Directorate of Distance Education

Master of Business Administration
I - Semester

317 13

MANAGERIAL ECONOMICS
Reviewer

Dr. V. Balachandran

Professor & Head,
Department of Corporate Secretaryship,
Alagappa University, Karaikudi

Authors:

D.N. Dwivedi, Professor of Economics, Maharaja Agrasen Institute of Management Studies, Delhi
Units (1-9, 11-13)

Dr. Suman Lata, Lecturer, Ginni Devi Modi Girls College, Modinagar, Ghaziabad
Unit (10)

Aditi Sharma, Freelance Author
Unit (14)

"The copyright shall be vested with Alagappa University"
<table>
<thead>
<tr>
<th>BLOCK I: BASICS OF MANAGERIAL DEVELOPMENT</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>BLOCK II: PERSONAL AND SUPPLY MANAGEMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit 5: Supply Analysis: Determinants of supply- Elasticity of Supply-Measures and Significance - Derivations of market demand – Demand Estimation and Fore casting- Demand and Supply equilibrium – Giffen Paradox</td>
</tr>
<tr>
<td>Unit 8: Strategic Behaviour of the firms and Game Theory - Nash Equilibrium: Implications – Prisoner’s Dilemma: Types of strategy – Price and Non price competition – Relation to the firm behaviour.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BLOCK III: COST AND BREAK FROM POINT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit 10: New Product Penetrative Decision and Skimming the cream Pricing- Government control over pricing - Concept of Profit- Types and Theories of Profit by Knight (Uncertainty), Schumpeter</td>
</tr>
<tr>
<td>Unit 11: Cost and Return (Pages 147-168)</td>
</tr>
<tr>
<td>Unit 10: Penetration and Skimming Strategies (Pages 169-189)</td>
</tr>
<tr>
<td>Unit 9: Cost and Return (Pages 147-168)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SYLLABI-BOOK MAPPING TABLE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Syllabi</strong></td>
</tr>
<tr>
<td>---------------</td>
</tr>
<tr>
<td><strong>BLOCK I: BASICS OF MANAGERIAL DEVELOPMENT</strong></td>
</tr>
<tr>
<td>Unit 4: Demand analysis: Meaning, Functions - Determinants of demand-Law of Demand – Demand Estimation and Forecasting - Applications of demand in analysis - Elasticity of Demand: Types, Measures and Role in Business Decisions.</td>
</tr>
<tr>
<td>Unit 5: Supply Analysis: Determinants of supply- Elasticity of Supply-Measures and Significance - Derivations of market demand – Demand Estimation and Fore casting- Demand and Supply equilibrium – Giffen Paradox</td>
</tr>
<tr>
<td>Unit 8: Strategic Behaviour of the firms and Game Theory - Nash Equilibrium: Implications – Prisoner’s Dilemma: Types of strategy – Price and Non price competition – Relation to the firm behaviour.</td>
</tr>
<tr>
<td>Unit 10: New Product Penetrative Decision and Skimming the cream Pricing- Government control over pricing - Concept of Profit- Types and Theories of Profit by Knight (Uncertainty), Schumpeter</td>
</tr>
</tbody>
</table>
(Innovation), Clark (Dynamic) and Hawley (Risk) - Profit maximization – Cost volume profit analysis – Risk and Return Relationship.  

**Unit 11:** Profit and Investment Analysis: Meaning – Measurement of profit – Theories of Pricing- Profit planning and forecasting- Profit and Wealth maximization – Cost volume profit analysis – Investment analysis and Evaluation: IRR, NPV and APV techniques.

**BLOCK IV: MACRO ECONOMICS AND REGULATIONS**


INTRODUCTION

BLOCK I: BASICS OF MANAGERIAL DEVELOPMENT

UNIT 1  BASICS OF MANAGERIAL ECONOMICS  1-15

1.0 Introduction
1.1 Objectives
1.2 Economics: Introduction
1.3 Meaning and Nature of Managerial Economics
   1.3.1 Scope of Managerial Economics
1.4 General Foundations of Managerial Economics
   1.4.1 Economic Approach
   1.4.2 Working of Economic System and Circular Flow of Activities
1.5 Economics and Business Decisions
   1.5.1 Relationship between Economic Theory and Managerial Economics
1.6 Answers to Check Your Progress Questions
1.7 Summary
1.8 Key Words
1.9 Self Assessment Questions and Exercises
1.10 Further Readings

UNIT 2  DECISION MAKING AND BUSINESS DECISIONS  16-27

2.0 Introduction
2.1 Objectives
2.2 Role of Managerial Economics in Decision Making
   2.2.1 Decision Making Under Risk and Uncertainty
2.3 Business Decision Rules
   2.3.1 Concepts of Opportunity Cost
   2.3.2 Production Possibility Curve
   2.3.3 Incremental Concepts
   2.3.4 Cardinal and Ordinal Approaches to Consumer Behaviour
   2.3.5 Time Value of Money
2.4 Answers to Check Your Progress Questions
2.5 Summary
2.6 Key Words
2.7 Self Assessment Questions and Exercises
2.8 Further Readings

UNIT 3  CONSUMER BEHAVIOUR  28-44

3.0 Introduction
3.1 Objectives
3.2 Marginalism, Equilibrium and Equi-Marginalism and their Role in
   Business Decision Making
   3.2.1 Equi-Marginal Principles
3.3 Utility Analysis
   3.3.1 Total and Marginal Utility
3.4 Cardinal Utility Approach or the Marshallian Approach
   3.4.1 Law of Diminishing Marginal Utility
3.5 Ordinal Utility Approach or the Hicks Approach
   3.5.1 Indifference Curve Analysis
3.6 Answers to Check Your Progress Questions
3.7 Summary
3.8 Key Words
3.9 Self Assessment Questions and Exercises
3.10 Further Readings

UNIT 4 DEMAND ANALYSIS

4.0 Introduction
4.1 Unit Objectives
4.2 Meaning, Functions and Applications of Demand Analysis
   4.2.1 Law of Demand
   4.2.2 Determinants of Demand
   4.2.3 Demand Estimation and Forecasting
4.3 Elasticity of Demand: Types and Measures
   4.3.1 Price Elasticity of Demand
   4.3.2 Cross Elasticity of Demand
   4.3.3 Income Elasticity of Demand
   4.3.4 Advertisement or Promotional Elasticity
4.4 Answers to Check Your Progress Questions
4.5 Summary
4.6 Key Words
4.7 Self Assessment Questions and Exercises
4.8 Further Readings

BLOCK II: PERSONAL AND SUPPLY MANAGEMENT

UNIT 5 SUPPLY ANALYSIS

5.0 Introduction
5.1 Objectives
5.2 The Law of Supply and Determinants of Supply
   5.2.1 Elasticity of Supply: Measures and Significance
   5.2.2 Derivations of Market Demand
   5.2.3 Demand Estimation and Forecasting
5.3 Demand Supply Equilibrium
5.4 Giffen Paradox
5.5 Answers to Check Your Progress Questions
5.6 Summary
5.7 Key Words
5.8 Self Assessment Questions and Exercises
5.9 Further Readings

UNIT 6 PRODUCTION FUNCTIONS

6.0 Introduction
6.1 Objectives
6.2 Production Functions: Short-Run and Long-Run Production Function
   6.3 Theory of Production and Managerial Uses of Production Function
      6.3.1 Short-Run Laws of Production
6.3.2 Isoquants
6.3.3 Long-Run Laws of Production
6.4 Cobb-Douglas and Other Production Functions: Empirical Estimation
6.5 Answers to Check Your Progress Questions
6.6 Summary
6.7 Key Words
6.8 Self Assessment Questions and Exercises
6.9 Further Readings

UNIT 7 FORMS OF MARKETS

7.0 Introduction
7.1 Unit Objectives
7.2 Meaning and Characteristics of Markets
7.2.1 Market Equilibrium and Changes in Market Equilibrium
7.2.2 Market Equilibrium: Practical Importance
7.3 Pricing Functions under Different Market Structures
7.3.1 Market Structures and Pricing Decisions
7.4 Price and Output Decisions under Different Competitive Conditions
7.4.1 Monopoly
7.4.2 Monopolistic Competition
7.4.3 Oligopoly
7.5 Answers to Check Your Progress Questions
7.6 Summary
7.7 Key Words
7.8 Self Assessment Questions and Exercises
7.9 Further Readings

UNIT 8 STRATEGIC BEHAVIOUR OF THE FIRMS AND GAME THEORY

8.0 Introduction
8.1 Unit Objectives
8.2 Strategic Behaviour of the Firms: Price and Non-Price Competition
8.3 Game Theory: Types of Strategy and Relation to the Firm’s Behaviour
8.3.1 Prisoner’s Dilemma
8.3.2 Application of Game Theory to Oligopolistic Startregy
8.3.3 Nash Equilibrium and Implications
8.4 Answers to Check Your Progress Questions
8.5 Summary
8.6 Key Words
8.7 Self Assessment Questions and Exercises
8.8 Further Readings

BLOCK III: COST AND BREAK FROM POINT

UNIT 9 COST AND RETURN

9.0 Introduction
9.1 Objectives
9.2 Relevant Costs for Decision-Making
9.3 Traditional Theory of Cost: Cost Function and Cost Output Relationship
9.3.1 Short-Run Cost-Output Relations
11.9 Self Assessment Questions and Exercises
11.10 Further Readings

BLOCK IV: MACRO ECONOMICS AND REGULATIONS

UNIT 12 MACRO-ECONOMIC FACTORS 209-232
12.0 Introduction
12.1 Objectives
12.2 Meaning and Nature of Macro-Economic Factors
12.2.1 Importance of Macro-Economic Factors
12.3 Economic Growth and Development
12.4 Business Cycle: Phases and Business Decisions
12.5 Inflation
12.5.1 Factors Causing Inflation
12.5.2 Factors Causing Deflation
12.5.3 Control Measures
12.6 Balance of Payment Trend and Its Implications on Managerial Decisions
12.7 Answers to Check Your Progress Questions
12.8 Summary
12.9 Key Words
12.10 Self Assessment Questions and Exercises
12.11 Further Readings

UNIT 13 NATIONAL INCOME 233-255
13.0 Introduction
13.1 Objectives
13.2 Introduction to National Income
13.2.1 Methods of Measurement
13.3 Theory of National Income Determination
13.3.1 Two-Sector Model
13.3.2 Three-Sector Model
13.3.3 Four-Sector Model
13.4 Sectoral and Population distributions
13.5 Per Capita Income: Definition, and Calculations
13.5.1 GDP and GNP
13.5.2 Uses and Limitations
13.5.3 Recent Developments in Indian Economy
13.6 Answers to Check Your Progress Questions
13.7 Summary
13.8 Key Words
13.9 Self Assessment Questions and Exercises
13.10 Further Readings

UNIT 14 ECONOMIC REGULATIONS OF BUSINESS 256-273
14.0 Introduction
14.1 Objectives
14.2 Antitrust Theory and Regulations
14.2.1 Antitrust Theory
14.2.2 Activities of Business which Require the Use of Antitrust Laws
14.3 The Structure – Conduct – Performance Paradigm
  14.3.1 Elements of the Paradigm
  14.3.2 Criticisms of the SCP Paradigm
14.4 Concentration: Overview and Measurement
  14.4.1 Measuring Industrial Concentration
  14.4.2 Monopoly Vs Concentration of Economic Power
14.5 Regulation of Externalities
14.6 Answers to Check Your Progress Questions
14.7 Summary
14.8 Key Words
14.9 Self Assessment Questions and Exercises
14.10 Further Readings
INTRODUCTION

The natural curiosity of a student who begins to study a subject is to know its nature and scope. Such as it is, a student of economics would like to know ‘What is economics’ and ‘What is its subject matter’. Surprisingly, there is no precise answer to these questions. Attempts made by economists over the past 300 years to define economics have not yielded a precise and universally acceptable definition of economics. Economists right from Adam Smith—the ‘father of economics’—down to modern economists have defined economics differently, depending on their own perception of the subject matter of economics of their era. Thus, economics is fundamentally the study of choice-making behaviour of the people. The choice-making behaviour of the people is studied in a systematic or scientific manner. This gives economics the status of a social science.

However, the scope of economics, as it is known today, has expanded vastly in the post-World War II period. Modern economics is now divided into two major branches: Microeconomics and Macroeconomics. Microeconomics is concerned with the microscopic study of the various elements of the economic system and not with the system as a whole. As Lerner has put it, ‘Microeconomics consists of looking at the economy through a microscope, as it were, to see how the millions of cells in body economic—the individuals or households as consumers and the individuals or firms as producers—play their part in the working of the whole economic organism.’ Macroeconomics is a relatively new branch of economics. Macroeconomics is the study of the nature, relationship and behaviour of aggregates and average of economic variables. Therefore, the technique and process of business decision-making has of late changed tremendously.

The basic functions of business managers is to take appropriate decisions on business matters, to manage and organize resources, and to make optimum use of available resources with the objective of achieving the business goals. In today’s world, business decision-making has become an extremely complex task due to the ever-growing complexity of the business world and the business environment. It is in this context that modern economics—howsoever defined—contributes a great deal towards business decision-making and performance of managerial duties and responsibilities. Just as biology contributes to the medical profession and physics to engineering, economics contributes to the managerial profession.

This book, Managerial Economics has been divided into fourteen units. The book has been written in keeping with the self-instructional mode or the SIM format wherein each Unit begins with an Introduction to the topic, followed by an outline of the Objectives. The detailed content is then presented in a simple and organized manner, interspersed with Check Your Progress questions to test the student’s understanding of the topics covered. A Summary along with a list of Key Words, set of Self-Assessment Questions and Exercises and Further Readings is provided at the end of each Unit for effective recapitulation.
Managerial Economics has emerged as a separate branch of economics. The emergence of managerial economics can be attributed to at least three factors: (i) growing complexity of the business environment and decision-making process; (ii) increasing application of economic logic, concepts, theories and tools of economic analysis in the process of business decision-making; and (iii) rapid increase in demand for professionally trained managerial manpower with good knowledge of economics. The growing complexity of the business world can be attributed to rapid growth of large scale industries, increasing number of business firms, quick innovation and introduction of new products, globalization and growth of multinational corporations, merger and acquisition of business firms, and large-scale diversification of business activities. These factors have contributed a great deal to the inter-firm, inter-industry and inter-country business rivalry and competition, enhancing uncertainty and risk in the business world.

Business decision-making in this kind of complex business environment has become a very complex affair. There was a time when family training, personal experience and business acumen were sufficient to make good business decisions
and run an organization successfully. In today’s business world, however, personal experience, knowledge and family training are no longer sufficient to meet the managerial challenges of the modern business world, though one can find a number of reputed businessmen with no management training.

Before we proceed to discuss the nature and scope of managerial economics, let us first know “what economics is about”.

1.1 OBJECTIVES

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

- Prepare an introduction to economics
- Discuss the meaning, nature and scope of economics
- Explain the general foundations of managerial economics
- Analyse the relationship between economics and business decisions

1.2 ECONOMICS: INTRODUCTION

Economists of different generations have defined economics in different ways according to their perception and subject matter of economics. For example, according to Adam Smith (1976), the “father of economics”, economics is “an inquiry into the nature and causes of the wealth of nations”. According to Alfred Marshall, (1922), an eminent economist of the neo-classical era, “Economics is the study of mankind in the ordinary business of life; it examines that part of individual and social actions which is most closely connected with the attainment and with the use of the material requisites of well-being.”

Economics can, thus be defined as a social science that studies economic behaviour of the people, the individuals, households, firms, and the government. **Economic behaviour** is essentially **economizing behaviour**. Economizing behaviour is the effort of the people to derive maximum gain from the use of their limited resources—land, labour, capital, time and knowledge, etc., which have alternative uses. Technically, the term ‘economizing’ means deriving maximum gains from a given cost and alternatively minimizing cost for a given gain. This is economizing behaviour—a natural behaviour.

Why do People Economize? People tend to adopt economizing behaviour because of the following facts of economic life of human beings.

1. **Human wants, desires and aspirations are endless.** Human wants are endless is the sense that they go on increasing with the availability of new kinds of goods and services and increase in ability to pay.

2. **Resources available to the people are scarce.** Resources (labour, land, capital, time and knowledge, etc.) available to the people at any point on time are scarce and limited; though they have alternative uses.
3. **People are by nature economizers.** Attempt to economize is a natural behaviour of the people. For example, if metro service is easily available, one would not like to hire a taxi, and if a pizza is available at ₹50 in college canteen and for ₹100 in a nearby restaurant, students would prefer to eat pizza in the canteen. This is economizing behaviour.

In order to maximize their gains, people allocate their resources between their competing wants in such a way that their total gain is maximized. Thus, economics as a social science studies how people allocate their limited resources to their alternative uses with the objective of deriving maximum possible gains from the use of their resources. For analysing economic behaviour of the people, economists use certain specific concepts, logic, tools of analysis and maximization techniques. The ultimate result of this kind of analysis is the formulation of economic theories.

**Microeconomics and Macroeconomics**

Economics as a social science has two major branches—*microeconomics* and *macroeconomics*. Microeconomics is the study of the economizing behaviour of the individual economic entities—individuals, households, firms, industries and factory owners. For example, microeconomics studies how individuals and households with limited income decide ‘what to consume’ and ‘how much to consume’ so that their total utility is maximized. In other words, microeconomics studies how individual consumers make choice of goods and services they want to consume and how they allocate their limited income between the goods and services of their choice to maximize their total economic welfare.

Macroeconomics, on the other hand, studies the economic phenomena at the national aggregate level. Specifically, macroeconomics is the study of working and performance of the economy as a whole. It studies what factors and forces determine the level of national output or national income, rate of economic growth, employment, price level, and economic welfare. Besides, macroeconomics studies how government of a country formulates its macroeconomic policies—taxation and public expenditure policies (the fiscal policy), monetary policy, price policy, employment policy, foreign trade policy, etc., to resolve the problems of the country.

### 1.3 MEANING AND NATURE OF MANAGERIAL ECONOMICS

Managerial economics can be defined as the study of economic theories, logic, concepts and tools of economic analysis applied in the process of business decision-making. In general practice, economic theories and techniques of economic analysis are applied to diagnose the business problems and to evaluate alternative options and opportunities open to the firm for finding an optimum solution to the problems.
Managerial economics is an integration of economic science with decision making process of business management. The integration of economic science with management has become inevitable because application of economic theories and analytical tools make significant contribution to managerial decision-making.

As we know, the basic managerial functions are planning, organizing, staffing, leading and controlling business related factors. The ultimate objective of these managerial functions is to ensure maximum return from the utilization of firm’s resources. To this end, managers have to take decisions at each stage their functions in view of business issues and implement decisions effectively to achieve the goals of the organization. As we will see later, almost all managerial decision issues involve economic analysis and analytical techniques. Therefore, economic theories and analytical tools are applied as a means to find solution to the business issue. This is how economics gets integrated to managerial functions and gives emergence of managerial economics. The integration of economic theories and concepts with quantitative methods creates *managerial economics*.

Some other definitions of managerial economics are given below:

According to *Spencer and Siegelman*:

“The integration of economic theory with business practice for the purpose of facilitating decision-making and forward planning by management”.

According to *McGutgan and Moyer*:

“Managerial economics is the application of economic theory and methodology to decision-making problems faced by both public and private institutions”.

In the words of *T.J. Webster*, "Managerial economics is the synthesis of microeconomic theory and quantitative methods to find optimal solutions to managerial decision-making problems.”

In the words of *Hirschev and Pappas*, "Managerial economics applies economic theory and methods to business and administrative decision making”

According to *Mansfield*, "Managerial economics provides a link between economic theory and decision sciences in the analysis of managerial decision making.”

*Brigham and Poppas* believe that managerial economics is “the application of economic theory and methodology to business administration practice.”

*Hague* on the other hand, considers managerial economics as “a fundamental academic subject which seeks to understand and to analyse the problems of business decision-making.”

It may be added at the end that economic science has a very wide perspective. All economic theories are neither applicable nor are applied to business decision-making. Most business management issues are of internal nature and a significant part of microeconomics deals with internal decision-making issues of
the business firms—what to produce, how to produce, how much to produce, and what price to charge, etc. That is why most microeconomic theories and analytical tools are generally applied to managerial decision-making. Therefore, managerial economics is treated as applied microeconomics. Macroeconomics deals with environmental issues—how is the economic condition of the country; what is the likely trend; what are government’s economic policies; how government policies might affect business environment of the country; what kind of business policy will be required, and so on.

1.3.1 Scope of Managerial Economics

The scope of managerial economics is comprised of economic concepts, theories and tools of analysis that can be applied in the process of business decision making to analyse business problems, to evaluate business options, to assess the business prospects, with the purpose of finding appropriate solution to business problems and formulating business policies for future. As noted above, economic science has two major branches, viz., microeconomics and macroeconomics. Both Microeconomics and Macroeconomics are applied to business analysis and business decision making depending on the nature of the issue to be examined. Managerial decision issues can be divided broadly under two broad categories: (a) Internal managerial issues and (b) External environmental issues.

Microeconomic Theories Applied to Internal Issues

Internal managerial issues refer to decision-making issues arising in the management of the firm. Internal managerial issues include problems that arise in operating the business organization. All such managerial issues fall within the purview and the control of the managers. Some of the basic internal management issues can be listed as follows.

- What to produce—choice of the business
- How much to produce—determining the size of the firm
- How to produce—choice of efficient and affordable technology
- How to price the product—determining the price of the product
- How to promote sale of the product
- How to face price competition from the competing firms
- How to enlarge the scale of production—planning new investment
- How to manage profit and capital.

The microeconomic theories and tools of analysis that provide a logical basis and ways and means to find a reasonable solution to business problems constitute the microeconomic scope of managerial economics. The main microeconomic theories that fall within the scope of managerial economics are following: Theory of Consumer Demand, Theory of Production, Theory of Cost, Theory of Price Determination and Theory of Capital and Investment Decisions.
Macroeconomics Applied to Business Decision

Macroeconomics is the study of economic conditions of the economy as a whole whereas a firm is a small unit of the economy. As such, macroeconomic theories are not directly applicable to managerial decisions. However, business managers, while making business decisions, cannot assume the economic conditions of the country to remain the same for ever. As a matter of fact, economic conditions of the country keep changing. Changing economic conditions change the economic environment of the country, and thereby business environment and business prospect. And, as management experts Weihrich and Koontz point out, “... managers cannot perform their task well unless they have an understanding of, and are responsive to the many elements of the external environment—economic, technological, social, political, and ethical factors that affect their areas of operations.” Therefore, while making business decisions, managers have to take into account the economic environment of the country. The factors which, in general, determine the economic environment of a country are (i) the general trend in national income (GDP), saving and investment, prices, employment, etc., (ii) the structure and role of the financial institutions, (iii) the level and trend in foreign trade, (iv) economic policies of the government, (v) socio-economic organizations like trade unions, consumer associations, and (vi) political environment.

It is far beyond the powers of a single firm, howsoever large it may be, to determine the course of economic, political and social conditions of the country. But the environmental factors have a far reaching bearing on the functioning and performance of the business firms. Therefore, it is essential for business decision-makers to take in view the present and future economic environment of the country. It is essential because business decisions taken ignoring the environmental factors may not only fail to produce the result but may also cause heavy losses.

Macroeconomic Factors

The major macroeconomic environmental factors that figure in business decisions, especially those related to forward planning and formulation of strategy, may be described under the following three categories: Trend in the Economy, International Economic Conditions and Government Policies.

1.4 GENERAL FOUNDATIONS OF MANAGERIAL ECONOMICS

Managerial economics is defined as science which deals with the application theory of economics in managerial practice. It is the study of allocation of resources available to an enterprise. In simple words, Managerial economics is economics applied in decision-making.

1.4.1 Economic Approach

What is a Positive and a Normative Science? A positive science studies the phenomena as they actually are or as they actually happen. It does not involve any
value judgement on whether what happens is good or bad, desirable or undesirable. A normative science, on the other hand, involves value judgement on whether what happens is socially desirable or undesirable, and if undesirable, how it can be made desirable. As J.N. Keynes puts it, “…a positive science is a body of systematized knowledge concerning what is [and] a normative or regulatory science is a body of systematized knowledge relating to criteria of what ought to be and is concerned therefore with ideal as distinguished from actual.” Friedman has defined ‘positive science’ more elaborately and clearly. In his own words, “The ultimate goal of a positive science is the development of a ‘theory’ or ‘hypothesis’ that yields valid and meaningful (i.e., not truistic) predictions about phenomena not yet observed.” Judged against these definitions of positive and normative science, economics as a social science deals with both positive and normative economic questions: ‘what is’ and ‘what ought to be’. Thus, economics is both a positive and a normative science. Let us look at positive and normative character of economic science in some detail.

**Economics as a Positive Science**

Economics as a positive science seeks to analyze systematically and explain economic phenomena as they actually happen; find common characteristics of economic events; brings out the ‘cause and effect’ relationship between the economic variables, if any; and generalizes this relationship in the form of a theoretical proposition. One of the main purposes of economic studies is ‘to provide a system of generalization’ in the form of economic theories that can be used to make predictions about the future course of related events. It means that economics has a positive character. Economics explains the economic behaviour of individual decision-makers under given conditions; their response to change in economic conditions; and it brings out the relationship between the change in economic conditions and economic decision of the people.

**Economics as a Normative Science**

Economics as a normative science is concerned with ideal economic situation, not with what actually happens. Its objective is to examine ‘what actually happens’ from moral and ethical points of view and to judge whether ‘what happens’ is socially desirable. It examines also whether economic phenomena like production, consumption, distribution, prices, etc. are socially desirable or undesirable. Desirability and undesirability of economic happenings are determined on the basis of socially determined values. Thus, normative economics involves value judgement and values are drawn from the moral and ethical values and political aspirations of the society. In simple words, normative side of economics deals with such normative questions as ‘what ought to be?’ and whether ‘what happens’ is good or bad from society’s point of view? It not, then how to correct it.
1.4.2 Working of Economic System and Circular Flow of Activities

If you look around, you find people busy in some kind of economic activity. Farmers, firms and factories are busy in producing goods and services; buyers and sellers in the shops are busy in buying goods; some persons are busy in ferrying the people from one place to another; some persons are busy in offices in finalizing their deals and recording the transactions; and so on. All these people are performing some kind of economic activity. Any activity that produces goods and services is productive activity and any activity that creates goods and services of value is called economic activity. The basic objective behind all economic activities is to make income, the source of livelihood.

An important feature of economic activities is that they are interrelated and interdependent in the sense that producers produce what consumers want to consume and consumers can consume only what producers produce and they produce only as much as consumers are willing to consume. Similarly, sellers can sell only what buyers are willing to buy and buyers can buy only what is offered for sale; and so on. This interrelatedness and interdependence of economic activities are carried out in a self-operated system. Given this brief description of people’s interrelated and interdependent, we may now attempt to define economy:

An economy is a social organism in which people act, interact, cooperate and compete in the process of production and consumption to make their living. An economy is constituted of interrelated and interdependent economic activities of the economic players. Economic players include individuals, households, firms, farms, factories, financial institutions and government. These economic players participate in economic organism in different ways. Individuals and households use their resources (land, labour, capital and skill) either themselves to produce goods and services, or sell services of their resources to other producers (firms, farms, factories and the government) to make their living. Producers hire the resources to produce goods and services which they sell in market at a profit. Financial institutions, e.g., banks, financial corporations, mutual funds, insurance companies, collect savings from the households and make it available to the producers on interest. Obviously, people of a society are constantly busy in some kind of these economic activities—production of commodities, buying and selling, transporting men and materials, saving and investing, borrowing and lending, and so on.

Government is an important institution in the modern economy. The government performs both non-economic (administrative) and economic functions. It taxes people’s income and hires factors of production and produces certain goods and services for the people. In addition, it intervenes with the economic activities of the people, as it controls, regulates and guides their economic activities with the purpose of achieving certain social and economic goals. The level of intervention and participation of the government in overall economic activities of the people determines the nature of economic system, i.e., whether an economy is
a capitalist or free enterprise economy, a socialist or command economy or a mixed economy. The basic features of these kinds of economic systems are discussed below.

**How an Economy Works**

Working of a modern economy is extremely complex. Millions of persons participate and contribute to its working in different ways and different capacities—as producers, traders, workers, financiers, and consumers. Thousands of goods and services are produced and consumed and millions of persons are engaged in production and distributions of a single commodity. All those who are involved in economic activities act and interact with different interests and motivations. The various forms and nature of cross-section interdependence and interrelatedness of their economic activities add to the complexity of the economic system. To present a complete picture of economic system showing the role of each individual participation in respect of each commodity is an extremely difficult task, rather impossible. However, we present below working of an economy in a simplified model.

**A Simple Model of an Economy**

A simple model of the economy consists of two sectors: (i) households, and (ii) business firms. The households and business firms are, in fact, the two main decision makers in an economy. The functions and the roles of these economic units in the model economy are described below.

The households play two major roles in the economic system: (i) households supply all the factors of production, viz., land, labour and capital, to the firms which constitute the production sector, and (ii) they consume all the goods and services produced by the business firms.

Business firms include all firms, farms, factories and shops engaged in production and distribution of goods and services. Business firms perform two functions: (i) they hire factors of production from the households and transform them into final goods and services, and (ii) they supply all the goods and services to the households, the consumers.

**Interaction between the Households and Firms**

The functions and the mode of interaction between the two kinds of economic entities and working of the economic system are exhibited in Fig. 1.1. Households and firms interact in two ways: (i) as sellers and buyers of inputs, and (ii) as buyers and sellers of output. The sale and purchase of inputs creates factor market where factor prices are determined, and sale and purchase of final goods and services creates product market where product prices are determined.

As Fig. 1.1 shows, factors of production (land, labour, capital, etc.) flow from the households to the factor market. The interaction between households...
Basics of Managerial Economics

NOTES

1.0 Material

The interaction between the input suppliers (the input suppliers) and business firms (the input demanders) determines the factor prices. Once factor prices are determined, inputs move to business firms. In return, factor payments (wages, rent and interest, etc.) flow to the households.

![Fig 1.1 Working of a Simple Economy and Circular Flows](image)

The business firms transform the factor inputs into finished products. Finished products flow to the product market. The interaction between the business firms, the suppliers, and the households, the buyers, determine the product prices. Once product prices are determined, products flow to the households. In return, the payments made by the households for their purchases flows to the firms. They use their receipts to hire inputs again and the process continues. In this process, two circular flows are generated: (i) real flows, i.e., flows of inputs and final products as shown by the outer circle, and (ii) money flows, show by the inner circle.

1.5 ECONOMICS AND BUSINESS DECISIONS

The primary function of managers is to take appropriate decisions and implement them effectively to achieve the objective of the organization to maximum possible extent, given the resources. Application of economics contributes a great deal to managerial decision-making as it provides guidance in finding an appropriate solution to the business problem. Just as biology contributes to medical profession and physics to engineering, economics contributes to managerial functions. As such, a working knowledge of economics is essential for managers. Managers are, in fact, practicing economists.

Let us now see how economics contributes to managerial decisions. All the areas of managerial decisions, as noted in Figure 1.1, have economic perspective. Therefore, economic theories, concepts and tools of analysis are applied as roadmap to find solution to business problems. It has been found empirically that
application of economic theories and tools of analysis makes significant contribution to the process of business decision making in many ways.

According to Baumol, a Nobel laureate in economics, economic theory contributes to business decision making in three important ways.

First, ‘one of the most important things which the economic theory can contribute to management science’ is providing framework for building analytical models which can help recognize the structure of managerial problem, determine the important factors to be managed, and eliminate the minor factors that might obstruct decision making.

Secondly, economics provides ‘a set of analytical methods’ which may not be directly applicable to analyse specific business problems but they do widen the scope of business analysis and enhance the analytical capability of the business analyst in understanding the nature of the business problems.

Thirdly, various economic terms are used in common parlance, which are not applicable to business analysis and decision making. Economic theory offers clarity to various economic concepts used in business analysis, which enables the managers to avoid conceptual pitfalls. For example, in general sense, ‘demand’ means quantity demanded at a point of time. But, in economic sense, ‘demand’ means the quantity people are willing to buy at a given price and they have ability and willingness to pay.

Apart from providing analytical models and methods and conceptual clarity, economics contributes to business decision in many other ways also. Most business conditions are taken under the condition of risk and uncertainty. Risk and uncertainty arise in business because of continuous change in business conditions and environment, and unpredictable market behaviour. Economics provides models, tools and technique to predict the future course of market conditions, ways and means to assess the risk and, thereby, helps in business decision making.

It is because of these important contributions of economics to business decision making that economics has been integrated with managerial decisions. Managerial decision making without applying economic logic, theory and analytical tools may not offer a reasonable solution.

1.5.1 Relationship between Economic Theory and Managerial Economics

We have noted above that application of theories to the process of business decision making contributes a great deal in arriving at appropriate business decisions. In this section, we highlight the gap between the theoretical world and the real world and see how managerial economics bridges this gap.

Theory vs. Practice

It is widely known that there exists a gap between theory and practice in all walks of life, more so in the world of economic theory and behaviour. A theory which
This gap between theory and practice has been very well illustrated in the form of a story by a classical economist, J.M. Clark. He writes:

‘There is a story of a man who thought of getting the economy of large scale production in plowing, and built a plow three times as long, three times as wide, and three times as deep as an ordinary plow and harnessed six horses [three times the usual number] to pull it, instead of two. To his surprise, the plow refused to budge, and to his greater surprise it finally took fifty horses to move the refractory machine… [and] the fifty could not pull together as well as two’.

The gist of the story is that managers—assuming they have abundant resources—may increase the size of their capital and labour, but may not obtain the expected results. Most probably the man in Clark’s story did not get the expected result because he was either not aware of or he ignored or could not measure the resistance of the soil to a huge plow. This incident clearly shows the gap between theory and practice.

Economic theories are, no doubt, hypothetical in nature but not away from reality. Economic theories are, in fact, a caricature of reality under certain specified conditions. In their abstract form, however, they do look divorced from reality. Besides, abstract economic theories cannot be straightaway applied to real life problems. This should, however, not mean that economic models and theories do not serve any useful purpose. ‘Microeconomic theory facilitates the understanding of what would be a hopelessly complicated confusion of billions of facts by constructing simplified models of behaviour, which are sufficiently similar to the actual phenomenon and thereby help in understanding them’. Nevertheless, it cannot be denied that there is apparently a gap between economic theory and practice. This gap arises mainly due to the inevitable gap between the abstract world of economic models and the real world.

How Managerial Economics Fills the Gap

There is undeniably a gap between economic theory and the real economic world. Therefore, economic theories do not offer a custom-made or ready-made solution to business problems. However, economic theories do provide a framework for logical economic model-building and systematic analysis of economic issues. The need for such a framework arises because the real economic world is too complex to permit considering every bit of relevant facts that influence economic decisions. In the words of Keynes, ‘The objective of [economic] analysis is not to provide a machine, or method of blind manipulation, which will furnish an infallible answer, but to provide ourselves with an organized and orderly method of thinking out particular problem…’. In the opinion of Boulling, the objective of economic analysis...
is to present the ‘map’ of reality rather than a perfect picture of it. In fact, economic
analysis equips us with a road map; it guides us to the destination, but does not
carry us there. This is how managerial economics bridges the gap between
economics and business decision-making. As an example, managerial economics
can also be compared with medical science. Just as the knowledge of medical
science helps in diagnosing the disease and prescribing an appropriate medicine,
managerial economics helps in analyzing the business problems and in arriving at
an appropriate decision.

Let us now see how managerial economics bridges this gap. On one side,
there is the complex business world and, on the other, are abstract economic
theories. ‘The big gap between the problems of logic that intrigue economic theorists
and the problems of policy that plague practical management needs to be bridged
in order to give executive access to the practical contributions that economic thinking
can make to top management policies’. Managerial decision-makers deal with the
complex, rather chaotic, business conditions of the real world and have to find the
way to their destination, i.e., achieving the goal that they set for themselves.
Managerial economics applies economic logic and analytical tools to sift wheat
from the chaff. Economic logic and tools of analysis guide business managers in
the following ways.

(i) Identifying their problems in achieving their goal,
(ii) Collecting the relevant data and related facts,
(iii) Processing and analyzing the facts,
(iv) Drawing relevant conclusions,
(v) Determining and evaluating the alternative means to achieve the goal, and
(vi) Taking a decision.

Without application of economic logic and tools of analysis, business
decisions are most likely to be irrational and arbitrary, which may often prove
counter-productive.

Check Your Progress
1. Name the two branches of economics.
2. What is the ultimate objective of the managerial functions?
3. State the main sectors comprising a simple model of an economy.

1.6 ANSWERS TO CHECK YOUR PROGRESS
QUESTIONS

1. The two branches of economics are microeconomics and macroeconomics.
2. The ultimate objective of the managerial functions is to ensure maximum
return from the utilization of firm’s resources.
NOTES

3. A simple model of the economy consists of two main sectors: (i) households and (ii) business firms.

1.7 SUMMARY

- The emergence of managerial economics can be attributed to at least three factors: (i) growing complexity of the business environment and decision-making process; (ii) increasing application of economic logic, concepts, theories and tools of economic analysis in the process of business decision-making; and (iii) rapid increase in demand for professionally trained managerial manpower with good knowledge of economics.

- Economists of different generations have defined economics in different ways according to their perception and subject matter of economics.

- Economics can, thus be defined as a social science that studies economic behaviour of the people, the individuals, households, firms, and the government.

- Internal managerial issues refer to decision-making issues arising in the management of the firm. Internal managerial issues include problems that arise in operating the business organization.

- Economics has two major branches—microeconomics and macroeconomics. The main economic theories and tools of analysis of both microeconomics and macroeconomics constitute the subject matter of managerial economics.

- Working of a modern economy is extremely complex. Millions of persons participate and contribute to its working in different ways and different capacities—as producers, traders, workers, financiers, and consumers.

- The primary function of managers is to take appropriate decisions and implement them effectively to achieve the objective of the organization to maximum possible extent, given the resources.

- It is because of these important contributions of economics to business decision making that economics has been integrated with managerial decisions. Managerial decision making without applying economic logic, theory and analytical tools may not offer a reasonable solution.

1.8 KEY WORDS

- Microeconomics: It is the study of the economizing behavior of the individual economic entities—individuals, households, firms, industries and factory owners.
• **Managerial economics**: It can be defined as the study of economic theories, logic, concepts and tools of economic analysis applied in the process of business decision-making.

1.9 **SELF ASSESSMENT QUESTIONS AND EXERCISES**

**Short Answer Questions**
1. Prepare an introduction to economics.
2. What is the subject matter of microeconomics and macroeconomics?
3. Write a short note on the scope of managerial economics.

**Long Answer Questions**
1. Discuss the general foundations of managerial economics.
2. Explain the working of a simple economy.
3. Analyse the relationship between economics and business decisions.

1.10 **FURTHER READINGS**


**Websites**

https://blogs.economictimes.indiatimes.com/et-commentary/heres-how-indias-widening-income-distribution-can-be-redressed/
UNIT 2 DECISION MAKING AND BUSINESS DECISIONS

2.0 INTRODUCTION

Managerial decisions are taken at different levels of sophistication. While some business decisions require only ‘rule-of-thumb’ technique, others involve the use of sophisticated techniques. The ‘rule-of-thumb’ technique, evolved out of the traditional business management practices, is used where routine type of business decisions are involved. Sophisticated techniques of business decision-making are used where business decisions involve the problem of handling complex business issues. There are certain fundamental economic concepts which are used—explicitly or implicitly, consciously or unconsciously, deliberately or otherwise—in business decisions of complex nature. This unit presents a brief explanation of some major economic concepts and their use in business decision-making.

2.1 OBJECTIVES

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

- Explain the role of managerial economics in decision making
- Define opportunity cost
- Describe production possibility curve with the help of a diagram
- Discuss the cardinal and ordinal approaches to consumer behaviour
- State the time value of money
2.2 ROLE OF MANAGERIAL ECONOMICS IN DECISION MAKING

Managerial economics, also known as business economics, lays emphasis on applying economic theory directly to business organizations. The application of economic theory using statistical methods assists business organizations in taking decisions and deciding on a strategy related to pricing, operations, risk, investments and production. The entire objective of managerial economics is to enhance the efficiency of decision-making in enterprises to enhance profit.

- **Pricing**: Managerial economics helps business organizations in determining the pricing strategies and suitable pricing levels for their products and services.
- **Elastic Vs. Inelastic Goods**: Economists can decide price sensitivity of products using a price elasticity analysis. This will ensure that it becomes easy to take decisions related to marketing and pricing of goods.
- **Operations and Production**: Quantitative methods are used by managerial to analyze production and operational efficiency by means of schedule optimization, economies of scale and resource analyses.
- **Investments**: Many managerial economic tools and analysis models are used to assist in making investing decisions both for corporations as well as individual investors. These tools are used to make stock market investing decisions and decisions on capital investments for a business.
- **Risk**: Uncertainty exists in every business and managerial economics can help reduce risk through uncertainty model analysis and decision-theory analysis. Statistical probability theory is largely used to provide potential scenarios for business enterprises while making decisions.

2.2.1 Decision Making Under Risk and Uncertainty

The concept of risk and uncertainty can be better explained and understood in contrast to the concept of certainty. **Certainty** is the state of perfect knowledge about the market conditions. In the state of certainty, there is only one rate of return on the investment and that rate is known to the investors. That is, in the state of certainty, the investors are fully aware of the outcome of their investment decisions. For example, if you deposit your savings in ‘fixed deposit’ bearing 10% interest, you know for certain that the return on your investment in time deposit is 10%, and FDR can be converted into cash any day. Or, if you buy government bonds, treasury bills, etc. bearing an interest of 11%, you know for sure that the return on your investment is 11% per annum, your principal remaining safe. In either case, you are sure that there is little or no possibility of the bank or the government defaulting on interest payment or on refunding the money. This is called the **state of certainty**.
In reality, however, there is a vast area of investment avenues in which the outcome of investment decisions is not precisely known. The investors do not know precisely or cannot predict accurately the possible return on their investment.

**NOTES**

### Meaning of Risk

In common parlance, risk means a low probability of an expected outcome. From business decision-making point of view, risk refers to a situation in which a business decision is expected to yield more than one outcome and the probability of each outcome is known to the decision makers or it can be reliably estimated.

There are two approaches to estimating probabilities of outcomes of a business decision, viz., (i) *a priori approach*, i.e., the approach based on deductive logic or intuition and (ii) *posteriori approach*, i.e., estimating the probability statistically on the basis of the past data. In case of *a priori probability*, we know that when a coin is tossed, the probabilities of ‘head’ or ‘tail’ are 50:50, and when a dice is thrown, each side has 1/6 chance to be on the top. The *posteriori* assumes that the probability of an event in the past will hold in future also. The probability of outcomes of a decision can be estimated statistically by way of ‘standard deviation’ and ‘coefficient of variation’.

### Meaning of Uncertainty

Uncertainty refers to a situation in which there is more than one outcome of a business decision and the probability of no outcome is known nor can it be reliably estimated. The unpredictability of outcome may be due to lack of reliