful product variation move. The use of market research in such decisions is therefore gaining wider acceptance. The attempt is to have an idea of what features the customer wants before introducing any variations. Similarly, product related research gets used in offering something new to the customer which he may not have perceived earlier.

All such efforts are costly. Whether product related or market oriented, research entails expenditure. That pushes up costs. The effect of incurring these costs is quite analogous to the effect of promotion. The actions by different firms vary according to their perception of what is valued by the customers. Some firms adopt product variations which appeal to price-conscious customers (Nirma). Other firms would prefer to cater to the market segment desirous of high quality goods (Surf). Still other firms may pursue an intermediate strategy and have a product in between the two ends. Product variation is essentially a differentiation exercise related to the market segments to be serviced and the perception of the customers in each of them.

Before we go to next section, it will pay you to recapitulate that in terms of revenue and cost conditions, though a perfectly competitive market differs from a monopolistically competitive market, yet the adjustment process is more or less similar in both markets. Profit encourage entry as loss induces exit; eventually a situation of "no loss no profit" or what is called "normal profits" emerge in the long-run. Though the process is same, the outcome is different depending upon the nature of market, perfect or monopolistic. In Figure 7 and Figure 8, we have illustrated the long-run situation in perfectly competitive and monopolistically competitive markets. You may contrast the two situations in terms of Figure 9 where the relevant feature of Figure 7 and Figure 8 are combined.

Figure 7: Long-run Equilibrium in Perfect competition

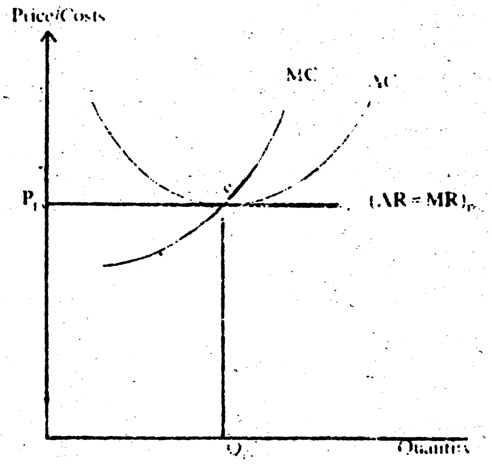


Figure 8: Long-run Equilibrium in Monopolistic Competition

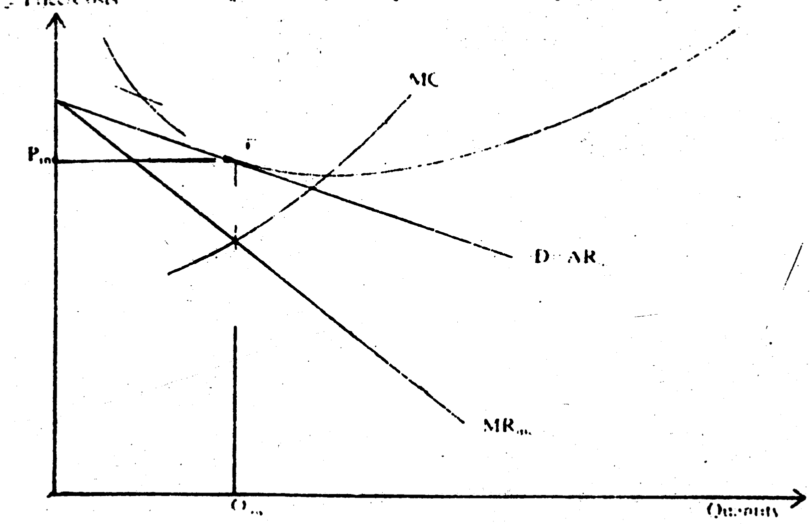
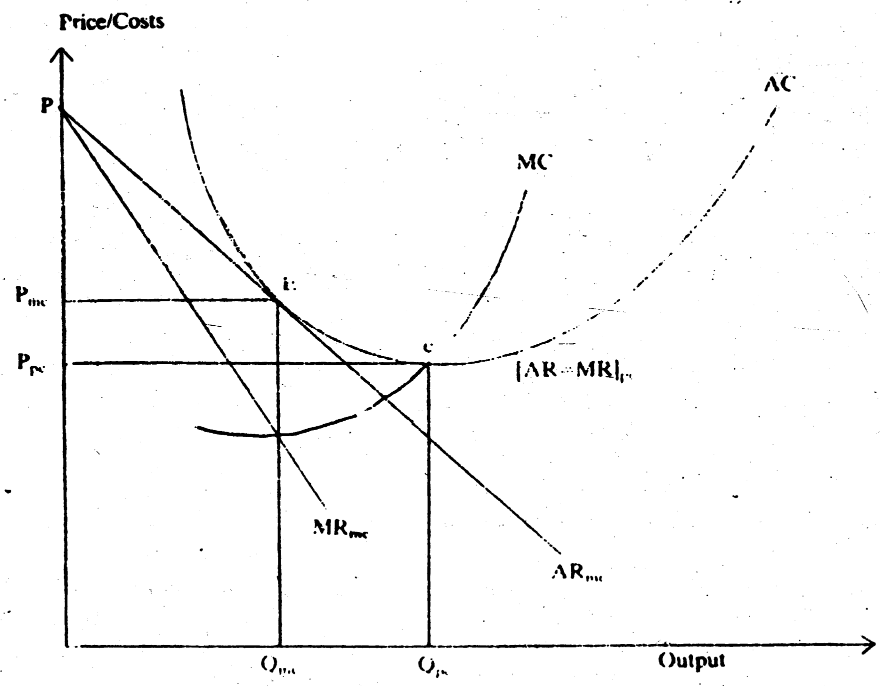


Figure 9: Long-run situation: Perfect vs. Monopolistic Competitions



You may note that in the case of a perfectly competitive situation, (Figure 7) both firm equilibrium condition (MR = MC) and the industry's equilibrium condition (AR = AC) are satisfied at single point C, whereas in a monopolistically competitive situation, (Figure 8) the "double conditions" are satisfied at two different points, (MR = MC) and (AR = AC). Comparing the two (in Figure 9), you may observe that price is higher (by $P_{mc} - P_{pc}$) and the scale of output, smaller (by Q_{mc} - Q_{pc}) under monopolistic competition than under perfect competition. The volume ($Q_{mc} - Q_{pc}$) is the measure of "excess capacity" associated with monopolistic competition. Such excess capacity is zero under perfect competition. Also note, under perfect competition, (P=AR=MR=MC), but under monopolistic competition, [(P=AR)>(MR = MC)]. Similarly, under perfect competition, [P=AC=MC], but under monopolistic competition [P=(AC>MC)]; the price-marginal cost difference measures the "degree of imperfection".

12.7 MONOPOLISTIC COMPETITION IN INDIA

The previous unit referred to the aspect of government policy acting as an entry-barrier in several industries. Beside, the growth of entrepreneurship is also a crucial element in the Indian context.

Until a decade or so ago, even products like soaps and toothpastes were characterised by oligopolies. For some reason, new firms just did not enter into several product lines despite favourable government policy. It is only since the 80s that one finds competition hotting up in the country's markets. Product variations, aggressive promotional campaigns, and easy entry of new firms are now commonly encountered in several consumer goods industries.

Activity 4

Identify two product groups where monopolistic competition can be found in India. Describe the competitive behaviour of the firms in these over the last one year (short-run) and over the last three years (long-run). What shape the competition would take in the near future? Please feel free and write a speculative note, based on present facts.

12.8 SUMMARY

In this unit, we have made an attempt to make you understand the market forces operating in perfect competition and monopolistic competition. The perfect competition is no doubt an idealised market environment. Though rarely found in reality, a systematic analysis of this market form does offer insight which can be used in formulating policies. Besides a careful study of perfect competition also helps in assessing market imperfections. It is more like a reference point fo the economic analysis of the real world markets which are often less than perfect.

The focus of managerial action is not on price-volume alone when one talks about monopolistic competition. Many sellers, differentiated products, existence of close substitutes are some of the distinguishing features of monopolistic competition. Each firm is moreworried 'bout differentiating its products to improve its distinctiveness and thereby gaining some competitive advantage. Unlike perfect competition, no clear-cut solutions or prescriptions can be recommended.

Even the concept of an industry is a little vague because no two firms are producing similar products. Naturally, anticipating market and understanding customer perception assume greater importance in monopolistic competition. Keeping this in mind you may undertake a comparison between two market situations and thereby comment on the optimality of the equilibrium attained.

12.9 KEY WORDS

Normal Profit is defined as a rate of return on capital just sufficient to attract the investment necessary to set up and operate a firm.

Economic Costs include normal profits.

Economic Profit represents on above-normal profit situation.

Equilibrium of a Firm (MR = MC) represents profit maximising price-output combination. In a situation where maximum profits mean a loss, the equation gives loss.

Equilibrium of an Industry is stated in terms of the condition of normal profit AR = AC such that the size and structure of the industry in terms of member of firms are strictly defined.

Comparing Price with AVC generally helps in reacting to a given market situation. (P = AVC) is the measure of "shut-down point".

Non-price Competition is a crucial parameter in understanding the bahaviour of firms under monopolistic competition.

High Cross-elasticities of demand is a characteristic feature of monopolistic competition.

Excess Capacity is typical of monopolistic competition, because even in the long-run firms do not produce least-cost output such that Price = AC, but price MC>AC<MC at the equilibrium level of output.

12.10 ADDITIONAL READINGS

Thompson, Arthur Jr., A.1977. *Economics of the Firm, Theory and Practice,* Prentice Hall.

Sawyer, Malcolm C and Milton Keynes. *Economics of Industries and Firms,* Second Edition, Croom and Helm.

Adhikary, M. *Managerial Economics,* 1987, Khosla Publisher, Ch. VI, section: 1,2,9,10.

12.11 SELF-ASSESSMENT TEST

1 a) Define monopolistic competition and give a few examples and
   b) identify the competitive and the monopoly elements.
   c) Why is it deficient or impossible to define the 'industry' in this case?

2 Discuss the long-run efficiency implications of a situation of monopolistic competition with respect to (a) utilisation of plant, (b) allocation of resources, (c) advertising and product differentiation.

3 Compare and contrast 'monopolisitc competition' with 'perfect competition' with respect to (a) adjustments and (b) outcome in terms of equilibrium.

4 Consider a product lilke "essential drugs". How will the price and output (both quantity and quality) of such products be affected if the Goverment decides to abolish the system of brand labelling?

5 Distinguish clearly between:
   a) Firms's equilibrium and Industry's equilibrium
   b) Short-run and long-run
   c) Equilibrium with and without excess-capacity
   d) Product differentiation and product variation
   e) Production costs and selling costs (like advertisement)

6 If perfect competition is too ideal to be real, then why should you study it? After all, you are all practicing executives concerned with real world.

7 Given the following information about demand and costs, determine the optimum levels of output (Q), advertisement (A), price (P):

$$P = 100 - 3Q + 4A$$
$$TC = 4Q^2 + 10Q + A.$$

(Answer: Q = 15, A = 900, P = 175)

UNIT 13 ANALYSIS OF MARKET STRUCTURE-SMAL GROUP CASE

Structure

13.0 INTRODUCTION

In this unit we shall continue our study of imperfect competition but we shall now move from the case of large group to the case of small group of sellers. We shall attempt an in-depth analysis of such market structures where a small number of sellers operate. Such markets may assume a variety of form like monopoly, duopoly, monopsony, duopsony and oligopsony, bilateral monopoly etc. An extreme case is just one firm. More commonly encountered situations have a number of firms, each large enough to have some control over the market, selling either differentiated or similar products. This part of imperfect competition is very complex and, therefore, offers a variety of situations. There are, however, a few clear-cut equilibrium solutions. A study of these situations would also help in appreciating the reality around us, particularly the present status of monopoly regulation in the country. Pricing strategy is at the heart of many business decisions. The discussion proposed in this unit should prepare us to take up this interesting issue in the next unit.

13.1 OBJECTIVES

On reading this unit, you should be able to:

- **identify** a variety of markets where small sellers (and buyers) dominate,
- **appreciate** the real world market situation in terms of an analytical framework,
- **analyse** the price-output decisions undertaken by a single/a few sellers (and buyers),
- **discover** the situation of market disequilibrium in some cases,
- **question** the profit maximising principle and its relevance in practice, and
- **prepare** yourself for discussing pricing practices and methods.

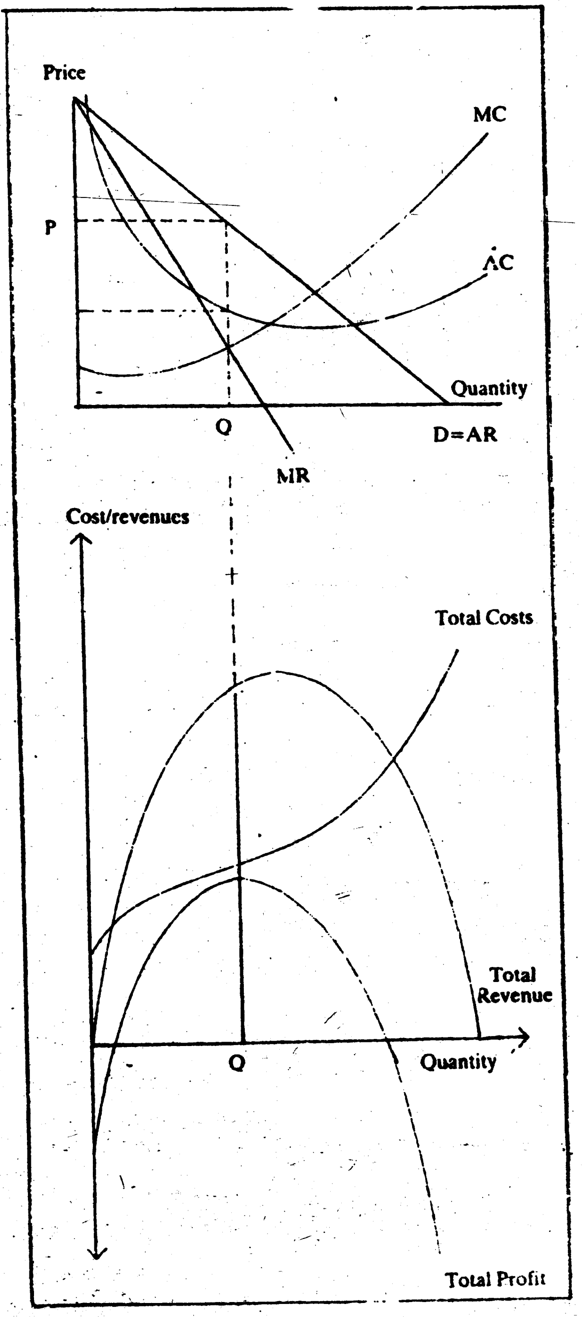
13.2 MONOPOLY

If perfect competition is at one extreme and of the market structure universe, the other end is characterised by monopoly. It exists when just one firm is the sole producer of a product which has no close substitutes. Just as perfect competiton is rare, monopoly is also rare in less regulated market economies. The public sector in India has significant monopoly elements. Analytically public sector monopolies have a different place in managerial economies and we shall not deal with them here.

Although monopoly is an extreme form of market concentration, its study helps us in analysing less extreme cases. Many of the economic relationships found under monopoly can be used to estimate optimal behaviour in the less precise but more prevalent, partly competitive and partly monopolistic market structures that dominate the real world.

Under monopoly, the firm is the industry, naturally, a monopolist faces a downward sloping demand curve. The fact that just one firm constitutes the industry imposes a crucial constraint on a monopolist. He can set either the price or the quantity but not both. Given a demand curve, if the monopolist decides to change the price, he has to accept the volume that it will accompany. Similarly, with the volume determination, the price gets automatically established through the demand curve. What will he do? He will operate at that level where his profits are maximum that is where marginal revenue equals marginal cost. You will notice from Figure

Figure 1: Profit Maximisation—Monopoly.



Source: Samuelson, p. 495.

that the graphical representation of the price-volume decision under monopoly is quite similar to the one seen under monopolistic competition (Unit No. 12, Figure 6). Also note that both price and quantity are simultaneously determined; hence the control over only one parameter.

13.3 APPLICATIONS OF MONOPOLY THEORY

The discriminating monopolist

The monopoly firm need not sell its entire volume at the price dictated by the demand curve. In some situations there is a possibility of selling in two or more markets at different prices and thereby increase profits. This is possible only if buyers in the two markets are isolated and do not have the chance to buy in a cheaper market and resell it in the 'high priced' one. A monopoly firm in the service industry will often manage to discriminate since a service cannot be resold. Further, price discrimination is also prossible between 'domestic' and 'export' markets

Consider a monopolist facing the following demand and cost curves.

$$P = 100 - 4Q, \; C = 50 + 20Q$$

Suppose the firm is able to separate its customers in two distinct markets with the following demand functions.

$$P_1 = 80 - 5Q_1, \; P_2 = 180 - 20Q_2$$

It can be easily verified that the aggregate demand curve remains unchanged at

$$P = 100 - 4Q$$

The two demand equations can be written in terms of quantities.

$$Q_1 = \frac{80-P_1}{5}$$

$$Q_2 = \frac{180-P_2}{20}$$

The total demand at any price P will be the summation of the two quantities.

$$\therefore Q = Q_1 + Q_2 = \frac{80-P}{5} + \frac{180-P}{20}$$

$$= 16 - 0.2P + 9 - 0.05P$$

$$= 25 - 0.25P$$

Solving this equation for P, we get

$$P = 100 - 4Q$$

For each market, the marignal revenue will be obtained as under:

$$P_1 \times Q_1 = R_1 = 80Q_1 - 5Q_1^2$$

$$MR_1 = 80 - 10Q_1$$

$$\text{Similarly } MR_2 = 180 - 40Q_2$$

In each market, the respective marginal revenues must be equal to the MC for maximum profits. Thus,

$$MC = \frac{dC(Q)}{dQ} = \frac{d(50+20Q)}{dQ} = 20$$

$$80 - 10Q_1 = 20 \text{ and } 180 - 40Q_2 = 20$$

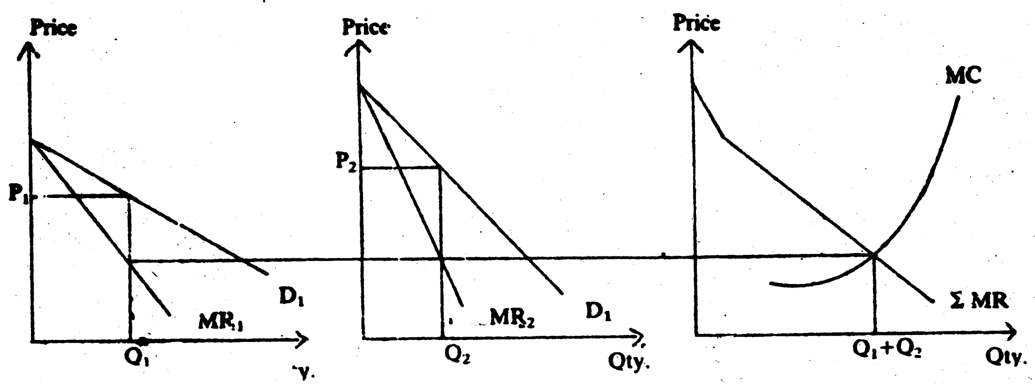
Besides, the combined marginal revenue (CMR) must also equal MC. The prices and quantities in the respective markets can be seen as

$$\text{Market} = 1 : P_1 = 50, Q_1 = 6$$

$$\text{Market} = 2 : P_2 = 100, Q_2 = 4$$

The typical discriminating monopolist is depicited in Figure 2. The MR curve is plotted by adding the respective MR curve horizontally.

Figure 2: Discriminating Monopolist



Market 1 Market 2 Combined

The maximum profit is calculated as

$$\pi = TR - TC = 300 + 400 - (50 + 20 \times 10)$$
$$= 700 - 250$$
$$= 450$$

Now, let us see what would happen if the firm were to face just one market. The MR equation in this case, is given by 100-8Q, the MC = 20 and hence profit maximising combination is P = 60, Q = 10. Profit is only 350.

Why has this happened? Let us compute the elasticities of demand at the equilibrium outputs in the two markets to understand the improvement in profit in the first case. The market which faces a demand curve $P_1 = 80 - 5Q$ exhibits an elasticity of demand which is 1.67 at $Q_1 = 6$ and the other market has only 1.25 at $Q_2 = 4$. You may calculate these coefficients by using the measure for elasticity σ = AR/AR-MR. This means the price is lower and the quantity higher in the market with greater demand elasticity. The monopoly firm has been able to exploit the differences in elasticities of demand to its advantage. As an extension of the two market cases, we can imagine that the firm will gain immensely if it were to subdivide its markets further on the basis of differences in the elasticities of demand.

Taxation and Monopoly Output

Now, consider a sales tax of Rs. 8 per unit on this monopolilst. Can he pass on the entire burden to the customer? The profit function of the monopoly firm will now be

$$\text{Profit, } \pi = (P \times Q) - C$$

$$= (100 - 4Q)\,Q - (50 + 20Q) - 8Q$$

$$= 100Q - 4Q^2 - 50 - 20Q - 8Q$$

$$= 72Q - 4Q^2 - 50$$

Set $\frac{d\pi}{dQ} = 0$ check for second order and $\frac{d^2\pi}{dQ^2} < 0$.

You can easily verify that his maximum profit in this case turns out to be 274, Price = 64 and Quantity = 9. Sales go down by 1 unit, price increases by only 4 units but profit declines by 76 units. The exchequer collects a tax of 72 units. Suppose, this same quantity is collected as lumpsum tax on profits. The monopolist will have to pay it but his profits would go down only to that extent. Since demand and cost conditions remain unchanged, price-volume would be the same as in the case when no tax was levied. Clearly, **a lumpsum tax is perferable to a tax per unit of sales in monopoly industries.** Customer does not pay a higher price, the profits of the monopolist decline by a lower amount and the tax revenue will remain the same.

Activity 1

Compare the taxation problem solved in Unit No. 12 (perfect competition) with the case discussed above.

Now let us change the objective constraint of our monopoly firm.

Illustration

A monopolist sets his goal as revenue maximisation. Yet, since he cannot ignore profitability he a minimum limit on his profits. His demand and total cost functions are:

$$P = 304 - 2Q$$
$$C = 500 + 4Q + 8Q^2$$

He must earn at least 1500 as profits and only after that, he will try and maximise his revenue. His profits are determined by the usual equation ($\pi = TR - TC$) - you may do this in either of the two ways:

i) Set it as a problem of constrained maximisation i.e., maximise TR subject to $\pi = 1500$. Thus the lagrangian function is:

$$L = [304\text{-}2Q]Q + \lambda\,[\pi\text{-}1500],$$

where π = TR-TC

$$= [304 - 2Q]\,Q\text{-}500\text{-}4Q\text{-}8Q^2$$

$$= 300Q - 10Q^2 - 500$$

and then check for $\frac{dL}{dQ}$ and $\frac{d^2L}{dQ^2}$

ii) The easiest way is to set $\pi = 1500$ whereby you will set two solution values for Q.

Verify that he will achieve his primary profit goal of at least 1500 at two levels of outputs Q = 10 and Q = 20. Which of the two will he choose given his revenue objective?

Activity 2

A monopoly firm uses only one input X to produce a given product Q. The input-output relation (called the production function) is given as

$$Q = 2\sqrt{X}$$

The demand function is P = 85 - 3Q

The input is purchased at a constant price of '5' per unit. What is the profit maximising price-volume combination?

(Hint: Convert the production function into a total cost function by using the price of input.)

....

....

....

....

....

....

....

....

....

....

....

13.4 DUOPOLY-COURNOT FORMULATION

As early as in 1838, a French economist Cournot analysed a special case of competitive business behaviour with only two firms in an industry.

The assumptions are quite strict but considering the time at which this formulation was developed, they cannot be faulted with too much. It is assumed that each member in this two-firm industry produces a homogeneous product, treats the rivals' output as given and maximises profit. We shall illustrate the equilibrium price-volume combination for each firm by taking a simple example. The rival firm's output behaviour with respect to one firm's output is called conjectural variation. Cournot assumed a zero conjectural variation.

Suppose, the total industry demand function was $P = 100 - 0.5\,(Q)$. Since the entire output is shared by just two firms, this can as well be written as $P = 100 - 0.5\,(Q_1 + Q_2)$. Firm number I for example has a constant cost function represented by $C_1 = 5Q_1$. Firm number II is having an increasing cost function $C_2 = 0.5Q^2{}_2$.

Each firm strives to maximise profits and therefore we can write the profit functions for them as:

Firm I's profit = π_1 = Total Revenue - Total Costs

$$\begin{aligned} &= PQ_1 - 5Q_1 \\ &= [100 - (0.5)(Q_1 + Q_2)]Q_1 - 5Q_1 \\ &= 100Q_1 - 0.5Q^2{}_1 - 0.5Q_1Q_2 - 5Q_1 \\ &= 95Q_1 - 0.5Q^2{}_1 - 0.5Q_1Q_2 \end{aligned}$$

For maximum profits we must have

$$\frac{d\pi_1}{dQ_1} = 0 \text{ and } \frac{d^2\pi_1}{dQ^2{}_1} < 0$$

Now remember that firm I makes variations in its quantities assuming that the quantity of firm II remains at a given level. This means that in computting $\frac{d\pi_1}{dQ_1}$ we must treat Q_2 as constant. When we do that, it can be readily seen that

$$\frac{d\pi_1}{dQ_1} = 0 \text{ gives us } 95 - Q_1\ 0.5Q_2 = 0$$

$$\text{or } Q_1 = 95 - 0.5Q_2$$

Cournot formulation calls this the **'reaction function'**. Reaction functions express the optional output of firm I as a function II's output and vice versa.

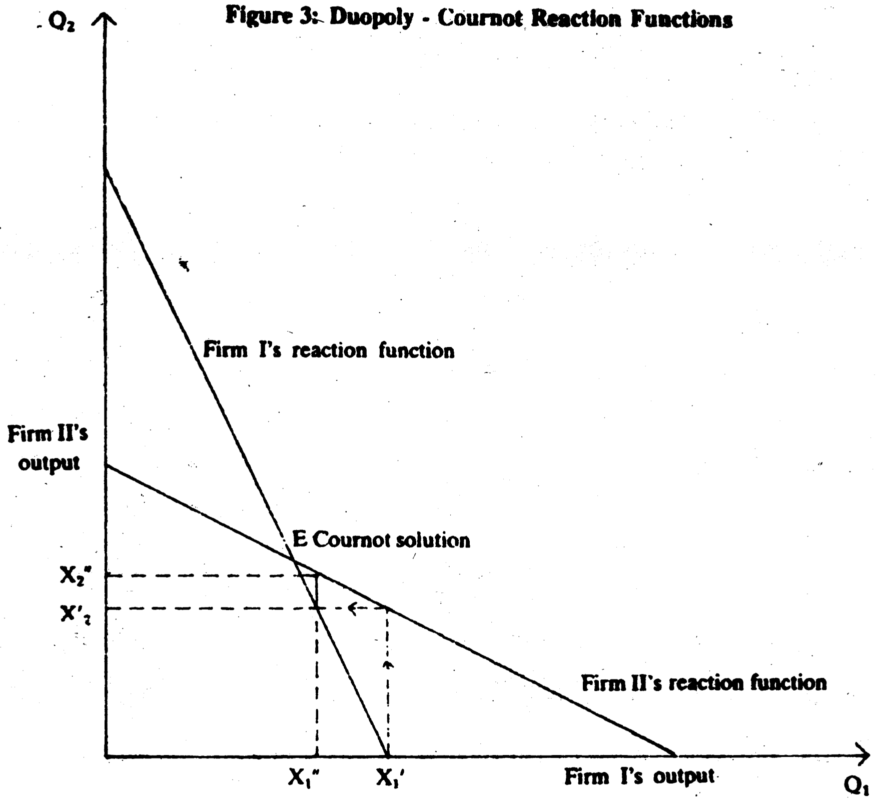
Given the value of $\frac{d\pi_1}{dO_1}$

we also note that $\frac{d^2\pi_1}{dQ^2} = -1$ which is negative as required for the maximisation condition. Similar algebraic manipulations for firm II should give us its reaction function as

$$Q_2 = 50 - 0.25Q_1$$

The typical reaction functions are depicted in Figure 3.

Figure 3: Duopoly - Cournot Reaction Functions



The solution of a duopoly equilibrium crucially depends on the nature of the reaction function of each duopolist. The equilibrium is reached when the values of Q_1 and Q_2 are such that each firm maximises its profit, given the output of the other and neither desire to alter the respective output. This is a very critical condition. From the reaction function that we have, it can be seen that if one firm increases its output, it will cause a reduction in the optimum output that the other firm can have. However, for a common solution, both the firms must achieve maximum profits and at the same time have no incentive for changing respective output levels. Such a solution is obtained at the intersection point of the two linear reaction functions (Please see Figure 3.)

Let us therefore solve the two reaction function equations in order to get the equilibrium solution.

Given $\quad Q_2 = 50 - 0.25Q_2; and Q_1 = 95 - 0.5Q_2$

Then, $\quad Q_2 = 50 - 0.25(95 - 0.5Q_2)$

Or, $\quad Q_2 = 50 - 23.75 + 0.125Q_2$

Therefore, $\quad 0.875 \quad Q_2 = 26.25$

Or, $\quad Q_2 = 30 \quad$ and $Q_1 = 80$

You can easily verify that P = 45 and firm I earns a profit of 3200 while that of the firm II is only 900.

Activity 3

Suppose both the duopolists in our case were to face the same constant function C = 5Q.

Under this assumption compare the price-volume relationships under perfect competition and also monopoly with that of duopoly. Your answer will look like this.

Decision Variable	Perfect competition	Duopoly	Monopoly
Industry Volume	190	$\left(\frac{190}{3} + \frac{190}{3}\right)$	95
Equilibrium	5	$\frac{110}{3}$	52.5

Notice that the duopoly equilibrium values fall in between those obtained under perfect competition and monopoly. Each duopolist exercises some control over the price but the degree of control is less than that of a monoporist. Further since the costs are assumed as identical and constant, an interesting relationship can be seen between the three solutions. What is it?

The Cournot Process

The assumption that the conjectural variation is zero appears to be quite arbitrary. A more important question pertains to the reactions of the rival firms. Is it that the two firms instantly discover the equilibrium point at the intersection of the two curves? Certainly not. Suppose that initially firm II does not produce anything and I produces an output equal to X_1'. Then firm II expecting I's output to remain at that level will set output at X_2' as indicated by its reaction function. (You will have to see Figure 3 as you read this). Now, firm I will respond according to its own reaction function taking X_2'. This action implies reducing the output from X_1' to X_2''. Similarly II will respond by increasing output to X_2''. The process will continue until the equilibrium is reached at E. Until then, each firm is in fact changing its output. Furthermore, the process may take a very long time to reach an equilibrium. You will get a feeling that this assumption of 'no change in the firm's output is rather shaky. The issue is still being debated and we must leave it at that for the present.

13.5 DUOPOLY–OTHER MODELS

Cournot assumed zero conjectural variation on output levels. One can imagine a situation with zero conjectural variation or price. This means, each firm takes the rival's price as given and sets its own price. Naturally every firm would like to set it as high as the market can bear. Soon each firm will realise that by cutting the price a little bit it can snatch the whole market. Once price cutting starts, it can go on till the firms reach competitive cost levels. A variation of this theme assumes capacity constraint. There is an upper limit above which each firm cannot increase its output in the short-run. It can be shown that under these conditions, no equilibrium solution is possible. The price oscillates between the monopoly price and some lower price for an indefinite period. The lower limit for the price is at least that obtained under the perfect competition but need not be so. This model is due to Edgeworth.

Stackleberg Model

A third type of duopoly analysis is suggested by the German economist **Stackleberg**. This is popularly known as leader-follower analysis. In this version, each firm has the option of either becoming a leader or remaining as a follower. A follower in this set up will behave like the firm in Cournot model, treating the leader's output as given. A leader knows that the follower is going to treat his (leader's) output as given and then proceeds to maximise profits given this assumption. Each firm would calculate its profits in both the alternatives-as a leader and as a follower-and then choose that role which gives greatest profit. You will realise that when both the firms do this they will naturally figure out that it does not pay to be a follower. If each firm decides to take on the role of a leader, no equilibrium can be reached. This is the symptom of "Stackleberg Disequilibrium"

Most models of duopoly contain an element of speculation of what the other firm will do. Indeed, all these can be extended to cover oligopolistic markets but the essential features will not change. The failure of the market to reach an equilibrium price-volume combination appears as the most striking feature in these models. It is quite natural to imagine that the firms would try to cooperate in some way in order to reduce the uncertainty of rival's actions. The cooperation may not be formal or clearly expressed as such. We shall have more to say on this in section 13.7 where some oligopolistic situations are analysed.

13.6 OLIGOPOLY - THE KINKED DEMAND CURVE HYPOTHESIS

We now come to probably the most intriguing part of the market structure analysis. When there are a few firms who sell either differentiated products or a homogenous product we say that the market is oligopolistic. Of these, differentiated products case offers interesting behaviour patterns amongs firms. The particular theory that we shall deal with in this section was simultaneously but independently developed by Paul Sweezy in the U.S.A. and Hall & Hitch in the U.K. around 1939. These researchers observed that oligopolistic situations lead to rigid prices. The price changes are infrequent. Besides, they are guided more by competitors' behaviour than by the objective demand and cost conditions. This is quite a difference between monopoly pricing where a change in demand and cost curves can be instantaneously matched by a price adjustment.

In oligopoly with differentiated products each firm has to make some intelligent guess about the competitors' response to a given action by the firm. The **Kinked demand** curve hypothesis states that rivals behave one way when a firm cuts its price, viz. match the cuts, but behave another way when a firm raises its current price viz. hold price constant at the current level. The behavioural assumption behind this theorisation is quite easy to appreciate. When one firm cuts its price, rivals do not want that firm to unfairly gain in market shares and therefore they will follow suit. On the other hand, when the price is increased, rivals think that it has given them an opportunity to grab more sales since they now are more competitive on the price front. Hence, a price increase is not followed by the rivals. The prices therefore tend to change infrequently, flexible downwards but inflexible upwards.

Figure 4: Oligopoly

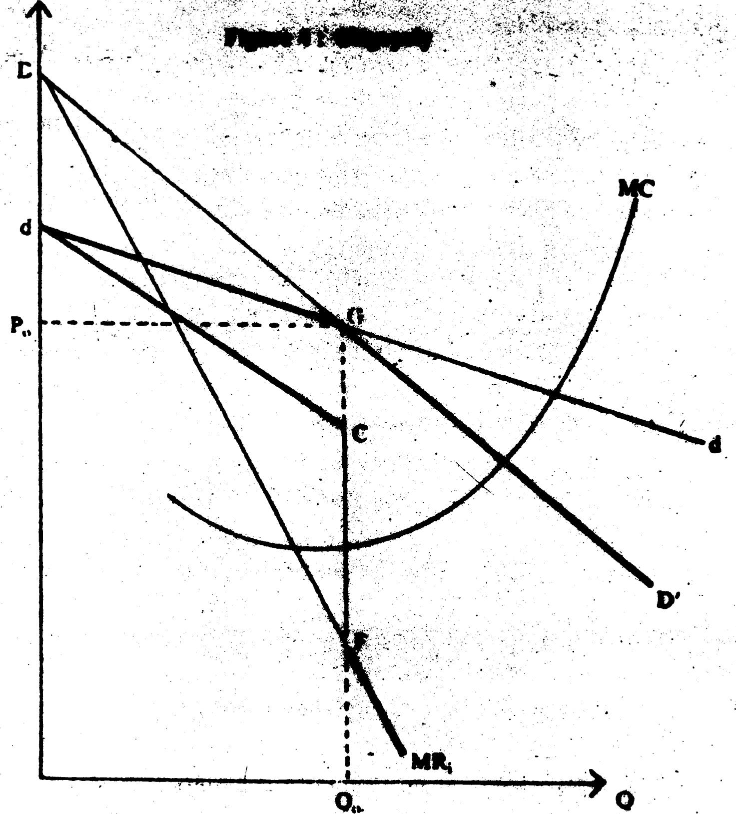
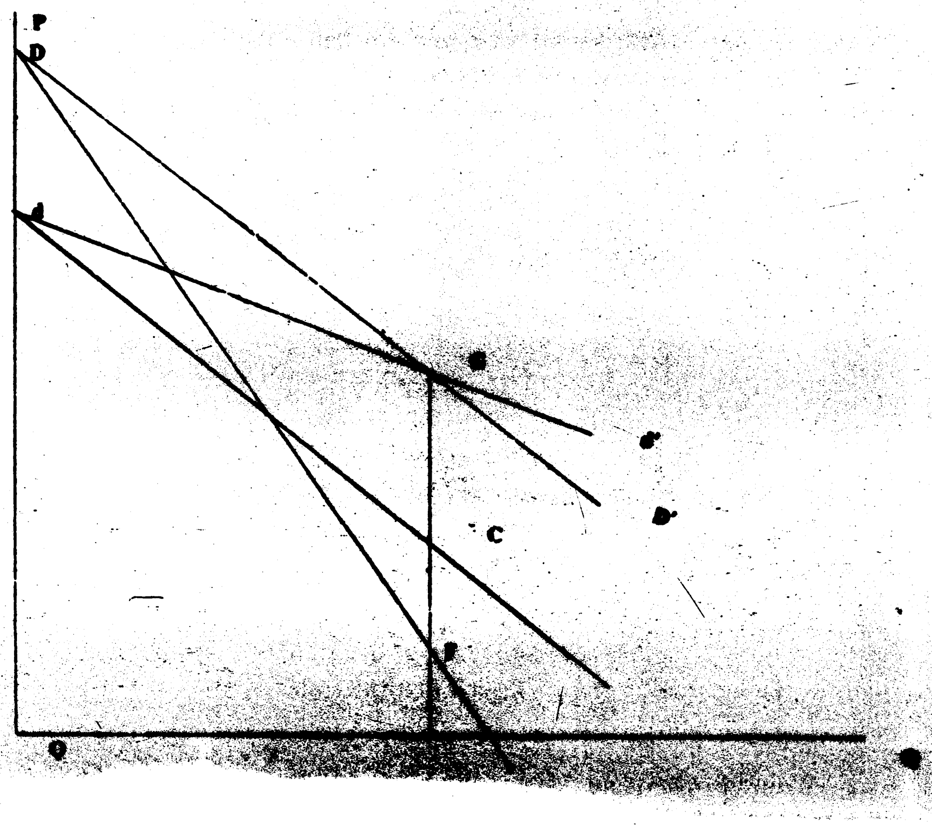


Figure 5: The difference in elasticities is less in this case and hence the inflexible price band narrower than in Figure 3.



Consider the situation depicted in Figure 4. Curve DD' represents the demand curve for a typical oligopolist on the assumption that rivals match its price changes both ways. The curve dd' is a more elastic demand curve for the oligopolist if rivals held their prices constant while the typical firm changed its prices. The dd' has to be more elastic because if rivals are keeping quiet on the firm's price increase they will get more sales and the firm would lose its share faster than when the rivals followed suit. Accordingly dd' is a flatter curve and DD' is a steeper curve. The 'Kinked' demand curve dGD' thus incorporates the assumption that price cuts are followed while price raises or not. Suppose the current price is Po. Below that level the typical firm would face the steeper curve (Portion GD') since all firms matching a price drop may not affect its sales too much. Above Po however, a more elastic demand curve is realistic since rivals would snatch away a larger cake from our friend. This portion is dG. There is thus a nodal point at the current price Po and demand curve takes the shape dGD'.

How would his marginal revenue curve behave? Corresponding to the steeper DD' we have DMR_1 and corresponding to the flatter dd' we have 'dG'. But, the demand curve has a corner at G. Therefore, if the price is below Po the firm would try to remain at the corresponding marginal revenue curve i.e. FMR_1. Above the price Po it will be on the MR curve relevant to more elastic portion i.e. dC. The marginal revenue curve is thus dC-gap-FMR_1. The gap or the kink or discontinuity is as large the vertical distance CF. Now, let the marginal costs be anywhere between C and F, the firm will have no reason to change its current price and quantity. In short, over a wide range of the cost curves where the MC cuts MR at some point in between the 'gap' of the MR curve (dC-gap-FMR_1), the firm will perceive the current price to be optional.

Otherwise, it will follow a path dictated by the respective portions of the kinked demand curve. This is often referred to as the phenomenon of 'sticky' or 'inflexible' prices. If the difference in the elasticity of the two demand curves is large then the width of the kink will also be large and prices would tend to be inflexible over a wide zone. Converse will be the case if the elasticities vary only slightly. Compare Figure 3 with Figure 4 and you will appreciate the importance of this observation.

Real Life Cases

You may ask: So far so good but do firms really behave this way? Empirical research done abroad offers mixed evidence. Whereas Hall & Hitch found some support for the 'stickyness' in the prices of actual oligopolies studied in the U.K. a study by Stigler showed that in two monopoly industries (aluminium and nickle) in the U.S.A. the prices were more stable than some oligopoly industries. In fact, Stigler's observations are that his study does not support Kinked demand hypothesis at all. Rather than review all the literature in this regard, we shall conclude this section by saying that while serious empirical work continues, this hypothesis provides a good starting point for the study of oligopolistic pricing situations.

13.7 OLIGOPOLISTIC COORDINATION, CARTELS AND PRICE-LEADERSHIP

Characteristics of Oligopolistic Markets

While discussing the duopoly case, it has been pointed out that in many such situations, the market fails to reach an equilibrium. This may be so technically. Yet, we do come across market structures which are quite stable. How is that markets function at all, if no equilibrium is possible? Or, is it that they are constantly in the state of disequilibrium? The nature of competition in an oligopolistic market is such that the rivals' actions are constantly weighing on the minds of the decision-makers in any firm. The uncertainty of future demand and customers' responses add to the complexity of the decision making process. As a result, in an oligopolistic market structure, there is no neat, simple and clear-cut equilibrium position towards which all firms tend to move such as those in a perfectly competitive market. Many variables are at the command of a firm—product features, price, service, promotion, to name a few. Secondly, given a competitive situation, several different and feasible courses of actions are open for the firms. Firms rely upon differences in price, quality, reliability, service, design, product development, advertising and product image to promote sales and increase profits. In view of this complexity, a number of plausible competitive situations can prevail in the market. Quite naturally oligopoly theory consists of dozens of models, each depicting certain features of oligopolistic conduct and performance but at the same time none telling the complete story of what constitutes competition among a small number of firms.

We have just seen one such formulation in the form of 'Kinked demand' theory. Presently, we shall discuss a few more models, which have got practical relevance. The economists have also developed sophisticated abstract models like Game Theory (See Annexure 3) as an explanation of strategy and tactics employed by an oligopoly/duopoly firm.

Dominant Firm Model

Recall the problem posed at the beginning of section 12.1. If there is one dominant firm in an oligopolistic market and the rest of the firms act as followers we will have a mixture of monopoly and perfect competition. The followers take market price as given and set their MC's to that price in order to maximise profits. The dominant firm acts as price leader and maximises profit by taking the supply curve of the followers as given. The dominant firm acts as monopolist constrained only by the supply of the rest of the lot(called the fringe firms). It can be shown that the presence of some firms which offer products at competitive prices dampens the degree of the dominant firm's control over the market price. If the market share of the followers goes up, the monopoly power of the leader suffers accordingly. Thus, in this formulation the equilibrium price is lower than what would be obtained by a pure monopolist. What are methods by which dominant firm can maintain its dominance in the market? Some well known responses are:

1) try and keep the industry price low enough to deter entry and also make expansion of fringe firms unattractive.
2) innovate on 'non-price competitive areas' - promotion, distribution, after sales service etc.
3) a defensive strategy involving confrontation with the aggressive fringe firms.

In order to illustrate a real life competitive situation we reproduce below a case. It is taken from '**Arthur A. Thompson,** *Jr. Economics of the Firm, Theory and Practice*' (pages 422-423) (enclosed as Annexure A).

Market Share Models

Other models of oligopolistic behaviour centre around the concept of market share variations. In all these models some reasonable assumptions are made about the cost conditions. As it happens, pricing conflicts emerge no sooner than firms start manipulating their prices to gain market shares. Depending on the degree of product differentiation and the price elasticity of market demand, the pricing strategies will be worked out. In general, you will observe that firms are not excited about cutting prices for gaining market shares. In the Indian market, the initial introduction price of the new variety of two wheelers like Kinetic Honda was higher than the popular Bajaj Scooter. After a while, Kinetic Honda reduced its price but also introduced a variation in the product features. It thus started selling two brands. Similarly, Bajaj Scooters are now more aggressively sold on non-price features like ruggedness, easy service, time-tested vehicle and so on.

Around 1986-87 it appeared that a price war may emerge in the two-wheeler market but eventually the fear did not materialise. Firms have relied more on product and promotion related competition than price related factors. You can identify several instances in the Indian industries where oligopolists have shield away from price competition and shifted the focus to some other variables.

Oligopolistic Coordination

Considering the realities of the market and the need to earn minimum acceptable profits, oligopolists are better in a 'cooperative mode' rather than in a competitive one. This does not mean that they do not compete. They do but with the understanding that there is a greater incentive in coordinating their actions. The cooperation may be subtle, non-formal and manifestly unnoticeable. There are many clues to this phenomenon. Often, the prices of competing brands in an oligoplistic market tend to move in a restricted range. This is so because no firm can set its price without any regard to those of the competitors. Secondly, the price revisions are not arbitrary and their timings has some well behaved patterns. Thirdly, most oligopoly markets have powerful 'industry associations' through which firms discuss issues of common interest, influence public policy and interact with the customers. If you carefully analyse the functioning of these associations the presence of such factors can be observed. For example, we have in India, fairly vocal associations of tyre firms, synthetic fibre manufacturers and cement units.

In the extreme case, oligopolistic coordination can be so perfect that all the firms may be able to act as a monopolist and maximise joint industry profits. This is called a cartel where all the firms 'administer' price output decisions jointly. The OPEC is the most prominent cartel in the world today. A case is reproduced here, taken again from the same book referred to above (pages 415 to 417) (enclosed as Annexure B).

The extent of olipolistic coordination in a market is likely to depend on a variety of factors. Legal framework of the business (MRTP Act, Companies Act, etc.) nature of demand and cost conditions level of entry barriers, attitudes of the managements are some of the important variables that influence coordination. For example in the Indian commercial vehicle industry, the entry barriers are quite high and a single firm controls about 50% of the market. Hence, not much coordination is observed. Similarly, if the market demand is booming and there is enough room for at least some firms to expand, coordination will be difficult. The ideal conditions for coordination are provided by a combination of weak demand, excess capacity, low entry barriers and rather weakly differentiated products. The reason for a well coordinated oligopoly in the Indian tyre, cement and synthetic fibre industries can be found in these conditions.

Activity 4

Tyre industry has come under severe attack for forming a cartel. Analyse the price volume, profitability and other variables for each firm and the industry as a whole to ascertain the validity of this criticism.

....

....

13.8 MONOPOLY REGULATION

The need for monopoly regulation is felt in every country. Oligopolies with dominant firms or cartels can have damaging effects on the markets. Prices can go up without any check or volumes can be restricted. It is therefore found necessary to put a regulatory control over the power of large firms so that customers do not suffer. In most countries, actions like takeovers, mergers and selling practices are open to scrutiny by some regulatory authorities. The MRTP Act in India regulates much more. It has under its umbrella all decisions pertaining to expansion, diversification, new locations, marketing practices and pricing apart from mergers and takeovers.

A complete review of the regulatory aspects will be made in a subsequent unit. Suffice it to say for the present that firms in oligopoly and near monopoly situations have to reckon with government regulation in any country. Considering the possible price-volume outcomes, regulation seems inevitable. The degree and mechanism of control however changes from country to country.

13.9 SUMMARY

In this unit we made an attempt to understand the competitive market structures where a few firms operate. As long as the number is small enough so that each firm is of a size which enables it to have some contorl over the price-volume mix, oligopolistic market is said to be present. We saw that apart from monopoly and some restricted versions of duopoly, the market structure with a few firms is not amenable to clear-cut and neat equilibrium solutions. There are many possible outcomes and each can be equally likely depending on the particular situation in the market. The existence of coordinated behaviour under oligopolistic situations was also observed. In the presence of considerable uncertainty about the competitor's actions and consumers' reactions firms find a natural incentive in cooperation. However, some objective conditions like price elasticity, demand, cost, entry barriers and regulation restrict the scope of such cooperation. In the process of analysis here, we have reviewed quite a few models. In a subsequent block, we shall make reference a few additional modles - Managerial and behaviour models and see if we can replace or reinforce the profit maximisation hypothesis.

13.10 KEY WORDS

Conjectural variation refers to the view that one firm takes towards the output/price variations of the rival firms given a certain action on its own output/price.

Reaction function is a relationship expressing one firm's optimal output as a function of the other firm's optimal output in the Cournot model.

Kinked demand curve theory assumes that firms in an oligopoly market follow price cuts but maintain their prices at existing levels with respect to a rival increasing price.

Kinked demand curve theory suggests that oligopoly prices tend to be sticky at the existing levels.

Oligopoly situations by and large offer incentive for firms to compete within a broadly cooperate framework. Cut-throat competition may not be typical.

Degree of monopoly can be measured by price-marginal cost deviation, e.g.

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

Dead weight loss is the triangular area which is shaded.

13.11 ADDITIONAL READINGS

Pappas, J.L. and Hirschey, M., Saunders Holt. *Fundamentals of Managerial Economics*, (International Editions, Second Edition)

Clarke, Roger, *Industrial Economics*, Blackwell Basil.

Henderson, J.M. and R.E. Quandt, 1971. *Microeconomic Theory*, (International Student Edition, McGraw-Hill).

Thompson Jr., Arthur A., 1987. *Economics of the Firm*, Theory and Practice (Prentice Hall, Second Edition)

13.12 SELF-ASSESSMENT TEST

1 a) "A monoplist can always make profit, be it under rising costs or falling costs or constant costs." Draw a set of three diagrams to illustrate this statement.

b) In theory, a monopolist should make profit, but in practice some monopolists (like the State Electricity Board or Delhi Transport Corporation or some other public sector units) are often found to incur huge losses. How would you explain this situation?

2 Prices tend to be rigid in an oligopoly market, despite fluctuations in demand and costs. Attempt of explanation of this rigidity using the concepts of 'Kinked demand curve' and 'price-leadership'.

3 Review your understanding of the following concepts; a) Stackleberg disequilibrium, b) Deadweight loss under monopoly, c) Degree of monopoly, d) Price equilibrium, e) Price-follower, f) Non-price competition, g) Oligopolistic cartels, h) Reaction function.

4 Assume that a) the two identical firms in a purely oligopolistic industry agree to share the market equally, b) the total market demand function is Q = 240 - 10p, and c) cost

schedules of each firm are given in the table and factor prices remain constant. Show that this market-sharing cartel reaches the monopoly solution. What are the total profits of the cartel? Is this solution likely to occur in the real world? (Attempt graphically or set MR = IMC)

(Ans. Q = 80, P = 16, π = 240)

	Firm 1	Firm 2	Total
Q	40	60	80
MC	8	12	16
AC	13	12	13

5 Given the demand function $P = (10 - Q)^2$ and the cost function, $TC = 55Q - 8Q^2$, find the maximum profit. What would be the effect of an imposition of a tax of Rs. 9 Rs. per unit quanuty on price? [Ans. $\pi = 54$: Due to tax, price will go up by Rs.15(64 - 49)]

6 Suppose the market demand is $P = 140 - 0.6Q$ and the total cost functions of the duopolists are $C_1 = 7Q_1$ and $C_2 = 0.6Q_2$

(Hint: $Q = Q_1 + Q_2$: derive reaction functions and solve simultaneously)
(Ans. $Q_1 = 93\frac{1}{3}$ and $Q_2 = 35$)

7 Given two isolated markets supplied by a monopolist, let the two corresponding demand function be $P_1 = 12 - Q_1$ and $P_2 = 20 - 3Q_2$, and the total cost function be $C = 3+2(Q_1 + Q_2)$. Find price, output and profit under, (a) price discrimination and (b) no discrimination.

[Hint: For, b) take $P_1 = P_2$ as the constraint and set of a profit maximising Lagrange function.]

(Ans.

variables	with discrimination	without discrimination
Q_1	5	4
Q_2	3	4
P_1	7	8
P_2	11	8
π	49	45

Annexure 1

The Life of the Dominant Firm: Procter & Gamble Versus Colgate.

The competitive life of a dominant firm is not always a bed of roses. Its very bigness and success not only attracts competition but also creates areas of vulnerability.

The situation of Procter & Gamble in trying to defend its number-one position in soaps, toiletries, and food products is a case in point.

Procter & Gamble virtually invented the laundry detergent business in1946 when it introduced Tide, "the revolutionary washday miracle" that could wash everything from clothes to dishes. Within three years P & G had a quarter of the laundry detergent business. Soon thereafter it introduced Cheer and, then, "a low sudser", Dash, telling housewives that while Tide, a high sudser, was best for top-loading machines that Dash was better for front-loaders. Other companies followed P&G's lead in detergents, but neither Colgate nor Lever Brothers could match P&G's headstart and ability to capitalise upon volume to achieve production efficiency.

The difference in cleaners and detergents, from product to product, lies in the technical formulation of essentially similar basic ingredients. Regardless of special qualities, each product is made using the same fundamental chemical manufacturing equipment. To a manufacturer, therefore, it makes little difference how much of each particular product is sold; the key is **total** volume. Except for chemical ingredients, it costs a firm little more to operate at 100% than at 50% of capacity. This accounts for why P&G, with its big volume, was and still is the most profitable detergent-maker.

However, the boom in detergents has long since matured. No longer is the market wide open; it is crowded. A new detergent that becomes a market success has to do so by taking sales away from brands already on the market. At the same time, older products need bigger advertising budgets toprotecttheir market shares from erosion.

According to Colgate's president, "Proportionately, it costs all of us the same to get extra business with advertising and promotion, but at this point we have more to gain by attacking than what it costs Procter to defend".

Colgate being the underdog in most products, can pick and choose where it wants to hit the giant; the giant, by contrast, must defend itself everywhere-from whichever firms elect to challenge. In this regard, a **Forbes** articles alleged:

> It's not unlike the situation in guerilla warfare, where the guerillas, with a lot less to defend, can concentrate their forces on one or two points and take a heavy toll. This is what Colgate has been doing; attacking first one P&G strong point, then another, hoping to keep P&G off balance. In fact, Colgate brass thinks its giant rival may be spreading itself too thin. Seeking to accommodate its expansion instincts and heavy cash flow, P&G has moved products outside the soap and toiletries field: food, paper products, coffee. In so doing, P&G is not only dispersing its energies, it is also moving against well-entrenched competitors like General Foods and Scott Paper. Thus....Procter is fighting on additional fronts at a time when it is under heavy pressure on the home front.
>
> "Let Procter fight it out with a rough customer like Scott. I (Colgate's president) want them to take on General Foods in coffee. It is also costing them plenty to get started overseas where we'have been for years. As we continue to put pressure on at home,

Procter will have to make a decision on whether to push back at us in a big way or go deeper into things like paper and coffee. I dont's think, big as they are, that they can do it all."

* "Colgate vs. P&G," Forbes (February 1, 1966), pp.26-29

OPEC: A Case Study in Joint Profit Maximisation

Although created in 1960, it took the Organisation of Petroleum Exporting Countries (OPEC) 13 years to achieve its two main goals: to raise to taxes and royalties earned by member governments from crude oil production and to take control over oil production and exploration away from the major international oil companies. Since 1973 OPEC, in closely controlling the world market price of crude oil, has emerged as perhaps the most successful *cartel* in world history. Cartels are one species of oligopoly; what distinguishes a cartel from "looser" oligopolistic market structures is the existence of a formal, explicit, and detailed plan for market-sharing and coordination of production levels and prices."**

OPEC consists of 12 countries (in order of the size of their estimated oil reserves): Saudi Arabia, Kuwait, Iran, Iraq, Libya, United Arab Emirates (Abu Dhabi, Dubai, five others), Nigeria, Venezuela, Indonesia, Algeria, Qatar and Ecuador. The top six OPEC countries have over 50% of the estimated world reserves of crude oil, and all 12 countries together account for more than two-thirds of world oil reserves. The OPEC countries are too small to use more than a friction of their own oil reserves and thus are major exporters. Because they account for more than 85% of world trade in oil, they can literally control the world market. Europe and Japan are totally dependent on OPEC oil, and the United States imports about one-third of its oil from OPEC sources.

Theoretically, the purpose of a cartel is to maximise the earnings of its members. This is accomplished by ignoring internal differences and preferences among members and setting a price that maximises the profits of the group as a whole—joint profit maximisation. When consumers are unable to substitute readily for the cartel's commodity or do without it, the price can be raised quite high without—in the short run—a great loss in volume of sales, with the result that total revenue is increased, perhaps dramatically. Indeed, this is precisely what OPEC accomplished when in 1973-74 it established a world oil price that gave participating governments over $10 a barrel net revenue—a ninefold increase in four years. The OPEC countries caused the price of crude oil to jump by raising the excise tax on each barrel of oil produced in their oil fields. These taxes are treated as a cost of production by the oil companies drawing oil from OPEC wells so that increases in the excise tax effectively raise the price which the oil companies must receive in order to cover production costs plus the tax.

OPEC's dramatic profits were made possible by the fact that the short-run price elasticity for oil has been estimated at −0.15. With demand so price inelastic in the short-run, the OPEC price increases have been successful in raising OPEC revenues to over $1000 billion per year.

** Cartels flourished between World War I and World War II, especially in Germany where the legal system condoned and helped to enforce market-sharing agreements. Such industries as chemicals, explosives, glass, steel and pharmaceuticals were prone to cartel organisations because of the large scale of operations required and the strong tendency to vertical integration from raw materials to distribution of the product. After World War II, industrial cartels virtually disappeared, partly because of the influence of U.S. antitrust regulations but mainly because of the rapid growth in world demand which made cartel organisations unnecessary. There did exist a number of "international commodity agreements" in such areas as tea, coffee, sugar, tin, copper, coco and bauxite, but most of these were only marginally successful and were generally undermined by the availability of substitutes or by the impossibility of preventing cheating by countries whose economic welfare was crucially dependent on the export of a single commodity.

Nonetheless, with the passage of time, perhaps five to ten years, the OPEC cartel faces some longer-run elasticity problems. The emergence of new forms of energy, coupled with energy conservation measures, give consumers more opportunity to curtail their usage of petroleum-based energy. Additionally, high oil prices and increased profits have stimulated new efforts to expand exploration and production of crude oil, to the extent that alternative sources of crude oil supply are developed, OPEC's control of the market is undercut. Long-run price elasticity of demand for petroleum has been estimated as close to -1.0, where higher prices are completly offset by loss of sales volume and total revenue is unchanged.

OPEC has recognised the danger to its position and has commissioned a number of studies to help it determine the profit-maximising price level and the pattern of price change that it should impose in the future. It has also initiated studies to determine when and how OPEC can cut back on production if demand weakens as expects. Some estimates project that OPEC will have as much as 25 to 33% excess production capacity by 1980. How OPEC would handle a surplus of this magnitude is a major concern because if some of the member countries should become dissatisfied with their assigned quotas, the seeds of discontent and dissension could break up the cartel. Venezuela has long suggested a pro-rationing scheme that would formally assign each OPEC country a rate of production based upon population, economic needs of the nation, and oil-producing capacity; however, several other OPEC nations have oppossed such criteria.

There are two major groups of countries in OPEC, and their interests tend to diverge on price and production policies. On one side are such countries as Iran, Venezuela, Iraq, and Algeria with large populations and ambitious economic development plans. These countries want maximum revenues now and are not overly concerned about the erosion of OPEC's market by high prices over the long term. Moreover, they have substantially fewer years of reserves left at current production rates. The second group of countries, which includes Saudi Arabia, Kuwait, and the United Arab Emirates, now has more money coming in then it can use and is more concerned about maintaining the long-run viability of OPEC's market control.

A number of observers are of the opinion that OPEC's price of oil in the mid-1970s overshot by a wide margin the price of oil which will prevail in the 1980s. They believe that OPEC is following a very sophisticated strategy of price discrimination based upon time. In their view, OPEC has calculatedly chosen to charge high prices now because alternative energy sources are virtually nonexistent and because the alternatives on the horizon have long lead times and require massive capital investments. Supposedly, OPEC will lower slowly over the future not only to discourage development of new energy alternatives but also to create uncertainty over where the price of crude oil is going to end up and confusion over just how profitable alternative energy sources will ultimately be in comparison to oil. By deliberately trying to increase the profit risk inherent in developing energy alternatives, the cartel hopes to forestall the emergence of substitutes for oil.

If this view is correct, then the internal strains on the cartel of cutting back oil production in order to keep prices propped up may be avoided. Nonetheless, others profess confidence that the internal differences within OPEC in terms of oil reserves, population, economic development plans, and military ambitions will eventually cause the cartel to break up. Although acknowledging that success of the cartel has probably made the participating countries more alert to the dangers of each country trying to go its own way by cutting price in order to reach its desired market share, they still regard cartels as inherently unstable whenever underlying economic turn unfavourable.

Experts have generally used a three-step analysis to determine the internal strain which may be placed on OPEC countries. Step 1 involves predicting world demand for oil each year to 1985.

Step 2 is to estimate how much oil will be available from non-OPEC sources, including increased production capacity within consuming nations (such as oil from offshore United States and the Alaskan North Sea slope). Step 3 is to subtract the estimated non-OPEC supply from estimated world demand to give a figure for the world market open to the OPEC countries. By comparing this to the total productive capacity the OPEC nations will have in place on an annual basis, one gains an indication of how great a problem OPEC would face in allocating output among his members.

Will OPEC be a model for others to copy? When OPEC succeeded in gaining control over the world price of crude oil, there was widespread speculation that what could be done with oil could also be done with other staple, basic commodities such as copper, coffee, and bauxite. In fact, an association of copper-producing nations was formed. But in 1975 copper prices declined about 65% over prevailing 1974 levels. The copper cartel met in late 1975 to discuss how to stop this price erosion; however, the consensus which emerged was that little could be done by the cartel to stop it. The falling copper price was attributed to a surplus of copper-producing capacity resulting from the 1974-75 worldwide recession.

Questions for Discussion

1. What arguments can you give to support the view that cartels are *inherently* unstable when underlying economic conditions will not support the cartel's price?

2. Why do you suppose the oil cartel has been successful, but other cartels have in recent years been generally ineffective? Is agreement among the cartel's members sufficient to guarantee success, or does it take something more, i.e., strong, inelastic demand for the product?

UNIT 14 PRICING STRATEGIES AND TACTICS

Structure

14.0 INTRODUCTION

The previous three units of this block were mainly concerned about developing an appropriate theoretical framework for the pricing decisions. It would now seem proper to see how firms in the real world set their prices. Do they rely on the marginalist principle (that is, equality of marginal revenue and marginal costs) while pricing their products? What role is played by the industry environment demand, capacity, competition etc?

Without stating so categorically, all along we had been assuming that product and sell only one product. Quite a number of firms produce two or more products, either at the same plant or at different locations. We begin our study in this unit with a simple case where a firm produces two products. We then go on to analyse the various aspects related to the pricing strategies of firms. Pricing of services (like software development, consultancy, medical care etc.) has a different set of problems. We try our hand at this case also. A brief review of the pricing methods applied by the apex statutory body on pricing in our country will round off our study on pricing strategies.

14.1 OBJECTIVES

The objectives of this unit are help you to:

- **understand** the pricing policies following by firms in actual practice,
- **analyse** the role of various factors in the pricing decisions of the firms,
- **state** clearly the relationship between the pricing objectives and the goals of firms, and
- **appreciate** the role of statutory pricing bodies in the Indian context.

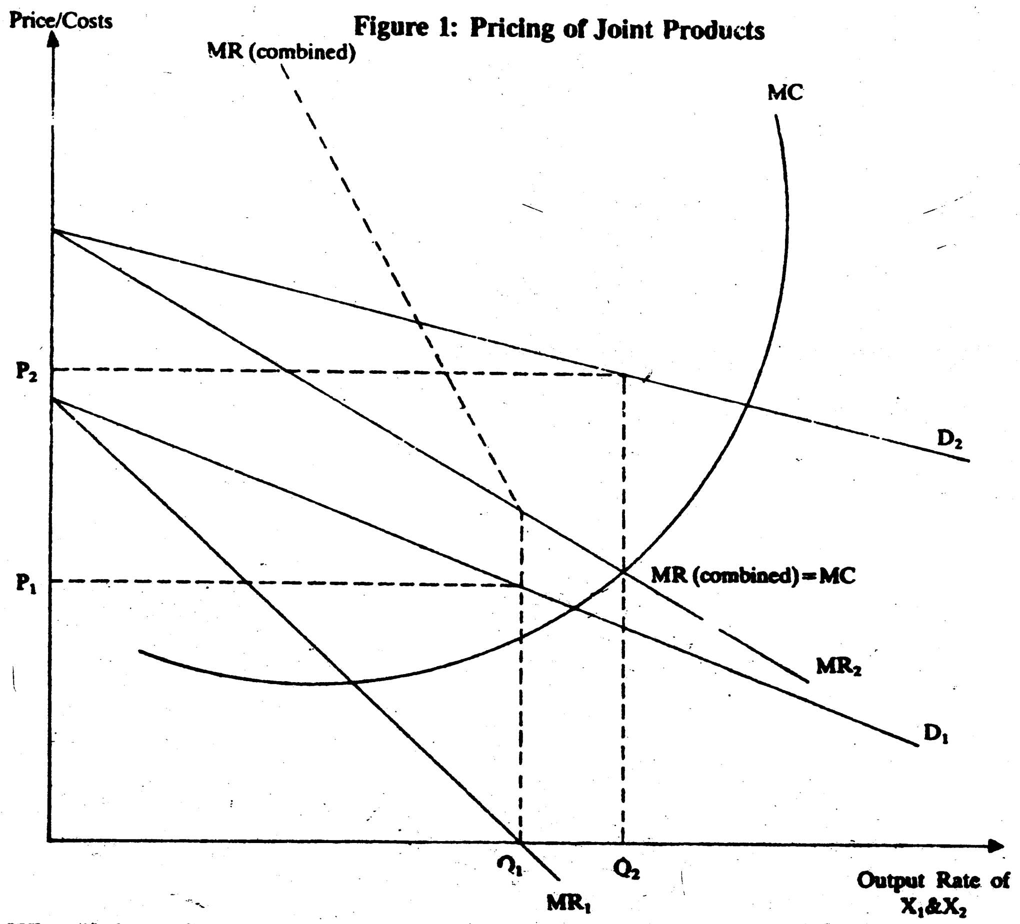
14.2 MULTIPRODUCT PRICING

Even if a firm is popularly classified as a single-product firm, it always has various models, sizes or styles of the same product. For pricing strategy, different models have to be treated as different products. Illustratively, 'Bharat Forge' manufacturers forgings as a broad product category but has different types of products, like the crank-shaft forging. Generally, we come across firms which manufacture two or more product groups. These may be inter-related from the demand side or may have common production facilities. Premier Automobiles Limited makes 'Padmini' line of cars and also the '118 NE'. Both are substitutes in a broad sense but are meant for different market segments. In a crude oil refinery some other products like petrol and diesel are jointly produced in some proportion.

Joint Products with Fixed Proportions

We shall take the simplest case here to illustrate the pricing decision in a multiproduct set-up. We assume that two products are produced in fixed proportions as a package deal. The same manufacturing set-up turns out these products in fixed proportions and therefore one set of cost curves will be sufficient for both. Indeed, it may not be possible to separate the costs of the two products. Figure 1 depicts the demand, MR and MC curves for two products X_1 and X_2. The MC curve represents the marginal cost of a unit change in the joint production of X_1 and X_2. Similarly MR_1 and MR_2 represent the respective marginal revenue curves for X_1 and X_2. In order to get the combined marginal revenues curve we shall have to add MR_1 and MR_2 vertically. You will observe that we are depicting 'rate of output' of joint products. Hence a unit change in the joint level will yield a change in both the outputs in the given fixed proportions. Vertical summation therefore appears to be the proper method of combining the two curves which is indicated by MR. The dotted line MR stops at the output rate Q_1 and thereafter the combined MR is identical to MR_2 since MR_1 is negative beyond Q_1. It is obvious that no firm will seek a rate of output which gives negative MR for one of the two products. The intersection of combined MR and MC will give the rate of output 'Q' and price levels P_1 and P_2 will be obtained by the respective demand curves D_1 and D_2.

Figure 1: Pricing of Joint Products

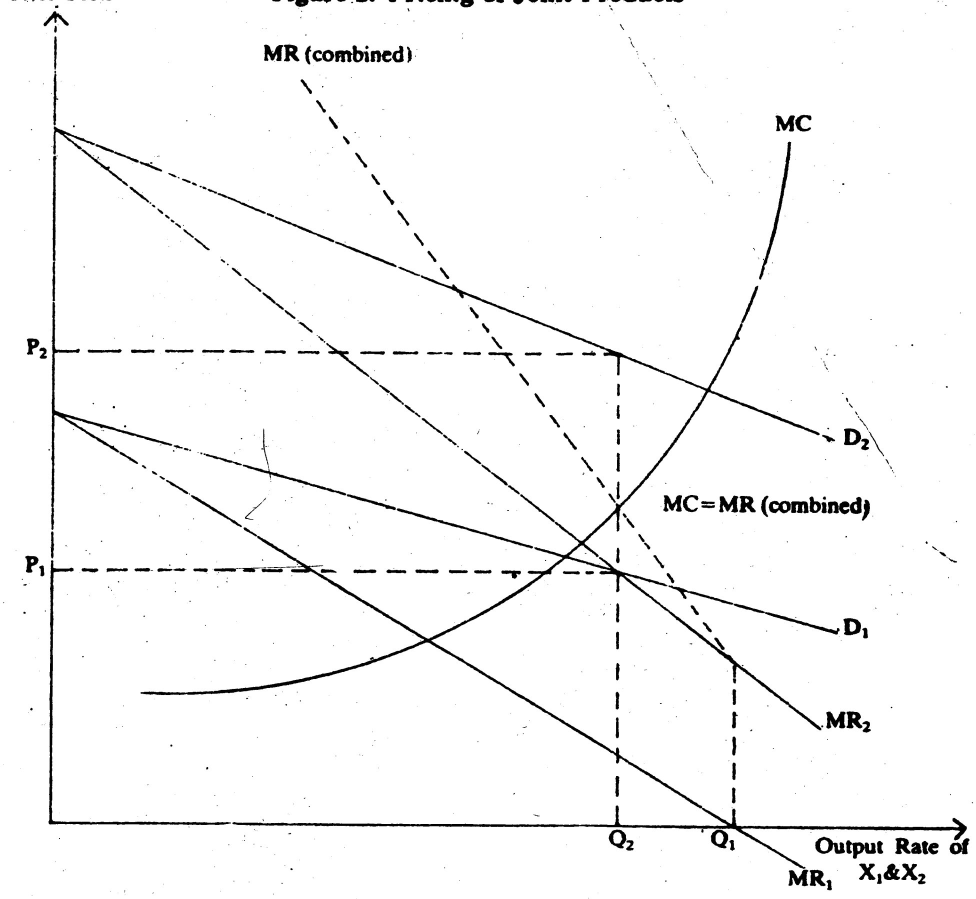


What if the MC curve happens to cut the MR in that part where MR_1 is negative? The firm will stop at the rate of output where MR_1 is zero. This situation is depicted in Figure 2. At output rate Q_1, $MR_1 = 0$ and although at Q_2 the MR is equal to MC the firm will not produce that combination. Thus, there seems to be excess production potential for the product

X_1. The firm should try and create a market condition (say by sales promotion) which results in a shift in D_1 in the upward direction. However, in the short-run it would be best to restrict the output rate at Q_1.

In actual practice, many processes display a variability in the proportions of the two products. The variability may be limited over a range of proportions. Thus, the number of alternative combinations are many and along with each combination we have the associated cost curves. The optional pricing decisions are quite complex when the firm has some control over the proportions in a joint-product set up. We shall not go into these complexities and move on to the next iten on our agenda.

Figure 2: Pricing of Joint Products



14.3 SITUATIONS REQUIRING MAJOR PRICING DECISIONS

Although pricing is a crucial managerial decision most companies do not encounter it in a major way on a day to day basis. Of course, a salesman, while offering a discount over the list price is very much taking a pricing decision. But, often he does so within the norms set by the management. Pricing becomes a major problem in essentially the following types of situations.

a) New product introduction or setting the price for the first time
b) Modification of an existing product
c) Response to a competitor's price or product change
d) Changes in costs
e) Shifts in demand pattern
f) Changes in government policy concerning taxes and duties.

When setting the first price, the decision is obviously a major one. When the company launches its product for the first time, the entire future depends heavily on the soundness of the initial pricing decision. Everytime a new product is added to the line or an existing one modified, the firm has to carefully look at the same issues. The nature of modification, its relation with the present product and several other aspects are to be considered. In case the new product is a close substitute for the company's own existing products there is always the danger of the existing products getting 'cannibalised' (or killed) by a faulty pricing of the new one. If the firm desires to replace a new product for an existing one it will be quite happy if this happens.

Pricing decision assumes special importance when one or more competitors change their prices or products or both. Sometimes, the competitor may introduce a new brand without altering the price of an existing brand. If the new brand is perceived to compete with a given brand more effectively then the firm in question may have to think on its pricing policy once again. Cost and demand changes provide obvious reasons for price revisions, but if you recall your oligopoly knowledge, they may be less important than competitive situation. This is amply illustrated by the pricing behaviour in the Truck and Bus market. For almost a three years period between 1982 and 1985 the manufacturers did not increase their prices despite cost increases. They perceived the market to be so stagnant and competition so severe that the industry accepted a decline in profits. For a leading firm in this industry the ratio of operating profits to sales declined from about 10% in 1983 to 6% in 1985.

It is thus clear that the area of pricing decisions is full of complexities. It may appear that MR = MC is too simplistic a way of looking at the pricing problem. Most empirical work indeed confirms this viewpoint. Seldom, if at, all managers go about equating their marginal costs and revenues. It is not even clear whether they are in a position to compute their MC and MR from the kind of information that is available. Yet, most managers do take decisions at the 'margin'. A typical case is a **product modification.** You will add a feature to an existing product. Find out the 'extra' cost of fitting it say a side 'fairing' to a bike. Then increase the price by such an amount as to cover that extra cost. You have therefore equated MR and MC assuming that the existing price without the additional feature was an optimal one. If not, you may increase the price by an amount which is more than what the extra fitment would cost the company. With this new price you may be nearer the optimality than what you were before.

Since precise estimation of marginal costs can become quite tricky, real life situations are dealt more on the criteria of pricing goals and consequently we have a variety of pricing situations. These are discussed in the next section.

14.4 VARIOUS PRICING STRATEGIES–MAIN APPROACHES

If firms do not adopt the principle of MR = MC in pricing their products, what methods do they employ? Various studies in this field have indicated that ften firms use a pricing method which enables them to earn a given rate of return. The rate of return is usually a relationship between the profits and suitably defined measure of invested capital. Thus, the desired profit target can be 5% on sales or a 10% on the total assets or 15% on the 'net worth' (share capital + reserves). For each firm the profit target may be different and may be influenced by several factors. The important point to note is the departure from profit maximisation.

Rather than choosing to maximise profits, firms are happy to earn a given return on the resources they employ. This method of pricing is referred to as **target return pricing.**

What are the various factors that may affect a firm's profit target? We give below a summary.

a) The most important factor is the management's perception of a 'fair' or 'reasonable' return considering the risks that the business entails. Some influence of the firm's past performance and its future aspirations can also be expected. If you read the Chairman's statement of most large companies, you will be able to draw some inferences about these aspects. Firms in risky business like computers may consider at 25% return on net worth as reasonable whereas engineering firms may be happy with 18% return.

b) This brings us to the industry related factors. The nature of the industry, the level of competition, the rate of technological change are very crucial in deciding the target rate of return.

c) The nature of product, its unique characteristics, the research and development inputs that go into making it and similar product related factors also have a bearing on the profit target.

d) Lastly, the firm's overall goals in regard to sales, growth in operations, market share, position in the industry, competitiveness etc. do influence the profit target since it is a part of the overall strategy of the firm.

Activity 1

If you are employed in a manufacturing or service organisation, can you identify the profit target for your firm? How does it compare with the industry behaviour? What specific factors had an influence on the profit target?

....

....

....

....

....

....

....

....

....

....

It must be noted that the profit target need not be expressed as a fixed number. It could be stated as 'above the industry average rate of return'. In such a case, the same firm may use a return on investment of 5% for the pricing decisions in a bad year and upgrade it to 12% in a better business period.

Illustration

Omega Steel Strips Ltd. (Omega for short) manufactures galvanised corrugated and plain steel sheets (GC/GP sheets). It has a plant with installed capacity of 10,000 tonnes per annum. The investment in the plant, machinery and other facilities is Rs. 12 crore. The profit target set by the management is 15% rate of return before taxes on the investment. This means that it must earn profits of Rs. 1.8 crore from its operations in a normal year. The average cost per tonne of finished GC/GP sheets is Rs. 20,000. This includes costs of raw materials (steel sheets of various sizes), processing, consumables like fuel, power, lubricants etc., working capital interest and overheads including wages. The per tonne profit will depend on the quantity sold but for announcing a market price the firm chooses a figure of 80% capacity utilisation. Omega has observed that it can utilise its capacity to the extent of 90%, but considering uncertainties in the input supplies and demand constraints, it is prudent to plan for a lower utilisation. If pricing is done on the assumption of 80% capacity utilisation, any increase in actual utilisation will be beneficial. On this assumption the per tonne profit will be Rs. 2,250. This is worked out as [1,80,000 ÷ (10,000) (0.8)]. Thus the end price of GP/GC sheets will be set at Rs. 22,250/tonne. Now, suppose the market demand for its products was 9500 tonnes. Omega therefore earns a pre-tax profit of Rs. 2.1375 crore that is 33.75 lakhs more than its target and consequently the return on investment is nearly 18%.

Cost-plus Pricing

There is one more method of pricing which is also widely prevalent. It uses average total cost as the basis for pricing of products. The price is calculated by adding a predetermined percentage mark-up to the estimated unit cost of the product. Generally all the normal variable costs and fixed costs are taken into consideration while computing the average cost. The mark-up reflects the management's belief regarding what constitutes a 'reasonable relation to the cost' of production. Again the mark-up has some bearing on the target profit that the firm desires to achieve. As in the above illustration, in the cost-plus method also some assumption has to be made about the 'normal' working—capacity utilisation, deployment of various overheads, selling and administration effort etc. The mark-up which represents the 'plus' factor in the 'cost-plus' formula may be fixed on the short-term considerations or long-term goals. Trading organisations are known to employ the mark-up method extensively. For various products traded by a departmental store, the mark-up may vary from 15% to as high as 60% in items with high inventory carrying costs. In manufacturing organisations selling products like TV sets, refrigerators, scooters etc. the final product may be priced by one method and spare parts may be priced the cost-plus method. Those parts with high 'use-rate', known as 'fast moving items are marked up' at a lower percentage over the manufacturing cost.

Concept of Cost and Return

While pricing the products one must be clear about the definition of 'cost' and 'return'. Some organisations use 'profit before depreciation, interest and taxes' as the definition of return. If this profit concept is used then it must be evaluated in relation to the associated concept of 'capital employed'. Similarly, if manufacturing costs imply those costs that are incurred in manufacturing the product then the mark-up has to cover all other costs and then leave sufficient balance to earn the targetted return. Since different concepts are used by different firms as yard sticks for measurement of performance there is no unique figure for a rate of return that is 'reasonable'.

New Products

Suppose a firm is contemplating introduction of a new product. There are a few existing firms selling similar products. The firm in question will therefore take the competitors' prices as starting point. The design of its product, manufacturing facilities must be dictated by what the others are doing. Suppose all the refrigerators in the market are selling in a range of Rs. 4,000 to Rs. 5,000 for a 165 litre model. A new model with similar features has to be priced accordingly. If the design is so unique that its manufacturing cost is say Rs. 5,200 per unit. Then, the firm must differentiate its product accordingly and communicate the unique features to the customer so that a higher price is properly justified. Thus, in a new product one works backward taking the competitors' prices as benchmark range. If this is not done, one may end up organising a product which costs more and has no features to go along with it.

14.5 DEMAND, CAPACITY AND PRICING BEHAVIOUR

It is now clear that costs, target rate of profit and competitor's price play a crucial role in determining the pricing strategy of a firm. But, this strategy is always placed in the overall industry context. The objective conditions of demand and supply at the industry level are equally important. Let us construct a case using some published information about the light commercial vehicles (LCV) industry. The price behaviour of the most well established firm in this industry 'Bajaj Tempo' will be examined. The aim is to see how industry demand and capacity (which represent potential supply) may have had an impact on its pricing strategies during different time periods.

Pricing Behaviour of Bajaj Tempo LCV's

For a number of years up to about 1985, the light commercial vehicle industry in the country had only 5 firms, Bajaj Tempo being the market leader. Other prominent firms were Mahindra & Mahindra and Standard Motors. Two firms namely, Hindustan Motors and Premier Automobiles also marketed LCV's but the volumes were not significant. In the wake of the modernisation process of the Indian automobile industry during the early eighties, four new firms entered the business, all with Japanese collaborations. These firms started bringing their vehicles into the market around 1985 and by the end of 1987 there were 10 firms in the LCV market. In the intervening period, an existing manufacturer of medium and heavy duty commercial vehicles diversified its product line by adding an LCV model. Some of the existing firms also increased their capacities.

The growth in capacity of this industry therefore outstripped the growth in demand. Whereas the annual industry sales went up by an average rate of around 12-15% the capacity growth was nearly 20%. At present, the level of competition in the LCV industry is quite fierce compared to the situation that prevailed around 1980. There are many more models offering various features to the customers. Today, the industry has the capacity to manufacture over 1.5 lakh vehicles but the annual demand is no more than 45,000 numbers.

Let us now look at the prices of a typical LCV produced by the company during the period 1979-87. In order to put the price changes in a proper perspective we must also see the general inflationary trend which exerts a push on the company's costs. The Table below gives some data in this regard.

TABLE

PERIOD	*WHOLESALE PRICE INDEX FOR*		*NET DEALER PRICE OR A.B.T.DELIVERY VAN (24WB) EXCLUDING TAXES & DUTIES Rs.*
	MACHINERY AND TRANSPORT EQUIPMENT	*BASIC METALS ETC.*	
	(Index No.1970-71 = 100)		
1979-80	215.9	251.9	42,619
1980-81	239.4	272.1	47,460
1981-82	265.1	317.1	57,100
1982-83	279.9	354.6	59,350
1983-84	289.4	381.0	62,050
1984-85	303.6	419.8	66,300
1985-86	337.9	477.1	72,600
1986-87	354.8	556.9	72,600

If you carefully analyse the table it will be noticed that in the period 1979-84 when the impact of competition was low, the firm could effect ~~price~~ increases with due regard to cost increases but as the competition became fierce, the ability of the firm to pass on the cost increases to the end price considerably reduced. The profitability must have suffered but not necessarily. The efficiency of management of costs must also be considered. As an exercise you can study this aspect by analysing the annual reports of the company over the same time period.

(The data and relevant material for this case was provided by the *Association of Indian Automobile Manufacturers.)*

It is possible to learn some principles of a general nature from this case? Though you find cases which do not conform to this behaviour, one thing can be stated with a fair amount of generality. That is, if an industry is faced with a situation of excess capacity, slow growth in demand and keen competition between a few powerful firms, the ability of the firms to pass on the cost increases to the customers gets substantially restricted. If however, the industry has one powerful firm, then despite excess capacity that firm may be able to exercise considerable control over its price.

Activity — 2

Bajaj Auto is the undoubted leader in the two wheeler industry. What has been the pricing behaviour of its various products? Please examine this aspect as an exercise.

....

....

....

....

....

....

14.6 PRICING OBJECTIVES AND CORPORATE GOALS

Pricing decisions are not taken independent of what the company as a whole desires to achieve. Product pricing is a part of a broader decision area called product policy which again is a part of the overall corporate strategy. Therefore, if you make an attempt to set prices without having due regard to the complete product policy and the corporate goals it will not be a proper method. Nothing can illustrate this maxim better than Alfred P. Sloan's approach at the General Motors during the early part of the growth of that company. We reproduce below selected paragraphs from his famous book *My Years with General Motors.* These extracts not only illustrate the importance of a product policy but also give a clear exposition of how pricing decisions are to be approached in different environments.

Extracts from Chapter 4 'Product: Policy and Its Origins'

(My Years with General Motors)

"After the two great expansions of 1908 to 1910 and 1918 to 1920—perhaps one should say because of them—General Motors was in need not only of a concept of management but equally of a concept of the automobile business. Every enterprise needs a concept of its industry. There is a logical way of doing business in accordance with the facts and circumstances of an industry, if you can figure it out. If there are different concepts among the enterprises involved, these concepts are likely to express competitive forces in their most vigorous and most decisive form.

Such was the case in the automobile industry in 1921. Mr. Ford's concept of a static model at the lowest price in the car market, expressed in the Model T, dominated the big volume market then as it had for more than a decade. Other concepts were present, such as the one implied in about twenty makes of cars calculated to have low volume and very high price, and those behind the various cars in intermediate price brackets. General Motors then had no clear cut concept of the business; there was no established policy for the car lines as a whole.

We had no position in the low price area. Chevrolet at the time being competitive with Ford in neither price nor quality. In early 1921, the Chevrolet was priced about $ 300 above the Model T (when an adjustment is made for comparable equipment) hence out of sight from the viewpoint of competition. The fact that we were producers of middle and high price cars, so far as I know, was not a deliberate policy. It just happened that no one had figured out how to compete with the Ford, which had then more than half the total market in units. It should be observed, however, that no producer at that time presented a full line of cars, nor did any other producer present a line as broad as General Motors' line.

The spacing of our product line of ten cars in seven lines in early 1921 reveals its irrationality. Our cars and their prices at that time (priced from the roadster to the sedan, F.O.B. Detroit) were as follows:

Chevrolet '490' (four cylinder)	$ 795 - $ 1375
Chevrolet 'Fb' (four cylinder)	$ 1320 - $ 2075
Oakland (six cylinder)	$ 1395 - $ 2065
Olds (four cylinder 'Fb')	$ 1445 - $ 2145
(six cylinder)	$ 1450 - $ 2145
(eight cylinder)	$ 2100 - $ 3300
Scripps-Booth (six cylinder)[1]	$ 1545 - $ 2295
Sheridan (four cylinder FB)	$ 1685
Buick (six cylinder)	$ 1795 - $ 3295
Cadillac (eight cylinder)	$ 3790 - $ 5690

[1]Six cylinder engine made by Oakland.

Superficially this was an imposing car line. From the inside the picture was not quite so good. Not only were we not competitive with Ford in the low price-field—where the big volume and substantial future growth lay—but in the middle, where we were concentrated with duplication, we did not know what we were trying to do except to sell cars, which in a sense, took volume from each other. Some kind of rational policy was called for. That is, it was necessary to know what one was trying to do, apart from the question of what might be imposed upon one by the consumer, the competition, and a combination of technological and economic conditions in the course of evolution. The lack of rational policy in the car line can be seen especially in the almost identical duplication in price of the Chevrolet 'FB', Oakland and Olds. Each division, in the absence of a corporation policy, operated independently, making its own price and production policies, which landed some cars in identical price positions without relationship to the interest of the enterprise as a whole.

The hard fact was that all the cars in the General Motors line, except Buick and Cadillac were losing money in 1921. The Chevrolet Division that year lost about half of its 1920 volume. Its dollar losses at one point in 1921 reached approximately \$ 1 million a month and for the year as a whole it lost nearly \$ 5 million.

This situation reflected in good part the poor quality and unreliability of the other cars in the line, as compared with the high quality and reliability of Buick and Cadillac; the effect of these factors was intensified by the stress of the general economic slump. Given the fact of the slump and the unavoidability of a general decline in sales, the relative decline of one division as compared with another was a question of management.

Thus the new management took the opportunity that comes rarely in the initial stage of a business, to stand back and review aims and deal with the matters at hand both in particular and with a considerable degree of generalisation. It was not going to be easy to get willing agreement on specific and immediate issues. For example the idea of the revolutionary car was very much entrenched in the Executive Committee, and I wanted to broaden the concept of the product to the concept of the business. I believe it was for this reason that we on the special committee first idealised the problem. We started not with the actual corporation but with a model of a corporation, for which we said that we would state policy standards.

The product policy we proposed is the one for which General Motors has now long been known. We said first that the Corporation should produce a line of cars in each price area, from the lowest price up to one for a strictly high-grade quantity-production car, but we would not get into the fancy—price field with small production; second that the price steps should not be such as to leave wide gaps in the line, and yet should be great enough to keep their number within reason, so that the greatest advantage of quantity production could be secured; and third that there should be no duplication by the Corporation in the price fields or steps.

Having set forth these concepts, we then approved the resolution of the Executive Committee, which had been passed on to us to study, to the effect that a car should be designed and built to sell for not more than \$ 600, and that another car should be designed and built to sell for not more than \$ 900. The special committee further recommended four additional models, each to be kept strictly within the price range specified. It also recommended that the policy of the corporation should be to produce and market only six standard models, and that as soon as practicable the following grades should constitute the entire line of the cars.

a) \$ 450 – 600
b) \$ 600 – 900
c) \$ 900 – 1200
d) \$ 1250– 1700
e) \$ 1700–2500
f) \$ 2500–3500.

This brand-new, hypothetical price structure, when compared with General Motors, actual price brackets listed earlier in this chapter, will be seen to reduce the car lines from seven to six (or ten to s.. cars if the Chevrolet 'FB' and the Olds '6 and '8' are considered separate cars, as they pretty much were). It opened up one new classification on the low end of the list where we had none. And where we had eight cars in the middle, above the lowest price and below the highest, we now had only four classifications. The new set of price classes meant that the General Motors car line should be integral, that each car in the line should properly be conceived in its relationship to the line as a whole. Having thus separated out a set of related price classes, we set forth an intricate strategy which can be summarised as follows: We proposed in general that General Motors should place its car at the top of each price range and make them of such a quality that they would attract sales from below that price, selling to those customers who might be willing to pay a little more for the additional quality, and attract sales also from above that price, selling to those customers who would see the price advantage in a car of close to the quality of higher-priced competition. This amounted to quality competition against cars above the price tag. Of course, a competitor could respond in kind, but where we had little volume we could thereby chip away an increase from above and below, and where we had volume it was up to us to maintain it. Unless the number of models was limited, we said, and unless it were planned that each model should cover its own grade and also overlap into the grades above and below its price, a large volume could not be secured for each car. This large volume we observed, was necessary to gain the advantages of quantity production, counted on as a most important factor in earning a position of pre-eminence in all the grades.

Alternative Pricing Strategies

Having seen the crucial link between overall corporate goals and the pricing strategy through the example of General Motors, U.S.A. we now discuss some more situations. Different occasions call for different approaches to pricing. Market environment, the firm's immediate and long-term objectives, the peculiar features of the product or customer behaviour have an important role to play in these decisions.

Penetration Pricing

While introducing new products or entering in new geographical markets, firms may deliberately set a relatively lower price in the hope of 'penetrating' into the market. The idea is to establish a market share first and then gradually move to a price which is more 'desirable' from the profit angle. However, this strategy works well provided the demand is highly price elastic and the nature of product differentiation is such that many customers are in a position to get attracted by the low price. At times low price may be set as a deterrent for the potential entrants. When a lower price is charged to a geographically separate but new market the firm is testing out the elasticities of demand before taking a final decision on the level of price discrimination.

Price Skimming

This again is a variation of the price discrimination theme, not over two markets but over a period of time. Typically, the firm starts with a high price appealing to those customers who are willing to pay a premium for better quality or because they put some additional value on the product. A new variety of 'cheese' which is popular abroad can be introduced in India at a very high price because some westernised Indians may buy it just for its prestige value. However, at a slightly lower price some more 'not-so-westernised' people can also be attracted. Essentially, a firm manufacturing that brand in India is 'riding-down' the demand curve over a period of time.

Each successive price drop hopefully brings in some more customers. But, there is a danger in letting the price drop beyond a point because of the 'perceived' correlation between quality and price.

These two situations are in the nature of short-term pricing tactics and the firm is taking a calculated chance before it discovers the long-term price which the product must bear. If the experiment with prices fails, the product introduction may come into serious trouble. Thus, quite often firms do not undertake such tactics and set the desirable price straightway on a cost-plus or whatever basis.

Loss-leader Pricing

A firm selling both razor blades and razors can price the razors below average variable cost if it is confident of selling large volumes of the blades which it claims are particularly suitable for the razors. The firm incurs loss on razors but more than makes it up in blades so that the total profits of the product line 'shaving system' are improved. In a sense therefore the so-called loss-leader is really a profit leader. Because of the loss making product, customers are induced into buying other complementary items in the line and the whole set becomes profitable. This tactic is feasible in complementary products where a 'one time purchase' of a high value item (Xerox machine) leads to repeat purchase of a low value but high volume item (toner).

Such a pricing strategy is feasible even in capital goods with heavy requirement of replacement parts and consumables. A firm making capital goods along with the parts and consumables can safely incur a loss on the main machine and recover it through the sale of other products. The parts and consumables must however be proprietory or dedicated. Otherwise the customer can buy the main machine from you and replacement parts from some competitor.

In almost every pricing tactic that any firm adopts there are some assumptions about the customer behaviour, elasticity of demand and competitors' reactions. It goes without saying that the efficiency of the tactics depends on the soundness of those assumptions. Firms therefore conduct extensive research before making any tactical move.

14.7 PRICING OF SERVICES

How does a service organisation go about the pricing decision? Technical consultancy, project management, market research, advertising and a host of such enterprises do not have a product which requires materials, equipment and processes. Of course they do need experts, some office equipment and may be computers. But, the principal resource that a service organisation has its manpower. The cost of that resource is most important in pricing the ultimate service—be it a project report or an advertising campaign. Principles of marginal costing will be less important and the concept of opportunity cost may be most relevant in pricing the services.

The delivery of a market research document typically involves great amount of field work and expenses related to hiring of investigators, their travel and other costs. Considerable time is spent in designing the research, conducting it, analysing the results and then presenting the same to the client. Each exercise requires unknown quantities of all these items (before it is taken-up) and hence the research agency estimates the time required in undertaking various activities. The time is translated in costs on the basis of the salaries of the respective persons carrying out the task. Unlike a product like the GP/GC sheets where the technology determines input-output relationships, a service input-output relation is very vague.

It is therefore feasible to justify large variation between the service prices of competing organisations on the 'quality' considerations. Unlike a product the quality of a given service can be tested on a consistent basis only in a limited sense. Human skills are therefore the principal technical inputs that need to be priced. Over a period of time, costs of these skills do get standardise-manhour cost of a programmer, a copywriter and so on; but differentiation of services becomes far more easy in view of the 'quality' angle.

The price of a service often gets influenced by the demand and supply in the 'market for skills' — labour market. If programmers are in excess demand, their opportunity cost goes up; hence a service firm has to continuously increase their salaries to retain talented and experienced people. That gets reflected in a higher manhour cost of that skill which ultimately makes the service more expensive. The reverse may happen when there is excess supply.

Activity 3

Study the market for computer software services over the last five years and explain the price movements. Some sources of data are given below.

Sources of Data: *Directory of Computer Manufacturers,*

a) Traders and software houses, brought out by a Delhi based firm. "Computer Directory of India 1986"
b) Field interviews with the owners of software houses.
c) Journals like Dataquest, Computer World, etc.

....

....

....

....

....

....

....

....

14.8 BUREAU OF INDUSTRIAL COSTS AND PRICES (BICP)

We already know that prices of many products like cement, steel, aluminium are regulated in our country. The apex body which supervises the setting of the prices of these products is called the BICP under the charge of the Ministry of Industry, which directly sets prices of some products and gives guidelines for the pricing of a host of other products. Besides, it also examines from time to time the 'fairness' of the prices of certain products on the basis of references from government.

The BICP adopts a 'profit-target pricing' strategy. It has a notion of a fair return which is of course different from industry to industry. For fertilizer industry, the post tax return used to be 12% on net worth at 80% capacity utilisation. Thus, the retention price for every manufacturer is so calculated as to give him that return.

Given a notion of a return on net worth or capital employed, the BICP then undertakes detailed cost audit of the firms in the industry. Based on the actual costs at some predetermined level of capacity utilisation the average manufacturing cost is arrived at. Depreciation, interest and overheads are allocated in some fashion depending on the nature of product, its working capital needs and other resource requirements.

Sometimes, the BICP uses a mixture of 'target-profit pricing' and 'cost-plus pricing'. You must have read a lot in the newspapers about the control of drug prices. The regulation of drug prices is a case in point for illustrating the exact method of computing prices used by BICP in some cases. Since it is a very important document, some paragraphs (number 2, 6, and 7) from the "Drugs (Price Control) Order 1987", are reproduced below.

You will notice that the drug prices are regulated by two layers of control. The retail price is determined by formula which links various items of the cost to the final price. At the second level, the overall profitability is regulated by the upper limit on the return on investment. The intention is to ensure that firms earn 'adequate' profits and the customer pays a 'reasonable' price.

"2 While fixing the price of a bulk drug the government may take into consideration a post tax return of 14 per cent on net worth or a return of 22 per cent on capital employed or in respect of a new plant an internal rate of return of 12 per cent based on long-term marginal costing depending upon the option for any of the specified rates of return that may be exercised by the manufacturer of a bulk drug.

Provided that the option with regard to the rate of return once exercised by a manufacturer shall be final and for any change in the said rate of return prior approval of the government shall be necessary.

6 Calculation of retail price of formulations:

The retail price of a formulation shall be calculated in accordance with the following formula namely:

$$RP = (M.C. + C.C. + P.M. + P.C.) \times (1+MAPE/100) + E.D.$$

where

R.P. means retail price.

M.C. means material cost and includes the cost of drugs and other pharmaceutical aids used including coverages, if any, plus process loss thereon specified as a norm from time to time by notification in the official gazette in this behalf.

C.C. means conversion cost worked out in accordance with established procedures of costing and may be fixed as norm from time to time by notification in the official gazette in this behalf.

P.M. means cost of the packing material used in the packing of concerned formulation and included process loss, a norm fixed from time to time by notification in the official gazette in this behalf.

P.C. means packing charges worked out in accordance with established procedures of costing and may be fixed as a norm from time to time by notification in the official gazette in this behalf.

MAPE means maximum allowable post-manufacturing expenses including trade margin referred to in paragraph 7.

E.D. means excise duty.

Provided that in the case of an imported formulation, in landed cost shall form the basis for fixing its price along with such margin to cover selling and distribution expenses including interest and importer's profit which shall not exceed 50 per cent of the landed cost.

Explanation: For the purpose of above proviso 'landed cost' shall mean the cost of import of drug inclusive of customs duty and clearing charges.

MAPE

MAPE referred to in paragraph 6 means the maximum allowance post-manufacturing expenses including trade margin and shall not exceed.

a) Seventy five per cent in the case of formulations specified in Category I of the third schedule.

b) One hundred per cent in the case of formulations specified in Category II of the said schedule.

Explanations

a) For the purpose of this paragraph 'post-manufacturing expenses' means all costs incurred by a manufacturer from the stage of ex-factory cost to retailing and includes trade margin.

b) For the purpose of categorisation of a formulation it shall be deemed as:

i) Category I formulation, if it contains any bulk drug either individually or in combination, specified for Category I formulations.

ii) Category II formulation, if it contains any bulk drug eigher individually or in combination, specified for Category II formulations.

Provided that, in case the formulation contains bulk drugs specified in both categories I and II it shall be deemed as a Category I formulation."

You would have noticed the use of the concept of 'long-run marginal cost' in paragraph 2 of the above mentioned document. Most manufacturers would find it difficult to estimate this because of the uncertainty pertaining to the size of the plant in the long-run and the product mix that the firm adopts which has to change from time to time depending on the demand conditions.

14.9 SUMMARY

We are now in a position to summarise the main ingredients of various pricing strategies and tactics adopted by firms in actual practice. While very few firms actually equate MR to MC in pricing their products, concepts embodied in these ideas are very much used. Whether guided by a minimum profit target or prompted by the urge to maximise profits, average cost and marginal cost play crucial role in determining the prices of most products. The margin or mark-up or profit depends on many factors starting from the attitudes of the top management to the objective demand and supply conditions. Competition is at the heart of any market. Naturally, competitive forces mould the price behaviour of firms in more than one way. It affects the market structure, shapes the objectives of firms and above all, it sets the directions in which prices may tend to change.

Indeed, the most important aspect of the behaviour of firms that emerges out of our study carried out in this block (Unit 11 to 14) is the importance of market structure and competitive forces within it. We made use of certain stylised though ideal situations like perfect competition and monopoly to bring home this point. The previous section on BICP illustrates a rather interesting point. Even a regulatory body uses the same concepts in determining the prices. The only difference is that the regulatory authority specifies an upper limit on the price/return. Without it the market may exercise that control if there was sufficient competition. Whenever market mechanism fails to achieve certain results, regulation is considered desirable. Since we know that oligopoly situation often lead to uncertain and hence in some cases, socially undesirable outcomes, in case of important products like drugs, government thinks it fit to regulate prices rather than leave them open to the market forces.

14.10 KEY WORDS

A **target return (or profit) price** is a price designed to yield the firm a predetermined profit from the sale of specific product groups.

In **cost-plus pricing method**, price is calculated by adding a predetermined percentage mark-up to the estimated 'per unit cost' of the product.

The strategy of deliberately setting a low price while introducing a new product or entering a new territory is called **penetration pricing**.

By **Price skimming** firms try to take advantage of the customer behaviour reflected in their willingness to buy a product at a high price since for some reasons it offers high value to them.

Loss-leader pricing can be defined as pricing an item at a level which gives less than normal or even negative contribution but is aimed at reaping higher total profit through either increased future sales of the same item or greater sales of the entire product line.

Pricing is a sub-set of the total product policy and has to be linked to the overall corporate strategy.

14.11 ADDITIONAL READINGS

Dorfman, R. 1965. *The Price System*, Prentice Hall of India: New Delhi.

Thompson, Arthur A. 1977. *Economics of the Firm. Theory and Practice*, Prentice Hall: Englewood-Cliffs.

Sloan, Alfred P. 1986. *My Years with General Motors*, Penguin: New York.

14.12 SELF-ASSESSMENT TEST

1 Examine the price-cost relationship from the standpoint of (a) theory and (b) practice.

2 State and explain the factors which you would normally consider while pricing a new product.

3 How is penetration price different from a skimming price? Under what conditions, would you recommend this kind of pricing policy?

4 "Pricing decisions are constrained by product-policy decisions and considerations of corporate planning". Comment, quoting appropriate examples from the real world.

5 What is 'cost-plus pricing'? Does it neglect the demand factors influencing pricing decisions?

6 Explain each of the following observations:

a) Products are priced differently depending upon not only the life cycle of the product but also the phase in the business cycle. If the business is under recession, you cannot charge high prices.

b) Pricing a service is different from pricing a product.

c) In deciding about the price, the firm may use profit as a constraint rather than as an objective.

7 We can apply the equi-marginal principle in the context of multi-product pricing. True or False? Explain.

BLOCK V THE FIRM IN THEORY AND PRACTICE

Managerial Economics is an applied branch of economic thought that extends abstract economic theory to business decision-making. Any branch of study has its foundations in theory (or theories), from which practitioners draw ideas. The purpose of this block is to provide an overview of some well known theories to understanding the firm alongwith their applications to business decision-making. Each theory has its own basic assumptions and conclusions about the actions of a firm. This block has four units.

Unit 15 deals with the Economic Theory of the firm. In terms of philosophy and outlook managerial economics draws more extensively upon Economic (or Micro Economic) Theory of the firm formulated by Marshall and others. The economic concepts of market, firm and industry are explained. This is followed by a criticial evaluation of the theory, especially its motivational and cognitive assumptions.

Unit 16 looks at the Behavioural Theory of the firm, an alternative to Economic Theory. The main thrust of the theory is that decision-making in a firm is guided by multiple objectives rather than a single objective. Hence, the firm in its decision-making tends to be a satisficer and not a maximiser. The unit begins with a discussion of the key concepts and assumptions underlying the behavioural model. The nature of decision-making process in a firm operating under uncertainty with multiple goals and multiple products is explained. The behavioural theory is contrasted with the economic theory. The usefulness and limitations of the main contributions in this model are discussed.

Unit 17 focuses on Managerial Theories of the firm which can be regarded as extensions or refinements of the Behavioural Theory. The firm is viewed as a **coalition** where the main objectives of top management, who weild the real power, is to harmonise the conflicting interests of various parties associated with the firm. In this lies the top management's own vested interests to assure their continuance in power. The unit thus begins by examining the behaviour of a manager. Three main contributions in the managerial model have been selected for discussion. Their underlying assumptions and differences have also been examined. The managerial model is then compared with the economic model of the firm. Some comments of general nature on managerial theories have also been offered.

Unit 18 examines profits, common thread found in all firms trying to survive in the modern market place. The concept of profit and its importance are explained. Then, measurement problems are reviewed to highlight the difficulties in conceptualising the ideas of profit in an environment that involves inventories, depreciation and price fluctuations. Alternative explanations of profit are discussed. The need for a profit policy in the context of profit planning and control is emphasised.

UNIT 15 ECONOMIC THEORY OF THE FIRM

Structure

15.0 INTRODUCTION

An overview of the price-output decision analysis in the preceding block brings out clearly the significance of concepts like markets, firms and industry. The decision process depends on the nature of decision environment characterised by factors like the nature of market, the number of firms and the size of industry. In analysing this set of factors, the economists have made a host of assumptions and have constructed a veriety of models. It is now time to take stock of all those assumptions and models together in the framework of a theory and then examine its emperical relevance. After all, in managerial economics, we hope to get an appropriate blend of theory and practice.

15.1 OBJECTIVES

On reading this unit, you should be in a position to

- **understand** better the economic concepts of market, firm and industry
- **state** the assumptions underlaying the economist's theory of firms
- **develop** a critique of that economist's theory or model
- **suggest** the need for alternative models
- **refine** your analytical tools to approximate real world business situations.

15.2 CONCEPTS OF MARKETS, FIRMS AND INDUSTRY

These concepts are crucial to the understanding of the economist's theory of firms. Let us recall these concepts with a bit of analytical rigour.

Markets

You must have noted that traditional economic theory distinguishes the following market structure:

i) **Perfect Competition** refers to a market situation where there is a large number of firms in the industry, producing homogeneous products, enjoying freedom of movement, having neither transportation costs nor information costs, and selling products at a single price. Although competition is **perfect**, there is no **rivalry** among the individual firms.

ii) **Monopoly** refers to a situation where a single firm constitutes the entire industry. Entry is blockaded. The monopolist **makes** rather **takes** the price.

iii) **Monopolistic Competition** exists in a market with a large number of firms which individually sell **differentiated products** . Thus products are close substitutes, entry is free and easy in the industry. Each firm enjoys some degree of monopoly power subject to competition from the rivals.

iv) **Oligopoly** exists in a market with small number of firms, grouped together, producing either homongenous or differentiated products. Individual sellers are conscious of their interdependence and therefore, take care of rivals actions and reactions in a variety of ways, before designning their market strategy and tactics.

Based on the description above, we may now suggest a market classification based on certain criteria:

a) **Degree of product substitutability**

This may be measured by the conventional price cross elasticity of demand (ep_{ji}) for the commodities produced by any two firms.

$$ep_{ji} = \frac{dQj}{dpi} \cdot \frac{Pi}{Qj}$$

It measures the degree to which the sales of the j-th firm are affected by changes in the price charged by the firm in the industry. If this elasticity is high, the products of the j-th and the i-th firm are close substitutes.

b) **Degree of interdependence of firms**

This may be measured by the unconventional quantity cross-elasticity for the products of any two firms

$$eq_{ji} = \frac{dPj}{dQi} \cdot \frac{Qi}{Pj}$$

It measures the degree to which the price of the j-th firm is affected by change in sales of the i-th firm. The higher the value of this elasticity, stronger is the degree of interdependence.

c) **Condition of entry**

The ease of entry may be measured by

$$E = \frac{Pa-Pc}{Pc}$$

Where E = condition of entry
Pc = Price under price competition
Pa = Actual price charged by firms

The condition of entry is a measure by which the established firms in an industry can raise their Pa above Pc without attracting entry.

The market classification based on the above criteria can now be reproduced in the following table.

Market classification criteria

Type of market	Product substitutability $e_{p,ji} = \frac{dQi}{dPi} \cdot \frac{Pi}{Qj}$	Seller's interdependence $e_{p,ji} = \frac{dPJ}{dQi} \cdot \frac{Qi}{P_J}$	Ease of entry $E = \frac{Pa-Pc}{Pc}$
Pure Competition	$\rightarrow \alpha$	$\rightarrow O$	$\rightarrow O$
Monopolistic Competition	$O < e_{p,ji} < \alpha$	$\rightarrow O$	$\rightarrow O$
Pure Oligopoly	$\rightarrow \alpha$	$O < e_{p,ji} < \alpha$	E>O
Heterogeneous Oligopoly	$O < e_{p,ji} < \alpha$	$O < e_{p,ji} < \alpha$	E>O
Monopoly*	$\rightarrow O$	$\rightarrow O$	Blockaded entry

*This is from *Modern Microeconomics* (2nd ed.) 1979 by A. Koutsoyiannis, p.7.

Activity 1

a) **Bring out implications of the table. Be sure to work out the logic of reasoning in detail.**

b) **Distinguish between the concepts of price-cross-elasticity and quantity-cross elasticity.**

c) **Redefine Bain's concept of "Barriers to entry" in terms of E.**

Industry

The concept of an industry has been developed to include firms which are in some form of close relationships with one another. These fimrs belonging to a group are behaviourally interdependent. The economists talk about **'individual equilibrium'** in the case of firms in contrast to **'group equilibrium'** in the case of an industry. You may recall-the condition of equilibrium in case of an individual firm is MR=MC, but that for an industry is AR=AC. The łouble conditions of equilibrium are generally satisfied only in the long-run. In the short-rur the firm may be in equilibrium without the industry being in equilibrium; and this will lead tc all sorts of adjustments, depending upon the nature of market.

The concept of an industry serves a lot of purposes. Firstly, it helps us to group the firms in terms of some criteria. Secondly, the concept of industry makes it possible to derive a set of general rules from which we can predict the behaviour of competing members of the group that constitute the industry. Thirdly, the concept of an industry provides the framework for the analysis of the effects of entry on the behaviour of the firm and on the equilibrium price and output. Fourthly, the businessman designs his strategy and tactics in view of the industry they belong to. Lastly, the Government policy is designed with reference to industry; most policies are industry-specific. Since the concept of an industry has a lot of operational use, it is meaningful to classify industry on the basis of some critieria. Normally, there are two **criteria of industry classification:**

i) **Product Criterion:** The firms are grouped in an industry if their products are close substitutes. You can talk of both "economic" substitutability (in the sense of similar price change). Such substitutability can be measured in terms of price cross-elasticity of demand, ep_{ji}.

ii) **Process Criterion:** The firms are grouped in an industry on the basis of similarity of process-technology, use of raw materials, methods of production, channels of distribution etc.

In general, all decisions of firms (price, output, product, style, advertisement, investment etc.) are taken in the light of actual as well as the potential competition by new entrants. This suggests that product criteria as well as process criteria should be integrated in analysing the market behaviour of firms.

Activity 2

a) Restate the following concepts in terms of the concept of marginals and averages:

i) Price-cross-elasticity of demand:

....

....

....

ii) Quantity-cross-elasticity of demand:

....

....

....

b) Explain the operational use of each of the above concepts of elasticity.

....

....

....

....

c) Name the criteria by which the following industries can be characterised.

Industry	Criteria
Textile	
Heavy chemicals	
Capital Goods	
Engineering	
Small Scale	
Electronics	
Iron & Steel	
Fertiliser	
Petro goods	
Automobiles	

d) Can the same industry be classified in terms of both product and process criteria? Give examples quoting other than listed in the above table.

....

....

....

....

e) Can the size of the firm be related to the structure of the industry?

....

....

....

15.3 RATIONALE FOR THE THEORY OF FIRM

It is now clear that the firm is a constituent element of an industry. Firms of different sizes, location, ownership, efficiency constitute an industry. The immediate environment of the firm is thus limited to what happend to the industry it belongs to. Industrial environment and its scanning explain the economic behaviour of the firm. However, we often observe that different firms belonging to same industry, facing the same market environment, behave differently. That is why the theory of firm is necessary .

The purpose of any theory is two-fold. First, it must explain what exists. Second it must enable us to predict what does not exist. Thus a theory must have both explanatory and predictive values. **Explanation** presupposes a **logical reasoning** used for analysis-starting from a set of assumptions to derive a set of **generalisations**. **Prediction** implies the possibility of forecsting based on observations. The validity of a **theory** (or the models it constructs) may be judged on the basis of several criteria: its **predictive power, its consistency**, realism of its **assumptions, its generality** and **application**, and its **simplicity**. There is no general agreement regarding which of the above criteria is more important. For example, milton Friedman emphasises the predictive value, while Paul Samuelson emphasises the explanatory power and realism of the assumptions.

The purpose of the theory of the firm is to provide models for the analysis of the decision making in the firm in various market structures. A theory of the firm needs to explain the whole range of price-output decisions-how the firms set their price, decide their product line, advertisement expenses and sales promotion efforts, research & development expenditure and so on-strategic as well as tactical decisions.

It is expected that a theory of the firm must have a minimum degree of generality so as to be applicable to the behaviour of a "group" of firms rather than a particular firm. Individual case studies are important and relevant, but several such case studies are required before a theoretical model of the behaviour of firms may be constructed

Over the years, markets have grown in terms of their complexity and so have grown the industries in terms of their variety. In their attempt to approximate the real world, the theorists have also developed their models at different levels of aggregation detail and sophistication depending upon their purpose. As a result we have today a variety of models explaining and predicting the behaviour of firms. These models can be classified into three sets of theory:

- Economic Theory of the firm
- Behavioural Theory of the firm
- Managerial Theory of the firm

And the interesting point is that these theories are still in the making, because real world practices still defy the existing theoretical constructs.

In this unit, we hope to limit ourselves only to the economists theory of firms, as other theories will be dealt with in subsequent units in this Block.

15.4 ECONOMIST'S THEORY OF THE FIRM

In the previous Block, you have been already exposed to the details of this theory. You have seen how the firms take or make price-volume decisions under different forms of market structure like perfect competition, monopoly, monopolistic competition-large group and small group i.e., oligopoly. There are other forms of market as well as such as monopsony, bilateral monopoly and so on. Such analysis includes both product markets and factor markets. In what follows, an attempt is made to provide an overview of all these models and thereby summarise the economists theory of the firm in terms of the following propositions:

1. **The firm is a transformation unit.** The economists use the term "firm" in a very general sense. Any unit which transforms input into ouput is called a firm; it may include an agricultural farm, a manufacturing unit like factories, mines, a commercial enterprise, a consultancy firm, bank, a hospital or an educational institution and what have you. The only requirement is that it must transform a set of inputs like men, machines and materials into a flow of output-finished or semi-finished, consumer goods or capital goods, goods or services. Thus every firm has a "transformation function" which defines its economic activity and technological process.

2. **The firm undertakes transformation of factors into product to create surplus value.** Transformation is economically meaningless if it does not add to value. In the process of transformation

the value of output must exceed the value of input, thereby creating a "surplus". In other words, there is no point in transforming log of wood into a set of furniture,if furnitures donot command a higher value in the market, compared to wood. Through transformation, the firm can add "form-utility" just as through transformation, the firm can add "place utility" to its product, adding such utilities enhances the market value. In modern terminology, the 'surplus value' is renamed as **profit** i.e., the surplus of revenue (value of output) over costs (value of input).

3. **The firm not only makes profit, rather it aims at maximum profit.** The basic objective behind a firm's transformation activity is profit maximisation. Profit reflects the firm's operational efficiency; profits distributed by way of dividends satisfy the shareholders; retained profit is the source of internal finance necessary for the firm's expansion and replacement programme. Considering all these, larger the profit, better it is for the firm, its survival and growth. Therefore, all firms aim at maximum level and maximum rate of profit. Profit maximisation thus guides the firm's activities.

4. **In order to maximise profit, the firm follows the equi-marginal principle, MR = MC.** You may recall that profit is the accounting difference between total revenues and total costs.

$$\pi = R - C \qquad \text{..... (i)}$$

Since revenues R and costs C - both depend on output Q, we may rewrite

$$\pi(Q) = R(Q) - C(Q) \qquad \text{.... (ii)}$$

It follows $\Rightarrow \frac{d\pi}{dQ} = \left[\frac{dR}{dQ} - \frac{dC}{dQ}\right]$ ..... (iii)

and $\frac{d^2\pi}{dQ^2} = \left[\frac{d^2R}{dQ^2} - \frac{d^2C}{dQ^2}\right]$ ..... (iv)

By the first order (equilibrium condition), we set (iii) equal to zero and thus get $\left[\frac{dR}{dQ} - \frac{dC}{dQ}\right]$

MR = MC (marginal revenue = marginal costs)

By the second order (stability condition), we should have (iv) < O $\Rightarrow \frac{d^2C}{dQ^2}$ > O i.e. rising MC

These conditions of a firm's equilibrium hold irrespective of the nature of market structure to which the firm belongs. In contrast to the firm's equilibrium condition (MR = MC), the industry's equilibrium condition is (AR = AC) i.e., what economist's call "normal profit"; unless all firms earn this "normal profit" (no profit, no loss), the structure of the industry continues to be unstable because of the firm's movement in or out of the industry.

5. **The market environment of the firms is known completely.** There is complete knowledge about the firm's size, product and factors, also about demand, supply and prices, the structure of the industry, rival's number and moves. Finally, there is also some difinite knowledge about the time perspective–short–run or long–run–in which the firm is operating. In other words, the firms can identify its market, perfect or imperfect. The firms knows the market conditions and understands them so as to design its market policies. The firm does not have to incur any information costs to obtain knowledge about the market environment. In other words, the firm operates in the world of perfect knowledge and certainty. Most of the economic models have been developed under the assumption of certainty, though lately some models have been specifically developed to deal with uncertainty and imperfect or incomplete knowledge about the market.

5. **Based on the knowledge about the market, the firm decides on its operational variables to maximise profit.** The firm knows its objectives and constraints; and accordingly the firm follows optimum decision rule to operate on the relevant variables. For example, under perfect-competition, the firm is a price taker. Under imperfect competition, the firm is a price maker. If demand elasticities differ between markets, the firm is price- discriminator. Sometimes, the firm acts as price-leader or price-follower, depending upon the market situation. Sometimes, the firm works in terms of non-price factors like advertisement and selling costs, typical of monopolistic competition. The firm may operate on real/physical variables like quantity or quality of the product rather than price variable. The choice of operational decision variables ultimately depends on the nature of the market.

Thus the economists' theory explains the adjustment process of the decision variable of the firm; it also helps in predicting the firm's behaviour under a set of given assumptions.

Activity 3

a) State the basic propositions of the economic theory of the firms.

....

....

....

....

....

....

....

....

....

....

....

b) Take each proposition and comment on its real world applications.

....

....

....

....

....

c) What is the purpose behind the theory of the firms?

....

....

....

....

d) Identify the objectives and the constraints for the following firms.

Firms	Objectives	Constraints
ONGC		
NTPC		
NTC		
SBI		
DCM		
Escorts		

e) Can a firm control price and output simultaneously? Can you think of any other operational variable?

....

....

....

....

....

....

15.5 A CRITIQUE OF THE ECONOMISTS' THEORY

We will begin by identifying the basic assumptions underlying the economists' theory. We would like to restate them candidly so as to comment on them and their real world validity. In the process, we may indicate alternative approaches (hypotheses) to the economists' theory.

If you examine the propositions which we have stated in preceding section, you will find that there are two basic assumptions underlying the economists' theory :

i) Motivational Assumption

ii) Cognitive Assumption

The **motivational assumption** is summarily stated in the term "profit maximisation". The motive or the driving force behind the firms activity is profit. The objective of the firm is, therefore, to maximise profit. This assumption, in its term, has two aspects . First, the firm maximises profit and no other decision variable. Second, the firm maximises profit and maximisation describes the behavioural intention of the firm.

The **cognitive assumption** is stated in terms like "complete knowledge" or "the world of certainty". The firm has perfect and complete knowledge about its decision environment and that is why it can act rationally to operate on a decision variable as per the optimum decision rule. In other words, the firms are able to recognise the market environment they are in. This assumption implies zero information cost i.e., there is no cost involved in obtaining perfect and complete knowledge. This is also means that the market environment is a risk-free world; there is no uncertainty; everything there is known, definite, calculate and predictable.

Both assumptions have **attacked** by several authors. Let us review them categorically.

With regard to the motivational assumptions, they have raised two related but fundamental questions:

a) Do firms really maximise **profit**?

b) Do firms only **maximise** profit?

The real world observation suggests that profit is not the sole variable in the objective function of a firm. The firms may have other goals also such as production, inventory market share, consumer consciousness, social responsibility and so on. Also, the firms need not always "maximise", but only "satisfy"; they may be only deliberative in their approach.

In view of this general observation and criticism, a set of **alternative hypotheses** have been suggested by authors, some of which are quoted below:

i) Firms intend to maximise not a single variable profit function but a general preference function, considering the profit interests of various groups like owners, shareholders, managers, workers, consumers, government etc. **(Popandreou)**

ii) In view of long-run survivals, the firms aim at organisational security and stability. **(Rothschild)**

iii) The firms are interested in 'safety margin' particularly when market competition is confined among the few. **(Fellnex)**

iv) Firms sacrifice profit objective because of reasons of financial control; firms go for internal financing out of retained profits so that they can keep financial control over their operations. **(Reder)**

v) Business firms (like banks) attempt to keep liquidity reserve sufficient to assure a sound financial position and retention of control. Liquidity Considerations are combined with safety and economy consideration. **(Cooper)**

vi) Firms often trade-off between profit and leisure. More profits you seek, less leisure you have at your disposal; firms therefore, aim at an ideal combination of profit and leisure in view of organisational safety and security. **(Scitovsky)**

vii) Business firms often are found to maximise sales revenue subject to the constraint imposed by a target profit set exogeneously. In the dynamic context, the firms aim at stable growth of revenue and profits are then endogeneously determined. **(Baumol)**

viii) Firms do not maximise profit, they only maximise the utility involved in managerial discretionary power. **(Williamson)**

ix) Firms aim at balanced growth subject to financial and managerial constraints they face. **(Marris)**

x) Oligopoly firms are not interested in maximum profits, they are only interested in stable market shares and reasonable regular flow of profit attached to it, they may thus share profits and share markets so as to occupy slowly and gradually the position of a market leader. **(Stackleberg)**

xi) Firms' behaviour is often 'satisficing' rather than 'maximising' satisfaction depending on the match between their aspiration and actual performance. **(Simon)**

This is just a suggested list. There are many more hypotheses. For example, it is sometimes observed that firms may sacrifice profit in view of their social responsibilities; Sometimes they are concerned about only 'customer service', 'pollution control', vulnerable sections of the society—'mass welfare', 'self sufficiency', 'autonomy', etc; some of the considerations may run counter to profit motive. This is often true in case of our public sector units.

Let us now come to the other assumption. The cognitive assumption is also not tenable in reality. Business decisions are meant to cope with changes. Most of these changes are unknown. The order, direction and impact of the changes can never be definitely known. Thus the decision environment is full of risk and uncertainty. To assume perfect knowledge is, therefore, highly unrealistic. In other words, the firms operating with imperfect knowledge cannot hope to work for maximum profit. The economists have been mostly considering production costs and selling costs in calculating their profits. The neglect of information costs is not justified. There is the cost of collecting, processing, storing and disseminating information in view of firm's own interest. Thus computer based management information system has developed and no modern firm can assume zero information costs. The firm has to search for information flow and choose the relevant information to direct its course of economic activity. In view of this, a few more hypotheses have been suggested:

i) The theory of choice and the theory of search thus become two essential parts of the theory of business firm. **(Savage and Small)**

i) The firms are "deliberative" rather than "maximising" in their attitude. **(Gordon and Margolis)**

) A theory which assumes knowledge of what cannot be known is clearly defective as a guide to actual behaviour of the firm. **(Boulding)**

) In the absence of knowledge concerning entrepreneurial horizons and expectations, the profit-maximising construction becomes an empirically irrelevant tautology. **(Popandreou)**

) Firms do not maximise profit; in fact, they cannot maximise profit, because they do not know what is the maximum to aim at. In reality, they follow short-cuts and rules of thumb rather than the marginal principle of profit maximisation. **(Hall and Hitch)**

If you consider the last three propositions above, you will find that they constitute a vehement attack on the economists' theory of the firm. You may wonder why the economists do not defend their theory. This is exactly what some economists have done. They have tried to rehabilitate the econmists' theory. **Milton Friedman**, for example, has argued that the test of a theory depends on its explanatory and predictive value. The soundness and significance of a theory depends on its operational use, its underlying assumptions cannot be the relevant point of attack. **Machlup** has defended the orthodox economic theory of the firm as reasonable; the profit maximisation considerations do guide business decisions; it is a subjective consideration, not capable of being measured objectively. The marginal principle, if necessary, can be modified by the more general principle of incrementalism. Incremental concept and reasoning often guide business decisions. **Horowitz** has reconstructed the economic theory of the firm under uncertainty. Thus profit maximisation hypothesis can be extended to take care of risk and uncertainty in the real world. The recent developments like game theory, econometric theory, theory of probability and economic decision models—all go to strengthen the profit-maximisation hypothesis of the traditional theory. In fact, alternative developments such as behavioural theories and managerial theories of the firms also do not substitute but only supplement the economic theory of the firm. In subsequent units, we will examine this observation. For the time being, we can say, based on empirical observation of "excellently managed firms" that **firms may not always maximise profit, but they do have a profit policy.** Thus profit may or may not enter explicitly the objective function; profit may be in the form of a constraint rather than an objective, yet profit is an index of operational efficiency of the firm and therefore, firms do have a definite policy about the size, rate, sources and allocation of profits. No firm can do without a profit policy. In the long-run profit remains as a motivational force. though in the short-run other goals may assume importance. And these goals may also be related directly or indirectly to the profit goal. In realistic cases, the firm has multiple goals and multiple constraints. The goal programming technique can still provide a dependable tool for analysis of economic behaviour of the firm towards profit maximisation.

Activity 4

a) Evaluate Friedman's position with respect to the economic theory of the firm. (Read the relevant reference)

....

....

....

....

....

b) Why is the cognitive assumption underlying the theory of the firm not tenable?

....

....

....

....

c) List a set of Indian firms under the following two categories:

Profit-miximising	Not profit maximising
....	....
....	....
....	....
....	....
....	....

15.6 SUMMARY

The firm is the microlevel decision-making unit existing within an industry and a market environment. The decision-making process of the firm is very complex; it depends on the nature and efficiency of the firm itself. It is very difficult to generalise on the economic behaviour of the firms. Yet the economists have attempted this generalisation at the cost of over-simplification. The economists have put forward the profit maximisation hypothesis which they think, holds good irrespective of the nature marked and its perfection, irrespective of the time-perspective, short-run or long-run, irrespective of the differences among the firms and the industry to which they belong. There may be questions on the realistic nature of the underlying assumptions of the economists' theory, but its explanatory and predictive value is immense. Quite a few alternative hypotheses, alternative to profit maximisation, have been suggested, but some how they still remain hooked to the concept of profit. In fact, even behavioural and managerial theories seem to supplement rather than substitute altogether the economic theory of the firm. This is what we will examine in the next two units before returning to the concept, measure and status of profit in our managerial economics.

15.7 ADDITIONAL READINGS

Adhikary, M., 1987. *Managerial Economics,* Khosla Publishing House: New Delhi. (Ch. III, Section 3& 4).

Anthony, N.R., 1960. "The trouble with Profit Maximisations "*Harward Business Review,* Nov. & Dec. 1960. pp 126-134.

Friedman, M., 1953. "The Methodology of Positive Economics" in *Essays in Positive Economics,* pp. 3-43.

15.8 SELF-ASSESSMENT TEST

1) State and explain the assumptions underlying the economists' theory of the firms.

2) "In theory, we assume perfect competition and profit maximisation, in reality,we find neither" —Comment.

3) What is a theory? What is its purpose? How can you judge the validity of a theory? Illustrate your answers to these questions, with reference to the economist's theory of the firm.

4) The actual economic behaviour of the firm cannot be explained with or without the concept of profit. —Comment

5) Make a list of firms. Classify them using source criteria like product or process. Now examine their objective and constraints. Finally comment on the usefulness of the profit-maximisation hypothesis.

UNIT 16 THE BEHAVIOURAL THEORY OF THE FIRM

Structure

16.1 INTRODUCTION

The traditional economic theory has been challenged on a number of accounts and in its place. Some authorities have come up with the behavioural theory of the firm. The behavioural theory re-examines the concept of firm and its environment; it analyses its decision-making process and consequences. In the process, it provides an alternative to the traditional thinking about the firm and its activity. For example, it suggests that the firm is not a single-goal decision-making unit, rather it is a multiple-goal operating unit such that in addition to profit-goal, the firm has also other goals like production goal, inventory goal, sales goal and market share goal. Such goal formation is an interesting process. In the process of forming demand-goal or what is called 'aspirations', the firm has to satisfy different groups of conflicting interests. Thus the firm emerges as a 'satisficer' rather than a 'maximiser'.

The behavioural theories started developing in the early 1950s. Some of the seminal work may be traced in Simon's classic article, "A Behavioural Model of Rational Choice", published in the **Quarterly Journal of Economics** in 1955. This theory has subsequently been elaborated by Cyert and March, with whose names it has been associated. In what follows, we shall reproduce a simple model of behaviourism as constructed by Cohen and Cyert.

16.2 OBJECTIVES

On reading this unit, you should be able to

- **understand** a few key concepts of the behavioural theory
- **state** clearly the assumptions underlying a behavioural model of the firm
- **describe** the decision-making process in a multi-goal, multi-product firm under uncertainty, operating in an imperfect market
- **have** an insight into Simon's model as well as of Cohen and Cyert
- **contrast** the 'satisficing' behaviour with the 'maximising' behaviour of a firm
- **take** stock of the contributions and limitations of the behaviour theory of the firm

16.3 SIMON'S SATISFICING THEORY

Inadequacy of the economist theory of the firm prompted writers in other disciplines to develop alternate models explaining the economic behaviour of the firm. One of the first of these was Simon's in 1955.

Simon argues that businessman must always have imperfect knowledge on which to base decisions, that if full knowledge was available, the calculations involved in decision-making would be too complex to be practicable; and that given this and other uncertainties surrounding decison-making in reality, businessmen can never know whether they are maximising profits or not. Instead businessmen only "satisfice", they do not "maximise", they aim merely at satisfactory profits. According to this line of reasoning, the firm; are satisfiers rather than maximisers. These firms are satisfied if their expectations or "aspirations" are achieved.

According to Simon, organisation behaviour and individual behaviour are comparable. Like an individual every organisation has its own aspiration, achievement, success or failure record and its aspiration level mechanism is set to work in view of its need, drive and attainment of goals. Firms make periodic review of their goals. Firms may face three atlernative situations:

i) The actual performance is less than the aspirations

ii) The actual performance is equal to the aspirations

iii) The actual performance is greater than the aspirations.

The first situation may be due to incomplete information about the future and the firms initiate the process of improving the quality of information. It is possible that the aspiration level has been pitched too high. The achievement level may lag behind the aspiration level because (a) there are wide fluctuations in economic activity, or (b) there is a qualitative deterioration in the performance level. Either way, there is flaw in information flow.

The second situation, when achievement matches the aspiration (targets) the firm is more or less satisfied. No action will be taken except to review that the aspired targets have not been pegged too low and that the firm's potential performance has been rightly estimated.

The third situation speaks of commendable performance by the firm. The firm is satisfied but there is a need to ensure that, the quantitative achievement is not the outcome of decreasing quality of performance. In this line of reasoning, it is suggested that we should question success rather than failure; questioning failure is normal, but questioning success is rare. When the firm critically analyses its success, it is in a position to detect the quality of its performance, this may enable the firm to design its future policy more efficiently.

It follows that except for the first situation, the firms remain satisfied. Faced with the first situation, the firms will have to organise "search activity" and "choice" to improve information flow. If neither the search behaviour nor lowering of aspiration level results promptly enough in the achievement of "satisfactory" situation, then the firm's behaviour pattern will become one of apathy or of aggression.

16.4 SOME COMMENTS ON SIMON'S MODEL

In analysing the economic behaviour of the firm, Simon has based his analysis on the analogy between individual psychology and organisational psychology. Simon's theory seems plausible because it is consistent with psychological theories of motivation which suggests that human action stems from drives, and that these actions terminate once the drives are satisfied. Thus Simon's model envisages a marriage between economics and psychology. Secondly Simon's model is quite consistent with empirical observation of business behaviour such as tendency of businessmen to set prices on the basis of costs plus a mark-up designed to generate "reasonable profit" or the tendency of big companies to state their profits targets in terms of earning a "satisfactory" return on capital employed.

The main drawback of satisficing theory is the difficult of making an operational statement of what is to be regarded as a "satisfactory" level of performances. Some critics argue that this approach seems less satisfying than the profit maximising model which suggests an optimum level of profit consequent upon the operation with the decision variables. Instead of one "optimum level" there may be many "satisfactory levels" depending upon the groups which take interest in the firm's activity. In this case, the operational value of Simon's model is nil or negligible.

The firms activity has to satisfy a number of groups. Unless good dividends are paid, the shareholders will be dissatisfied. Unless fair wages are paid, the workers will be dissatisfied. Unless low price is charged, the consumer will be dissatisfied. Unless attractive returns result from investment the owner will be dissatisfied. Unless good pay and perquisites are given, the salaried managers will be dissatisfied. Unless tax returns are paid, the Government will be dissatisfied. Thus, the firm, has to choose a particular profit level and rate, which will "satisfy" one and all. This may not be operationally feasible.

Activity 1

a) Recall the list of concepts, common to the analysis of both individual behaviour and organisational behaviour.

....

....

....

....

b) Baumol has a hypothesis that the firms maximise revenue subject to the constraint imposed by a target profit. To what extent can this hypothesis be incorporated within the framework of Simon's model? Give reasons for your answer.

....

....

....

....

....

c) Visit some firms personally and interview a few selected executives. Whom and how do their firm's "satisfy"? Examine the response against Simon's model.

....

....

....

....

....

16.5 BASIC CONCEPTS IN CYERT AND MARCH MODEL

In their book A Behavioural Theory of the Firm (1963), Cyert and March have focused on the decision-making process of the large multi-product firm under uncertainty in an imperfect market. In such firms, ownership is divorced from management. Cyert and March show their interests more in the **decision-making process** than in the **motivation**, profit maximising or satisficing, of such firms.

There are certain concepts, basic to the understanding of Cyert and March Model. Some of these concepts are:

Firm as an adaptive rational system has the following properties:

i) There exists a number of states of the system. At any point, the system in some sense 'prefers' some of these states to others.

ii) There exists an external source of disturbance or "shock" to the system. These shocks cannot be controlled.

iii) There exists a number of decision variables internal to the system. These variables are controlled according to some decision principle.

iv) Each combination of 'external shock' and 'internal decision variable' changes the state of the system. In other words, a new state is decided by the existing shock from the external environment and by the decisions taken and executed from within the firm.

v) Any decision rule that lead to a preferred state is more likely to be used in future than it was in the past.

In view of these propositions, we can say that the rationality on the part of firm's behaviour suggests that the firm, operating with its internal decision variables, adopts itself to changing external shocks and thus a new state results and the system continues through changing states.

Firm as a Coalition is another view about the firm's behaviour. You have already seen that the firm has to satisfy various groups—consumers, workers, shareholders, owners, managers, suppliers, financiers (institutions), Government (Bureaucrats and Ministers) and so on. If a firm is operating successfully, it means that the Coalition of groups of conflicting interests is working satisfactorily. It is this **Coalition** within the organisation which has to be maintained so that the firm survives and grows eventually. The growth of the firm may strengthen or endanger the Coalition depending upon how the vested interest-groups are satisfied.

Organisation Goal has reference to the objective of the firm. Like individuals, organisations also set their goals to direct their respective business activity. The Coalition members together decide on the goal and the goal-directed activity of the firm. Several points regarding organisational goals should be noted:

i) Goals are decided upon by the interaction among the Coalition members.

ii) All resolutions of goals within the Coalition are not made by money. Rather many side payments are made to Coalition members in the form of policy commitments.

iii) Some goals are stated in the form of normative dictum.

iv) Some goals are stated in non-operational form.

v) When goals are stated in the form of operational targets, it is possible to compare "aspirations" with "achievements". Such comparisons, particularly periodic review, help the Coalition to revise goals and review its aspiration level mechanism.

vi) The firms mostly have multiple goals, rather than a single profit goal.

Organisational slack: A Coalition is viable if the payments made to various Coalition members are adequate to keep them together in the organisation. If enough resources exist to meet all demands, the Coalition is a feasible one. However, because of frictions in mutual adjustment of payments and demands (aspirations), there is ordinarily a disparity between the resources available to the organisation and the payments required to maintain the Coalition. This difference between available resources and required payments is called **"organisational slack"**. Slack consists in payments to members of the Coalition in excess of what is required to maintain the organisation.

When firms operate under market imperfections, many forms of slack exist. The shareholders may be paid dividends in excess of those which are required to keep them within the organisation. The workers may be paid wages in excess of what is needed required to keep them committed to the firm. The executives may be paid salaries more than what is needed to prevent them from leaving the organisation. The customers may be given price discount in excess of what is required to purchase the product. All such excess payments are slack expenditures for the firm. In traditional economic theory, at least in equilibrium, such slack is zero, but in behavioural theory slack is positive and it has a constructive role. Slack payments keep the Coalition in existence. Organisational slacks enable firms to adjust themselves to gross shifts in external environment. Slacks represent a cushion to absorb the shocks. Thus slack play both a stabilising and an adaptive role.

Activity 2

a) The Firm is a Coalition of groups of conflicting interests — TRUE OR FALSE? Explain.

....

....

....

....

b) The Firm is not a single-goal decision unit. State the multiple goals of a firm.

....

....

....

c) Explain the process of formation of organisational goals.

....

....

....

....

d) Bring out clearly the role of the following payments as means for resolution of group conflicts in an organisation:

....

....

....

....

Money payments:

....

....

....

Side payments:

....

....

....

Slack payments:

....

....

....

e) Slack payments increase during periods of flourishing business and decrease during periods of bad business; Explain quoting real world examples.

....

....

....

....

....

16.6 DECISION-MAKING PROCESS IN CYERT AND MARCH MODEL

Organisational goals set by top-management are to be implemented by decision. Decisions are taken/made at various levels of hierarchy within the organisation. From this standpoint, two levels of decision-making can be distinguished-one at the level of top management, and the other at the lower level of administration.

At the top management level

Given the organisational goals and resources, the allocation of these resources to the various departments is decided by the top management and is implemented by the budget. Each department is allocated a certain portion of the budget, the share of the budget depending mainly on the bargaining power and skill of the head of each department. The bargaining power is to a large extent determined by the past performance of the particular department. In this process of internal allocation of most of the resources, the top management always retain some funds for allocation at its own discretion at any point of time to the best of its judgement and ability. The top management examines and decides on the proposition (project) presented to it by the managers. Two crude criteria are set for the evaluation of projects:

i) **Budgetary (financial) criterion:** Availability of funds for the project

ii) **Improvement criterion:** Contribution of the project towards modernisation etc.

In the above decision process, information is required to facilitate the decision-maker. Information is not without cost. Search has to be undertaken to locate and collect information. Search is a resource-absorbing activity, but it is not made on marginalistic rules (as suggested by the traditional economics thoery). Search is directd to the particular area in which some problem appears. Search is undertaken when and where a problem arises so that minimum resources are spent in that activity. Information flow has a very important role. Information determines the aspirations (demands) of the groups, which, in turn, determines the setting of goals by top management. Information flow has to be monitored to bridge the gap between aspirations and achievements of the firm. But information flow often suffers from a **position bias;** every manager provides such information as would safeguard his individual position. Also information may get distorted, diluted and delayed depending on the channels of information flow. Thus the decisions taken by top management may not always depend on adequate and appropriate information.

At lower levels of administration

The decision process at lower levels involves various degrees of freedom of action. Once the budget-shares are allocated each manager has normally considerable "discretion" in spending the funds allocated to his department. For example, the sales force allocation is decided by the sales manger; the organisation of labour on the shop floor is decided by the production manger. Day-to-day decisions are simplified by delegation of authority within each section and by simple rules of convention which form the "blue print" of the organisation. The administration staff at lower levels of the hierarchy "learn" by experience and are helped by the "blue print" rules in making their decisions. Based on their experience, the lower level staff clearly know how, when, where and how much of top management decisions to implement. Individual staff officers adapt themselves in the context of developments within and outside the firm which itself is an "adaptive rational system".

16.7 COHEN AND CYERT MODEL OF BEHAVIOURISM

In this section, we briefly present a simple model used by Cyert and March as an illustration of the decision-making process within a large corporation. Model refers to the case of a duopoly, with homogeneous output selling a single price ultimately. No changes in the inventory are allowed in this model. The steps outlined here are based on Cohen and Cyert (1965).

i) **Forecast of rival's reaction:** The forecast is a straight forward extrapolation of the past observed reactions.

ii) **Forecast of firm's demand:** Future demand is an extrapolation of past sales.

iii) **Estimation of costs:** The current costs are assumed to be the same as the historical (past) costs, with necessary adjustments (increase) of 'slack payments'.

iv) **Specification of goals.** The firm specifies its profit goal, based on revenue and cost calculations. The aspiration level of profits may be taken as some average of past profits.

v) **Evaluation of achievements vis-a-vis goals (aspirations): The actual performance of the firm** is compared with the aspired goals as translated in terms of the target variables like price-and-output. If achievement matches aspiration,the firms is satisfied; if it is not, then the firm is dissatisfied and starts a fresh decision-making process sometimes seeking an improvement in the information system.

vi) **If profit goal is not achieved the firm re-examines the cost estimate** because some costs in the form of slack expenditure are under direct control of the firm.

vii) **If the new solution with downward adjustments of costs leads to target profit it is adopted.**

viii) **If cost adjustments do not yield target profit, the firm re-examines its demand estimates:** An upward adjustments in initial demand calls for a revision in sales strategy.

ix) **Evaluation of new solution:** If new solution with revised demand and cost estimates attains the profit goal, it is adopted. If not, the firm proceeds to the next step.

x) **Downward adjustment of aspirations:** The firm makes downward adjustment in its profit goal (aspiration levels).

In actual practice, the firm has multiple goals. Profit goal is just one among them. Other goals are: production goal, inventory goal, sales goal, market share goal etc. When the firm works with multiple goals, in a goal programming framework, it works as a "satisficer" rather than a "maximiser". The goals change overtime depending on past achievements, aspirations, demands of coalition groups, expectations, shocks and slacks in the system etc. When failure occurs, search is undertaken in that direction. Search activity yields market information which may be used towards improvement in the firm's performance. By information searches, the firm wants to overcome **market uncertainty** (like changes in tastes, preferences and techniques). On the other hand, **Uncertainty of rival's reaction** is overcome by creating a "negotiated environment".

Activity 3

Compare the behavioural theory with the traditional economic theory of the firm with reference to the following parameters concepts.

No.	Concepts/parameters	Economic theory	Behavioural theory
1)	Concept of firm		
2)	Motivational objectives		
3)	Decision rules		
4)	Decision process		
5)	Decision variables		
6)	Slacks		
7)	Degree of uncertainty		
8)	Information costs		
9)	Search activity		
10)	Behavioural feature		

16.8 A CRITIQUE OF BEHAVIOURAL THEORY

The behavioural theory has made positive contributions towards the development of the theory of firm in several respects. Its main contributions are:

i) **It provides an insight into the process of goal information.** It takes a very realistic view of the firm as a coalition of groups of conflicting interest. The resolution of conflicts is a step towards formation of agreed goals to be implemented through managerial action.

ii) **It provides a practical view of the process of decision-making.** Setting targets in view of multiple goals, choosing instruments to attain the targets, specifying the decision variables to be operated upon, evaluating the achievements vis-a-vis aspirations, organising search activity to improve information flow towards better performance-these are all elements in decision-making process and the behavioural theory has built them together in a logical framework which holds in reality.

iii) **It deals with resources allocation within the firm, from behavioural angles** hitherto unexploited by the traditional economic theory. The resource allocation process reveals the group dynamic process operating within the firm. But the internal environment is not emphasised at the cost of external environment. Environmental 'shocks'are adapted by the firm.

iv) **The concepts of 'slacks' and 'shocks' have tremendous operational use.** Slack payments, in particular, play a stabilising role for the firm and its orderly function. (You may note that the behavioural concept of 'slack' is equivalent to the economic concept of 'rent'. However, Cyert and March deal with only one form of slack, the managerial slack, but other forms of organisational slack are equally important and operationally useful).

Despite these contributions, the behavioural theory has a number of shortcomings:

i) The behavioural theory basically provides a **simulation approach** to the complexity of the mechanism of the modern multi-goal, multi-product corporation. Simulation, however, is a predictive technique; it **does not explain the behaviour of the firm.** It predicts the behaviour without explaining any particular action of the firm.

ii) **The behavioural theory does not deal with industry equilibrium.** It does not explain the interdependence and interaction of firms, nor the way in which interrelationship of firms leads to a stable equilibrium of price-output at the industry level. There is no discussion as 'condition of entry' or effects of a threat of potential entry on the behaviour of established firms.

iii) **The behavioural theory concentrates on short-run situation** when search activity is organised on a selective basis in problem areas only. Such search activities cannot explain dynamic aspects of invention and innovation which have long-run implications

iv) [illegible] **(readjustment) of the aspiration levels downwards** whenever the targets are not attained, **deprives the theory of any objective criterion of measuring the satisfactory performance of firms.**

v) No exact prediction about the firm's behaviour can be made by the behavioural theory. **The acceptance of satisficing behaviour renders practically the theory into a tautological structure;** whatever the firms do can be rationalised on the lines of satisficing.

vi) Cyert and March have based their theory on four actual case studies and two experimental studies conducted with hypothetical firms. Thus **the empirical base is too limited to furnish the details of theorising.**

On balance, the behavioural theory, despite some limitations, has enabled us to think of multi-goal firms, without rejecting the profit goal. In the present-day world, when the firms have to deal with so many constraints, market and non-market, within and outside the firm, it seems logical to suggest that their approach is towards satisfaction rather than maximisation. Of course, we do need more of empirical research to enrich the existing behavioural theory.

Activity 4

In India, we have a large number of small firms which are single-product, single-entity. Does that mean that the behavioural theory is irrelevent? Can you think of any concept or proposition in this unit which is applicable to such firms.

....

....

....

....

....

(You may like to organise a discussion on the subject: Are there no "shocks" for small firms? Do small firms pay no "Slacks"?)

16.9 SUMMARY

As an alternative to the traditional economic theory of the firm, the modern behavioural theory of the firm has been developed. The behavioural theory considers a large firm producing multi-product, having multiple goals, operating under uncertainty in an imperfect market. Thus there are considerable differences between the economic and the behavioural theory. Some of the key concepts in this theory are firm as an adaptive rational system, firm as a coalition of groups of conflicting interests, shocks, slacks etc. The economic firm maximises but, the behavioural firm satisfies in view of market uncertainties and uncertainties relating to rival's reaction. The behavioural firm aims at bridging the gap between aspiration and achievement by monitoring information flow through organised search activities in problem areas. Despite making a lot of contribution towards a refreshing a new analysis of firm's decision-making process and resource allocation process the behavioural theory suffers from a set of limitations because it concentrates on short-term perspective and it takes simulating approach which can predict rather thanexplain the firm's behaviour. However, the behavioural theory does not really replace the traditional theory, it only supplements and extends traditional theory. For example, in place of profit goal, it considers other goals relating to production inventory, sale and market share. More empirical research is needed to substantiate the theoretical hypotheses of behaviourism.

16.10 ADDITIONAL READINGS

Koutsoyannis, A., 1979. *Modern Microeconomics.* Second edition, The Macmillan Press Ltd., New York (Ch. 18).

Cyert R.M., and J.G., March 1963. *A Behavioural Theory of the Firm,* Prentice Hall

Choen K.J. and R.M. Cyert, 1965. *Theory of the Firm,* Prentice Hall: Englewood Clif

Simon, H.A., 1955. "A Behavioural Model of Rational Choice", *Quarterly Journal of E[illegible]* *(1955).*

16.11 SELF-ASSESSMENT TEST

1) Recall some of the key concepts and basic assumptions underlying the Behavioural Theory of the Firm.

2) Examine clearly the role of the following in the context of the Behavioural Theory:

   a) Slack payments
   b) Search activities
   c) External shocks

3) Explain the process of–

   a) goal formation
   b) conflict resolution
   c) decision-making
   d) achievement evaluation

4) Compare and contrast the Behavioural Theory with the Economic Theory of the Firm.

5) Comment briefly on each of the following statements:

   a) Taking a simulation approach, the behavioural theory can only predict rather than explain firm's behaviour in many cases, it cannot even accurately predict the firm's action.

   b) The behavioural theory is short-sighted. Without taking a long-run view, it is not possible for a firm to organise planning towards minimising uncertainties, typical of an imperfect world.

   c) The firm 'satisfices' rather than 'maximises'. It is a simple proposition, but it emphasises psychology rather than economics in explaining the economic behaviour of the firm.

UNIT 17 MANAGERIAL THEORIES OF THE FIRM

Structure

17.0 INTRODUCTION

Managerial theories of the firm can be considered as a subset of the behavioural theory that we have presented in the previous unit. As such, while analysing the economic behaviour of the firm, the starting hypothesis is the same: the firm is a **coalition** (of managers, workers, shareholders, suppliers, customers, tax collectors) whose members have conflicting goals which must be reconciled so that the firm survives. Additionally, it is hypothesized that the mos important member is the 'top management' because of its power in decision-making and access to information. The basic characteristic of 'managerial business' is the **divorce of ownership from management.** The owners are promoters and shareholders whose power lies in appointing the Board of Directors, which in turn appoints the top-management. The top-management enjoys considerable power in strategic and tactical decision-making, provided that the level of profits is 'acceptable' to the owners, the rate of firm's growth is 'reasonable' relative to the growth of other firms, and that the dividends paid out to the shareholders are 'sufficient' to keep them happy. If these conditions are satisfied there will be 'job security' for top management. The divorce of ownership from management permits the top management to deviate from profit-maximisation (which maximises utility to the owner) and pursue goals which **maximise their own utility subject to a minimum profit constraint** implied in the satisficing conditions stated in the last paragraph.

We will present three models of managerialism: Baumol's model of 'sales maximisation', Marris's model of managerial enterprise and Williamson's model of managerial discretion.

17.1 OBJECTIVES

On reading this unit, you should be able to:

- **draw** a line of contrast between organisation behaviour and individual manager's behaviour;
- **Understand** the basic assumption underlying three different models of manager's utility maximisation;
- **locate** the differences among these models of managerialism;
- **compare** the managerial theories with the traditional economic theory of the firm; and
- **appreciate** better the objectives, constraints and instruments (policy variables) in the context of modern firms where ownership is divorced from management.

17.2 BAUMOL'S MODEL OF SALES REVENUE MAXIMISATION

In his celebrated work **Business Behaviour, Value and Growth** (Macmillan, New York, 1959), W.J. Baumol suggest sales maximisation as an alternative to profit maximisation hypothesis. He has basic model: the first is a static single-period model, the second is a multi-period dynamic model. Each model has two versions, one without and one with advertisement expenditure. Some empirical evidence is available to verify Baumol's hypothesis of sales- maximisation and examine thereby the predictions of Baumol's model.

Static Models

Baumol's static models have the following assumptions:

1) The time-horizon of a firm is a single period.

2) During this period, the firm's objective is to maximise sales revenue (i.e., value rather than volume of sales).

3) The firm operates under the profit constraint; the critical minimum profit is exogenously determined by the demands and expectations of the shareholders and other members of the coalition.

4) Conventional U-shapes average cost curve and the downward-falling average revenue curve hold true.

Under the set of above assumptions, it is possible to identify four different models:

- A single-product model, without advertisement
- A single-product model, with advertisement
- A multi-product model without advertisement
- A multi-product model with advertisement.

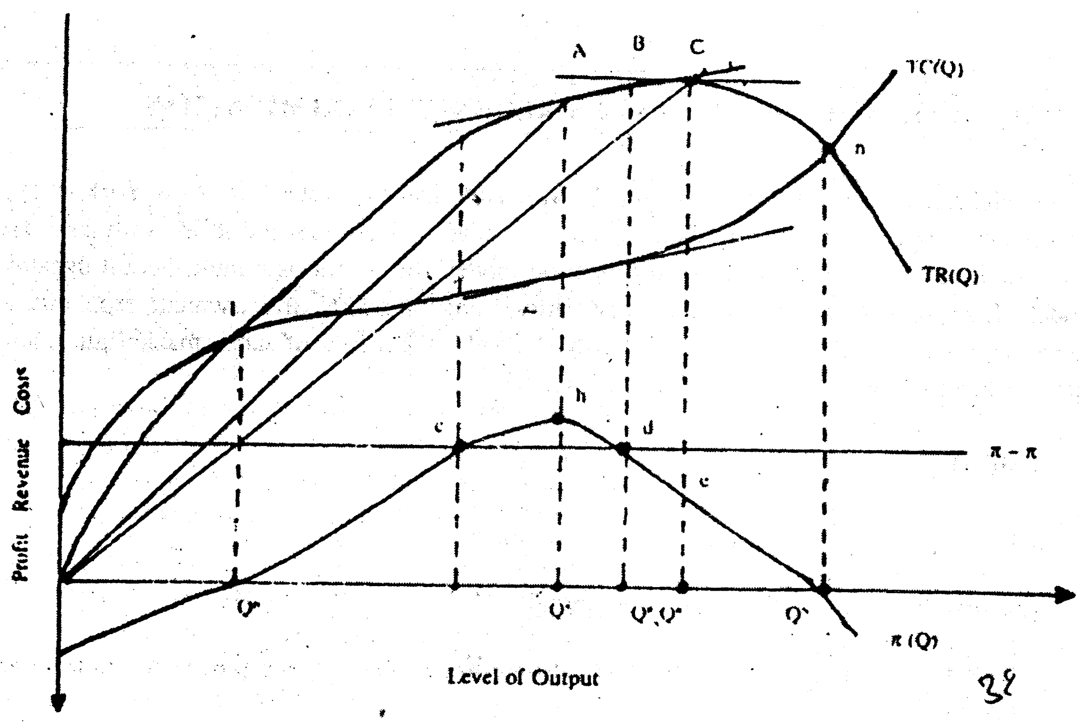
Let us consider a **single product firm without advertisement**. You may recall that the firm may maximise total revenue by setting the marginal revenue equal to zero. This is the **absolute activity level principle.** Thus a profit maximising firm follows the relative activity level principle, MR=MC, whereas for a sales revenue maximises, MR=0. The optimum decision rule for a revenue maximiser implies, at the level of equilibrium output, the firm has unitary price elasticity of demand for its product, because the elasticity,

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

Since MR = 0,

$$\Rightarrow e = \frac{AR}{AR} = 1$$

The implications of Baumol's static model can be brought out more clearly, if you concentrate on the following diagram:



38

You may note that TR (Q) and TC (Q) curves have been drawn, assuming conventional shapes and slopes; based on which the profit curve, $\pi(Q)$ has been drawn. OF measures fixed costs; if nothing is produced, then OF=OL measures loss. Corresponding to the "break-even" point b, we have profits to zero, just as at the 'normal profit' point, n, we have zero profit. The break-even output Q may now be compared to the unconstrained profit maximising output, $Q\pi$, at which the slopes of TR and TC are equal i.e.,

$\frac{d(TR)}{dQ} = \frac{d(TC)}{dQ}$ or MR = MC. The point h measures the top (maximum) of the profit bill. Q^R represents the unconstrained revenue maximising level of out at which $\frac{d(TR)}{dQ} = MR = 0$. In case, the firm has decided on a "critical minimum" (target) profit of OM, then you may find that this target is met either at c or d on the profit curve; of the two, we have relatively larger total revenue corresponding to the point d and hence $Q = Q^R_\pi$ measures the profit - constrained revenue maximising level of output which Baumol's firm chooses to produce.

Some predictions can now be made:

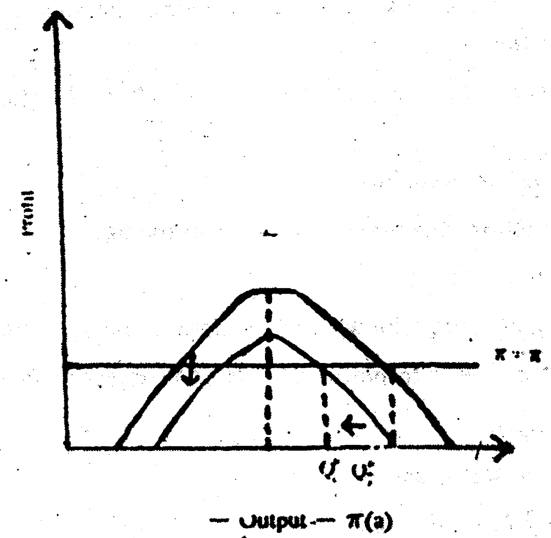
1) The sales (revenue) - maximiser will produce a higher level of output compared to a profit-maximiser. Note that $Q^R > Q^\pi$ or even $Q^R_\pi > Q\pi$

2) The sales-maximiser will sell at a price lower than the profit-maximiser. The price at any level of output is measured by the slope of the line through the origin to the relevant point on the TR curve, corresponding is that level of output. Note that the slope of OA> slope of OC or OB.

3) The sales-maximiser will earn lower profit than the profit-maximiser. Note that $[Q^R - e] < [Q^\pi - h]$

4) The sales-maximiser will never choose a level of output at which the price-elasticity is less than unity

5) An increase in fixed costs (say, the imposition of a lump-sum tax) will affect the equilibrium position of a sales maximiser: he will reduce his level of output by $\left[Q^R_\pi - Q^R_{f\pi}\right]$ and increase his price, since the increase in fixed costs shifts the total profit curve downward. By contrast, such changes in fixed costs leave the profit-maximising output unchanged.

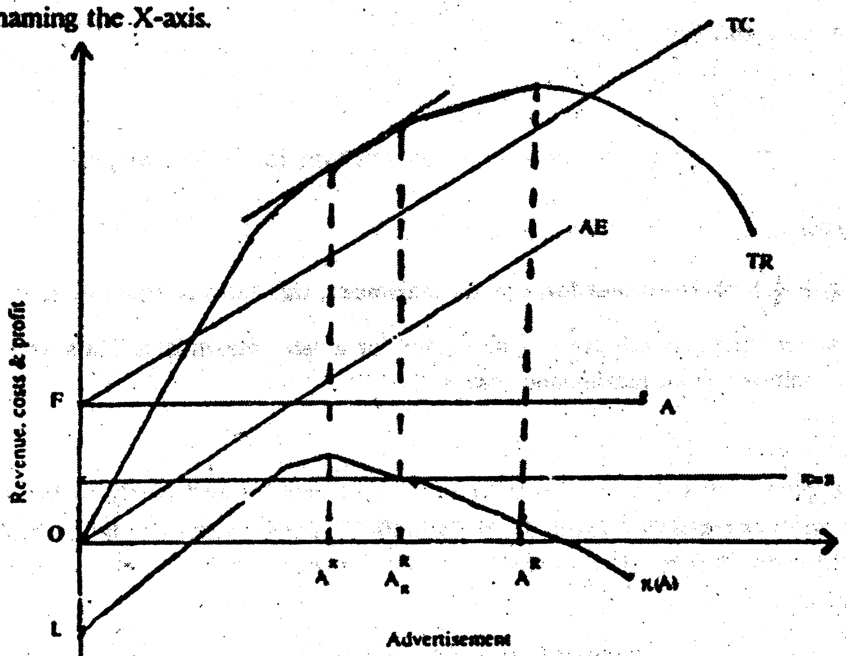


Let us now consider the case of a single product firm with advertisement. Baumol makes two crucial assumptions:

a) Advertising always increases sales revenue.

b) Total costs are independent of advertisement expenses.

Advertisement is a means of sales promotion. In a perfectly competitive environment, the profit-maximiser does not need to advertise his product because he is selling homogeneous product at a given price. However, in imperfect market where his product is differentiated and where he is a price-maker rather price-taker, advertisement has a role. The decision on the size of advertisement budget is a crucial decision. The sales-maximiser decides on optimum advertisement by examining its impact on sales revenue. You may readopt the earlier diagram by renaming the X-axis.



Using the same kind of logic as before, you may locate and find that revenue-maximising advertisement expenses are larger than the profit constrained revenue-maximising advertisement which is larger than unconstrained profit-maximising advertisement i.e., $A^R > A^R_{\bar{\pi}} > A^{\pi}$. You may note that - AE is the advertising expense curve; FA is the advertising curve; TC is the total costs including that advertisement.

You may now attempt a mathematical presentation of Baumol's model. Let us start by defining some of the relations:

TR = R (Q,A) $\Rightarrow$ Total Revenue Function

TC = C(Q) $\Rightarrow$ Total cost function, independent of advertisement

$\pi = \bar{\pi} \Rightarrow$ Critical Minimum Profit

AE = AE(A) $\Rightarrow$ Advertising cost function where Q is output and A is advertisement

The optimising problem of the firm is to maximise R (Q,A) subject to $\pi = \bar{\pi}$, given $\frac{d(TR)}{dA} > 0, \frac{d(TC)}{dQ} > 0$ and Q>0

The Lagrange function may be stated as

$$L = R(Q,A) + \lambda[TR - TC - A - \bar{\pi}]$$

The necessary conditions for a maximum are :

$$\frac{\partial L}{\partial Q} \leqslant 0, \frac{\partial L}{\partial A} \leqslant 0, \frac{\partial L}{\partial \lambda} \geqslant 0,$$

– Differentiating the L-function with respect to Q,

$$\frac{\partial L}{\partial Q} = \frac{\partial R}{\partial Q} + \lambda\left[\frac{\partial R}{\partial Q} - \frac{\partial C}{\partial Q}\right] \leqslant 0$$

Given Q>0, the above expression holds as an equality. Solving for $\frac{\partial C}{\partial Q}$, we get

$$\frac{\partial C}{\partial Q} = \frac{\lambda+1}{\lambda} \cdot \frac{\partial R}{\partial Q} = \left(1 + \frac{1}{\lambda}\right)\frac{\partial R}{\partial Q}$$

Given that $\lambda > 0$, it is obvious that

$\frac{\partial C}{\partial Q} > \frac{\partial R}{\partial Q} \Rightarrow MC > MR$ for Baumol's firm. Note that in case of a profit maximising firm, by contrast, MC = MR

Similarly differentiating with respect to A,

$$\frac{\partial L}{\partial A} = \frac{\partial R}{\partial A} + \lambda\left[\frac{\partial R}{\partial A} - \frac{\partial(AE)}{\partial A}\right] \leqslant 0$$

Given that A > 0, the above equation holds as an equality solving for $\frac{\partial(AE)}{\partial A}$, we get

$$\frac{\partial(AE)}{\partial A} = \left(1 + \frac{1}{\lambda}\right)\frac{\partial R}{\partial A}$$

or, $\frac{\partial(AE)/\partial A}{\partial R/\partial A} = \left(1 + \frac{1}{\lambda}\right) \Rightarrow$ Note that for a profit maximiser, this ratio is equal to 1.

Given $\lambda > 0$, the advertising expenditure will be higher for a sales maximiser. Thus, you may now argue that for a sales-revenue maximising firm,

$$\left[\frac{\partial C}{\partial Q} / \frac{\partial R}{\partial Q}\right] = \left[\frac{\partial(AE)}{\partial A} / \frac{\partial R}{\partial A}\right] = \left[1 + \frac{1}{\lambda}\right]$$

This implies that surplus profits will be devoted partly to advertising and partly to increased production. This is another prediction form Baumol's model. These models can be generalised with respect to multi-product firm. The product-mix of a revenue-maximiser will be different from that of a profit-maximiser.

Dynamic Model

The multi-period model of Baumol is based on the following set of assumptions:

1) The objective of the firm is to maximise the rate of growth of sales revenue over its lifetime.

2) Profit is not a constraint; it is the main source of financing growth of sales. Profit is thus an instrumental variable whose value is endogenously determined.

3) Demand and cost curves have traditional shape; average revenue is downward-falling and average cost is U-shaped.

Let us assume that sales revenue (R) grows at a rate of growth (g) per cent. Over its lifetime, the firm will have a stream of revenues,

R, R (l+g), R $(1 + g^2)$ = ,......... R $(1 + g)^n$

The present value of this stream of future revenues can be computed through the usual technique of discounting,

$$R, R\left(\frac{1+g}{1+r}\right), R\left(\frac{1+g}{1+r}\right)^2 ......, R\left(\frac{1+g}{1+r}\right)^n$$

Where 'r' is the rate of discount determined by the level of expectations and risk preferences of the firm.

The total present (discounted) value of all future revenues is

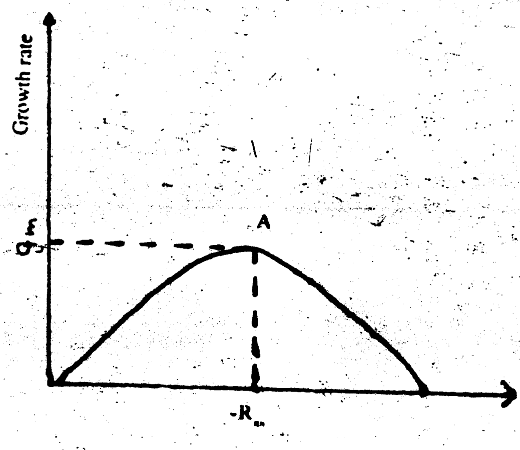
$$S = \Sigma_{i=0}^{n} R\left(\frac{1+g}{1+r}\right)^i$$

The firm attempts to maximise S by choosing appropriate current values of R and g. It is obvious that

$\frac{\partial S}{\partial R} > 0$ and $\frac{\partial S}{\partial g} > 0$

Also note that

g=g (π, R) $\Rightarrow$ growth function and $\pi = \pi$ (R, C, g, r) $\Rightarrow$ profit function. The growth function is actually derived from the profit function. Growth is mainly financed by retained profits which depend on current level of revenue (R), Costs (C), growth rate of sales ($\dot{g}$) and the discount rate (r). To maximise S, the firm can choose a particular combination of R and g out of a set of alternatives. These combinations are plotted along the growth curve, shown in the diagram. You may note that upto point A, which corresponds to maximum profit level, both R and g increase simultaneously. Beyond A, R increases but g falls. Thus beyond R πm, sales revenue level and growth rate become competing goals.



The most desirable combination of R and g may not be a feasible one and vice-versa. Actual choice depends on desirability as well as feasibility. The desirability may be defined in terms of the iso-present value curve. This curve shows all combinations of g and R which yield the same S.

You may recall the definition of S.

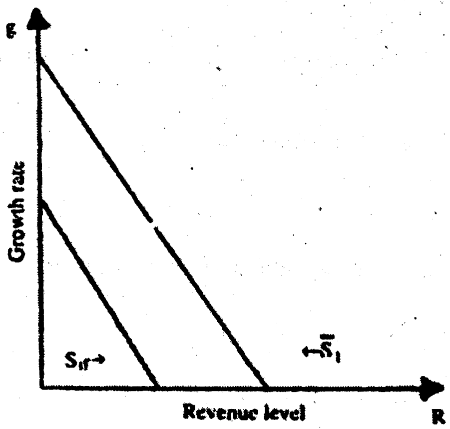
S, the aggregate discounted present value of revenue depends on R and g, given the discount rate r determined exogenously. Thus we can postulate that

$\overline{S} = a.g + b.R$ .... such that

$g = \frac{1}{a}\overline{S} - \frac{b}{a}R$

and $R = \frac{1}{b}\overline{S} - \frac{a}{b}g$

This is the simplest form you can have an equation of the iso-present value curve in the slope intercept form. Thus, you may have a family of such curves, the highest one representing the largest present value of S and the lowest one representing the least possible present value. The slope of this straight-line is given by $\frac{a}{b}$; along a given curve, the level of S remain same.

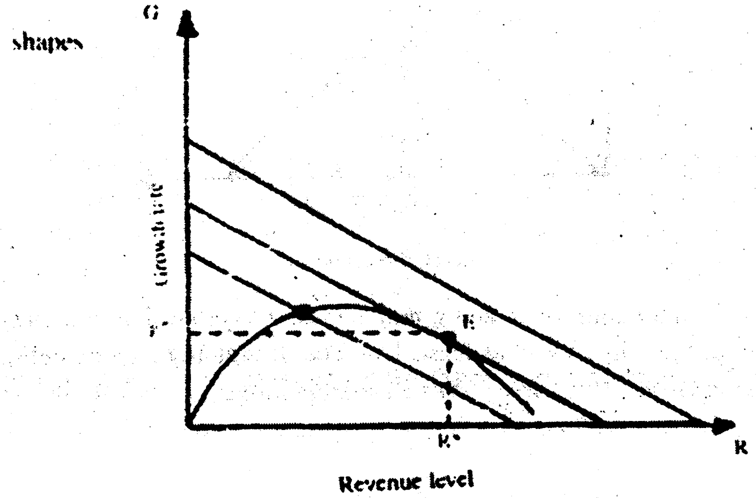


For deciding on the optimum combination of R and g, we have to combine both the preceding diagrams in one and design it as a case of growth-constrained iso-present value of revenue maximisation. In this case, the equilibrium solution is reached at point E which suggests the firm will choose a combination of R* and g* to get the highest possible level of S, given the growth function as a constraint.

The dynamic model described above can be modified to allow for exogeneously determined profit, advertisement and other non-price competition activities. The predictions of the dynamic model are the same as those of static models.

Empirical Evidence

Various studies have been conducted to test Baumol's hypothesis. However, the empirical evidence is not conclusive in favour of or against the sales-maximisation hypothesis. For example we may cite the findings of two studies.



McGvire, Chiv and Elling (see *Industrial Economic Review*, 1962) have attempted to test Baumol's contention that executive salaries to be far more closely correlated with the scale of operations of the firm than with profitability. Their results suggests that correlation between executive incomes and sales revenue is stronger than the correlation between executive income and profits. However, it is well known that such correlation may not necessarily imply any causation.

M. Hall (See *Journal of Industrial Economics*, 1967) in a comprehensive study has attempted to test the hypothesis, implicit in Baumol's model that if profits above the minimum constraint are earned, **ceteris paribus**, firms pursue policies (for example, cut prices, increase advertising and investment) in order to increase their sales revenue. The regression results of Hall do not justify this hypothesis.

Thus we have evidence both for and against Baumol's model. This calls for more empirical research towards testing of Baumol's hypothesis.

Activity 1

a) Compare and contrast Baumol's static models with dynamic model, with reference to the set of underlying assumptions and predictions.

....

....

....

....

....

....

b) Given the following conditions

$P = 12 - 0.4\ Q$ (Demand function)

$C = 0.6\ Q^2 + 4\ Q - 5$ (Total cost function)

$\pi \geqslant 10$ (Profit constraint)

Find out the price-output decision under (i) unconstrained profit maximisation, (ii) unconstrained revenue maximisation, and (iii) Profit- constrained revenue maximisation.

....

....

....

....

....

....

....

....

Ans:	Variables	Case (i)	Case (ii)	Case (iii)
	Price	10.4	6	10
	Output	4	15	5

c) Consider the effect of each on Baumol's profit constrained sales'maximising firm's behaviour Draw diagrams, if necessary, to illustrate

i) Imposition of a lump-sum tax :

....

....

ii) Imposition of a specific tax per unit of output:

....

....

iii) Decrease in Overhead expenses:

....

....

iv) Increase in variable costs:

....

....

v) Shift in demand:

....

....

d) Suggest economic terms for the following notations:

i) $\frac{\partial R}{\partial A} =$

ii) $\frac{\partial (AE)}{\partial A} =$

iii) $\lambda =$

iv) $\frac{\partial g}{\partial R} =$

v) $\frac{\partial g}{\partial \pi} =$

vi) $\frac{\partial S}{\partial g} =$

e) In the context of dynamic model of Baumol's we have assumed that the firm intends to maximise S subject to $g = \overline{g}$. Suppose you reverse the assumptions. The firm wants to maximise g subject to $S = \overline{S}$. Draw a diagram to illustate this new case. Would you agree that the optimum decision rule (stated in terms of equi-marginalism) remains unchanged even if we reverse the assumption?

....

....

....

....

f) Bring out clearly the role of advertisement in Baumol's model.

....

....

....

g) Under imperfect market conditions, the price elasticity of demand equals-one for a sales maximiser, but for a profit-maximiser, it is less than 1. TRUE or FALSE ? Explain.

....

....

....

....

17.3 MARRIS'S MODEL OF MANAGERIAL ENTERPRISE

Marris's model is found in two of his celebrated works:

- An article: "A Model of the Managerial Enterprise" in *Journal of Economics,* 1963.
- A book : *Theory of Managerial Capitalism* (Macmillan) 1964.

Marris has propounded a dynamic model of the firm, by stating clearly the objectives and the constraints of the firm.

Objectives and Constraints

The objectives of the firm is maximisation of the **balanced rate of growth** (G). The G itself is dependent on two factors : the rate of growth of the demand for firm's product (G^d) and the rate of growth of capital supply (G^s). Thus

$$G = G^d = G^s$$

In pursuing this balanced growth objective, the firm faces two constraints : managerial constraint and financial constraint. The **managerial constraint** is set by skill and efficiency of available manager's team. The **Financial constraint** is set by the desire of managers to attain the maximisation of their own utility function and their owner's utility function.

In modern organisation, ownership is separated from management. It is true that owners and managers have conflicting interests. But it is true that sometimes, their interests may coincide. "Balanced growth of the firm" is one such common area of interest. Both managers and owners are interested in this, because balanced growth ensures fair return on owner's capital and it also ensures continued trust and faith in managers who have achieved it. If the firm is not having balanced growth manager's job may be at stake just as owners capital is at stake. Thus the **goals of managers** may coincide with the **goals of owners.** In other words, despite ownership being divorced from management, owners and managers still may work together for a common cause or concern, viz, balanced growth of the firm.

In this context, Marris specifies two different utility functions - one for the manager and the other for the owner. The **utility functions of the manager** (U_m) includes variables like salaries power, status, job security, (s) etc. The **utility function of the power** (U_o) includes variables like profits, capital, output, market-share, public esteem etc. But most of these explanatory variables are ultimately related to the size and the steady growth of the firm. Thus Marris states:

$U_m = m(G^d, s)$

$U_o = O(G^s)$

To start with, Marris treats s as an exogenously determined constraint by assuming that there is a saturation level of job security. Above that level, $(\partial U_m/\partial_s) = 0$, while below that level, $(\partial U_m/\partial_s) = \alpha$. With this assumption the managerial utility function can be restated as $U_m = m(G^d)^{\bar{s}}$ where

$s = \bar{s}$ is the job security constraint.

We may now relook at the constraints in this model. First is the **managerial constraint.** Marris adopts Penrose's thesis that there exists a definite limit on the rate of managerial expansion such that "managerial ceiling" sets a limit to the growth of a firm. Second is the **financial constraint** which can also set a limit to the growth of a firm; and this constraint originates in the job-security considerations. In view of job-security, the managers become risk-avoiders by choosing a prudent financial policy which consists of determining optimum levels of some critical financial ratios such as:

1) Leverage or Debt Ratio r_1 = Value of debts/Total assets
2) Liquidity Ratio r_2 = Liquid assets/Total assets
3) Retention Ratio r_3 = Retained porfits/Total profits

These three ratios may be combined into a single parameter, $\bar{r}$ to represent the financial security constraint.

To affect balanced growth of Marris's firm, there are three **Instrumental variables.**

$\bar{r}$, the financial security co-efficient

d, the rate of product diversification

p, the average profit margin

Considering the set of objectives, constraints and instruments, Marris's model in the complete form may be presented as follows:

Structure of the Model

1)	Demand growth equation	$G^d = D(P, d)$
2)	Profit equation	$\pi = \pi(p, d)$
3)	Supply of capital equation	$G^r = \bar{r}[\pi(p, d)]$
4)	Security constraint	$\bar{r} \leq r$
5)	Balanced growth equilibrium condition	$G^d = G^s$

It may be noted that the level of profit, is endogenously determined, whereas the security constriant, $\bar{r}$, is exogeneously determined by risk-attitutdc of the managers. Given this, the balanced growth of the firm is effected through the operation of two instrumental variables, p and d. In balanced growth formulation we have in fact, one equation in two unknown:

$$D(p, d) = \bar{r}[\pi(p, d)]$$

Activity 2

a) State and explain the limits to the growth of a firm.

....

....

UNIT 18 PROFIT CONCEPTS AND ANALYSIS

Structure

18.0 INTRODUCTION

Profit-maximisation lies at the heart of traditional economic theory. Even behavioural and managerial theories have not dispensed with profit concept; rather they have come up with various other concepts of profit such as actual profit, reported profit, critical minimum (target) profit etc. Also, profit is sometimes seen as endogenous, sometimes as exogenous. Considering all these, profit is a very mixed and vexed concept. We need to relook at profit concept and its measurement, we need to recount the role that profit plays in the context of business decisions; we need to reexamine the sources and uses of profit. All these are necessary towards a better understanding of profit as a decision variable. Firms may not maximise profit, but all firms do have a profit policy. As such, we cannot neglect profit analysis.

18.1 OBJECTIVES

On reading this unit, you should be able to :

- **understand** different concepts of profit;
- **appreciate** the difficulties in measuring profit;
- **explain** various theories of profit;
- **locate** the determinants of a profit policy; and
- **distinguish** between the short-term and long-term issues in the context of profit planning, control and management.

18.2 CONCEPTS OF PROFIT

The concept of profit is not one, but many different people talk about different concepts of profit, sometimes their concept may not tally.

You may recall the distinction we have suggested in the introductory unit - a distinction between *accounting profit* and *economic profit*. An accountant looks at profit as a surplus of revenues over costs, as recorded in the books of accounts. An accountant is interested in accounting, auditing, planning and budgeting profit. The accountant does not take care of implicit or opportunity and non-operational, but these can be measured in terms of other variables. For example, manager's prestige and position is reflected in terms of the amount of emoluments or **slack** they receive in the form of expense accounts, luxury offices, company car and etc. Similarly staff expenses may also act as an indication of manager's power and position. Being the head of a large staff is a symbol of power and status. In the same way, the level of discretionary investment gives positive satisfaction to managers. Considering all these, the **utility function of the manager** may be written in the form of

$U = u(S, M, I_D)$ where

S = Staff expenditure

M = Managerial emoluments (slack)

I_D = Discretionary investment

Before we construct a model of the managerial utility-maximising firm, let us introduce some basic relations and definitions

Relations & Definitions

i) **The demand of the firm, P :**

The firm has a known downward sloping demand curve such that

$P = F(Q, S, E)$ Where P is the price per unit
Q is output
S is staff expenditure
E is a demand shift parameter

Thus, $\frac{\partial P}{\partial Q} < 0; \frac{\partial P}{\partial S} > 0 \; and \; \frac{\partial P}{\partial E} > 0$

ii) **The production costs, C :**

Costs depend on the level of output such that

$C = C(Q)$
$\frac{\partial C}{\partial Q} > 0$

iii) **Actual profit, πa :**

Actual profit is sales revenue less total costs including staff expenditure

$\pi a = R - C - S$

iv) **Reported profit, πr**

This actual profit less managerial emoluments is reported to the tax authority

$\pi r = \pi a - M = R - C - S - M$

v) **Minimum profit, $\pi 0$**

This is the critical minimum profit (after tax) which is required to pay reasonable dividends to the shareholders, without which the "job security" of the manager may be at stake. Thus

$\pi r \geq \pi 0 + T$ and
$T = \bar{T} + t\pi r$ Where $\bar{T}$ = Lump-sum tax
t = marginal tax rate

vi) **Discretionary investment, I_D**

This is the amount left from the reported profit after keeping the critical minimum profit and the tax obligations. Thus

$I_D = \pi_r - \pi_o - T$

vii) **Discretionary profit, π_D**

This is the amount of profit left after subtracting the minimum profits and the tax.

$\pi_D = \pi_a - \pi_o - T.$

The Model

We may now venture to construct Williamson's model in its complete structure. The optimisation problem in the model may be stated as:

Maximise $U = U(S, M, I_D)$

$= U(S, M\ \pi r - \pi_o - T)$

Subject to $\pi r \geq \pi_o + T$

It is assumed that the marginal utility of each of the component of the utiltiy function is diminishing but positive. Thus means that $S > 0$, $M > 0$ and $I_D > 0$. With this assumption, the constraint becomes redundant and we may treat this optimisation case as a simple case of unconstrained utility maximisation.

Substituting

$\pi r = \pi a - M = R - C - S - M$

and $T = \bar{T} + t(R - C - S - M)$

we obtain,

$U = U[S, M, \{(1 + t)(R - C - S - M) - 0\}]$

we may also substitute M as follows. Define P as the ratio of retained to actual profit:

$P\frac{\pi r}{\pi a} \Rightarrow \pi r = \pi a.P$

Thus, $\pi r = \pi a - M = \pi a.P$

Therefore, $M = (1-P)\ \pi a = (1-P)(R - C - S)$

Where (1-P) is the proportion of profits obsorbed by emoluments. Thus the managerial utility function in the final form, appears as :

$U - U[S, \{(1 - P)(R - C - S), P(1 + t)(R - C - S) - \pi 0\}]$.

Assuming that t and $\pi 0$ are being exogenously determined. We are left with three policy variables: Q, S and P. The manager has to choose such values of these variables which will maximise utility.

In the maximisation procedure, let us denote the first partial derivative of U with respect to S, M and I_D as U_1, U_2 and U_3, that is

$\frac{\partial U}{\partial S} = U_1, \frac{\partial U}{\partial M} = U_2$ and $\frac{\partial U}{\partial I_D} = U_3$

The total differential of the managerial utility function will be

$dU = U_1(dS) + U_2(DM) + U_3(d, I_D)$

Equality the partial derivatives of the managerial utility function to zero (by way of first order conditions) and taking into account the above total differential, we obtain:

i) $\frac{\partial U}{\partial Q} = U_2\left[(1 - P)\left(\frac{\partial R}{\partial Q} - \frac{\partial C}{\partial Q}\right)\right] + U_3\left[P(1 - t)\left(\frac{\partial R}{\partial Q} - \frac{\partial C}{\partial Q}\right)\right] = 0$

ii) $\frac{\partial U}{\partial S} = U_1 + U_2\left[(1 - P)\left(\frac{\partial R}{\partial S} - 1\right)\right] + U_3\left[P(1 - t)\left(\frac{\partial R}{\partial S} - 1\right)\right] = 0$

iii) $\frac{\partial U}{\partial P} = U_2[(-1)(R - C - S)] + U_3[(1 - t)(R - C - S)] = 0$

From (i), we obtain:

iv) $\left(\frac{\partial R}{\partial Q} - \frac{\partial C}{\partial Q}\right)[U_2(1 - P) + U_3 P(1 - t)] = 0$

The only way this equation (iv) to be satisfied is

$$\left(\frac{\partial R}{\partial Q} - \frac{\partial C}{\partial Q}\right) = 0 \Rightarrow \frac{\partial R}{\partial Q} = \frac{\partial C}{\partial Q} or MR = MC$$

⇒ Thus output decision is based on traditional equi-marginalism.

From (ii) we obtain:

v) $\frac{\partial R}{\partial S} = \left[1 - \frac{U_1}{U_2(1-P)+U_3(P)(1-t)}\right]$

Given the assumption that all elements appearing in the fraction are positive, it follows that $\frac{\partial R}{\partial S} < 1 \Rightarrow$ the optimum level (i.e., beyond the point where MR = MC). Thus there is a tendency for managerial firm to overspend on staff or to employ more administrative staff than a profit–maximising firm.

Lastly, from equation (iii), we obtain

vi) $(R - C - S)[-U_2 + U_3(1 + t)] = 0$

Given (R - C - S) > 0, for (vi) to be zero, we must have

$[-U_2 + U_3(1 + t)] = 0$

or $U_2 = (1+t)\, U_3$ = > This implies that some amount of profit will be obsorbed as emoluments, the amount depending on the tax rate: higher t, the smaller U_2/U_3, the smaller the marginal rate substitution of emoluments for discretionary investment, and the more will be spend on M and the less on I_D.

Implication of the model

We may now attempt a comparison of the utility-maximising firm with the profit maximising firm as follows :

Points of comparison	utility maximising firm	Profit maximising firm
Equilibrium condition	$\frac{\partial R}{\partial Q} = \frac{\partial C}{\partial Q}$	$\frac{\partial R}{\partial Q} = \frac{\partial L}{\partial Q}$
	$\frac{\partial R}{\partial S} < 1$	$\frac{\partial R}{\partial S} = 1$
	p < 1	p = 1
Decision variables of equilibrium	M>0	M=0
	S>0	S=0
	$I_D > 0$	$I_D = 0$
Shifts in market demand	$\frac{\partial Q}{\partial \epsilon} > 0; \frac{\partial S}{\partial \epsilon} > 0; \frac{\partial P}{\partial \epsilon} > 0;$	$\frac{\partial Q}{\partial \epsilon} > 0; \frac{\partial S}{\partial \epsilon} > 0; \frac{\partial P}{\partial \epsilon} = 0;$
An increase in the tax rate	$\frac{\partial Q}{\partial t} > 0; \frac{\partial S}{\partial t} > 0; \frac{\partial P}{\partial t} > 0;$	$\frac{\partial Q}{\partial t} = \frac{\partial S}{\partial t} = \frac{\partial P}{\partial t} = 0;$
Effect of a lump-sum tax	$\frac{\partial Q}{\partial T} < 0; \frac{\partial S}{\partial T} < 0; \frac{\partial P}{\partial T} > 0;$	$\frac{\partial Q}{\partial T} = \frac{\partial S}{\partial T} = \frac{\partial P}{\partial T} = 0;$

Activity 3

a) Suggest economic term for the following, as indicated below:

$\frac{\partial U}{\partial S}$ = Marginal Utility of staff expenditure

$\frac{\partial U}{\partial M}$ = ...

$\frac{\partial U}{\partial I_D}$ = ...

$\frac{\partial R}{\partial S}$ = ...

$\frac{\partial Q}{\partial E}$ = ...

$\frac{\partial R}{\partial t}$ = ...

$\frac{\partial P}{\partial T}$ = ...

b) Why is P = 1 for a profit-maximising firm, but P < 1 for an utility-maximising firm?

....

....

....

....

d) Compare the behaviour of a sales-maximising firm with that of an utility-maximising firm.

....

....

....

....

....

....

....

....

d) Recall the various concepts of 'profit' in Williamson's model and define each one of them.

....

....

....

....

....

....

....

....

....

17.5 SOME COMMENTS ON MANAGERIAL MODELS

Williamson's model is a realistic one. He has tested the hypothesis that managerial discretion influences the expenditures for which managers have a strong "expense preference" (staff expenditure, emolumnets, discretionary investment). Some of his expirical evidence do support the managerial discretion model:

1) Increase in S and M in boom, and drastic cut of these expenditures in recessions

2) Reaction of firm to changes in taxation

3) Changes in Q, S and M in response to changes in fixed costs

4) Drastic cut in S by newly appointed top management, without affecting the productivity of the firm.

However, it is felt that such available evidence is still not sufficient for verification of a theory. More empirical research is needed to establish the propositions of these new theories.

The managerial theories have got some basic limitations. For example, these theories fail to explain oligopolistic interdependence in non-collusive markets. Such theories seem suitable for large firms where there is scope for 'product diversification' and 'discretionary investment' for a small firm, managerial discretion is limited. Also, the managerial models do not explain how price is determined in the market; their focus is on output as a decision variable and the impact of output decision on revenue, costs, profit etc. The managerial models do not capture all the constraints which the firm faces. Williamson's model, for example, assumes away the role of constraints in a very convenient way. Marris's model talks of only two constraints - managerial and financial. In real life, the firm has to face many more constraints - social, cultural and physical (environmental). Finally, the managerial (and the behavioural) models do not really replace the traditional economic theory, rather they re-inforce the profit-maximising behaviour of a firm. For example, Williamson's model can be treated as a general case which can incorporate traditional economic theory's profit-maximisation hypothesis. Williamson's model yields identical results, compared to traditional theory, if the ratio of reported to actual profit, P = 1 and the marginal utility of staff expenditure, $U_1 = 0$.

The point remains that the traditional economic theory and its profit maximisation hypothesis need not be dispensed with. The behavioural models and the managerial models are not free from limitations. For example, these models do not focus on "industry equilbrium" situation. Some of the concepts of these models do have their counterpart in traditional economic theory. For example, the concept of "slack" exists in traditional theory, in the name of "economic rent" - actual price exceeding the transfer price (opportunity cost) of a factor. Similarly "Managerial ceiting" is nothing but traditional theory may describe as a forms of 'managerial diseconomies' - some sort of 'internal diseconomies of scale'. Thus, the traditional economic theory is not totally useless.

17.6 SUMMARY

In this unit, we have moved from behaviourism to managerialism as an alternative to the traditional economic theory. In the process, we have focused on the behaviour of an utility maximising firm in contrast to that of a profit-maximising firm. In particular, we have specified the utility function of managers in contrast to that of owners. Thus, the focus has been more on the behaviour of a group of **individuals** than on the behaviour of a group of **firms**, i.e., industry situation. Such an approach has uses as well as limitations. More we examine these uses and abuses, more we may discover the strength of traditional economic theory. In fact, despite its inadequacies, the profit-maximisation hypothesis still remain a fundamental construction which has a lot of applicability in real world business situation. You must have observed by this that the behavioural and managerial theories do only supplement rather than substitute the traditional theory. If so, we need to revisit the concept of profit and its measurement; this is the subject matter of our next unit.

17.7 ADDITIONAL READINGS

Wildsmith, J.R., 1972, *Managerial Theories of the Firm* (Martin - Roberston).

Gilbert M., (ed.), 1973. *The Modem Business Enterprise* (Penguin, article by Williamson).

Marris. R., 1964. *Theory of Managerial Capitalism* (Macmillan).

Baumol. W.J., 1967. *Business Behaviour, Value and Growth* (Revised ed., Harcourt, Brace & World, Inc.)

Koutsoyiannis, A. 1979. *Modern Microeconomics* (2nd ed.) (Macmillan).

17.8 SELF-ASSESSMENT TEST

1) Attempt a comprehensive comparison between:

   a) Managerialism and Marginalism

   b) Managerialism and Behaviourism

2) Examine the policy-implications of the principles underlying the models of :

   a) Baumol

   b) Marris

   c) Williamson

3) Explain, with reference to the context, each of the following propositions:

   a) Profit may be exogenous or endogenous depending upon the time-horizon of a firm' activity.

   b) Marginal utility of discretionary investment must be proportional to marginal utility of managerial, emoluments (slack), the proportionality factor being dependent on exogenously determined tax rate.

   c) The growth of a firm's activity is desirable, but balanced growth may not be always feasible in view of limits to firm's expansion.

4) Critically examine the models of 'managerial enterprise' and 'managerial discretion'. In particular, comment on the existing empirical evidence towards verification of these models.

5) Comment on each separately :

   a) If the resources (and costs) are not given, the multiproduct firm will reach a different product-mix, depending on whether it is a profit-maximiser or a sales-maximiser.

   b) "Sales-maximisation is not incompatable with the goal of profit-maximisation" (M.H. Peston)

   c) Product diversification and Advertisement: Both have revenue and cost implications in a single-period and over multi-period.

....

....

b) **How does the demand for a firm's product get affected by diversification overtime?**

....

....

....

....

c) **Examine the role of Profit as a variable in Marris's model. Is it endogeneous or exogenous? Is it a target or an instrument?**

....

....

....

....

d) **In the context of the theory of firm, compare and contrast the dynamic models of Baumol and Marris.**

....

....

....

....

e) Attempt to collect some empirical evidence towards testing Marris's hypothesis. For example, in order to tone up **managerial efficiency** the Indian Public Enterprise Selection Board is locating Chief Executives from the private sector. Or, say, in order to overcome **financial problem** many of our public enterprises are being allowed to float bonds. Can you quote many more of such instances?

....

....

....

....

17.4 WILLIAMSON'S MODEL OF MANAGERIAL DISCRETION

In his article "Managerial Discretion and Business Behaviour" in *American Economic Review* (1963), Williamson argues that managers have discretion in pursuing policies which maximise their own utility rather than attempting the maximisation of profits which maximise the utility of owner-shareholders.

Managerial utility function includes such variables as salary, security, power, status, prestige and professional excellence. Of these variables, only salary is measurable, the rest are non-pecuniary costs. From the accounting profit, he takes out these implicit costs to compute his economic profit. Thus

$$\pi = R - C$$
$$\pi^* = \pi - OC$$
$$= R - C - OC$$

Where π = Accounting profit

R = Revenues

C = Explicit Costs

OC = Opportunity Costs

Since every decision involves a sacrifice of alternatives, the opportunity costs are implied in any decision-making. Therefore, the economists would like to consider them in computing his profit. On the other hand, for an accountant, the opportunity costs are subjective; those are not accurately measured always and therefore, those costs are ignored. But if opportunity costs are considered then economic profit (π^*) would tend to be less than accounting profit (π) of couse, π may be equal to π^* only under the assumption of zero opportunity costs; but in real world decision situation, this is a highly unrealistic assumption.

Think of a simple realisitc situation. A trained mechanic has started a repair shop by investing his own capital of say Rs. 5000. His shop is located in the garage of his own house such that he pays no rent. Annually, he makes a net income of Rs. 10,000; this is the difference between his revenue earned (Rs. 25,000) and his materials and all other explicit costs (Rs. 15,000). Had he not started and run his business, the mechanic himself could have earned a monthly wage of Rs. 500 i.e., an annual salary of Rs. 6,000. Similarly, had he invested his capital of Rs. 5,000 in the bank, in course of a year, he would have got a 10% return on it, i.e., Rs. 500. In the same way, his garage would have fetched a monthly rent of Rs. 200 i.e., an annual rent of Rs. 2,400. The information can now be presented in a tabular form:

Total Revenue earned	=	Rs. 25,000
Total Costs of Operation	=	Rs. 15,000
Accounting Profit (π)	=	Rs. 10,000
Opportunity costs of		
• Own labour (wages)	=	Rs. 6,000
• Own capital (interest)	=	Rs. 500
• Own building (rent)	=	Rs. 2,400
Total opportunity costs	=	Rs. [illegible]
Therefore, economic profit (π^*)	=	Rs. 1,100

You may note that economic profit is less than accounting profit by the amount of opportunity costs. Sometimes, there may be 'accounting profit' but 'economic loss' in the same operation; much would depend on the order and computation of opportunity costs. In the example, all our opportunity costs were neatly calculable, but in reality it may not be so, because some opportunity costs may be in real rather than monetary terms; they may be subjective rather than objective; they may not be always quantifiable and measurable. This is what makes computation of 'economic profit' a difficult task. This is not to deny that calculation of economic profitability even if it is just approximate rather than accurate is a useful guide to decision-making. The business firms hardly prepare data on economic profit. The balance-sheet data that you have often access to records only accounting profit. Such accounting profit reflects the performance of business unit. In fact, the operational efficiency of the business unit can be judged in terms of some charts. There are some charts which may record accounting profit data in such a way that it can help the firm in controlling and managing its activities in specific areas. Let us reproduce some of these charts below, because those can provide you an insight into various concepts of accounting profit. You may be already familiar with some of these charts.

Profit Performance Chart

Sales Revenue
– Direct Costs (on labour, materials, variable expenses)

Contribution
– Fixed or Period Costs

Gross Operating Profit
– Depreciation

Net Operating Profit (EBIT)
– (Interests + Taxes)

Net Profit
– Dividends

Net Retained Earnings

Management Achievement Chart

TMP	=	Total Management Performance
OMP	=	Operating Management Performance
FMP	=	Financial Management Performance
S	=	Sales Revenue
I	=	Investment or Capital employed
C	=	Contribution
OP	=	Operating Profits
NP	=	Net Profit
NW	=	Net Worth (i.e., Assets - Liabilities)

[TMP] = [OMP] × [FMP]

$$\frac{NP}{NW} = \left[\frac{S}{I} \times \frac{C}{S} \times \frac{OP}{C}\right] \times \left[\frac{NP}{OP} \times \frac{C}{NW}\right]$$

$\frac{NP}{NW}$	$\frac{S}{I}$	$\frac{C}{S}$	$\frac{OP}{C}$	$\frac{NP}{OP}$	$\frac{C}{NW}$
↑	↑	↑	↑	↑	↑
Net profit to net worth Ratio	Turnover Ratio	Profit Volume Ratio	Margin of Safety	Financial Operation Ratio	Financial Lerverage Ratio

DU-pont Control Chart

Rate of return on investment (ROI)

⇓

⇓	⇓
Earnings as percent of sales	Sales turnover
⇓	⇓
Earnings/Sales	Sales/Total Investment
⇓	⇓
Sales Revenue less Costs	Fixed Capital+Working Capital
⇓	⇓
Cost of sales + Selling expenses + Administrative expenses	Cash + Inventories + Receivables

When you examine these charts carefully, you may encounter a host of accounting concepts of profit. Which particular concept will be used by an accountant in a given situation depends partly on "accounting convention" and partly on "accounting convenience".

In the same way, when an economist uses the concepts of profit, he has a host of concepts like actual profit, reported profit, a critical minimum profit. Sometimes profit is treated as a constraint i.e., target profit (mark-up); sometimes profit is the goal variable. Sometimes profit target is exogenously determined outside the system; sometimes it is endogenously determined. For example Earning Before Interest and Taxes(EBIT) and dividends together determine the net retained earning of the firm. The various economic concepts mentioned here have already been referred earlier. In case you do not remember them, you may like to go back to some of the previous units.

Activity 1

a) Construct an example with imaginary data to show that despite making an accounting profit, a firm incurs an economic loss.

B) Distinguish briefly between :

i) Economic profit and accounting profit:

....

....

....

ii) Gross profit and net profit :

....

....

....

....

iii) **Contribution and Gross operating margin:**

....

....

....

iv) **EBIT and net retained earnings :**

....

....

....

v) ROI and net ratio $\frac{NP}{NW}$:

....

....

....

c) An accountant judges total *management performance* through the concept of "averages", whereas an economist uses the concept of 'marginals'. TRUE or FALSE? Explain.

....

....

....

....

....

....

....

....

....

....

....

(Hint : For an economist, profit maximisation by a firm must satisfy the condition, Marginal profit=zero. For an accountant, management performance chart has all parameters in the form of ratio or averages.)

d) **Despite an increase in sales revenue, a firm's rate of return on investment is showing a downward trend. How can this be? How would you control this situation?**

....

....

....

....

....

....

....

18.3 FUNCTIONS OF PROFIT

Based on the discussion above, particularly the concepts and the charts, you will find that profit has a very important role in decision situations. We may categorically state the functions of profit:

1) **Profit is a measure of operational efficiency of the firm.** An increasing level profit, other things remaining the same, is an indicator of effective business strategy and efficient business tactics. It reflects management performance, operational and financial.
2) **Profit is the premium that covers up the cost of staying in business** - replacement, obsolescence, risk and uncertainty.
3) **Profit is the source of internal finace.** The net retained earnings is used to finance expansion and replacement programme. Profit reflects the security and solvency of the firm.
4) **Profit, its level and rate, is the basis, of financial decision analysis.** Commerical profitability is an established criteria of project evaluation or appraisal.
5) **Profit** reflects the income-earning capacity of the firm and thus **determined if the firm is capable of discharging social responsibilities** (like adoption of a village) or not. This is the reason why public enterporises are also required to generate profit as a reinvestible source.
6) **Inter-firm comparison, inter-industry comparison or product-line comparisons are all facilitated by the index of profit.**
7) **Profit and its sources also indicate the areas of planning, control and management.**

You may go on enlarging the list. The point remains that no business firm can function without having a profit policy, because profit plays a crucial role to help decision-making.

Activity 2

a) We, often hear of terms like 'private profit' and 'social profit'. Give examples. Why is this distinction useful?

....

....

....

....

b) **Profit** is not bad, what is ethically unjustified is **profiteering**. Comment

....

....

....

....

c) Name some profit-based financial ratios which you will choose to indicate the financial health of a company.

....

....

....

....

....

....

18.4 MEASUREMENT OF PROFIT

The difference between accounting profit and economic profit makes the measurement of profit a difficult task. The difficulty arises due to the difference in cost concepts. The economists 'opportunity costs' are not easily identifiable and measurable; on the other hand, the accounting costs, both direct and indirect, are easily identified and recorded.

A fundamental accounting principle is
Assets = (Liabilities + Proprietorship)
(Assets - Liabilities) = Net Assets - Proprietorship

The *balance sheet* of the firm indicates the value of the firm's assets corresponding to claims of creditors and owners at some given time. The *income statement* or profit and loss statement records the changes in these items resulting from business transactions over course of a year. The *funds flow statement* records the sources and uses of funds. In preparing these statements, the accountant normally shows assets valued at original (historical) costs. By contrast, the economists value the assets at replacement costs. Ideally, the economists try to assess the present value of future cash flows which existing assets will bring. Traditional financial accounting finds it unsatisfactory, because it is some what speculative.

There are three specific aspects of profit-measurement where the use of accounting profit and of economic profit give different results

1 **Depreciation**

An accountant measures the cost of depreciation by several methods.

i) **Straight-line Method** whereby a fixed percentage of original value of the asset is deducted annually over the working life of the asset. Thus

$$D = \left[\frac{F-S}{n}\right]$$

Where F = Cost of the fixed asset
S = Salvage value
n = Life of the asset
D = Annual depreciation charge

ii) **Diminishing-balance Method** whereby a constant percentage of an annually diminishing written-down value is deducted such that

$$D = \left[1-\left(\frac{S}{F}\right)^{\frac{1}{n}}\right]$$

iii) **Annuity Method** requires the cost to be covered equal the fixed costs of the asset, F, plus an interest rate, r, equal to the cost of capital, covering the annual fixed instalment over the estimated life, such that

$$D = \left[\frac{F+r}{n}\right]$$

iv) **Service-unit Method** is appropriate where the life of the asset depends on its use rather than time. Here the difference between fixed asset costs and its salvage value is divided by the life time capacity, Q, such that

$$D\left[\frac{F-S}{Q}\right]$$

For an economist, the above accounting methods are of no use. The economist looks at depreciation in terms of opportunity costs and uses the asset replacement costs (R) rather than the original or historical costs of the asset (F). R is the difference between the new investment (I) and the salvage values (S) of the old machine. Thus

R = I - S and

R > F during periods of inflation

R < F during period of falling prices.

The replacement investment is needed to keep the "capital stock intact". The opportunity costs of not taking timely replacement is increasing level and rate of depreciation and/or obsolescence.

2 **Inventory Valuation**

This is another area of profit measurement where accounting conventions and economic concepts give different results. Inventory or stock refer to goods in pipe-line difference between production and consumption. When production exceeds consumption, the stocks file up. Such inventory-building or stock-filing would have posted no problem of valuation had prices remain constant. In reality, prices do not remain stable, materials costs change and, therefore, the valuation of stocks must change. The accountant uses his standard methods some of these methods are :

i) **First-In-First-Out (FIFO)** : Materials are withdrawn from stocks in the same order they are acquired such that current manufcturing costs are based on the oldest materials in stock.

ii) **Last-In-First-Out (LIFO)** : The most recently purchased materials are withdrawn first from the stock such that the current manufacturing costs are based on the costs of the newest material in stock.

iii) **Weighted Average Costs (WAC)** : A weighted average of different costs, corresponding to different lots of materials purchased at different periods is used for valuation of stock.

The economist feels that the recorded value of business income in different periods may differ considerably depending upon the methods of valuation chosen. For example, FIFO will show unrealistically high profits, during inflation and low profits during deflation.

The economist argues that neither of the accounting methods reflect the change in the real value of profit (business income), because the accountant consider only nominal prices, past or present. For a true measure of valuation, the net business income should be measured at constant prices. The use of a constant deflation enable us to make intertemporal comparison of net business income and there a true economic profit is calculable without neglecting the opportunity cost of time when price level may change.

iii) **Annuity Method** requires the cost to be covered equal the fixed costs of the asset, F, plus an interest rate, r, equal to the cost of capital, covering the annual fixed instalment over the estimated life, such that

$$D = \left[\frac{F+r}{n}\right]$$

iv) **Service-unit Method** is appropriate where the life of the asset depends on its use rather than time. Here the difference between fixed asset costs and its salvage value is divided by the life time capacity, Q, such that

$$D\left[\frac{F-S}{Q}\right]$$

For an economist, the above accounting methods are of no use. The economist looks at depreciation in terms of opportunity costs and uses the asset replacement costs (R) rather than the original or historical costs of the asset (F). R is the difference between the new investment (I) and the salvage values (S) of the old machine. Thus

R = I - S and

R > F during periods of inflation

R < F during period of falling prices.

The replacement investment is needed to keep the "capital stock intact". The opportunity costs of not taking timely replacement is increasing level and rate of depreciation and/or obsolescence.

2 Inventory Valuation

This is another area of profit measurement where accounting conventions and economic concepts give different results. Inventory or stock refer to goods in pipe-line difference between production and consumption. When production exceeds consumption, the stocks file up. Such inventory-building or stock-filing would have posted no problem of valuation had prices remain constant. In reality, prices do not remain stable, materials costs change and, therefore, the valuation of stocks must change. The accountant uses his standard methods some of these methods are :

i) **First-In-First-Out (FIFO)** : Materials are withdrawn from stocks in the same order they are acquired such that current manufcturing costs are based on the oldest materials in stock.

ii) **Last-In-First-Out (LIFO)** : The most recently purchased materials are withdrawn first from the stock such that the current manufacturing costs are based on the costs of the newest material in stock.

iii) **Weighted Average Costs (WAC)** : A weighted average of different costs, corresponding to different lots of materials purchased at different periods is used for valuation of stock.

The economist feels that the recorded value of business income in different periods may differ considerably depending upon the methods of valuation chosen. For example, FIFO will show unrealistically high profits, during inflation and low profits during deflation.

The economist argues that neither of the accounting methods reflect the change in the real value of profit (business income), because the accountant consider only nominal prices, past or present. For a true measure of valuation, the net business income should be measured at constant prices. The use of a constant deflation enable us to make intertemporal comparison of net business income and there a true economic profit is calculable without neglecting the opportunity cost of time when price level may change.

3 Unaccounted Value Changes

There may be certain items of business expenditure which may not have any impact on current business income, but which may increase future income of the firm. The accountant does not consider the future value of the present expenditure on items like Research and Development, advertisement, recruitment of skilful managers etc. In the process, the accountant may understate current profit and overstate future profit. Thus, 'accounting profit' may not be a true measure of 'economic profit'. The economist would like to take care of intangible assets and liabilities and the opportunity gain and costs involved therein so that a time measure of economic profitability is estimated.

Activity 3

a) Depreciation is the factor which accounts for a difference between Gross and Net Profit. Explain.

....

....

b) How does the economist's view of 'depreciation' differ from that of an accountant?

....

....

....

....

....

c) Inventory valuation of an accountant does not reflect the true economic profit. Why?

....

....

....

....

....

d) Quote some examples of "intangible assets" of a firm. How does the economist's view of such assets differ from the accountant's view?

....

....

....

....

e) Accounting data show that during a certain period, the profit of a firm has *increased*, while an economic analysis of the same data shows that profit has *decreased*. How can that be? Give a very precise answer.

....

....

....

....

18.5 ECONOMIC THEORIES OF PROFIT

We have considered the meaning and measurement of economic profit. Let us now consider the explanation of economics profit : Why does profit arise? Who produces it? Answers to such questions throw further light on the nature of economic profit — its sources and uses, its accrual growth.

In this context, we may briefly review various theoretical propositions suggested by the economists from time to time. You may find that most of these propositions are still relevant today. For your convenience, let us present, to begin with, a summary view:

Economic Theories of Profit

Traditional views	Modern views	Recent views
- Adam Smith, Ricardo - Karl Marx - Walker - Marshall	- J.B. Clark, Schumpeter - Hawley - F. Knight - Keirsteed	- Kalecki - Kaldoer - Passinetti

Traditional Views

The classical economists like *Adam Smith* and *Ricardo* regarded profit and interest together as a kind of business income which was left after paying labour (wages) and the indestructible powers of the soil. In this view, there was no distinction between ownership of capital and management of enterprises.

Karl Marx aggregated rent, interest and profit in the category of "unearned income" (from property) in contrast to wages as the "earned income" (from labour). According to him, labour produces more than what is needed to maintain labour in the productive process, this "surplus value" is pocketed by the capitalists. Thus profit as 'surplus' value' originates in exploitation of the proletariat (labourer) by the bourgeois (capitalists) In this view, profit is not socially justifiable.

The neo-classical economists tried to develop functional theories to justify profit as legitimate business income. *Walker* developed the "rental ability" theory of profit on lines of Ricardian theory of rent. According to Ricardo, the difference in the level of productivity between two grades of land determines the rent on the A-grade land; rent is thus a "differential-surplus" over the costs of production of the marginal land. Walker argued; the Entrepreneurs is a superior grade of labour; the difference in the level of productivity between ordinary labourers and efficient entrepreneurs entitles the later to earn a profit; thus profit is a payment for rental ability. — the capability and capability of an entrepreneur.

Marshall developed a similar functional notion of profit. He categorised all productive factors into four—land, labour, capital, and organisation and their respective incomes as rent, wages, interest and profits. Marshall also spoke of normal profit part of costs prevailing in the long-run against "frictional profit," prevailing in the short-run, such profits are functionally related to the efficiency (productivity) of the entrepreneurs.

The neo-classical economists developed this functional notion into the ***marginal productivity theory of profit.*** The argument run, if all factors are paid according to their respective marginal production, then the value of total factor payments i.e., no Adding-up problem,

Thus, Rent = marginal productivity of land [L]
Wages = marginal productivity of labour [N]
Interest = marginal productivity of capital [K]
Profit = marginal productivity of entrepreneurship [E]

There may be difficulties in the measurement of marginal productivity of an entrepreneur, but one can follow an accounting rule :

$$\text{Profit} = \begin{bmatrix}\text{Value of the} \\ \text{total product}\end{bmatrix} - \begin{bmatrix}\text{Payment to factors other} \\ \text{than entrepreneur}\end{bmatrix}$$

$$\text{or, } \pi = [P.Q] - \left[L\frac{\delta Q}{\delta L} + N\frac{\delta Q}{\delta N} + K\frac{\delta Q}{\delta K}\right]$$

$\Rightarrow \text{Total } \pi = E.\frac{\delta Q}{\delta E} \Rightarrow$ This holds under perfect competition and constant return to scale i.e., given product force $P = \overline{P}$ and production factor, $Q = Q(L, N, K, E)$ homogeneous of degree one.

Eventually, the concepts of marginal product have been refined further into concepts like marginal value product and marginal revenue (value) product and it has been shown that even under imperfect competition, the theory of marginal productivity may hold good, provided factor payments are made according to marginal revenue product of factors. The greatest achievement of this neo-classical view was to challenge the view of profit as a parasitic income.

Modern Views

J.B. Clark develops the ***dynamic theory of profit*** suggesting that profit is a "dynamic surplus" – surplus over production costs in an economy which experiences some "generic changes" like

- change in the size of population
- change in the stock of capital
- change in the technique of production
- change in the form of business organisation
- change in consumers' wants and preferences.

An economy without such changes is a static economy where every factor gets a "natural price". In a dynamic economy, some factors like entrepreneurship gets more than the natural price, as and when they try to capitalise on these changes. Thus successful entrepreneurs in a dynamic economy grasp profit, though they may not be able to hold it for long because of competition.

J.Schumpeter considers a very specific type of change in the form of "innovations" —new ideas, new techniques, new resources, new markets, new products, etc. Entrepreneurs who innovate make profit in a dynamic economy. Such innovation results through a distinct process : scientific discovery - invention – application - spread - innovation. Unless there is a successful commercial adoption of invention, the innovation does not yield profit. The profit motive induces the flow of invention and innovations and thereby activises the process of economic development in a capitalist economy.

Hawley has advanced the risk theory of profit. An innovating entrepreneur makes profit, because he undertakes risk. Supervision and coordination are routine jobs of management. An entrepreneur undertakes a more challenging job of risk-bearing. Unless there is some reward in the form of profit, nobody would undertake risk. No risk, no gain. Profit is a premium for risk bearing. The person who invests or innovates, undertakes great risks—his investment may or may not yield him any return, his process, pattern or product may or may not be commercially viables, so on and so forth. One who accepts greater risks, expects larger profits. That is how profit averages in a dynamic economy; changes involve risks.

Frank Knight has developed a hybrid theory by combining the view points of Clark and Hawley. Knight makes a critical analysis of deicision environment and finds that neither changes nor risks are homogeneous and that profit cannot be associated with all types of changes and all types of risk.

Changes may be known or unknown. Known changes may have definite or indefinite outcome. If the outcome is definite, when the change in known, then there is not risk involved, it is a world of certainty. If the change is known, but the outcome is indefinite, then the acturial probability of the outcome may be calculated; this is risky environment. If the change is unknown such that the probability of any outcome cannot be calculated, it is a world of uncertainty. One who takes incalculable risk, typical of the world of uncertainty, is entitled for profits positive or negative. The calculable risk can always be insured; interest rate is such an insurace premium. Thus, you may note that all risks do not result in profit (only incalculable, indefinite, uncertain risks do), and that all changes do not result in profit (only unknown changes do), and that profits and interest cannot be clubbed together profit is associated with uncertainty, whereas interest is associated with calculable risk.

Keirsteed goes a step further. It is true that pr ... is associated with incalculable risks, typical of uncertainity prevailing in a dynamic economy characterised by unknown changes. Such changes and risks may relate to market or non-market factors, financial or non-financial factors and so on. Whatever may be the form of a risk, risks are essentially measured by way of "stand deviation between the actual outcome and expected outcome". Expectations are of two types: static and dynamic. Expectations about future outcome is based on observations regarding present outcome. When the actually observed phenomena today are expected tomorrow, it is static expectations. If expectation takes the course other than what is observed, it is dynamic expectations. Profit is associated with incalculable risks and dynamic expectations. An innovating entrepreneur is rewarded profit becuase he undertakes incalculable risks while making dynamic expectations.

Recent Theories

Recently some economists have taken a macro-view of profit. Kalecki starts with a few micro-assumptions but derive a profit proposition having macro-implications. On the other hand. Kaldon starts with a few macro- assumptions to derive a proposition about profit which has micro-implications. Kalecki has developed a monopoly theory of profit. His starting point is cost plus pricing such that

$$P = r + w + \pi$$

where P = price per unit of output

r = average cost of raw material

w = average cost of labour

π = average rate of profit

$$\Rightarrow \; = p - (r + w)$$

$= p - AVC$ ∴ labour & raw materials are variable factors

$= p - MC$ ∴ AVC is assumed constant i.e., linear cost function.

$$\Rightarrow \frac{\pi}{w} = \frac{p-MC}{w}$$

$$= \frac{p-MC}{w} \cdot \frac{MC}{MC}$$

$$\frac{\pi}{w} = \frac{p-MC}{MC} \cdot \frac{MC}{w}$$

This reads that profit to wage ratio depends on two factors (a) degree of monopoly as measured by $\frac{p-MC}{MC}$ and (b) cost structure as measured $\frac{MC}{w}$. Other things remaining equal, higher the degree of monopoly, larger is the share of profits to wages. Thus profits can be treated as a monopoly net revenue. If there is no market imperfection in the world of changes and risks, there can be no supernormal profit in the form of a surplus over costs.

Kaldor **develops a** ***macro theory of profit*** **incorporating Karl Marx and Keynes. He starts with a few identities:**

i) $Y = \pi + W$

ii) $I = S$

iii) $S = S_1 + S_2$

iv) $S_1 = s_c.\pi$

v) $S_2 = s_w.W$

Where Y = National income

π = Profit income

W = Wage income

I = Investment

S = Savings

S_1 = Total savings of the capitalist

S_2 = Total savings of the wage earners

s_c = Propensity to save of the capitalist $\frac{S_1}{w}$

s_w = Propensity to save of the wage earners, $\frac{S_2}{w}$

Making necessary substitutions, we may write :-

I = Sc.π+ Sw.W

I = Sc.π+ Sw (Y $-\pi$)

I = Sc.π+ SwY $-$ Sw.π

I = (Sc$-$ Sw) π+ SwY

$$\pi = \frac{I - S_wY}{Sc - S_w}$$

$$\frac{\pi}{y} = \frac{I}{y} \quad \frac{1}{(Sc - Sw)} - \frac{Sw}{Sc - Sw}$$

This reads that the share of profits to national income depends on investment rate $(\frac{I}{Y})$ and propensities to save of the capitalist and the wage earners, i.e., (sc and sw).

To sum up, there are both micro and macro-economic views about profit. Not a single view by itself may suffer to after a satisfactory explanation of profit. To explain inter-firm or inter-industry difference in profit, you may have to use a number of theories simultaneously.

Activity 4

a) **The history of economic thought abounds in profit theories, but most of them fall into three major categories suggested below. Write down the basic proposition against each category. Also function the list of authors.**

Group of Theories	Proposition
Compensatory and functional theories	
Friction and monopoly theories	
Technology and innovation theories	

b) Why can't all 'inventions' be treated as innovations?

....

....

....

....

....

c) Marx would have argued "workers do not save, because they cannot save, capitalists save their entire income, because they survive on surplus value produced by labour". If so, how does Kaldor's proposition about the determinants of share of profit to national income change?

....

....

....

....

....

....

[HInt : Assume sw = 0, but sc = 1]

18.6 PROFIT PLANNING AND CONTROL

Profit planning refers to operating decisions in the areas of product line, volume of productions, and pricing, whatever the firm's profit goals may be, profit planning must take into account the expected demand for firm's products, its capacity to meet the demand, and all of its costs. A good profit should identify objectives, set targets and specify the instruments for attaining these. Profit planning, control and management is a short-run as well as a long-run operation.

There are a number of approaches to profit planning and control as used by the accountant and the economist. Some of these may be briefly discussed here:

i) Profit Budget
ii) Break-even Analysis
iii) Regression Analysis

Profit Budget

A profit budget serves the function of planning, coordination and control of firm's activities. The budgeting process enables managers to anticipate their needs, and forces coordination of firm's activities that otherwise might not take place. The control function is not automatic, but by means of suitable reports, it enables management to maintain a systematic check on the company's operations in terms of planned *verse's* actual results.

Break-even Analysis

It is an adjunct to profit planning that graphically or algebraically portrays the probable effects of alternative courses of action. You may recall the graph in a preceding unit where we have illustrated the break-even point and the break-even output (Q^B). Also in the context of cost application, we have derived the measure of break-even output volume (Q^q) and value (Q^v). You may recall :

i) $Q^B = \frac{TFC}{P-AVC} = \frac{TFC}{ACM}$

ii) $Q^v = P\,Q^B$ You may also add now:

iii) $S^B = \frac{TFC}{1-(AVC/P)} = \frac{TFC}{1-(TVC/TR)}$

iv) $\%_B = \frac{TFC}{(P-AVC)Q_{max}}.100$

Where	TFC	= Total fixed costs
	TVC	= Total variable costs
	AVC	= Average variable costs
	P	= Price
	ACM	= Average contribution Margin
	S^B	= Break-even sales
	$\%_B$	= Break-even percentage of capacity

An Example

Suppose an airline carry a maximum of 10,000 passengers per month on one of its routes at a fare of Rs. 850. Variable costs are Rs. 100 per passenger and fixed costs are Rs. 30,00,000. In this case,

$$Q^B = \frac{Rs.30,00,000}{Rs.(850-100)} = 4000 \text{ passengers}$$

$$Q^V = Rs.850 \times 4000 = Rs.34,00,000$$

$$S^B = \frac{RS.30,00,000}{1-Rs.(100/850)} = Rs.34,00,000$$

$$\%_B = \left[\frac{Rs.30,00,000}{Rs.750(10,000)} \times 100\right] = 40 \text{ per cent}$$

In order to project the production and sales necessary to the planned profit $(\bar{\pi})$ is treated as an additional increment of fixed costs. Thus, the basic equation now change :

$$Q^B = \frac{TFC+\pi}{P-AVC}$$
$$Q^V \text{ or } S^B = \frac{TFC+\pi}{1-(AVC/P)}$$
$$\%_B = \frac{TFC+\pi}{(P-AVC)Q_{max}}$$

Example Continued

Continuing with the preceding example, suppose a profit target is fixed at Rs. 20,00,000 Then using the above formula, you may find that now

$Q^B = 6667$ passengers

$Q^V =$ Rs. 56,66,670

$\%_B = 66.67$ per cent

It follows that in order to reduce the break-even point, there are either of the three alternatives open:

- Reduce fixed costs
- Reduce average variable costs
- Increase price

In pursuing an alternative, the management is required to keep margin-of-safety (MS) which may be defined as :

either, $MS = \frac{\pi}{TFC}$

or, $MS = \frac{Q^V - Q^B}{Q^B}$

You may observe that the break-even analysis is simple, easy to understand, and relatively inexpensive to perform. Its weakness is the assumption that profit is a simple function of output. In reality, profit is also affected by production process, selling effort, and other factors—internal and external to firm.

Regression Analysis

This is another method commonly used in profit planning; this procedure relate to profit forecasting. You may recall demand forecasting; it is the same econometric technique. The only difference is that instead of 'demand', you are now to consider 'profit' as a variable. The goal of regression technique is to discover a functional relationship between the company's profit and other explanatory variables. In practice, this approach to profit forecasting is greatly enhanced when some logical lead-lag relationship can be found, between company's profit and other variables, where such logical relationship cannot be found, the independent variables must themselves be forecast before a prediction of profit can be made. Of course, the limitations of regression as a statistical technique should also be taken note of.

When you move from planning to control of profit, you should remember that the best method of evaluation is simply to compare actual profit (performance) with budgeted profit (performance) and to undertake necessary steps to bridge the gap between achievements and aspirations.

18.7 PROFIT POLICIES

You may recall, while reviewing critically the economic theory of firm, we have made a remark, *Firms may not maximise profit, but they do have a profit policy,* we may now like to refer to various facets of profit policy decision. After all, profit planning and profit-policy must go together, the former is more technique oriented while the later is more strategy oriented. The firm has to consider a lot of short-run and long-run factors in desigining its profit polity.

There are two issues involved in profit policy decisions : (1) setting profit standards and (2) limiting the target profit.

Profit Standards

This involves a choice of *particular measure and concept of profit* with reference to which achievements and aspirations may be compared. You may refer back to the "profit performance chart" so that you may choose an appropriate concept and measure of profit: much would of course depend on your ultimate goal. In general, you have to choose from among concepts like net retained earning, net operating margin, rate of return on investment, etc.

The next task in profit policy decision is to decide on *an acceptable rate of profit* based on various considerations such as

i) Rate of profit earned by other firms in the same industry.

ii) Normal or historical profit rate earned by the firm itself in the past.

iii) Rate of profit sufficient to attract equity capital.

iv) Rate of profit necessary to generate internal finance for replacement and expansion.

Profit Limits

In addition, the firm considers a set of environmental factors to limit its rate of target profit. Profit target should be limited such that

i) the shareholders do not ask for higher dividends;

ii) the wage earners do not ask for higher wages;

iii) the Government (tax officials) do not ask for higher taxes;

iv) the consumers do not ask for lower prices and discounts;

v) the suppliers do not ask for enhanced rates;

vi) the goodwill of business is not affected.

In short, while deciding on the target rate of profit, the firm has to minimise the payments of "slacks" and to maximise the capacity to absorb external "shocks".

18.8 SUMMARY

We have questioned the profit-maximising hypothesis; but we have not been able to challenge the fact that all firms do have a profit policy. In view of this, we have undertaken a review of various concepts and measures of profit, we have compared and contrasted the accountant's view of profit with economist's view of profit. We have seen that basic difference between the two views lies with regard to calculation of opportunity costs, real value at constant prices, use of current price in valuation etc. We have also made reference to various theories of profit to explain the emergence and growth of profit. Various theories have been grouped into traditional, modern and recent theories. Alternatively we have also suggested another classification into functional and compensation theories, friction and monopoly theories, technology and innovation theories. Next, we have moved from theories to empiricism; we have tackled issues relating to profit policy decisions, profit-planning and control. You should now appreciate that profit is a mixed and a vexed concept, but it plays an important role in managerial decision making process.

18.9 ADDITIONAL READINGS

K.K. Seo, *Managerial Economics*, Richard. D. Irwin Inc., Sixth ed. (First Indian rep. 1988), Chs. 13 & 14.

David Solomon, "Economic and Accounting Concepts of Income", *Accounting Review*, July 1961, pp. 374-83.

J.L. Riggs, *Engineering Economics*, McGraw-Hill. N.Y. 2nd ed., 1982, Ch. 3.

18.10 SELF-ASSESSMENT TEST

1) Explain, with suitable examples, the conceptual and statistical problems involved in measuring economic profit rather than accounting profit.

2) Suave Hats, Inc., manufactures men's hats and sells them at Rs. 25 each. Its total fixed costs is Rs. 6000 per week and its average variable cost is Rs. 13 per unit.

   a) What is the company's break-even sales volume? Draw a break-even chart to illustrate.

   b) What would be company's profit be at its normal production capacity of 2000 hats per week?

3) "Firms may not maximise profit, but they do have a profit policy". Discuss, bringing out clearly the various facets of a profit-policy decision by a firm.

4) Bring out clearly your understanding of the following terms:

   a) Frictional profit

   b) Normal profit

   c) Monopoly profit

   d) Net profit

   e) Target Profit

   f) Reported profit

5) a) Why should management be interested in profit data?

   b) Why should management limit its target rate of profit?

   c) How can management undertake planning and control of profit?

BLOCK 6 INVESTMENT DECISIONS

In this final block, we are going to concentrate on investment decision Investment decision is the be-all and end-all of business decision. The appropriateness and adequacy of any business decision, be it price-output or input-output or sale-purchase or any other, is often judged in terms of the rate of the return on investment. In **Unit 19**, we focus out attention on various concepts and techniques relevant for private investment decision based on the criterion of commercial probability. In **Unit 20**, we rework with the conventional technique of capital budgeting to throw light on the analysis of public investment decisions based on the criterion of social desirability. Any investment decision, private or public, involves a good deal of risk and uncertainty. Therefore, in **Unit 21**, we present a formal economic analysis of risk and uncertainity with particular reference to investment decision, though this analysis is relevant for all types of business decision. To sum up, you must now appreciate that Managerial Economics gives us tools and techniques to arrive at optimal decision under any type of decision environment-certainty, risk and uncertainty.

UNIT 19 CAPITAL BUDGETING

Structure

19.0 INTRODUCTION

It takes a lot of monetary and non-monetary effort to set up a business. And it has to work hard to survive. In a protective market, survival can be ensured through non-market forces but once the market is opened up, competitive forces may play havoe on non -enterprising business organisations. Since consumers' tastes and preferences are ever-changing. Its dependence on one product forever would be suicidal. Needless to say, during the life of a product, changes in design and concept must be brought about along with the increases in productivity through learning experience, economies of scale etc. While this may ensure short to medium term survival, new products have to be launched for long-term survival, keeping in mind the dictates of the market.

All this would mean investing in new opportunities with changing time. To cater to this need, capital must be found either from own resources or borrowing from the market. And more often than not, capital would be scaree given the number of investment opportunities and possibilities. Accordingly, given that resources are scaree, their allocation needs to be biased towards those investments which show more financial prospects. In otherwords, firms have to select opportunities from a given range of possibilities to maximise return on the capital invested.

It must, however, be appreciated that capital being expensive, (because capitals scaree and 'scareity' generates a 'price' to be paid by the investor-say interest on borrowed capital), the investor has to work towards prioritising the scaree resources from a long-term point of view. It is more so because, the revenue and (capital) expenditure wider or cost considerations are spread over a time period. The basic objective of the investor is to maximise the net return i.e. revenue minus costs. Capital could, then, be invested only in those products (and accordingly, the plants and machineries to manufacture them) where the excess of revenue (arising from the sale of the products) over (Capital) expenditure or investment is the maximum over the period of that investment i.e, the life of the plant.

Is this process limited to products or plant and machinery only as has been mentioned above? It can, indeed, be applied to various other areas including service industry. In setting up a management-consultancy firm, for example, investment will be made in acquiring professionals. In most cases, they are very expensive. The product, here, would be the service provided by these professionals in solving a client's problem. Revenue will come from the sale of their services. Accordingly, capital would be required to set up such an organisation.

The critical questions in the context of capital investment are: What to select and how to select from the competing investment opportunities? What is the rational behind choosing one option against the others? Techniques are available to help solve these problems and these form the subject matter of this unit.

19.1 OBJECTIVES

The purpose of this unit is to :

- **explain** various concepts of capital budgeting,
- **explain** various techniques of capital budgeting,
- **examine** the uses and limitations of each technique in the context of real world business applications, and
- **prepare** you to evaluate public investment decisions.

19.2 CONCEPTS OF CAPITAL BUDGETING

Capital budgeting involves investment decision balancing the sources and uses of funds for acquiring fixed capital assets like machinery and equipment. As such the subject is also termed as "fixed asset management" or "equipment choice". Investment in plant and machinery implies the choice of a specific project. There are various projects which compete for the allocation of limited funds. Since funds are scarce, only a specific project can be chosen out of a given set of alternatives. The techniques of capital budgeting enable us to evaluate this choice of project. Therefore, the subject also passes under the title of "project evaluation" or "project appraisal". In this context, it may be mentioned that there are two aspects of project evaluation-technical and financial. An investor has to study both the "technically feasibility" as well as the "economic feasibility" of a project before deciding to invest in it. The study of technical feasibility is essentially concerned with considerations of engineering factors like plant load factor, reserve capacity, technology base etc. The economic feasibility takes care of financial considerations such as the return on investment and risk involved. Under capital budgeting, we are basically concerned with the economic and financial feasibility aspects of project management.

Project selection may imply various types of choices. It may be a choice between 'new investment' and 'replacement investment' or between, 'make or buy' or between 'buy or lease' of a machine, or between 'widening' (diversification) and 'deepening' (extension in the same line) of investment.

Project choice of this nature has to be made on certain objective criteria. These criteria are known as "techniques of capital budgeting". In what follows, these techniques will be discussed in detail.

Activity I

a) Why is investment decision an economic problem?

....

....

....

....

....

....

....

....

b) What is the essence of capital budgeting? Name some specific **sources** and **uses** of funds by an organisation.

....

....

....

....

....

Sourcess	Uses
....	....
....	....
....	....
....	....

c) How is 'economic feasibility' different from 'technical feasibility' of a project?

....

....

....

....

....

19.3 PAYBACK OR PAYOUT PERIOD

This is a very simple rule-of-thumb. As the name suggests, this technique seeks to answer an investor's searching question as to how long he has to wait before his invested capital is recovered. The answer, therefore, is given in units of time-years. months, etc. As cashflows start coming in and accumulate, there will naturally be a time when teh accumulated cash inflows will **equal** the original investment - it is at that point of time that the **payback** occurs and the time it has taken for it (the payback) to happen is the Payback or Payout Period. Let us see this in terms of an example.

Example

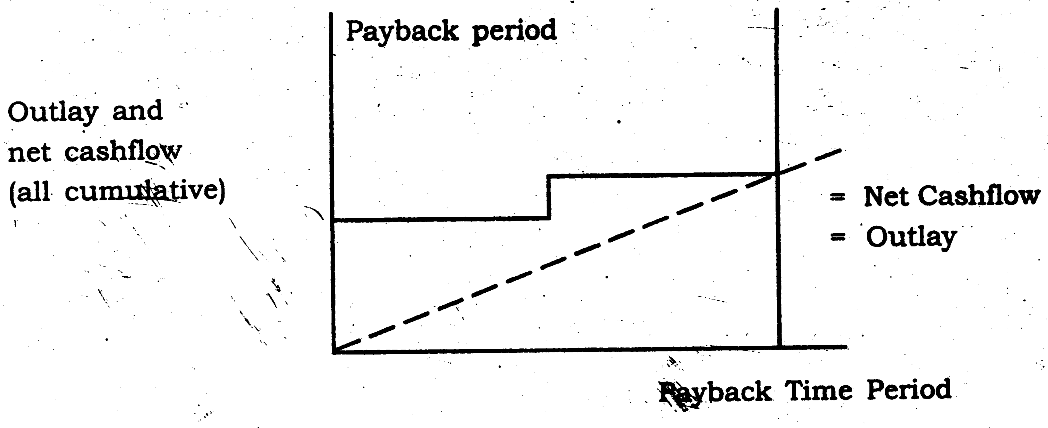
Pay Back Period Calculation

1	2	3	4	5	6	7
Year	Initial outlay	Recurrent outlay	Outstanding total outlay (Beginning of year)	Net cash flow	Balance yet to recover (end of year)	Payback period
1.	500,000	-	500,000	-	500,000	-
2.	-	-	500,000	200,000	300,000	-
3.	-	50,000	350,000	150,000	200,000	-
4.	-	200,000	150,000	50,000	-	-
5.	-	-	50,000	75,000	-25,000*	4 years 8 months

Notes : Col. (4) - Col. (5) = Col. (6)

*The minus sign here means that the Payback Period is already achieved there is now some surplus cash available after recovering the outlay. At the end of 5 years. this is 25,000.00 in this case.

Net cashflow occurs from second year onwards. At the end of the first year. outstanding capital outlay is 500,000 but it is reduced to 300,000 because of the 200,000 net cashflow. At the beginning of third year, the outstanding outlay increases by 50,000 to 350,000 because of an interm capital expenditure (there is nothing uncommon in this). By the end of fourth year. balance to recover is only 50,000. In year 5, net cashflow is 75,000. Assuming a uniform spread of this 75,000 over the 1-2 month period. the monthly figure amounts to 6,250. We have yet to recover 50,000. It will, therefore, take another 8 months (8 x 6.250= 50,000) that is 4 years and 8 months in total to recover all the initial and recurrent capital outlay. Figures are always welcome to depict financial situations and the present case can be shown in terms of a figure as next page.



The diagram is self explanatory - it depicts a situation as evidenced in the above example.

The following points are worth nothing in accepting or rejecting a project on Payback Period basis :

i) If a firm has a specific requirement that, say the investment has to be recovered within 3 years then any project with a payback period of less than 3 years would be preffered to those which have longer payback periods.

ii) If the projects are mutually exclusive, one with a shorter payback would be accepted. Mutually exclusive projects are alternative investments; if one project is taken on, the other must be dropped.

Some of its features are as follows :

a) It is quite popular because it is simple to calculate and easy to understand - a sort of Rule-of -Thumb approach.

b) It ends to place too much emphasis on quick return and liquidity. The company may shelve projects that bring in long-term growth in favour of quick return projects.

c) It ignores the returns beyond the payback period. The emphasis is on quickly recovering the outlay. So the subsequent inflows are ignored.

d) The method does not really show how profitabls the investment is.

e) It ignores the time value of money. This is its major defect. If this gets incoporated in the method. it will have some credibility particularly for high-technology industries where avoding obsolescence is part of their strategy.

19.4 ACCOUNTING RATE OF RETURN

This is just the reverse of the payback method. If we use

A = Annual Cashflow

I = Investment

There, I/A = Payback period.

But, if we reverse this ratio and calculate the same in percentage terms, then we get

$$\left[\frac{A}{I} \cdot 100\right] = \text{Rate of return on investment (ROI).}$$

It can be easily seen that the decision rule in applying payback method is: shorter the payback, better the project. In case of ROI method, the decision rule reads: Larger the ROI, better the project.

The ROI is an established accounting method to compute commercial profitability of a project. However, one may end up estimating different rates of return, depending upon the nature of measures used. For example, Annual cashflow 'A' for a number of years may be related to investment "I" such that you may compute.

$\frac{A_1}{I}, \frac{A_2}{I} \ldots \frac{A_n}{I}$ separately

or you may compute $\sum_{i-1}^{n} \frac{At/n}{I}$. In the second case,

you are computing average ROI which may be different from a single year ROI like $\frac{A_1}{I}$

Similarly when investment is spread over a number of years, you may compute

$\frac{A_1}{I_1}, \frac{A_2}{I_2}, \ldots \frac{A_n}{I_n}$ or alternatively

$$\frac{\sum_{t=1}^{n} I_{t/n}}{\sum_{t=1}^{n} I_{t/n}} \quad \text{or} \quad \frac{A_1}{\sum_{1-I}^{n} I_1}$$

In all these cases, you may end up getting different results. The point remains that the ROI method may be used by way of accounting convenience rather than accounting convention. Thus different ROIs may be computed for different purposes. More important, like the payback, the ROI, method also does not consider the time value of money.

Activity 2

a) Given the following data, calculate the payback and average rate of return for both projects X and Y. Which project would you choose and why?

	Cashflows		
Year	Project X	Project Y	
1	Rs.500	Rs.200	Payback ....
2	Rs.100	Rs.201	Average → ROI
3	Rs.100	Rs.301	
4	Rs. 200	Rs.200	

Investment = Rs. 700

b) What are the common limitations of the payback and ROI methods ?

....

....

....

....

c) If annual cashflow differs in size, can you apply $\frac{I}{A}$ to compute payback? Explain

....

....

....

....

....

19.5 TIME VALUE OF MONEY

Money saved over a period of **time** yields **interest** and herein lies the concept of time value of money. It has been suggested that the acceptability or otherwise of an investment should be judged from this principle. There are two methods which take account of this principle namely the Net Present Value and the Internal Rate of Return. Before we take them up, we should try to understand the importance of this concept and this is done in the following section on discounting.

Principle of Discounting and the Present Value Concept

Time value of money implies that today's Rs.100.00 is worth more than Rs.100.00 received a year hence. But by how much? This is provided by the market rate of interest. Supposing that this was 10% p.a Rs.100.00 saved today, say, with a bank, would grow to Rs.1 10.00 (Rs.100.00 + 10% interest) in a year's time. In other words, at this rate of interest, today's Rs.100.00 is the same as Rs.1 10.00 a year from now or Rs.121.00 after 2 years. In the first year, Rs.100.00

will grow to Rs.1 10.00. At the beginning of the second year, Rs.1 10.00 will be the principal and on it, there will be another 10% interest. At the end of second year, then, Rs.110.00 will grow to Rs. 110.00 + 10% of Rs.110.00 = Rs. 110.00+ Rs. 1.1.00 = Rs. 121.00. This goes on and this is the well-known principle of compound interest. Therefore, we can say, Rs.100.00 received today is the same as Rs.1 10.00 received after a year or Rs.121.00 received after 2 years all at 10% rate of interest. Rs.1 10.00 or Rs.121.00 are all **future** sums. Following from this, we can also say, assuming a rate of interest of 10%, the present value that is, value today of Rs.1 10.00 (receivable after a year) is Rs.100.00. This is how the phrase "Present Value of a **Future Sum**" is interpreted. The future sum after one year, in this case, is Rs. 10.00. Its **present value** is Rs.100.00. The future sum, after two years, isRs 121.00 - its present value is Rs. 100.00 and so on. In calculating the future sum, we compound a given sum at a given rate of interest.

If the future sum is given, to find out its present value, we discount, that is, calculate in the reverse direction.

Future sum, say, is given Rs.1 10.00 receivable after one year. Rate of interest at 10%. Here we divide the given future sum of Rs.1 10.00 by 1.1 which gives us Rs.100.00 as the present value of the future sum of Rs.1 10.00 receivable after one year. If the future sum was Rs.121.00 receivable after 2 years at 10% rate of interest, we would divide Rs.121.00 by (1.1 x 1.1) or (1.1)2 to get Rs.100.00 as its present value.

In terms of mathematical formula, to compute the present sum.

$F = P(1+r)^n$

Where F = Future sum

P = Present sum or the principal
r = Rate of interest
n = number of years.

in our example above

$F = 100(1+0.1)^1 = 100 \times 1.1 = 110.00$ after one year
$F = 100(1+0.1)^2 = 100 \times (1.1)^2 = 121.00$ for two years
Here $(1+0.1)^n$ is the compounding factor.
like $(1.1)^1 = 1.100$ first year
$(1.1)^2 = 1.210$ second year
$(1.1)^3 = 1.331$ third year and so on.

If we want to find out the **present value** of a future sum, the procedure is just the reverse. The formula above, after re-arranging, becomes

$$P = \frac{F}{(1+r)^n}$$ Discounting future cashflows to Present Value)

Now if the future sum F is given, say as Rs.1 10.00, r = 10% and n 1 then

$$P = \frac{110.00}{(1+0.1)^1} = \frac{110}{1.1} = 100$$

Similarly, $P = \dfrac{121.00}{(1 + 0.1)^2}$

i.e. the present value of Rs.1 21.00 receivable after 2 years at 10% is Rs.1 00.00

The above formula $P = \dfrac{F}{(1 + r)^n}$ can be re-written as

$$P = F \times \frac{1}{(1+r)^n}$$

The second part, that is, $\dfrac{1}{(1+r)^n}$

represents the discounting factor, such that at r = 10%

$\dfrac{1}{(1 + .01)^1} = \dfrac{1}{1.1} = 0.9091$ is the discounting factor for year 1

$\dfrac{1}{(1 + 0.1)^2} = \dfrac{1}{(1.1)^2} = 0.8264$ or year

$\dfrac{1}{(1 + 0.1)^3} = \dfrac{1}{(1.1)^3} = 0.7513$ for year 3 and so on.

Thus, if future sum is Rs.1 10.00 receivable after one year, the discounting factor to find its present value is 0.9191 (110.00 X 0.9191 - 100.00) **or** if future sum is Rs. 121.00 receivable after 2 years, the corresponding discounting factor will be 0.8264 so that

Rs.121.00 X 0.8264 - Rs.100.00 is the present value of that Rs.121.00.

To summarise the above, if the rate of interest 10%, Rs. 100.00 is the present value of Rs.110.00 receivable after 1 year. Rs.100.00 is the present value of Rs.121.00 receivable after 2 years

To arrive at future sum: Compound : $F = P(1+r)^n$
To arrive at present sum: Discount : $P = F \; 1 / (1+r)^n$
Compounding factors represented by $(1 + r)^n$

Discounting factors represented by $1 / (1+ r)^n$

Diagramatically, we can say,

Compound
Present sum -------------------- > ------------------ > ----------------- > Future sum
Present (Value) -------------------- < ------------------ < ----------------- Future Cashflow (or sum)
Discount

Readymade tables are available in Financial Management textbooks to find out the discounting factors for a given r over the years.

Before we conclude this sub-section, there are certain other features of the Present Value (PV) concept that we should look into. We have seen earlier, if the Cashflow receivable after 2 years is Rs.121.00, its Present Value (PV) would be Rs.1 00.00 and if the rate of interest (also at times referred to as discount rate) is 10%, the discounting factor will be 0.8264 $(1/(1+r)^2 = 1/(1+0.1)^2 =$ 0.8264 and 12LOO X 0.8264" 100.00). Naturally, if a higher r was used, that is, if the firm's cost of capital turned out to be higher, and if everything else i.e. investment, cash outflows, cash inflows, etc. remained the same, the discounting factor would be lower and the discounted value of the cashflows would be smaller, Rs.121.00 receivable after 2 years at r = 10% would have a PV of Rs. 100.00, the discounting factor being 0.8264. But with r = 12%, the same Rs.121.00 would have a PV ofRs.96.46,the discounting factor being.

0.7972 (($1/(1+r)^2 = 1/(1+0.12)^2 = 1/(1.12)^2 = 0.7972$) as against

0.8264above.

The oppositive results would prevail if r is put at say 8%. Therefore, higher the r, lesser is the discounting factor arid hence also the PV of a future cashflow. This gets reflected in the following table 1.

Table I

Discounting Factors under varying r and the Present Value of Rs.100.00 Received after I to 5 years.

Year	r8% $(1/(1+0.08)^n)$	PV of* 100.00	10% $(1/(1+0.1)^n)$	PV of* 100.00	12% $(1/(1+0.12)^n)$	PV of* 100.00
0	0.0000	100.00	0.0000	100.00	0.0000	100.00
1	0.9259	92.59	0.9091	90.91	0.8929	89.29
2	0.8573	85.73	0.8264	82.64	0.7972	79.72
3	0.7938	79.38	0.7513	75.13	0.7118	71.18
4	0.7350	73.50	0.6830	68.30	0.6355	63.55
5	0.6806	68.06	0.6209	62.09	0.5674	56.74

* Derived by multiplying 100.00 by the respective discounting factors.

Activity 3

a) What do you understand by "time value of money"?

....

....

....

....

....

b) How do you compute the present value of a future series of income?

....

....

....

....

....

....

....

....

....

....

....

c) What is "discount factor "? Is it right to say that the more distant is the income, heavier must be discounting? Justify your answer.

....

....

....

....

....

....

....

....

....

....

....

Cost of Capital

In the above examples, we have repeatedly used the expression, rate of interest - in short 'r' as used in the formulae. We have treated it as the market rate of interest. In the analysis that follows, in both the NPV and IRR, we shall treat 'r' as the cost of capital. Needless to say, most investments take place through equity and borrowed capital - either from banks and or other public sources.

It may be noted that each source of capital has its own cost. Accordingly, these costs need to be weighted depending upon how much is borrowed or raised from which source. If r for one source was 12% and 4% for another and the sums borrowed were Rs. 75 and Rs. 25 respectively, then the weighted r on weighted cost of capital would naturally by

$$\frac{75}{100}\times 12+\frac{25}{100}\times 4=10\%$$

It is along there Hnes that "r" will be used in the calculations that follow that is r = the average weighted cost of capital of the project.

It may be noted from the table above that at r = 10.0% the PV of the same Rs. 100.00 received after 2 years is 82.64 but goes down to 62.09 if received after 5 years. Intuitively. it can be said that under the discounted cashflow system investment with high net cashflows in earlier years will be favoured against those with high net cash flows in later years.

Activity 4

a) Take a book on Financial Management. Look for the PV tables. Fill in the information.

	PV for Re 1					
Year	5%	10%	15%	20%	25%	30%
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
15						
20						
25						

b) Note down your observations about the above information.

....

....

19.6 NET PRESENT VALUE (NPY)

Over the years, the NPV method of judging the acceptability or otherwise of an investment has come to stay. The basic reason is that through the process of **discounting**, as explained above, all the future cashflows - outflows or costs and inflows or revenues can be converted to their present value. Since capital is being raised and invested today it will be naturally convenient to the investor if he knows how the sum of all the future net cashflows look like in today's terms. That way, the profitability of the investment after taking into account the cost of capital could be judged instantaneously. This is exactly what the NPV (and in a way, also the IRR) seeks to do through the process of discounting the cashflows and bringing them down to a common base, that is, to their present values using a rate of interest which in this case would be the firm's cost of capital.

An example

Assume that an investment of Rs.25,000.00 is undertaken, the cost of capital being 10.0%. The life of the project is 5 years and in those 5 years, the net cash inflows are:

Year	1	10,000.00
	2	8,000.00
	3	6,000.00
	4	5,000.00
	5	5,000.00

We prepare the following table to calculate the discounted values of these sums:

Years	Net cash inflow 1	Discounting factor at r= 10% 2	Discounted PV of net cash inflows 3 = 1x 2
1	10,000.00	0.9091	9,091.00
2	8,000.00	0.8264	6,611.00
3.	6,000.00	0.7513	6,010.00
4.	5,000.00	0.6830	3,415.00
5.	5,000.00	0.6209	3,105.00
	36,000.00	Gross or Total PV	28,232.00
		Investment -	25,000.00
		NPV	= 3,232.00

The last column here expresses all the net cash inflows in present value terms. The sum of all these flows or the gross PV is put at 28,232.00. The investment, since it is being carried out now, is already in its PV (25,000 x 1,0000 for year 0 = 25,000.00) Therefore, the investor can see immediately that there is actually a surplus (of 3,232.00) even after deducting the investment from the sum of the PV of the net cash inflows over the 5 years. Everything being in present value terms. it is not difficult for our investor to say to himself that the project is profitable. The process, then, to judge the profitability (and acceptability) of an investment involves :

- given the inputs
  - Initial Investment
  - Net cash inflow
  - Cost of capital
  - Project life
  - convert the net cash inflows to other PV using the discount factors given by the cost of capital
  - sum up all these PVs
  - Deduct uinitial investment from this
  - This gives the NPV of the Investment

If it is positive, the project is accetable, Larger the NPV, better the project.
It is quite possible that the investor will use his own resources rather than going to the market. Even then, the profitability or otherwise of the investment should be judged. taking the market rate of interest as the guide. Let us look at another example. Supposing there is Rs. 10,000.00 available for investment. The market rate of interest is 10% and the net cash inflows for the project life is

Year 1	1000.00
Year 2	2000.00
Year 3	2000.00
Year 4	6000.00

The NPV calculation would look like :

Year	Net Cash inflows	Discounting factor @ 10%	PV of net cash inflows
1.	1000.00	0.9091	909.00
2.	2000.00	0.8264	1.653.00
3.	2000.00	0.7513	1.503.00
4.	6000.00	0.6830	4.098.00
		PV	8.163.00
		Investment	10.000.00
		NPV	-1.837.00

Here, of course, the NPV is negative, which means there of 10.0% is too high; that is with these cash inflows, to make NPV positive, the actual 'r' would have be to be much lower. Let us look at it from another angle. The net cash inflows are investible at the market rate of interest of 10% which means; given the project life of 4 years, Rs. 1000 at the end of year 1 can be reinvested at 10%

for the next 3 years, Rs.2000.00 of year 2 for 2 years, and so on. Therefore, the terminal value of the reinvestment of these cashflows would amount to:

Cashflow of	Reinvested for	Future Value	Compounding factor $(1 + r)^n$
1000 Yr. 1	3 years	1,331,00	1.3310
2000 Yr. 2	2 years	2,420.00	1.2100
2000 Yr. 3	1 year	2,200.00	1.1000
6000 Yr. 4	0 year	6,000.00	1.0000
		11,951.00	

Against the figure of 11, 951.00, if the original Rs.1 0,000.00 was straight away put in the bank earning a compound interest of 10% p.a., the total figure,
principal + interest would be $F = P(1+r)^n = 10{,}000\,(1+0.1)^4 = 14{,}641.00$

This shows a loss of Rs. 2,690.00 (14641.00 - 11,951.00) and when this is multiplied by 0.6830, the discounting factor of year 4, the figure to emerge is 1837.00, the earlier negative NPV figure.

Alternatively,
Invested in Bank : PV of 14,641 = $\dfrac{14{,}641}{(1+0.1)^4} = 10{,}000 \quad \left(P = \dfrac{F}{(1+r)^n}\right)$

Invested in Business: PV of 11, 951 = $\dfrac{11{,}951}{(1+0.1)^4} = 8{,}163 \quad \left(P = \dfrac{F}{(1+r)^n}\right)$

Difference = 1,837

This was the negative NPV derived earlier. This is how, recommendations can be done about acceptability or otherwise of the prodject.

In alogebra terms, the net Present Value formula then becomes

$$NPV = \frac{F_1}{(1+r)} + \frac{F_2}{(1+r)^2} + \frac{F_3}{(1+r)^3} + \frac{F_n}{(1+r)^n} - C_0$$

where $F_1, F_2, F_3 \ldots F_n$ = Net cash inflows for years 1,2,3,..... to n

r = Interest Rate (or the discount rate)
C_0 = Initial Investment

If the net cashflows are uniform, it becomes easier to calculate NPV. To take an example,

If the Net cash inflows amounted to say, Rs. 1000.00 each year over a 4 year period of a project at r = 10% and initial investment of Rs.3000.00, then the NPV calculation would be:

Method I
As above

Year	Net cash inflow	Discounting factor @ 10%		PV
1	1000.00	0.9091		909.00
2	1000.00	0.8264		826.00
3.	1000.00	0.7513		751.00
4	1000.00	0.6830		683.00
		3.1698	PV	3,169.00
			-	3,000.00
			NPV +	169.00

Method 2
Alternative

Net Cashflow p.a.= 1,000.00
NPV = 1000 X 3.1698 = 3169.80 - 3000.00 = 169.80.
(3.1698 is the sum of discounting factors for the 4 years).

Note : Difference in NPV due to rounding error.

We have talked about single projects above. If there is more than one project and they are mutually exclusive of one another, then the one with higher NPV will be chosen.

Example

	Project A			Project B		
Year	Cashflow	Discounting factor (12%)	PV of cash Flow	Cashflow	Discounting factor (12%)	PV of cash Flow
1	6,000	0.892	5,352	3,000	0.892	2,676
2	14,000	0.797	11,158	10,000	0.797	7,970
3	23,000	0.711	16,353	27,000	0.711	19,197
4	31,000	0.635	19,685	34,000	0.635	21,590
	Pv of inflows		52,548			51,433
	Less initial cost		50,000			50,000
	NPV		2548			1,433

The NPV of project is A higher so it should be selected.

Also note that Net Present Value Index (NPVI) = $\frac{\text{PV}}{\text{(Initial Cost)}}$ and that NPVI is higher in case of proejct A

Activity 5

a) Distinguish between NPV and NPVI. Would you agree that positive NPV impliesNPVI > unity illustrate with an example.

....

....

....

....

....

....

....

....

....

....

b) Distinguish between the concepts of discount factor and PV factor, if any

....

....

....

....

Year	Cashflow I	Cashflow A	PV factor at 10%	Present values of cash inflow
0	1500	0	1.000	
1		650	.909	
2		550	.826	
3		450	.751	
4		320	.683	
5		320	.621	
NPV =			NPVI =	

I = Initial, A = Annual

Complete the above table.

We shall now look into IRR and then compare it with NPV.

19.7 INTERNAL RATE OF RETURN

You would recall that under the NPV method, we had followed these steps:

- Find the sum of PVs of all future net cash inflows using the discounting factors as relevant to the r being used.

- Deduct the initial outlay from the sum of these PVs to find out what the NPV is

The resultant NPV would, in most cases, be either positive or negative. The IRR method goes a step further and raises this question:

What is that r which will make the sum of PVs of all the future net cash inflows **equal** to the initial investment?

This is another way of saying

What is 'that r which makes NPV = 0? Following the NPV analysis above it amounts to this that where NPV is positive, the value of r will increase to reduce the size of NPV to eventually make it zero and where the NPV is negative, the value of r is **too high** and accordingly, the cashflows would have to be discounted using a lower r so that the size of the negative NPV would be reduced and the process continued until it reaches zero. The rationale of using IRR, as in the case of NPV, is the same, that is, if everything **could** be interpreted in today's terms, taking a decision would be easier. If the resultant 'r' is greater than the cost of capital then the investment is worth undertaking. In terms of algebra, we can say, the present value of a series of future cashflows is given by -

$$\frac{F_1}{(1+r)} + \frac{F_2}{(1+r)^2} + \frac{F_3}{(1+r)^3} + \ldots\ldots + \frac{F_n}{(1+r)^n}$$

The internal rate of return is defined as that discount rate r which equates the present value of the expected future cashflows to the initial cost of the project., i.e., if the initial investment is C

$$C = \frac{F_1}{(1+r)} + \frac{F_2}{(1+r)^2} + \frac{F_3}{(1+r)^3} + \ldots\ldots + \frac{F_n}{(1+r)^n}$$

We know the value of C and the cash inflows. F., F2 etc. So we have to solve the above equation for r. In the NPV method, r the cost of capital is known and we determine the PVs and then the NPV. In the IRR method we are trying to determine what rate of return our initial outlay is earning for us during the life of the project taking into account the discounting principle.

From the above, it is immediately clear that the IRR is a percentage concept (rate of return) whereas the NPV is a sum, positive or negative at a given r.

With a uniform cashflow situation, calculating the IRR value creates little problem. But in real life, cashflows are rarely uniform and in such cases, real difficulty arises as soon as the life of a project is more than 2 years. Quadratic equation can be used in a situation where the cashflows extend over 2 years only. For instance, if an investment had an outlay of Rs. 500.00 and net cash inflows of Rs. 300.00 in the first year and Rs.400.00 in the second year, the IRR could be calculated in this way:

$$-500 + \frac{300}{1+r} + \frac{400}{(1+r)^2} = 0$$

Multiplying both sides by $(1 + r)^2$, we get

$-500 \; (1+r)^2 + 300(1 + r) + 400 = 0$

Since a quadratic equation takes the form

$ax^2 + bx + c = 0$ where

$$X = \frac{-b + / \; b2 - 4ac}{2a}$$

the above cashflow equation can be transformed into a quadratic equation shape in which case we have the following:

1+r = x
-500 = a
300 = b
400 = c

$$\text{Therefore } 1 + r = \frac{-300 \pm /(300)^2 - 4(-500 \times 400)}{2\,(-500)}$$

$$= \frac{-300 \pm \; /90{,}000 + 800{,}000}{-1000} = \frac{-300 \pm 943}{-1000}$$

In line with the solution for x, this too will have two values. Since in a project where cash inflows are greater than the cash outflows, a negative value for would not make any economic sense. Accordingly, we take the positive solution only, so that

$$1 + r = \frac{-300 - 943}{-1000} = \frac{+1243}{+1000} = 1.243$$

Therefore r or the IRR = 24.3% (Putting this value in

$$-500 + \frac{300}{(1 + r)} + \frac{400}{(1 + r)^2} = 0 \text{ satisfices the equation}$$

Alternate Method

Alternatively, this can also be solved through linear interpolation in this way that we try out a discount factor which yields a positive NPV and another one which yields a negative NPV. The smaller the NPV size, the better it is.

Year	Cashflow	$r_1 = 23.0$	NPV
0	-500	1.000	-500.0
1	+300	0.8130	243.9
2	+400	0.6610	264.4
			+ 8.3 NPV_1

Year	Cashflow	$r_2 = 26.0$	NPV
0	-500	1,000	- 500.0
1	+300	0.7937	238.1
2	+400	0.6299	252.0
			- 9.9 NPV_2

It follows that the resultant r should be a little more than 23.0% (NPV is still positive) but a little less than 26.0 (NPV is now negative).

The interpolating formula is

$$r = r_1 + \frac{NPV_1}{NPV_1 + NPV_2} (r_2 - r_1)$$

$$= \quad 0.23 + \frac{8.3}{18.2} (0.03)$$

= 0.2437 i.e. r = 24.4% which is near enough to the quadratic equation result.

This is near enough to what is generally known as the trial and error method. Even in this simple example, quite an amount of search would need to be carried out to find a small enough positive NPV on the one hand and a small enough negative NPV on the other. In a situation, where several years are involved, the process just goes on, as is shown in the following example:

The outlay is Rs. 50,000. The cash inflows are Rs.6, 000, Rs. 14,000, Rs. 23,000 and Rs.31,000 in years through 1 to 4. What is the IRR?

Assume IRR = 10%. Find NPV

Year	Discounting factor for 10%	Cash inflow	PV of cashflows
1	0.909	6,000	5,454
2	0.826	14,000	11,564
3	0.751	23,000	17,273
4	0.683	31,000	21,173
	Present Value		55,464
	Less outlay		50,000
	NPV		5,464

NPV is positive i.e. the Present Value is higher than the outlay. So we have to raise the discounting factor.

Assume IRR 13%:

Year	Discounting factor at 13%	Cash inflow	PV of cashflow
1	0.884	6,000	5,304
2	0.783	14,000	10,692
3	0.693	23,000	15,939
4	0.613	31,000	19,012
	Present Value		50,947
	Less cost		50,000
	NPV		947

Still NPV is positive. So the IRRhas to be raised. Assume IRR = 14%. In this case the PV of cashflow = 49,905. This is very close to the outlay of Rs. 50,000. So the IRR of the project is close to 14%.

The result has been p retty close to what it should be through the trial and error method, but this tiring process can be somewhat simplified with the help of graphs. The idea is to get an approximate picture first and then carry on with the trial and error method for still finer results. If we assume a two-dimensional system wherein the y-axis represents NPV and x-axis, r i.e. the discounting rates, and we plot at least two values of NPV and the corresponding r and joining them up, then the line so formed will, when extended, cut the x -axis some r where NPV will be read as "zero". The r at that cut-off poinf would approximate to IRR, although some more calculations may be necessary. This is because the line so formed will, in reality, not be a straight line but somewhat concave • to the origin. Because of this concavity, the IRR will only be approximate. So, make it a little more accurate, (a negative NPV, r) combination is desirable. Let us consider a project which would require an investment outlay of Rs.10,000 and is likely to generate a net cash inflow of Rs.5000.00, Rs-4,000.00, Rs.3,000.00 and Rs.1 000.00 through 1 to 4 years. We are to find out the IRR as accurately as possible.

Year	Net cash inflow	Discounted cashflow r= 10.0%	r = 15%	r=12%	r = 20.0%
1	5,000	4,545	4,348	4,464	4,167
2	4,000	3,306	3,025	3,188	2,778
3	3,000	2,254	1,973	2,135	1,736
4	1,000	683	572	636	482
	PV	10,788	9,918	10,423	9,163
	Outlay	10,000	10,000	10,000	10,000
	NPV	788	-82	423	-837

To try, please proceed with : r = 10%, r = 12%, r = 15%, r = 20%

The above table shows that the IRR should be lower than 15% around 14.5%. To dispense with this tedious process, computer programmes are available. A sample programme of IRR calculation is given in the appendix.

Which one iochoose-NPV or IRR?

Until now, we have discussed the technical aspects of the two DCF (Discounted cashflow) based methods, NPV and IRR. Both of them have this similar feature that they take account of the time-value of money. It needs to be hardly emphasized that this is their major virtue.

It must be said, that of the two, the NPV is a simpler method and yet provides a logical acceptance criterion. In the case of IRR, what we must not forget is that the cashflows may not be able to attract a re-investment rate equal to the 'r' and yet it is assumed in the IRR method that cashflows are to be re-invested at the IRR rate. This may be totally unattainable because of the magnitude of r. Secondly, the "r" may not be the true "r" because, if there are negative cashflows in between the years, the possibility, of multiple IRR emerges. As a rule-of-thumb, the number of times negative cashflows occur will determine the multiplicity of IRR. Given this difficulty in the interpretability of IRR, one would feel more at home using the NPV method. It is not that the NPV method does not have its limitations. For instance, we have to be very cautious indeed while selecting the cost of capital i.e. the discounting factor. It is quite possible that the firm's cost of capital will change over a period of time however representative the pool of borrowed funds might seem to be at a particular point of time. Thus, equity capital being much costlier than say debenture, raising the equity base during the course of the project might change the overall cost of capital. This might seem to be a plus point for IRR because there is scope for management to be flexible in fixing the cost of capital. To incorporate this flexibility in the NPV method, we might have to use the cost of capital over a range rather than one fixed for the whole of the project life.

In practice, however, businessmen are more at ease whenever there are references to rate of return as a measure of profitability. IRR is, therefore more widely recognised 6y them than NPV.

Activity 6

a) Distinguish clearly between ROI and IRR.

....

....

....

....

....

....

....

....

....

....

b) It is said that the basic difference between NPV and IRR techniques lies in the fact that in one case r is known and in the other case it is unknown.. Do you agree?

....
....
....
....
....
....
....

c) How is a PV table different from an Annuity table? (See a text on financial management for these tables).

....
....
....
....
....
....

d) Write down the steps to locate the IRR, through linear interpolation method.

....
....
....
....
....
....

19.8 SOME PROBLEM AREAS

Conflicts arise in ranking projects depending upon which method is being used. If two mutually exclusively projects are being considered, NPV method may signal preference for one which IRR may not. In other words, conflicting results may be generated by NPV and IRR. NPV of projects when drawn with NPV on the Y-axis and the discount rate on X-axis, may intersect each other creating what is generally known as the "Fisherian intersection" (owing its title to Irving Fisher who, in the early part of this century, made significant contribution in this area) thereby generating conflicting results.

Such a situation can arise because of size disparity between projects as also because of timing of cashflows with no size disparity in the initial outlays. Under such circumstances, the following steps are likely to lead to a reconciliation between NPV and IRR:

- Develop the NPV profiles of both the projects - X' and - Y at various discounting factor say, at 5, 10,20,30 until NPV becomes negative.

- Draw a third profile, call it (X-Y) differential, so that when plotted, the IRR of (X - Y) would be the cross-over point between Project X and Project Y. If the cost of capital is to the left of (X-Y)'s IRR, then accept project X; otherwise accept project Y.

The second problem area could be associated with projects with unequal lives. Some writers have suggested that under such circumstances, mere ranking by NPV might give misleading answers. To avoid this, a comparison of average NPVs between the projects should form the basis. This can be done in this way, assuming there are two Projects X and Y.

- Find out the NPVs of X and Y.
- Divide the NPVs by respective Projects' annuity factors at the cost of capital. (As noted above, there is the cumulation of discounting factors at each r). This will give the average net present values of the projects.
- The one (X or Y) wi'th the higher average NPV value would bemore attractive.

19.9 SUMMARY

In this unit you have been exposed to various techniques like payback, ROI, NPV and IRR. The last two take care of time value of money through discounting.

There are problem areas where the choice between NPV and IRR may pose difficulty. We have indicated that IRRs can specially cause confusion in the choice of projects. For one thing, sizes differ a project half the size of another with cashflows in the same proportion may generate equal IRRs. Secondly, there are questions such as: if Project X has an IRR of 20% and Project Y of 30%, what does this extra 10% in Project Y mean in money terms? Most problems like these, apart from the others mentioned in the previous paragraphs, are better taken care of by NPV method. Above all, the method itself is simple enough. There lies the attractiveness of NPV.

While discussing "Payback" method above, we noted that one of its major drawbacks was its neglect of the time value of money. To accept or reject an investment proposal, time value of money continues to be much popular amongst businessmen. A discounted "Payback Period" method has been suggested where the cumulative cost and revenues are discounted costs and revenues at the firm's, cost of capital. Under such circumstances, the payback period would necessarily be longer than that determined by the traditional approach. It can be denied that in fast changing industrial environment, 'payback period' is rather useful to know to plan ahead effectively. A discounted system of payback, coupled with an NPV analysis (to overcome the other drawbacks of payback method) would sound an effective combination.

We must not lose sight of this important fact that in the DCF calculations, all variables have their values in real terms. This is not to be taken as inflation.

Where figures are given at current market prices, they need to be adjusted through the use of appropriate index numbers.

In this unit, we have virtually left out **capital rationing** which, in simple terms, means that although there may be many "acceptable" projects, how many of them could be undertaken would depend on how much capital is available. This is an area where programming, i.e. linear programming, has made considerable contribution. Interested readers may probe further into this area. Finally, the reader may feel that these techniques of capital budgeting are relevant for only private investment decision. This is not quite correct. In the next unit, we would like to rework and extend these techniques to demonstrate their applicability in the context of public investment decision.

19.10 ADDITIONAL READINGS

Walker, E.W. (1978) Essentials ofFinancial Managements, 2nd ed. Prentice-Hall ol India Pvt. Ltd.: New Delhi. Ch.6.

Adhikary, M. (1987) Managerial Economics, Khosla Educational Publishers: Delhj Ch.VII.

19.11 SELF-ASSESSMENT TEST

1 What are the various aspects of investment decisions? How are they involved in opportunity cost calculation?

2 What is 'Project Evaluation'? How do you go about it? Suggest and explain some standard criteria on the basis of which you may select a project.

3 Suppose different techniques give you different ranking of a project vis-a-vis other projects. How would you select an investment project ia this case?

4 Consider the following data and analyse them to fill in the necessary data in the table below:

Method	Project A	Project B	Preferred Project
- Pay back period			
- Average rate of return on average invetment			
- Internal rate of return			
- Net Present Value			
- Net Present value Index			

Data Given : Investment = Rs. 1500

Annual Cashflows

public investment decisions must be understood by managers who have to deal with public agencies and are affected by their investment decisions. What are the different types of public investment? Given the complexity of factors influencing public investment, one has to understand different concepts required for the analysis of public investment decisions. How can governments make an appraisal of alternative (public) investment proposals or projects? What are the factors influencing the nature and magnitude of public investment? This unit intends to answer these questions.

20.1 OBJECTIVES

The present unit is intended to help' you to:

- **appreciate** the significance of public investment decisions
- **understand** the intricacies of the concept of public investment
- **distinguish** between public and private or non-governmental investment decisions
- **take** a total view of the organisational, constitutional and legal arrangements under which public investment decisions are made
- **classify** different types of public investment
- **have** a conceptual foundation towards the analysis of public investment decisions
- **locate** the determinants of public investment
- **learn** various techniques or methods of appraising public investment decisions.

20.2 SIGNIFICANCE OF PUBLIC INVESTMENT DECISIONS

Modern governments are a major generally the single largest - participant in and organiser of economic activities. This is true across the globe, both for the large and more developed countries and for the small, newly independent and developing countries. Not with standing the differences in the nature, objectives and socio-economic character of the state in different countries, they claim to act and speak on behalf of society in general. Hence it is natural that the state would have a great deal of interest in the future as well. After all, societies are not mortal like individuals. If one has to identity one major activity of the state which shows its concern for future and which enables it to link the present and future in indissoluble ties, it is the investment activity of the state.

As we shall see below, a large number of factors activate the State to undertake a variety of investments.

They all reflect the interests and responsibilities of the State, whether it be the building of school buildings or space shuttles. A large number of current and future wants of an economy are satisfied b y the State. The capacity to discharge these tasks requires investments in men, materials, machines and institutions. Many of the private and non-government activities and investment become possible on the basis of a whole range of public investments in roads, bridges, ports, airports, universities, armaments, etc. The present level of Gross National Product, its rate of growth, stability and its capacity to associate all those members of the labour force who are willing and capable of working depend in a good measure on the past and present public investments. Government policies whether they concern growth, social justice, self-reliance, full-employment, military strength or regional balance, depend for their success on the size, composition and characteristics of public investment. The pursuit of development by the Third World countries has involved their governments increasingly in public investment decisions. The budgetary operations of most governments are generally repetitive with marginal changes except for the elements of big changes and sharp departures introduced by public investment decisions with commensurate implications on the revenue side of the budget, growth of the economy, stability and employment level in the economy.

Public investment is a critical catalyst and stimulator of the level of private investment, demand for various inputs, labour power and skills, and affects its allocation among different branches, industries and services. External balance, i.e., Balance of Payments position of a country and its competitive strength in the comity of nations also bear the impact of public investment. Perhaps the most significant aspect of public investment, besides its contribution to the interests of the generations yet unborn, is the provision of **public goods,** whose range, variety and significance are steadily growing. These goods are critical for the continued reproduction of the economy and cannot be provided by any agency other than the public ones. These goods are complementary to the production of other goods.

Many of them are universal intermediates like means of transport, energy supply, communication, etc. No significant managerial decisions can afford to remain uninformed of the nature and extent of the influence of public investment decisions.

Activity I

Having read the above briefly answer the following questions:

a) Why is there so much interest in public investment decisions?

....

....

....

....

b) What are the government activities which require investment?

....

....

....

....

c) Mention some important aspects of the economy which are influenced by public investment decisions.

....

....

....

....

20.3 PUBLIC INVESTMENT Vs PRIVATE INVESTMENT DECISION

As you are aware investment is an activity which diverts a part of currently available national output (national production supplemented by foreign trade depending on the trade opportunities available to and availed of by a country) for the satisfaction of future wants. Such investment involves:

a) **Diversion** of income from the satisfaction of current wants, i.e., cut-down on current consumption. By and large, the bulk of such savings come out of incomes like profits, rent, interest payments, large salaries, etc.

b) **Transformation** of such savings into capital goods either in an existing production facilities or by setting up new productive capacity.

Governments are an important agency making such investments. As a part of making decisions regarding the public expenditure every year in the course of preparing the annual budget, governments have to decide on those schemes of spending or allocating resources which go to create facilities to satisfy social wants in future over the useful life-span of the facility so created. In countries like India which make public interventions in the framework of a national plan, these decisions have to form part of the budgetary policies as determined by the national plan. In India, by and large, such decisions are related to the Annual Plans of the Union and State governments as arrived at in consultation with the Planning Commission. However, public investment decisions in India also include the capacity creation for satisfying future needs or capital goods purchase and installation decisions of statutory corporations and government companies as well. The decision to take-over existing private capital is a transfer of resources and not a public investment decision.

In terms of budgetary classification, public investment decisions form a part of what is called capital budgeting. Public expenditure is divided into two parts: current account expenditure and capital expenditure corresponding to consumption and investment respectively. It is quite possible that some current expenditure like transfers in the form of pensions and cash subsidies may be used for investment purposes. However, public investment would not include these kinds of investments as they represent secondary and indirect effects of current public expenditure. Loans and advances made by the government to public enterprises, other than temporary ways and means accommodation, are classified as public investments. These public investment decisions are concerned with construction of capital assets like buildings, purchase of capital equipment, installation of machinery and subscription to the equity capital of public enterprises, expenditure on long-term socio-economic infrastructure like roads, bridges, power plants, telephone exchanges, building and equipment for educational, technical and training institutions, loans granted for financing investment projects, etc.; various socio-economic overhead facilities, that is, facilities which provide necessary services and support for a whole range of other activities and are not specific to one particular use or agency. These investments are generally 'lumpy', i.e relatively large and indivisible.

The concept of public investment can be put into a clear focus by contrasting it with private investment decisions. Both the decisions link present to the future in the sense that both involve postponement of current consumption in order to generate the capacity to obtain a larger flow of consumption in the future. However, since two different types of agencies undertake these investments, there emerge some notable dissimilarities between the two. First and foremost, the basic motives and objectives underlying public investment are different from those which give rise to private investment. In a private enterprise market economy (i.e., a capitalist economy) private entrepreneurs invest a certain amount of money capital (M) (i.e., financial resources)' with a view to obtain a larger sum of money capital (M) than the initial amount they started off with. The difference between the two (M_1-M) is the surplus or profit which is the basic motive force behind private investment decisions. You may recall from our discussion of private investment decision the various methods of discounting and investment criteria, etc. and the manner in which they are used in a specific, well-defined situation for giving a concrete form to the profit calculus. The nature of the product, technology, market structure, degree of risk and uncertainty, conditions in the financial market, international factors, etc., influence the calculation and operation of private profit motive. The public investment decision is not based on such criteria of private gain or loss because the State is a compulsory, coercive, **juridical** institution which generally acts and speaks on behalf of the entire community. It undertakes investment for common social purposes and can make a draft on the resources of community as a whole. Generally, these investments are such that they are complementary to a large number of private activities. If these investments are left to private entrepreneurs, either they may not be made at all, or made inadequately on such terms and conditions as may be considered harmful to society. For example, roads, bridges, railway lines, power plants, etc. tend to have a high degree of monopoly. Private investment in such projects may be undesirable on several counts: first, they require huge funds that no private investor would come forward with; secondly, it may lead to private control over vital public services; thirdly, private profit is adverse to the concept of social utility. Such investments are a social necessity as their down-stream effects and up-stream effects (i.e., backward and forward linkages) are crucial for economic development. Hence such public investment decisions are based on social objectives. Large **external economies** of these investments make them unsuitable from the angle of private profit.

Public investments are 'lumpy', i.e., they are large-sized and are characterised by indivisibilities, and are useful for a large number of activities. In this sense, public investments are treated as over-head facilities or socio-economic infrastructure. Private investments are usually directed to the satisfaction of individual wants for which a price or **quid pro quo** can be easily charged. Being overheads and lumpy, often public investments have a long gestation period-both in the sense of requiring a long construction/installation period and yielding benefits for a number of years. These facts, along with the difficulties in pinpointing the beneficiaries and the extent to which they derive the benefits (causing '**externalities**'), differentiate public investment from private investment. Such investments have to be undertaken from the point of view of society as a whole and have to be based on a long time horizon. As against this, private decisions are based on narrow considerations and tend to be myopic.

It means the criteria for public investment are in terms of national objectives. In countries like India which has adopted national planning, these decisions have to be in terms of social profitability defined in terms of the plan objectives. These differences, as we shall discuss in the following pages, give rise to decision- framework for public investment different from that of private investment decisions.

Activity 2

In the light of the discussion above, briefly answer the following questions.

a) **List five important examples of public investment from the experience of: a country like India.**

....

....

....

....

b) **Mention three important differences between public and private investment decisions.**

....

....

....

....

c) **Briefly explain the following:**
overheads, external economies, lumpiness, social profitability.

....

....

....

....

20.4 VARIOUS TYPES OF PUBLIC INVESTMENT DECISIONS

Given their broad-based nature, variety of purposes, large number of functions; different historical contexts of their emergence, one can come across many different types of public investments. Since a number of different types of public agencies like national government, state governments, local bodies, statutory agencies, government companies make public investment decisions, some of these decisions may be distinguished from each other in terms of the authority/agency which makes these decisions. An analysis of different types of public investment decisions should show their range, diversity and contexts.

Perhaps the largest number of public investments, accounting for the bulk of public investment in most countries and historically its most enduring variety is the called socio-economic overheads or infrastructure' like power plants, network of roadways, bridges ports, means of transport and communication. Universities, Research and Development outfits like laboratories, institutions of management, technology - education and training may be counted as components of public overhead investments. These investments go as inputs to other productive activities', have lumpiness' and external economies, and are generally non-tradeable internationally. The governments may follow a variety of discretionary policies regarding these products / services regarding pricing and norms of financial management. It is generally prudent to go in for a little excess production of overhead facilities, choosing a larger scale than to go in for short supply or 'knife-edge' demand and supply balance. This is because of large volumes of external economies (spread effects), economies of large-scale operations and non-importable nature of many of the overheads. In an underdeveloped economy, public investment in overheads has an advantage since this sector need not be restricted by the limited size of demand. This is because overhead supplies need not follow the commercial principles, particularly during the early phase of development. They have large external economies and, given their longer life, prior supplies are essential to increase their utilisation.

In contrast to public investment in socio-economic overheads, there are investments in "directly productive activities". These are public investments which are competitive and tend to restrict the scope available for private investment. Private enterprise often fears that such public investment may crowd or elbow them out. There are many economic, social, political and strategic reasons for public investment in the production of various consumer, capital and intermediate goods and services. In quite a few developing countries large public investments in directly productive activities have been made.

Public investment decisions may be classified according to Central, State and Local government bodies which make these investments. In India, there are certain constitutional provisions determining which level of government would make investments of what kind. For instance, investments in railways, airlines, defence, etc. are made by the Central government, while those in irrigation projects, health, education, forestry etc. are largely made by the State governments. Investments in municipal services are made by local bodies.

Various types of public investments may be distinguished from each other according to the organisational form of management and control of the assets and public investments. Public investments undertaken by government departments, statutory corporations and government companies are some of the major types. One may also mention public investments taken up by public agencies jointly with private parties or foreign governments.

Certain classifications applicable to any investment are applicable to public investments as well. Any given magnitude of investment may be divided into fixed and working capital investment. While certain public investments enable the use of natural resources, some investments are made for the development of 'human capital', by means of education, training, provision of cultural facilities, etc.

Activity 3

Give examples of various kinds of public investments discussed above.

a) Overhead Facilities

....
....
....
....
....
....
....

b) Directly Productive Activities

....
....
....
....
....
....
....
....

c) Central Public Investment Decisions

....
....
....
....
....

d) State Public Investment Decisions

....

....

....

....

....

e) Zila Parishad Public Investment Decisions

....

....

....

....

....

f) Municipal Public Investment Decisions

....

....

....

....

....

20.5 ORGANISATIONAL FRAMEWORK FOR PUBLIC INVESTMENT DECISIONS

In India the Central and the State governments, along with public enterprises under their control, account for an overwhelming part of public investment. Various local bodies at the level of districts, blocks, cities, towns and villages too make public investment creating a variety of assets. Thanks to constitutional power and responsibilities, particularly the power relating to taxation, borrowing, monetary policy, determination of external economic ties, and socio-economic planning, the Central government is the most important entity making public investment decisions. Through its control over resources and planning, it exercises inflence over the investment decisions of the states and local bodies.

India has adopted a system of medium-term planning. As a result, Five-Year Plans for the development of the economy are formulated. This responsibility has been entrusted to the Planning Commission which was set up in 1951 by a resolution of the Government of India. It has a number of full-time and part-time members and works under the Chairmanship of the Prime Minister of India. There is Deputy Chairman who looks after day-to-day work of the Commission. The Planning Commission decides, for each plan the overall rate of investment and he share of public sector in it. The Planning Commission makes sectoral and nter-state allocation of the investment and outlays determined as a part of the Plan. It also determines, in consultation with the Central Ministry of Finance, the Annual Plan. The Annual Plans of the State governments are also finalised in consultation with the Planning Commission.

The sectoral and inter-state allocation of public investment reflects the development strategy, scheme of priorities and quest for balanced regional development. These decisions are macro, **aggregate** decisions and cannot be subjected to any specific, ready-to-use technique or a set principles of decision making. Since planning is **essentially** a political process involving the entire economy, the scheme of sectoral and inter-state allocation reflects the politics of planning as much as it does various principles and theories of development. However, increasingly, the questions of inter-sectoral, inter-regional and inter-temporal consistency are being asked and tackled. This happens gradually as the planning process improves its data-base, technical expertise, theoretical finesse and computational abilities. Various devices like input-output analysis, macro modelling of the economy, sectoral planning physical and financial balancing, regional and sectoral planning and manpower physical may contribute to the improvement of the content of investment planning. These techniques are important in the domain of scientific planning. Generally the discussions of public investment decision do not concern themselves with these macro, sectoral and regional questions. Conventionally, it remains concerned with micro, project investment decisions.

Many government agencies, besides the Planning Commission, are concerned with project level public investment decisions. Our plans evolve various 'programmes' involving both current outlays and long-term asset-forming investment decisions. These decisions are generally taken by the concerned administrative ministries or departments or specific programme authorities in consultation with the Planning Commission and the Finance Ministry as a part of the exercise of preparing annual budgets.The Annual Budget has a fairly large, significant and growing Capital Budget. In fact, in so far as public investment decisions constitute the core of our medium-term development planning, we may call Indian planning medium-term capital budgeting. In most of the ministries there is an Expenditure Finance Committee to go into these commitments of resources for various programmes.

It is clear that investments in individual projects and programmes are arrived at within the parameters and overall limits set by the plants size and sectoral allocations. However, there exists a set of arrangements for a review of each public investment proposal prior to its sanction and inclusion in the plan budget. This is considered essential for detailed planning of investment on the basis of appropriate criteria of investment planning. It is an attempt to marry consistency planning with optimising planning.

A number of projects for expansion/diversification/modernisation, etc. are sanctioned by various public enterprises, particularly if the cost involved is within certain specified limits. A good deal of public investment is financed by loans, by various public financial institutions (like IDBI, ICICI, SFCS etc.) either directly to the public sector or to private units. These investments are reviewed from bankability and social profitability criteria by the respective financial institutions. Within certain ceilings, investment is sanctioned by the Expenditure Finance Committee of various ministries. Routine and smaller value schemes are considered by the Finance Ministry or Internal Financial Advisor. Nearly two-thirds of the total central investment is sanctioned as a part of routine administrative- financial procedures and without systematic project appraisal.

Two major agencies for appraising public investment projects are: Project Appraisal Division (PAD) of the Planning Commission and the Public Investment Board (PIB). The financial value of the projects Which are to be appraised by these agencies are fixed from time to time.

Activity 4

a) Explain the importance of the Central government in public investment decisions in India.

....

....

....

....

....

....

b) Bring out the role of planning in Indian public investment decisions

....

....

....

....

c) What are the activities of PIB, PAD, and EFC?

....

....

....

....

....

....

....

....

20.6 DETERMINANTS OF PUBLIC INVESTMENT

It is very difficult to understand or analyse the factors which determine the total size or specific form of private investment decisions. According to Lord Keynes, the 'animal spirits' play a part in these domains. Public investment decisions are not based on vagaries of individual motives, perceptions.

emotive structures and attitudes. This, however, is not to underrate the complexity of the forces which influence the size, pattern and speci... forms of public investment.

The nature of the forces determining various facets of public investment depend on the nature of the economic and social system, level of development of productive forces, physical resources, technical possibilities, international context, etc. The nature of the state, its ideology, administrative-managerial capabilities and its role in the socio-economic spheres too perform a role in determining public investment. The fiscal and monetary system too are relevant in this context.

More concretely, in a country like India which has adopted national development planning, public investment is a key variable of the development strategy. The adoption of growth of national income as the basic objective of Indian planning is sought to be attained mainly by the active use of public investment - by its steady increase and concentration in basic and heavy capital goods industries and in socio-economic infrastructure for industry, agriculture and social consumption, particularly for the poor. Public investment is also related, in a mixed capitalist economy like ours, to the tasks which dominant private investing classes do not or cannot perform. As a result, of planning both the overall rate of capital accumulation and the share of public investment have registered impressive increases over the planning era. During the mid-1980s total investment in India is almost one-fourth of the gross domestic product and the public investment component is nearly half of it. The total investment in Central government's non-departmental undertakings stands at about Rs. 70 thousand crores. The coming decades are likely to see a contiuance of these trends, not withstanding a lot of talk about privatisation.

Activity 5

a) List the main factors which determine the role of public investments in India.

....
....
....
....
....
....
....
....

b) How do the determinants of private investment differ from those bearing on public investment?

....
....
....
....
....
....
....

20.7 SOCIAL COST-BENEFIT ANALYSIS

Private investment decisions are generally made by those who have direct personal stakes in the project and its performance. Even then, it cannot be assumed that they can arrive at sound, let alone rational, decisions just on the basis of their 'horse sense'. As seen in the discussion on private investment decisions, various analytical tools and devices and methods of interpreting them are used for working out in a systematic and precise manner the implicationsthe costs, benefits and returns of alternative investment choices . A prudent and forward-looking investor would benefit and avoid coming to grief by proper use of various techniques of appraising (evaluation prior to committing resources to specific uses) alternative investment proposals and basing his/her decision on such an appraisal.

Such considerations favouring conscious, systematic decisions regarding public investment decisions are applicable with added force and for many reasons to public investment decisions. These are made by a representative agency which claims to speak and act on behalf of society. Various social processes (politics, planning, public debate, social mobilisation, perception of class or group interests, etc.) contribute to the determination of the purposes of public investment. They are a complicated set of goals, generally in their plurality ,and with a complicated system . of *inter se* ranking. It is rather difficult to make intuitive investment decisions which are faithful to this complicated hierarchy of social objectives. In any case, since a set of **agents** make long-term investment decisions on behalf of a **principal**, called country or society, the latter must have a means of occasionally examining whether the decisions by the former are true to their stated objectives. In order to enforce such accountability, explicit and announced appraisal of projects is a useful instrument. Since the investment decisions are a part of the process of national planning for development,, recoarse to various techniques of appraising investment projects come handy as a means of creating necessary coordination and mutuality between particular investment decisions and overall national planning. It follows from the above that the principles and techniques of appraising public investment has to be different from those relevant for private investment decisions. .The former have social ends to serve; such ends are many and have no fixed ranking; the people acting on behalf of society may impose or smuggle in their personal or sectarian ends in the exercise. Even though such decisions have to be placed in the broader national context, they too have to compare alternatives in terms of their costs, benefits and net social returns over time. They too have to take account of costs and benefits spread over a long period and, therefore, have to reduce them to a common comparable temporal stream. They too have to face the problems of identifying, quantifying, valuing and comparing the costs and benefits. They too have to arrive at methods of comparing the costs and benefits which have been computed and discounted. Thus there is a technical similarity between the format and some of the mathematical techniques used in the appraisal of public and private investment proposals.

As far as the appraisal of public investment is concerned, **social costs-benefit analysis** is the most useful, extensive and powerful analytical technique. This technique is now at a fairly advanced stage in its theoretical aspects with the emergence of many different manuals (like those of OECD, UNIDO, etc.) for carrying it out. Since it covers the life of the project, such analysis has to reckon with many uncertainties. To some extent sensitivity analysis and probability theory help to deal with such problems. Let us discuss some crucial aspects of social cost-benefit analysis in the following sections.

What is Social Cost-Benefit Analysis?

Social Cost-Benefit Analysis is the evaluation of investment proposals in terms of their estimated net impact on the economy. The estimated impact is evaluated by using parameters reflecting national goals and social objectives. It can be said to be a socio-economic or macro-economic appraisal of an investment proposal. It is concerned with the systematic appraisal of the effects of the project on the economy as a whole. It measures the economic, social and environmental costs and benefits to the society expected to arise from the implementation of the project. It is on the basis of a comparison of such social costs and benefits that a 'scientific' public investment decision can be made, if the comparison covers all the relevant alternatives. It is increasingly considered essential to undertake such an exercise prior to an investment decision committing social resources to a specific purpose. It is an attempt to evaluate the difference to an economy, **ceteris paribus**, as a result of a specific investment it portrays the difference between the economy **with** the project and the economy **without** the project.

How is it different from Financial Analysis?

The financial analysis of a project, particularly for a private investment decision identifies the cash **outflows** and **inflows** and, the **money profit** accruing to the **project operating entity**, whereas social cost-benefit analysis measures **the socio-economic effect of the project** on the entire economy evaluated in terms of the economy-wide objectives. Normally, the private investor, when deciding on whether or not to invest in a project, uses the yardstick of commercial profitability, for he is interested in maximising money profits subject to constraints relevant to his decisions. Hence only outputs and inputs that enter his objective functions are included in the financial investment appraisal and are valued at prevailing **market prices**, i.e., the ruling commodity prices, wage-rates and interest rates are used in computing profitability. This is because it is these market prices which he pays and receives. Given the physical-technical relationship, they determine the results or rewards that can be obtained by him.

The economist engaged in social cost-benefit analysis, is not, in essence, asking a different sort of question from those asked by the decision-maker or owner-controller of a private firm. The social cost-benefit analysis seeks to achieve for the society what financial analysis of a project does for the project operating entity. Social cost-benefit analysis seeks to make the maximum contribution to he objectives of a society. Assuming that the objectives of the government and the society are the same while making a public investment, it becomes imperative to use social-cost benefit analysis for appraising public investment proposals in terms of social profitability.

The point of departure for social cost-benefit analysis is that **commercial profitability** may not be the proper criterion for valuing a project from the point of view of national interests which is what the government is concerned with. This difference is reflected in way one interprets costs and benefits and in their different methods of valuation. Thus a money payment made by a project operating entity, for, say, wages is by definition a financial cost, but it will be an economic cost in social cost-benefit analysis. Commercial profitability by definition is concerned only with private costs and benefits, i.e. cash outflow and cash inflow respectively. *These need not and, except by means of a rare coincidence, would not coincide .with s.ocial costs and social benefits which need to be evaluated with respect to jMtdy.'s welfare and national objectives.*

The rejection of commercial profitability as a criterion for public investment does not imply that the concept of profitability itself is rejected. In fact, the use of social cost-benefit analysis in place of commercial profitability analysis, involves an extention of the concept of commercial profitability to social profitability. Social profits of an investment proposal during any period of time may be defined as the differences between the social costs and social benefits accruing/resulting from the proposal during that period of time. Social cost-benefit analysis is not concerned with the loss or benefit to any individual or some small groups of society as such, but to the **society as a whole.** If we consider the construction of a bridge on a river, there is a definite financial loss to the State exchequer and to a particular class of able to utilise the facilities and those who used to pay for crossing the river will be able to increase their income and diversify consumption. There would be saving of time etc. One has to weigh these gains against the costs (or losses) to the State exchequer and the loss of income to the ferry-owners. It involves intergroup comparisons of losses or gains. The cost of the bridge is not only its construction and maintenance costs but also the social cost of employment lost by the ferry- owners.

Social cost-benefit analysis is a useful device for the choice of projects, taking.into account the national objectives. Financial/cost-analysis is also an important device for purposes of **project control and management.** Projects are separate units having the same objectives as the objective of national public investment, e.g., the greatest possible increase in the standard of living of the people over a well-defined time-horizon. The units are separated for the purpose of operational efficiency and proper control. It would be out of the reach of one unit to take account of the total public investment and run it efficiently, given its diverse nature and large magnitude. Thus it becomes essential to separate public production units into smaller, manageable units and projects with their techno-economic and physical unity. Financial analysis of the project becomes an essential function as a part of project management for the purpose of control, accountability and efficient accomplishment of the project. Thus, financial analysis (various forms of which were discussed in an earlier unit), though it does not serve the role of project selection from the point of national objectives and social profitability, becomes essential for the efficient management of the project after it is approved on the basis of social profitability. It would be impractical and/or unnecessary to go into the social cost-benefit analysis of every decision in the management of the project. Financial analysis may also be needed from a banker's or financier's (even though it may be a public sector unit) point of view in order to see if the project will have the capacity to repay itself or will need to be subsidised. This is essential for determining the bankability of the project.

Why should Social Cost-Benefit Analysis be used?

A proper investment decision is of utmost importance because it sets a series of economic and social factors into operation towards the achievement of socio-economic goals, and the objectives of society, particularly for India which has been striving for development through planning. A mistake of misallocation at the point of investment decision acts in the opposite direction. Social cost-benefit analysis is a tool for making investment decisions best suited to the development strategy and objectives so that the scarce resources contribute most towards the national objectives.

Government's objective while making a public investment decision is the greatest possible attainment of the objectives of society over a chosen period of time e.g., increase in the standard of living of the people as a whole, employment, equity, self-reliance, regional balance, etc. Public investment becomes essential because of a number of reasons. To begin with, a case for public production arises in connection with certain goods and services whose inherent character is such that they cannot be left to the private suppliers. Such goods are known as public goods e.g., administration, police, military, etc. Their benefits are consumed collectively by all the members of the society, and each member's consumption does not reduce the consumption available to the rest of society. For example, the police protection benefits are enjoyed by each member of the society; the use of police service by one person does not reduce the services available to others (though for better and more efficient service an increase in the investment in policing may be required). This also implies that the marginal cost of production is zero or near zero. Since marginal cost has to equal price for efficient resource allocation it ultimately implies that it cannot be left in the hands of private control.

The second case of production by public control, as was pointed out earlier, arises in the case of **quasi-public goods**, such as postal services, electricity, city transport, roads, bridges, public utilities, etc. Such goods, by their very nature, are monopolies, where the supposedly regulating forces of competition are absent, and so cannot be left in the hands of private entrepreneurs without inviting adverse social effects. Furthermore, the heavy investment requirement would make such commodities beyond the reach of the average investor. The monopolistic pricing may deprive many persons of the use of these goods. Quasi-public goods being essentials, the consequences of such deprivation would be serious and far reaching. This would mean a decline in the level of social welfare, which is contrary to the principles of a welfare state. To keep the consumption of such commodities within the reach of the low income bracket consumers, it becomes imperative to keep the prices of such commodities low (the marginal cost of production of such commodities is also low). These factors make it essential to bring the investment for the production of such commodities under the sphere of the State.

There arises a third case of public investment in the field of infrastructure. Such investments like roads, railways, etc., and even certain quasi-public goods like electricity, are the foundations of a nation's economic development. In such cases it becomes essential to make public investment, because private investment is lacking due to heavy investments required in such fields and the low rate of return accompanied by high risk and a long gestation period. Even if private investments are allowed in such fields, the price of such goods and services will be so high as to create an adverse atmosphere for overall economic development. The secondary effect of the high price will raise the cost and price of other goods and services also. If an industry had to pay fully for all the overhead facilities like road, water, etc., i.e., at the full cost of production, the cost of output of the industry in turn will tend to be rather high. The other field where public investment becomes essential are the key industries like steel, banking, etc., because such industries are crucial to the economic development. If such fields are left to private investment, they will be used for the maximisation of private profit instead of social welfare.

In such cases, public investment which is justified in terms of social cost-benefit analysis, will contribute to development in terms of social profitability.

In any country, as can be seen in the case of India, there is a large public sector with a great deal of diversity, ranging from infrastructure and capital goods to consumer goods, and this keeps on increasing with new capital budgets and plans. Given the size and the diversity of public sector, it cannot be managed and operated as one single unit. It is essential to divide it into smaller corporations and projects, the basic purpose of public sector remaining the same the greatest possible contribution towards the achievement of national objectives and social welfare.

Given the diversity and subdivisions of public investment, various sectors and projects appear as competing alternatives, giving rise to the problem of choice among the feasible projects in order to achieve the maximum possible social welfare and contribution to national objectives. Social cost-benefit analysis is a tool for making the best choice from the given alternatives, keeping in view the national objectives and social profitability. The choice of one project rather than another must be viewed in the context of its total national impact, and this total impact has to be evaluated in terms of a consistent and appropriate set of objectives. Since the social cost-benefit analysis is concerned with the economy as a whole and with the welfare of a society *its project choice is subject to the general objectives of national planning and policy.*

Steps in Social Cost-Benefit Analysis

In order to appraise a public investment project in terms of social cost - benefit analysis, the detailed project report is prepared. Based on this report the following steps are taken:

- *Identifying costs and benefits*

The project cost in terms of financial outlays is taken as its financial cost. It has two components: fixed investment and working capital. The first covers the payment made for fixed capital goods: machinery, plant, building, equipment, etc. The other covers the running, operational and maintenance costs over the life of the project. The social costs are the costs of the foregone uses of the physical and human resources used in the construction, operation and maintenance of the project. One has also to identify the indirect costs (external diseconomies) and intangible costs of the project. The sum total of all these costs represents the sacrifice made by society for setting up the project in question. This is its total social cost. Both the fixed and working costs have to be annualised. The financial costs are not taken into account as they do not use up real, social resources. Among the direct social costs we include the social opportunity cost of land, labour and capital equipment. From among the current cost elements interest charges and depreciation allowance are excluded from social costs as they represent no real costs. The current costs too can be divided into physical and human resources costs. The indirect costs would flow from the economy-wide adverse socio-economic, ecological and agro-economic or industrial impact of the project. These costs are often intangible and many difficulties are involved in their quantification and valuation. Often they are taken care of by organisational, administrative and policy measures. The identification of such costs also enables one to form an overall judgment in decision-making about the project.

After identification of various elements constituting social cost of the project. For this purpose one may use market prices. Various physical measures can be used. For example, labour used during the project construction period can be added to the current labour inputs over the life of the project.

All these cost elements have to be added up in order to arrive at the total social cost. It means all the cost items have to be expressed in terms of a common numeraire. Market prices do not serve this purpose as they reflect the balance of existing demand and supply mediated by the ruling or dominant market structure. Such demand, supply and market structure reflect individual preferences, decisions and control over resources and, generally fail to reflect social objectives, preferences and valuation based on common and future demand. Owing to monopolitic structure of the market, the market prices distort even individual valuations and assign disproportionate weight to the decisions and preferences of some over the others. A society at early stages of its growth may not produce certain goods at all or in small quantities. A new investment project may totally alter the prevailing market valuation and hence it would be inappropriate to evaluate it in terms of pre-project prices. Market prices may include various transfer payments like taxes and subsidies which are based on considerations altogether different from the relative scarcity or the present or projected demand-supply balance in the economy. Use of such market prices would not reflect social valuation of the output/input under reference. Hence market prices are a misleading and irrelevant entity for use as numeraire either for aggregating the social costs or the benefits.

From the social point of view one has to take into account the social opportunity cost of using a certain input for the public investment project. This is the **accounting or shadow or social** price of a commodity. Since social costs and benefits have to be valued in terms of these prices, one may assign weights reflecting social objectives.

For example, the shadow price of labour used in a project would be not the money wage paid to a skilled or unskilled worker, but the sacrifice of output in the use from which labour is withdrawn or prevented from being allocated. The device of valuation in terms of shadow prices is an important means of strengthening a micro decision of the macro social perspective.

Social evaluation of various cost/benefit elements in terms of shadow prices does not enable one to obtain a correct, social picture of the implications of alternative project decisions unless account is also taken of the fact that various costs are incurred and benefits are obtained at various points of time. The cost and benefit values change according to their distribution over time, or the time period to which they refer. In fact, it would be incorrect to add up the costs/benefits of the year say 1980 and those of the year say 1988. After all, it makes a lot of difference as to when how near or far from the present a cost/benefit has to be incurred/obtained. Human beings have limited telescopic faculty. The risk and uncertainty increase as one thinks of increasingly distant future. Human beings generally consider a bird in hand better than two in the bush. Similar factors affect social valuation, especially when a society is underdeveloped at the moment and hopes to enjoy a higher standard of well-being in the future. This factor may be called the rate at which society prefers the present over the future. If is called the social rate of discount. The stream of social costs/ benefits calculated in terms of shadow prices have to be discounted using a **social rate of discount** in order to convert the quantities involved into a common temporal perspective.

Just as with social costs, social benefits of a project too have to be computed on the basis of the following sequential steps: identification of direct or indirect (externalities) benefits, their quantification or a listing of the non-quantifiable (intangible) benefits, social valuation in terms of shadow prices (if the commodities in question may involve external economic relations, shadow pricing of foreign exchange) and conversion of these quantities into a common time perspective by means of discounting at the social rate of discount. It may be mentioned that just as market prices of various inputs/outputs cannot serve the purpose of accounting prices, for a similar set of reasons, the structure of market interest rates cannot serve as the social rate of discount. In the determination of the social rate of discount, social values, objectives, attitude towards risk and uncertainty and political factors play their roles. The time-horizon adopted in the process of plan formulation and the concern for the interests of the future generations exert an important influence on the choice of the social rate of discount.

- *Comparing Cost and Benefit: Decision Rules*

The prime objective of investment project appraisal is to compare the costs and the benefits of different projects with a view to determine the project which gives a greater return for the total amount of inputs (goods, services and work) to be put into the project. We have seen why social discounting of costs and benefits is an essential part of this exercise. Important methods of computing discounted costs and benefits are: Benefit-cost ratio; Net present value and Internal rate of return.

As you will notice, these steps are the same which we discussed while studying private investment decisions. However, two factors have to be kept in mind while making an investment decision.

a) There is no one best technique for estimating the social worth of a project.

b) These criteria of investment are only tools of decision-making providing synoptic, systematic guidance. There are many non-quantitative and non-economic, non-tangible, criteria for making project choices. Hence no analytical formula or technique is a substitute for judgments involved in decision-making. These formula cannot take away the romance of decision-making.

Four major aspects of a project have to be taken into account prior to applying the investment criteria or the decision-guidance rules: technology, scale or size, location, timing of the project in terms of the current economic direction and the budgetary position of the government. All these form the context of national planning in a country like India.

$$\text{Benefit-cost Ratio} = \frac{\text{Present worth of Benefits}}{\text{Present worth of Costs}}$$

Benefit-cost Ratio: The Benefit-cost ratio is exclusively used for economic analysis of projects, or we may say, is related to the concept of social cost-benefit analysis only.

The practical application of benefit-cost ratio becomes very limited because of the interest character of the ratios. If we take two projects A and B and compare the Benefit-cost ratio of the two, it would clearly bring out the limitation of benefit-cost ratio.

Project	Total discounted social cost (Rs.in lakhs)	Total discounted social benfit (Rs.in lakhs)	Benefit-cost ratio	Net present worth (Rs. in lakh)
A	100	150	1.5	50
B	1,000	1,300	1.3	300

In this example we find that the B-C ratio of project A is higher than that of project B and the Net present value of project B is higher than that of project A. Hence the total benefit from project B is higher than the total benefit from project A but the benefit per unit of cost is higher in the case of A as compared to B. If the total investible resources are Rs.10 crores then we do not know the B-E ratio of other projects built at the cost of remaining 9 crores and hence we cannot make a comparison in terms of B-C ratio between projects A and B. Thus the use of B-E ratio is limited to seeing the return per unit of social cost. It cannot be used for the purpose of selecting from projects A, B,... .... Z.

The only case where Benefit-cost ratio can be used for ranking is when all the projects from A to Z have the same social cost.

Net present value: The most straightforward discounted measure of a project worth which aids in the choice is the Net present value. Net Present Value of a project is simply the value of the surplus generated by the project in 'to-days' term. It is given by the discounted values of the benefits occurring from the project minus the discounted values of the costs incurred on the project. The formula for its calculation has already discussed while studying private investment decisions.

Internal rate of return: This measure, used by World Bank and other international financial agencies, determines that rate of discount at which the present worth of the project is Zero. **Internal or Economic Return** is the same as Internal rate of return. This term is used to distinguish between Financial calculation and Economic appraisals of the project.

There is a direct relationship between the internal rate of return and the net present value of a project. Internal rate of return is that rate of discount at which the net present value of the project comes to zero.

Net present value of a project as a tool for making an investment decision seems to have some advantages over the internal rate of return from the following angles:

a) It takes into account a specific rate of discount (social rate of discount) which the internal rate of return fails to cover in its values. It is imperative to incorporate society's time preference for the flow of costs and benefits emerging at different points of time.

b) Present value rule premits the possibility of society's time preference changing at different points of time over the life of the project. If we take internal rate of return, the rate of return is given as uniform and no changes can be allowed for.

c) When choice has to be made between two projects which are mutually exclusive (i.e. the implementation of one rules out the other) the internal rate of return becomes very misleading.

However, internal rate of return has certain advantages over the net present value. It shows a clear picture of the maximum rate at which funds can be borrowed. What is of concern in social cost-benefit analysis is the maximum social benefit that can be obtained from the resources set aside for investment. Since the real rate of society's time preference is very difficult to calculate, the internal rate of return gains an advantage over the net present worth of the project because it also shares the controversy concerning the social rate of discount.

Activity 6
On the basis of reading the above, please attempt the following questions:

a) Define and illustrate, with specific examples, the concepts of:

- Social Cost:

....

....

- Social Benefit:

....

....

- Social Profitability:

....

....

- Social Rate of Discount:

....

....

- Externalities:

....

....

- Shadow Prices:

....

....

b) How does social cost-benefit analysis differ from financial analysis?

....

....

....

....

c) Why are market prices not found suitable for use in social cost-benefit analysis?

....

....

....

....

d) Explain the following in the context of public investment:

i) Benefit/Cost-Ratio

....

....

....

ii) Net Present Worth

....

....

....

iii) Internal Rate of Return

....

....

....

20.8 PUBLIC INVESTMENT APPRAISAL TECHNIQUES: USES AND ABUSES

The social cost-benefit analysis is an extension of the NPV technique of capital budgeting. An act of private investment yields a series of annuities over the future years, which when aggregated and discounted, gives us its present value. In the same way, a public investment project (say, a multipurpose project like construction of a dam) yields a series of social returns (like flood control, electricity generation, irrigation of land, creation of employment, increase in output and income etc.) such returns can also be aggregated and then discounted with reference to a social rate of discount so as to calculate the present value of social returns over the future years. Exactly in the same manner, the present value of investment costs social costs can also be calculated. And then either the **Benefit-cost deviation** (net social returns) or the **Benefit-cost ratio** (social net present value index) can be calculated to study the economic feasibility of a project.

An example

Let us take an example to illustrate the use of Benefit-cost analysis in project evaluation. Suppose, there is a public project which involves various categories of investment. The available data are represented as follows:

Capital Investment	Original Expenditure (Rs.)	Expected life (years)	Interest rate of 5%
Category A	18,000	30	Log 1.05 = 0.0212
Category B	12,000	50	Antilog of 0.6360 = 4.325
Category C	23,000	60	Antilog of 1.0600 = 11.48
Category D	20,000	120	Antilogof 1.2720 = 18.7056
Annual Operation cost 1,000			Antilogof2.5440 = 349.8995

Benefit:

Value of product before the Project = Rs. 33,000

Value of product after the Project = Rs. 52,000

Associated cost = Rs. 1,000

Costs: You may note that the project has several categories of investment cost. The estimated amount of each category is to be converted in annual cost by the formula

$$A = P\left[\frac{r+(1+r)^{n}}{(1+r)^{n-1}}\right]$$

Where A = Annual cost; P = Original investment cost; r = rate of interest; n = Expected life.

Category A

$$A = 18,000\left[\frac{\frac{5}{100}\left(1\ \frac{5}{100}\right)^{30}}{(1-5/100)^{30-1}}\right]$$

$$= 18,000\left[\frac{.05 \times (1.05)^{30}}{(1.05)^{30-1}}\right]$$

$$= 18,000\left[\frac{.05 \times 4.325}{4.325-1}\right]$$

Calculation

Letx = $(1.05)^{30}$

Log x = 30 Log 1.05

= 30 X 0.0212

= 0.6360

∴ x = Antilog of 0.6360

= 4.325

$$= 18{,}000\left[\frac{0.21625}{3.325}\right]$$

$= 1170.676$

Activity 7

Make similar kind of calculation for the remaining categories of cost:

Fill in the missing steps.

Category B

$$A = 12.000\left[\frac{.05(1.05)^{50}}{= (1.05)^{50-1}}\right] \qquad \text{Let } x = (1.05)50$$

$= 657.252$

Logx =

=

$\therefore$ X =

=11.48

Category C

A =

=

=

=

=

= 1214.951

Let x =

Log x =

=

=

x = Antilog of

= 18.7056

Category D

A =

=

=

=

=

Let x =

$\therefore$ x = Antilogof

The total annual cost of the oroiect then is equal to Category

A : Rs. 1170.676

B : Rs. 657.252

C : Rs. 1214.951

D : Rs. 1002.866

Annual operation cost : Rs. 1000.000

Total Annual Cost = Rs.5045.75

Benefits:

You may calculate the net present value of the prqj'ect as follows:

Value after the project		= Rs. 52,000
Value before the project		= Rs. 13,000
		19,000
	Less associated cost =	Rs, 1,000
	Net Annual Benefit =	Rs. 18.000

Cost-benefit Calculation

Now you may compute

i) Benefit-Cosi deviation Rs.18,000-Rs.5045.745 = 12954.255

ii) Benefit-Cost ratio Rs.18,000/Rs.5045.745 = 3.5

You need to repeat this process for other comparable projects, and then towards the choice of a project, you need to follow the decision principle: Larger the benefit-cost deviation or higher the benefit-cost ratio, better the project.

Activity 8

a) Compute both benefit-cost deviation as well as benefit-cost ratio of an irrigation project. Use the data given below.

b) Suppose you were to choose between two projects, one in the **example** above and the other in the **activity** here, which one would you choose and why?

Capital investment	Expected life (years)	Original investment	Rate of Interest = 4%	Other data
Category				
A	50	1,000	log 1.04-0.017	Anmial operation cost = Rs. 1,000
B	30	9,000	Antilog 0.850 "779 Antilog 0.510-3.236	Associated cost = Rs500
			Value of crops	
C	100	18,000	Antilog 0.425-2.661	Aftr irrigation Rs 13,000
D	25	2,000	Antilog 1.700-50.12B	efore Irrigation Rs.6,000

The cost-benefit analysis of the kind we have illustrated here raises a series of questions: Which costs and which benefits are to be included? How can these costs and benefits be measured? How can these be quantified in money terms? At what interest rate are these to be discounted? Are there no constraints?

We often talk about both **direct** (primary) **benefits** and **indirect** (secondary) **benefits.** For an irrigation project, water supply and electricity generation are direct benefits, whereas employment income-output generation are indirect benefits. In the same way, there may be both **direct costs** and **indirect costs.** Sometimes these costs and benefits may not be fully identified: even if identified, they may not be quantified; even if quantified, they may not be measurable in money terms; even if money-value is calculable, the actual value may not be reflected. The **market prices** do not reflect time or scarcity value due to (a) market imperfection and (b) presence of externalities. In such cases, the **shadow prices** have to be used; but there are difficulties in computing such 'social prices' or 'planning prices'. Also, the device of a **social rate of discount** presents difficulties; we may not have relevant data to compute such a social discount rate. The point remains that the calculation of social benefits and social costs involves a good deal of subjectivity and value judgments. This is where the cost-benefit analysis may be grossly misused. For example, an investment decision may be purely a **political** one; yet it may be forced to get justified through the useof economic techniques. It was observed: "**the economists** have developed the technique of cost-benefit analysis, but the politicians often use (rather misuse) it to suit their own convenience". In our country, a set of socio-political factors often compel the Government to take a number of investment decisions first and then ask the Planning Commission to justify those decisions on the basis of cost-benefit analysis. Political decision first, economic justification later. Ideally the process should be reversed. Investment decisions should be dictated by economic rationality. Unfortunately, most of the public investment decisions which cannot stand the criterion of "commercial profitability" are often justified on the basis of "social desirability" i.e. the social cost-benefit analysis.

Activity 9

Write a lucid note on the limitations of Social Cost-Benefit Analysis. Use real world examples to illustrate your arguments.

....

....

....

....

20.9 SUMMARY

In this unit, we have drawn a line of contrast between private investment and public investment on several counts. We havs also justified the analysis of public investment decisions in a mixed economy like that of ours. We have extended the standard techniques of capital budgeting such as NPV and IRR to evaluate public investment protects. In the process, we have comparable concepts to take care of: private and social returns, private and social costs, interest rate and the social rate of discount, market prices and shadow prices etc. Finally, we haw indicated the use as well as limitations of the social Cost-Benefit Analysis for appraisal of public investment project. One of the major limitations from which the present as well as the preceding unit suffers is that we have not considered the role of risk-and-uncertainty in all context of investment decisions. We felt that the element of risk-and-uncertalirty is involved in all decisions including the investment decisions. We would now like to proceed towards an economic analysis of risk and uncertainty factors in general and specifically, with reference to investment decisions.

20.10 ADDITIONAL READINGS

Adhikary, M. (1987) *Managerial Economics*, Khosia Educational Publishers: Delhi Ch. VII

Little, I.M.D. & J.A. Mirrlee, (1969) *Social Benefit Cost Analysis*, Parms OECD.

Meier, G.M. 4th ed. (1984) **Leading Issues in Economic Development**, Oxford University Press: London. Ch.X.

20.11 SELF-ASSESSMENT TEST

1. Review your understanding of the following terms. Give examples:
   a) Net social benefits
   b) Social cost: direct and indirect
   c) Social benefits: primary and secondary
   d) Social rate of discount
   e) Shadow prices

2. How do **public** investment decisions differ from private investment decisions?

3. State the steps involved in social cost-benefit analysis of a project. How is such an economic analysis different from the accounting methods of capital budgeting?

4. Write an essay on the strengths and weaknesses of social cost-benefit analysis.

. List separately the possible **social benefits** and **social costs** involved in thefollowing projects:

   a) Drought Relief Programme 1987
   b) ASIAD1982
   c) Clean Ganga Project
   d) HBJ Pipeline Construction
   e) Calcutta Metro-Rail

UNIT 21 THE ECONOMICS OF RISK AND UNCERTAINTY

Structure

21.0 INTRODUCTION

Investment decisions involve a great amount of risk and uncertainty. Risk is the difference between expected and actual returns on investment. In fact, for any decision, we can talk of this deviation between expected and actual outcome/pay-off/return. Thus an analysis of risk is crucial for decision-making. You may recall that there are two types of risk - calculable and non-calculable (uncertainty). In a world of certainty, the risk element is zero, there are neither market risks nor non-market risks. In reality, such certainty hardly prevails. Therefore, an analysis of risks and uncertainties is required before any decision is taken. In what follows, such an analysis is attempted. The analysis may be illustrated mostly with reference to investment decisions, but other decision areas also provide scope for applying risk analysis. It is not possible for a decision-maker to eliminate risk and uncertainty altogether; but he must try to reduce it as far as possible, business forecasting and forward planning are meant to reduce the element of risk and uncertainty involved in decision situations. Before you can think of reducing risks, you should be able to measure the degree of risks, there are statistical techniques to do the same. These techniques are not necessarily based on only objectivity, many times subjectivity is equally important. For example, the **attitude** towards risk is a very important parameter in analysing the decision environment characterised by risk and uncertainty. Also, many times, the objective techniques alone are not enough to evaluate risks; subjective value judgments also influence; appeal etc. are equally important for risk evaluation. The fact remains that risk and uncertainty are typical of any real world business situation. An economic analysis of risk and uncertainty is, therefore, required to aid the decision-makers. In conventional managerial economics, this analysis is assumed any way; most of the economic decision models are based on the assumption of certainty, typical of market perfection. In modern managerial economics, the risk-analysis is an important ingredient, because the manager often functions in the world of market imperfection accentuated by the lack of information; and this is what involves risk and uncertainty.

21.1 OBJECTIVES

On comoleting this unit. vou should be able to:

- **recall** the concepts of risk and uncertainty;
- **realise** the contribution of probability theory in risk analysis;
- **measure** the degree of risk and uncertainty involved in any decision;
- **analyse** categorically decisions under risk and uncerta'nty; and
- **appreciate** the role of both quantitative and qualitative techniques of estimating and evaluating risk and uncertainty.

21.2 CONCEPTS OF CERTAINTY, RISK AND UNCERTAINTY

The distinction among **certainty**, risk and **uncertainty** reflects differences in the degree of knowledge with regard to the decision environment. Investment climate is a very specific form of decision environment. If we conceive of an investor's state of knowledge as a continuum, then certainty (complete knowledge) would occupy one end and uncertainty (complete lack of knowledge) would occupy the other. You may recall that we have already introduced these concepts earlier. These concepts have special appeal in the context of investment decision, though these are applicable in any other decision environments such as production, marketing etc. In this section, we would like you to revise these concepts with special reference to investment decision environment because investment decision is the be-all and end-all of the decision-maker.

Certainty
It is a state of knowledge in which the investor knows in advance the specific outcome of alternative investment projects. The investor has the perfect knowledge of the investment environment such that he is definite about the size, regularity and periodicity of the flow of returns, there is no risk involved in this case. How realistic is this concept? At first sight, this state may appear theoretical and impractical, actually this is not so. There are many short-run situations in which the investor/manager has complete knowledge. For example, when we put our money in Fixed Deposit with a nationalised bank or in National Saving Certificates with our Post Offices, we know for certain, the exact return we are going to get on maturity. The annuity as well as maturity time are completely known.

It is, however, true that in reality, few things only remain certain for long. The outcome of a long-range investment, for example, is really impossible to predict, because in the long-run so many things may change: political climate, technological advancement, market competition, economic environment etc. Currently, these things do not change, but eventually they may change. Thus currently there is no risk, but eventually there may be. Risks may not be visible today, but you may imagine risk for to-morrow. In fact, most strategic decisions (including investment decisions) are made under conditions of less than perfect knowledge, we classify such conditions as either risk or uncertainty.

Risk

It is defined as a state of knowledge in which each alternative leads to one of a set of specific outcomes, each outcome occurring with a probability that is known to the decision-maker. Investment risk may be regarded, therefore, as the quantitative measurement of a return, positive or negative, such that the probability of the return (outcome or pay-off) can be calculated. Two approaches can be used to measure this probability: one is a **priori**, by deduction; the other is a **posteriori**, a statistical analysis of empirical data. In the **a priori** method, the decision-maker is able to determine the probability of an outcome without experimentation sampling or an analysis of fast experience. For example, we know that a coin has only two-sides. Because of this, a tossed coin must come up with either a head or a tail. The **posteriori** method, by contrast, assumes that past performances were typical and will continue in future, in case, we have to observe, collect, process and analyse statistical data. Risk is the actuarial probability of an event. Risk can be measured by way of standard deviation between the expected return and the actual return on investment. For this, we have to study "distribution" of data in statistical sense. There is a distinction between **frequency distribution** and **probability distribution.** The former is a tabulation of how many times certain events have occurred in the past, while the latter is a tabulation of the percentage of time that they are likely to occur in the future. For an investor, the event may be defined in terms of size and timing of return flows.

Uncertainty

Risks, if calculated, can be insured. Thus, insurance companies can predict with a high degree of accuracy the probabilities of deaths, accidents, and fire losses. These probabilities enable them to make decisions about premium levels and rates. But they cannot establish the probability that a particular individual will die or that a particular house will burn or a particular car will meet with an accident. When actual probability is not calculable and therefore, the element of risk is not measurable, we have the environment of uncertainty. Unlike risk, therefore, uncertainty is a subjective phenomenon; no two individuals who view the same event will necessarily formulate the same quantitative opinion because of a lack of knowledge.

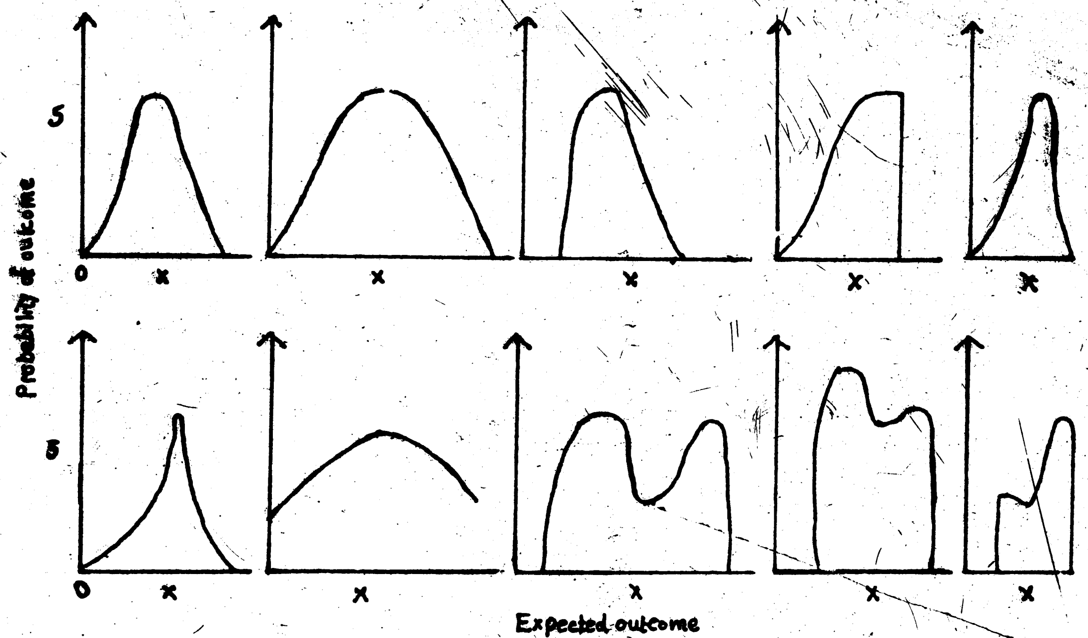
Since expectations are subjective, there will be degrees of uncertainty on the part of decision-makers. Two investors, reading the same investment opportunity, may have different expectations about the flow of net return. There are two sources of uncertainty.

a) Complete ignorance
b) Partial ignorance

In case of **complete ignorance**, the decision-maker makes no assumptions about the probabilities of the various states of nature. The investor may, for example, use any one of a number of rational criteria for making investment decisions; he may adopt a pessimistic or conservative outlook. In dealing with such situations, the investor may use, for taking investment decisions, different criteria such as 'maximum criterion'. We will discuss these criteria subsequently.

In case of partial ignorance, the decision-maker assumes the ability to assign a priori probability ~~distribution~~ ~~to~~ ~~possible~~ outcomes. With the development of "subjective" or "personal" probability theories, the element of risk can be guessed-and-estimated i.e. guesstimated, though it cannot be objectively estimated. Thus a decision problem under uncertainty can be developed nto a problem under risk.

Since a great many decisions are often taken under complete ignorance or ncomplete knowledge and imperfect information, a brief intuitive sketch of the notion of **subjective probability distributions** should be of interest.



Activity 1

a) Interpret the basic theme underlying the above set of diagrams.

....

....

....

....

....

....

b) What is risk ? How would you measure it ?

....

....

....

....

c) What is uncertainty? How is it different from risk? Distinguish between differen "degrees" of risk and uncertainty.

....

....

....

....

....

d) Review different concepts of probability?

....

....

....

....

....

21.3 THE PAY- OF MATRIX

The concept of "subjective probability distributions" was developed by LJ. Savage, and was further extended and put into operational form by authors like R, Schlaifer and H. Raiffa. These developments together have come to be known as "Decision Theory". Pay-off matrix is one of the basic concepts in Decision Theory. Originally this concept was developed to support the theory of gamesa an approach to oligop(You may recall this from your Block 4. However, the pay-off matrix concept has since been adopted to general decision-making under conditions of risk and uncertainty. The pay-off matrix or a decision matrix, as it is sometimes called, helps the decision-maker to conceptualise and formalise the decision process into

- statement of objectives
- selection of pay-offs
- evaluation of alternative pay-offs.
- selection of alternative strategies

An example of a pay-off matrix is shown below :

Table I

Alternative Strategies	State of Nature			
	N_1	N_2	N_3	N_4
S_1	6	6	6	4
S_2	25	7	7	-15
S_3	20	20	7	-1
S_4	19	16	9	-2
S_5	20	15	15	-3

In this matrix, the decision-maker conceives of four possible states of nature, i.e environment: $N_1 \ldots N_4$; the four **states of nature** could be economic environment associated with boo , stability, recession, or depression. Similarly, there are five alternative strategies; the **strategies** may involve different size of investment or selection of different investment projects. The **pay-offs** represent our ' investor's best estimate of outcome (returns) of each combination of strategy and state of nature. Three questions arise about the pay-offs

1 What do they represent?
2 How do we find them?
3 How do we evaluate them?

The first question has been answered above. The pay-offs represent returns on investment, if our reference is, to the investor. The pay-offs may refer to sales, if our reference is to marketing decision, and so on.

The second question also can be answered, provided we know the nature of decision. For an investment decision, , you may seek information from the finance manager. Like-wise for marketing decision, consult the sales manager, and so on. The source of information must be dependable. As for the question of evaluation the matrix makes it clear that the best strategy depends upon which state of nature occurs. Unfortunately, the decision maker cannot predict what state of nature will occur. Nevertheless, he must select a strategy. Under conditions of **certainty**, the decision process involves the optimisation of some known objectives (such as profits), given certain known constraints (such as costs). Under conditions of **risk**, the decision-maker may obtain objective posterior probability estimates. But under conditions of **uncertainty**, probability distributions and constraints are not known. This is where subjective probability distribution becomes relevant. In actual practice, most decision-makers follow the **Buyesian postulate**, which says that in the absence of meaningful knowledge of probabilities, equal probability should be assigned to each outcome (pay-off). Whether the probabilities are objective or subjective, they must add up to 1.0 since one of the outcomes is certain to occur. In symbols,

$P_1 + P_2 + \ldots\ldots\ldots\ldots\ldots\ P_n = 1.0$

Since probabilities add up to 1.0, they may be used as weights to calculate the mean (average) pay-offs, known as the **expected value.** In symbols;

$$E(X) = P_1X_1 + P_2X_2 + \ldots\ldots P_n X_n = \sum^{n} P_i X_i$$

Where

E(X) = The expected value or weighted mean pay-off.

P_i = The probability of the ith pay-off

X_i = The ith pay-off.

21.4 DECISION-MAKING UNDER RISK

To make a decision, the expected value is computed for each strategy and the one with the highest expected value is chosen. Continuing with the data presented in Table I, let us suppose that the decision-maker assumes a 20% chance of boom (N_1), a 65% chance of stability (N_2), a 10% chance of recession (N_3) and a 5% chance of depression (N_4). Note that these probabilities add up to 100% i.e. 1.00. The -expected value of each strategy is now calculated as follows:

S_1	2(6)	+ .65(6)	+ .10(6)	+ .05 (4)	=	5.90
S_2	2(25)	+ .65(7)	+ .10(7)	+ .05 (-15)	=	9.50
S_3	2(20)	+ .65(20)	+ .10(7)	+ .05 (-1)	=	17.65

S_4 2(19) + .65(16) + .10(9) + .05 (-2) = 15.0
S_5 2(20) + .65(15) + .10(15) + .05 (-3) = 15.10

These results alongwith our earlier data can **now** be reproduced in the following table.

Table 2

Alternatuve Strategies	State of Nature N_1	N_2	N_3	N_4	Expected Value E (S_i)
S_1	6	6	6	4	5.90
S_2	25	7	7	-15	9.50
S_3	20	20	7	-1	17.65
S_4	19	16	9	-2	15.00
S_5	20	15	15	-3	15.10

You may note that S_3 is the best strategy because it has got the largest expected value associated with it. The problem of choosing a strategy may arise, where two strategies have the same expected values. In that case, you cannot use just expected value criterion. You have to use some other yardstick. The yardstick is **degree of risk** which is considered a secondary or auxiliary measure of expected value.

Degree of risk

Intuitively, the degree of risk is indicated by the degree to which the actual outcome or pay-off of a strategy or investment project deviates from its expected (mean) value. This is indicated by the spread or variation in the probability distribution of possible outcomes for each proposal. Thus an accurate measurement of risk is provided by **standard deviation**. However, the standard deviation cannot be used it compare risk when expected values are different. Such comparison require the relative standard deviation, commonly called the **co-efficient of variation**, which is simply the ratio of standard deviation to mean, expressed as a percentage.

Activity 2

a) Review the statistical measures of
   i) Standard Deviation, δ =
   ii) Co-efficient of Variation ψ =

   (Write out the formula, consult a book on statistics)

b) Expected value of return being the same, lower the risk, better the project. True or False? Explain.

....

....

....

....

c) Tighter the probability distribution of possible outcomes, the smaller the risk of a given decision. True or False? Explain.

....

....

....

d) Smaller the standard deviation, lower the riskiness of the alternative. True or False? Explain.

....

....

....

e) Co-efficient of variation is better than standard deviation as a measure of degree of risk. True or False? Explain.

....

....

....

f) Consider the following data:

State of the economy	Probability of the state	Pay-ofT (Returns)	Expected Value
Project A			
Recession	0.2	Rs.400	
Stability	0.6	Rs.500	
Boom	0.2	Rs 600	
	1.0		
Project B			
Recession	0.2	0	
Stability	0.6	Rs. 500	
Boom	02	Rs. 1000	
	1.0		

i) Complete the table above by computing expected values.

ii) State the relationship between the state of the economy and project returns. Do it for both the projects. Which project is more risky?

....

....

....

iii) The project A should be preferred to project B. True or False? Explain

....

....

....

....

....

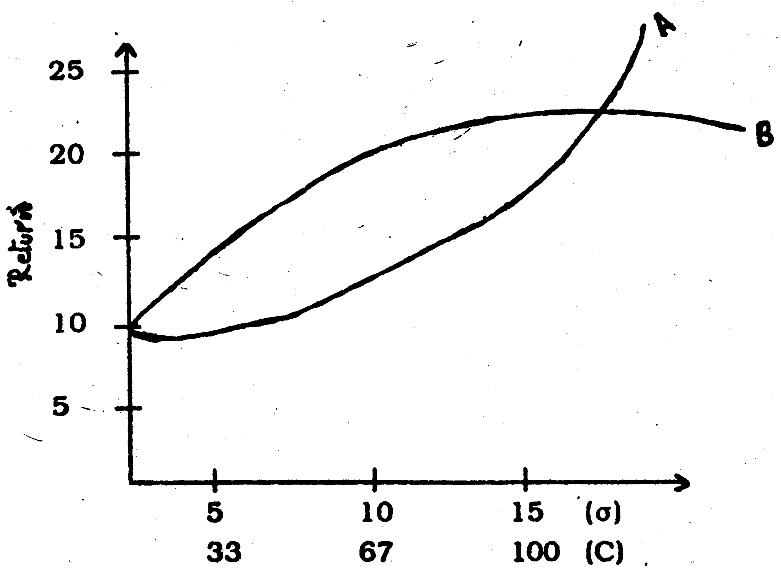
....

....

....

Risk-Return Trade-off

The final decision depends on consideration of risk as well as return. Returns is often taken as a function of risk. We hear for example, no risk, no gain. However different investors may take different views of the investment climate. As a result, we may find different **risk-return functions** (sometimes called, **market indifference curves**) for different investors. Consider the following diagram:



Absolute Risk (σ) and Relative Risk (C)

You may note that curve A reflects **risk aversion** because as risk increases, the required return increases at an increasing rate. Curve B reflects the attitude of a **risk-seeker**, because as risk increases, the required return also increases, but at a decreasing rate. The desirability of an expected return is measured by its vertical distance and direction from the decision-maker's risk return trade-off curve.

Risk attitudes and utility theory

In theory, we can define three possible attitudes towards risk.

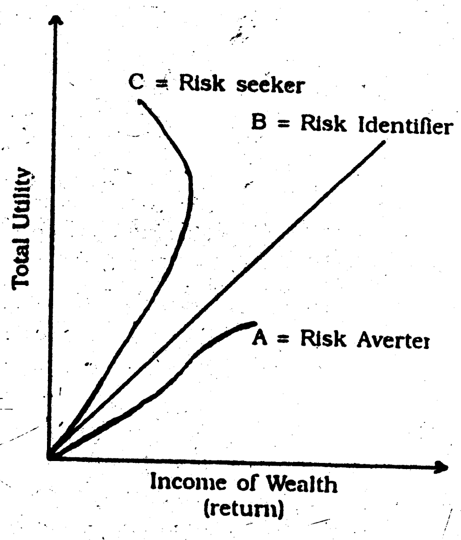
1 A **risk-seeker** is the one who prefers risk, given a choice between more or less risky investments, with identical expected money returns; he will select the riskier investment.

2 A **risk-averter**, faced with the same situation, will select the less risky investment.

3 A **risk-indifferent**, faced with the same situation, will be blind to the choice, for him, any investment is equally preferable to the other.

Activity 3

a) Consider the diagram and interpret it fully.



....

....

....

....

....

....

....

....

....

....

....

You may note that there is increasing marginal utility of income for a risk-seeker; decreasing marginal utility of income for a risk-averter and equal marginal utility of income for a risk-indifferent person.

21.5 ADJUSTING THE VALUATION MODEL FOR RISK

Diminishing marginal utility leads directly to risk aversion, and this risk aversion is reflected in the capitalisation rate applied by investors in determining the present value of future return. You may recall that the expected series of return (yields) from a given investment activity have to be discounted so that the future returns are comparable with present investment costs.

$$P = \sum_{i=1}^{n} \frac{\alpha F_i}{(1+r)^i}$$

Discounting helps computation ofthlie-value of money. i.e., the return. There may be a lot of risks involved investment decisions. Anumber of procedures have been developed for handling risk in this valuation context. Some of these are discussed here:

i) Certainty Equivalent Adjustments

Under this approach, the decision-maker must specify how much money he must be assured of receiving to make him "indifferent" between a "certain sum" and the expected value of 'risky sum'. To illustrate, suppose a rich eccentric offers you the following alternatives:

1 Flip a coin. If a head comes, he gives you Rs. 1 million. If a tail comes you get nothing. The expected value of the gamble is Rs.5,00,000 (0.5 X Rs. 100,000 + 0.5 x 0).

2 Rs. 1,00,000 cash.

If you find yourself indifferent between the two alternatives, Rs. 100,000 is your certainty equivalent for the risky Rs. 500,000. A certainty equivalent less than Rs. 500,000 indicates risk-aversion. We calculate the "certainty equivalent adjustment factor", for each level of risk, by taking a ratio of certain income, Y, and risky income, y.

$$\alpha = \frac{Y}{y} = \frac{\text{Rs.}10{,}000}{\text{Rs.}50{,}000} = 0.20$$

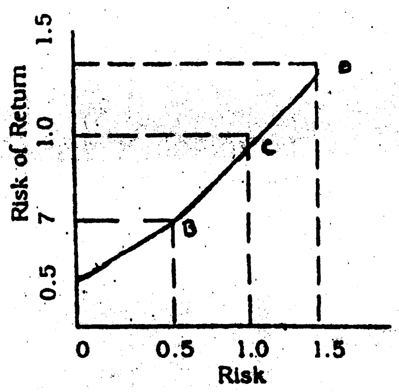
And, now the valuation model may be adjusted as:

$$P = \sum_{i=1}^{n} \frac{\alpha F_i}{(1+r)^i}$$

ii) Risk-adjusted Discount Rates

An alternative procedure for taking risk into account calls for making adjustments to the discount rate, r. Like the certainty equivalent method, risk-adjusted discount rates are based on investor's trade-off functions between risk and return. Suppose, an investor is indifferent to a riskless asset with a sure 5% rate of return, a moderately risky asset with a 7% expected return, and a very risky asset with a 15% expected return. As risk increase, higher and higher expected returns on investment are required to compensate investors for additional risk.

The difference between the expected rate of return on a particular risky asset and the rate of return on a riskless asset is defined as risk-premium on the risky asset. In the figure here, the riskless rate is assumed to be 5%, a 2% risk premium is required to compensate for a co-efficient of variation of 0.5; and a 10% risk premium is attached to an investment with a co-efficient of variation as high as 1.5. The average investor is indifferent between risky investments B,C,D and the riskless alternative A.



Since required returns are related to the level of risk perceived to be associated with a particular investment, we can modify the basic valuation model to account for risk through an adjustment of the discount rate. r. Such a modification results in the valuation model:

$$P = \sum_{i=1}^{n} \frac{F_t}{(1+r)^t} \quad \text{were } K = r + \delta$$

K, the risk-adjusted discount rate, is the sum of the riskless rate of return, and the risk premium δ and hence, a function of the variability of investor's returns. Thus for an investor whose co-efficient of variation on returns was 1.0, the appropriate discount rate would be 10%: a riskier investor with δ = 1.5 would be evaluated with a 15% discount rate.

We can illustrate the use of risk-adjusted discount rates for managerial decision-making

An example

The Walchand Watch Co. is considering to manufacture two mutually exclusive types of Watchbands . A and B. The expected investment outlay for design, engineering, production set-up and so on is Rs. 100,000 for each alternative Expected cashflows are Rs.20,000 a year for eight years for project/product A and Rs.23,000 a year for eight years for project/product B. Because of the captive market for A, the co-efficient of variation of the expected annual returns from the project is only 1.0, while that of project B is 1.5. In view of this risk differential. the management of Walchand Watch Co. decides that A should be evaluated with a 10% cost of capital, while the appropriate cost of capital for B is 15%. The question is:

Which project/product should be selected?
We can calculate the risk-adjusted value for each project/product as follows:

Project A

$$NPV = \sum_{i=1}^{8} \frac{\text{Rs. } 20{,}000}{(1.10)^t} - \text{Rs. } 100{,}000$$

$$= \text{Rs. } 20{,}000 \left(\sum_{i=1}^{8} \frac{1}{(1.10)^t} \right) - \text{Rs. } 100{,}000$$

$$= \text{Rs. } 20{,}000 \times 5.335 - \text{Rs. } 100{,}000$$
$$= \text{Rs. } 6700$$

Project B

$$NPV = \sum_{i=1}^{8} \frac{\text{Rs. } 23000}{(1.15)^t} - \text{Rs. } 3000$$

$$= \text{Rs. } 23000 \left(\sum_{t=1} \frac{1}{(1.15)^t} \right) - \text{Rs. } 100{,}000$$

$$\text{Rs. } 23000 \times 4.487 - \text{Rs. } 100{,}000$$

$$= \text{Rs. } 3200$$

Thus Project or Product A should be preferred to B because this choice maximises the value of the firm.

Activity 4

Review your understanding of the following terms:

a) Risk-discounting:

....

....

....

....

....

....

....

b) Creativity equivalent:

....

....

....

....

....

c) Risk-premium :

....

....

....

....

....

d) Market indifference curve:

....

....

....

....

....

....

21.6 DECISION TREE

When a sequential series of conditional decisions are to be made under risk, a **decision tree** enables the decision-maker to visualise and evaluate all possible options for action.

A **decision tree** is a graphic device that shows a sequence of strategic decision and expected consequences under each possible set of circumstances.

An example

A petroleum firm is considering the possibility of expanding into agricultural chemicals. It decides to proceed in stages.

1. Spend Rs. 100,000 for a market survey of supply-demand conditions in the agricultural chemical industry.

2. If the survey results are favourable, spend Rs. 2 million on a pilot-plant to investigate production methods.

3. Depending on the costs estimated from the pilot study and the demand potential from the market study, either abandon the project or build a large plant or build a small one.

Thus the final decision is made in stages, with subsequent decisions depending on the results of past decisions. The sequence of events can be mapped out to resemble the branches of a tree, hence the term "decision tree". With a set of imaginary data, we may present here an illustrative decision-tree. The reference is to the same petroleum film. Since the net present value of the large plant is higher, should we decide to construct it? Perhaps, but not necessarily since the investment costs differ depending upon the size of the plant. You must examine the co-efficient of variation of the net present value possibilities.

Activity 5

With regard to both plants above, estimate :

a) Standard deviation :

....

....

....

....

....

....

....

....

....

b) Co-efficient of variation :

....

....

....

....

....

....

The co-efficient of variation for the large plants present value is 4.3, while for the small plant it is only 1.5. Risk is greater, if the decision is taken to build a large plant. The decision-maker could take account of risk differentials in a variety of ways. He could assign utility values to cash flows given in column (4) and thus state column (5) in terms of expected utility. He would then choose the plant providing the greatest expeted value. Alternatively, he could use the method of certainty equivalent or the method of risk adjusted discount rate. The plant that offers the large risk- adjusted NPV would then be the optimal choice.

You should note that in this section, we have used a very simple decison tree to spell out the idea. In actual use, the decision trees are far more complex. Particulary, there may be "decision trees" with multiple "decision points".

Activity 6

Look up any other texts for illustrative and complex dicision trees. Then construct a decision tree illustrating the nature of decision problem that you normally face.

21.7 SIMULATIOIN

Another technique designed to assist managers inmaking decisions under risk and uncertainty is computer simulation. To illustrate, let us consider a specific decision situation.

An example
A new textile plant is to be built. The exact cost of the plant is not known (definitely). It is expected to be about Rs. 150 million. If the environment is favourable, the cost can be as low as Rs. 125 million; on the other hand, an unfortunate series of events strikes, technical problem, unprojected increase in material costs, and the like could result in an investment outlay as high as Rs. 225 million.

Revenues from the new facility, which will operate for many years, depend on population growth and income in the region, competition, developments in synthetic fabrics, research and textile import quotas. Operating costs depend on production efficiency, materials and labour cost trends, and the like. Both sales revenues and operating costs as well as annual profits are uncertain.

Assuming that probability distributions can be developed for each of the major cost and revenue determinants, a computer programme can be constructed to simulate what is likely to occur. In effect, the computer selects one value at random from each of the relevant distributions, combines it with other values selected from the other distributions, and produces an estimated profit and net present value or rate of return on investment. This particular profit and the rate of return occur only for the particular combination of values selected during the trial. The computer proceeds to select other sets of values and to compute other profits and rates of return for perhaps several hundred trials. A count is kept of the number of times each various rate of return is computed, and when the computer runs are completed, the frequency with which the various rates of return occurred can be plotted as a frequency distribution.

Finally, the technique requires probability distributions about a number of variables—investment outlays, unit sales, product prices, input prices, asset life etc.—and involves a fair amount of programming and machine-time costs. Full scale simulation is, therefore, not generally feasible except for large and expensive projects such as major plant expansions or new product decisions. In these cases, however, when a firm is deciding whether or not to accept a major undertaking involving an outlay of millions of rupees, computer simulation can provide valuable insights into the relative merits of alternative strategies.

Sensitivity Analysis
It should also be noted that a somewhat less expensive simulation technique is available as an alternative method of analysing the outcomes of various projects or strategies. Instead of using probability distributions for each of the variables in the problem, we can simulate the results by starting with best-guess estimates for each variable, then change the values of the variables (with reasonable limits) to see the effects of such changes on the rate of return. Typically, the rate of return is highly sensitive to some variables, and less so to others. Attention is then concentrated on the variables to which profitability is most sensitive. This sensitivity analysis, as a technique is considerably less expensive than the full-scale simulation and provides similar data for decision-making purposes.

Activity 7

a) How is **sensitivity analysis** different from **simulation approach?**

....

....

....

....

....

....

....

....

b) What are the **assumptions underlying simulation? What are the likely** limitations of this technique for investment planning?

....

....

....

....

....

....

....

....

21.8 DECISION-MAKING UNDER UNCERTAINTY

Decision-making under uncertainty is generally subjective. However, if the decision maker can identify the possible states of nature and estimate the resulting pay-offs for each available strategy, then two basic approaches are available.

i) Subjective probabilities', discussed earlier, may be used to convert the condition of **uncertainty** into a condition of **risk** for which **"expected value"** is the criterion of choice.

ii) The decision-maker may either **disregard probabilities or treat them as** equal. When this approach is taken, four **decision criteria are available** for evaluaticl of proposed strategies:

- The **Wald** decision criterion, also called **Maximin**
- The **Hurwicz alpha** decision criterion
- The **Savage** decision criterion, also called **Minimax Regret**
- The **Laplace** decision criterion, also called the **Bayes** decision criterion

The Wald Decision Criterion or Maximum

This is often described as the criterion of pessimism, the criterion of extreme conservatism, and an attempt to maximise the security level. It says: determine the worst possible outcome of each strategy and then pick the strategy yielding the best of the worst results. The maximum criterion can be illustrated by applying it to the example used in Table 1.

Table 3

Application of Miximin and Maximax Criterion

	State of Nature				Criterion	
Strategies	N_1	N_2	N_3	N_4	Maximin	Maximax
S_1	6	6	6	4	4*	6
S_2	25	7	7	-15	-15	25*
S_3	20	20	7	-1	-1	20
S_4	19	16	9	-2	-2	19
S_5	20	15	15	-3	-3	20

*Best strategy under stated criterion.

You may note that the most dismal pay-off from each row is chosen as the minimal security level associated with the strategy. The largest of these, a value of 4, implies that S_1 is the best strategy.

Is this a good choice? It all depends upon what you mean by "good". Note that if the state of nature N_4 should occur. S_1 is the only strategy that avoids a loss. On the other hand, if any other state occurs, S_1 yields the poorest return. However S_1 not only promises the smallest returns, it also involves the smallest risks. Because the criterion is fiscally conservative, it is particularly well suited to small business firms whose survival depends on avoiding losses.

Added to Table 3 is the anti-thesis of maximin, called **Maximax**. Here the decision-maker is completely optimistic and therefore chooses the maximum pay-off. The strategy that offers the best of the best is then chosen as optimal: S_2 represents that optimal strategy. This, of course, is non-sense. We have included it, because Maximin and Maximax represent the extreme of alpha in Hurwicz criterion.

The Hurwicz Alpha Decision Criterion

It proposes to create **decision index** (d) for each strategy, which is weighted average of its extreme pay-offs. The weighting factors are a co-efficient of optimism (d), which is applied to the **maximum pay- off** (M), and its **complement** ($1-\alpha$) which is applied to the **minimum pay-off** (m). Thus, the value of each strategy is

$d_i = M_i + (1 - \alpha) m_i$

The co-efficient of optimism ranges from 0 to 1, enabling the decision-maker to express his attitude towards risk-taking as a subjective degree of optimism. Thus $\alpha = 0$ means extreme pessimism and $\alpha = 1$ means extreme optimism. Let us assume the $\alpha = 0.7$ i.e. the decision-maker is on the optimistic side. His analysis of the current decision problem is shown in the table below.

Table 4

	Hurwicz Alpha			Solution to Decision Problem			
Strategy	M	α	αM	m	(l-α)	(l-α)m	d
S_1	6	0.7	4.2	4	0.3	1.2	5.4
S_2	25	0.7	17.5	1-15	0.3	-4.5	13.0
S_3	20	0.7	14.0	-1	0.3	-0.3	13.7*
S_4	19	0.7	13.3	-2	0.3	-0.6	12.7
S_5	20	0.7	14.0	-3	0.3	-0.9	13.1

*Best strategy under stated criterion

It can be seen that the highest weighted average pay-off results from selecting S_3.

The decision indicated by the Hurwicz criteria depends on the value of α, which in its turn depends on the decision-makers risk attitude. It is suitable for use by business firms; but if the decision-maker's degree of optimism proved unfounded, substantial losses are likely. Therefore, due caution must be taken in using this criterion.

The Savage Decision Criterion or the Minimax Regret Criterion

This criterion examines "regrets" which are opportunity costs of incorrect decision. Regret is measured as the absolute difference between the pay-off for a given strategy and the pay-off for the most effective strategy within the same state of nature.

The rationale for measurement of regret is quite simple. If any particular state of nature occurs in the future, and if we have chosen the strategy that yields the maximum pay-off for that state of nature, then we have no regret. But if we choose any other strategy, regret is the difference between what actually occurs and what we could have earned had we made the optimal decision. After determining the maximum regret for each strategy, the strategy with the smallest maximum regret is chosen.

We can rework with our previous example to construct a **regret matrix.**

A regret matrix can be constructed by modifying the pay-off matrix. Within each column (state of nature), the largest pay-off is subtracted from each pay-off number in the column. The absolute difference between them is the measurement of regret.

Table 5

A Regret Matrix

	Pay-off Matrix				Regret Matrix				
Strategy	N_1	N_2	N_3	N_4	N_1	N_2	N_3	N_4	**Maximin regret**
S_1	6	6	6	4	19	14	9	0	19
S_2	25	7	7	-15	0	13	8	19	19
S_3	20	20	7	-1	5	0	8	5	8
s_4	19	16	9	-2	6	4	6	6	6*
s_5	20	15	15	-3	5	5	0	7	7

*Best strategy under stated criterion.

You may note that the decision-maker who uses the Savage criterion explicitly abandons attempts to maximise pay-off in favour of a strategy to achieve a satisfactory pay-off with less risk. The Savage criterion is therefore particularly useful for evaluating a series of projects over a long span of time.

The Laplace Decision Criterion or Bayes Criterion

There is a Bayesian postulate that if the probabilities of occurrences are unknown, they should be assumed equal. The Laplace criterion uses this postulate to calculate the expected value of each strategy; hence it is also called "Bayes criterion". The strategy selected is the one with the greatest expected value that results from the assumed (subjective) probabilities.

For strategies S_1, S_2, S_3, S_4 and S_5 in our example, the expected values are 22/4, 22/4, 46/4, 42/4 and 47/4 respectively, and strategy S_5 would be selected. By assuming an equal probability, we convert 'Uncertainty' into a "risk" environment. Thus we can apply our previous analysis.

The Laplace criterion is a criterion of rationality, completely insensitive to the decision-maker's attitude. It is, however, sensitive to decision-maker's definition of states of nature.

To sum up, decision-making under uncertainty is essentially one of choosing a criterion and then performing the calculations necessary to establish a choice within that criterion. We have also seen that the four decision criteria, when applied to the same decision matrix, can lead to four different strategy selections.

Activity 8

a) Review your understanding of the following terms:

i) Minimax:

....

....

....

ii) **Maximin:**

....

....

....

iii) **Maximax:**

....

....

....

iv) **Regret:**

....

....

....

v) **Alpha:**

....

....

....

vi) **Bayes' postulate:**

....

....

....

b) **Consider the pay-off matrix given below and construct a regret matrix.**

Alternative Strategics	State of Nature	
	Oil Crisis	No. Oil Crisis
Investment in Govt. Bonds	\$5250	\$5250
Investment in Oil Venture	\$7500	\$2500

c) **Given the pay-off matrix, which alternative would be selected by applying different decision criteria that we have discussed? Assume that the co-efficient of optimism is 3/5. Show your calculations.**

Strategy	State of Nature			
	N_1	N_2	N_3	N_4
A	11	15	9	6
B	13	4	14	7
C	10	10	10	10
D	9	11	15	13
E	8	3	7	5

21.9 OTHER METHODS OF DEALING WITH UNCERTAINTY

There are also other methods of dealing with uncertainty in a non-quantitative way. Some of the popular methods in this context are:

i) **Referral to authority for guidance**
This is a very pragmatic approach to the reduction of uncertainty. In some cases, there is a literal authority such as Wage Boards, Industrial Relations Commission, Bureau of Industrial Costs and Prices. Whatever such authorities decide, management accepts them. Sometimes, there are also figurative authorities, like tradition, rule-of-thumb, convention, group pressure, professional ethics etc.

ii) **Control of the environment**
This approach usually takes the form of attempts to gain a monopoly by means of patents, exclusive dealership, contacting and influencing important officials who matter etc.

iii) **Hedging**
It is a very common method by which business executives can replace future uncertainty with the security of a present contract. Hedging takes many forms; it emerges most commonly in the writing of contracts for goods and services, and in trading of futures at commodity exchanges. Even in share market, sometimes hedging takes the form of forward trading.

iv) **Flexible investments**
The investor often takes care of flexibility while deciding on long run investment. For example, a general-purpose machine may be a costlier investment than a specialised machine. But when changes are frequent, it is less riskier to invest in a general-purpose machine.

v) **Diversification**
This is closely related to flexibility. There is an old adage "Don't put all your eggs in one basket". In view of this, to minimise risk and uncertainty, businessmen often, go for diversification of their products, source of raw-materials, markets etc.

vi) **Modification of goals**
In the face of complete uncertainty, an optimal decision might be impossible. Knowing this, many times a behavioural firm, you may recall, attempts modification of its goal its target variables. This is a practical approach to face the uncertain environment. Goals are constantly revised, along with it, the strategy and tactics are revised, depending upon the situation the decision-maker faces; in this way he commands a better control on his decision environment.

vii) **Acquisition of additional information**
Relevant, reliable information is the key to successful decision-making. A gain, you may recall, the behavioural theory suggests that to bridge the gap between achievements and aspirations, management information system must be properly monitored.

Presently in the days of computer, information technology is undergoing revolutionary changes such that-managerial success, in view of coping with risk and uncertainty, depends to a large extent on the timely availability of adequate and appropriate market information. If information is "too little and too late", management cannot get the better of risky and uncertain world. Monitoring market information regularly has, therefore, been accepted as a means to reduce market risk and uncertainty.

21.10 SUMMARY

Decisions are made within the context of a decision-maker's state of knowledge, which defines the condition of certainty, risk and uncertainty. Each of these conditions calls for a different set of decision-making tools and techniques. The condition of **certainty** is often associated with optimising models. The conditions of **risk** (actuarial probability) and **uncertainty** (non-calculable risk) are typical of business decision environment. The decision-maker has to estimate at least guesstimate the element of risk and uncertainty. The degree of risk and uncertainty can be measured either a **priori** or a **posteriori**. Under conditions of risk and uncertainty, a **pay-off matrix** provides a useful decision model, showing possible states of nature, available strategies and the pay-offs (returns) associated with each strategy. How the decision-maker chooses to deal with risk depends upon his attitude whether he is **risk-seeker, risk-averter** or **risk-indifferent**. The decision-maker will accept risk if there is a commensurate risk premium. There is also a trade-off between risk and return, which is depicted in the market indifference curve. The profit-maximising model can be simultaneously adjusted for both risk and the true value of money by several techniques. Two of the most common are the **risk-adjusted discount rate** and **certainty equivalent approach.** Under conditions of risk, the decision criterion for selecting the optimum strategy is **expected value.** There can be two measures of risk - an absolute measure (standard deviation) and a relative measure (co-efficient of variation). Under conditions of uncertainty, there are different criteria for decision-making: Wald/Maximin, Hurwicz/alpha, Savage/Minimax regret, Laplace/Bayes criteria. Different criteria may indicate different choices of strategy. This is where the subjective considerations of the decision-maker become relevant. There are also **non quantitative methods** of dealing with uncertainty: **referral to authority for guidance,** control of environment, flexibility, hedging, diversification, goal modification and information acquisition and monitoring. Finally, the analysis of risk and uncertainty must be followed by **business forecasting.** Forecasting such as demand forecasting and forward planning are designed to reduce the degree of risk and uncertainty. There are no techniques which can completely eliminate risk and uncertainty involved in business decisions.

21.11 ADDITIONAL READINGS

Seo, K.K. (1988) *Managerial Economics* (First Indian Reprint) Surjeet Publications: Delhi. Chs. 2 & 3.

Brigham E.F & J.L. Pappas. (1976) *Managerial Economics* 2nd ed. The Dryden Press: Illinois, Ch. 3.

Baumol, WJ. (1978) Economic *Theory and Operations Analysis* Prentice-Hall of India; New Delhi. Chs. 23 & 24.

21.12 SELF-ASSESSMENT TEST

1 "Uncertainty is an important consideration in decision-making, but there is no unique way of dealing with it". Discuss.

2 How is the concept and measure of probability useful in analysing decisions under risk and uncertainty ?

3 In researching a new product, a marketing consultant has come up with four alternative brand names, five package designs and three advertising campaigns.

a) How many strategies must the management consider?

b) What states of nature might affect management's choice? Give examples.

c) How can management take into account the reaction of competitors?

4 What is your idea of a "decision tree"? Illustrate your idea with the help of a hypothetical example of a bookseller who has the choice of buying 1,2 or 3 thousand books. Assume imaginary data on probabilities, profit etc. for different situations.

5 Imagine the situations and interpret the diagrams below: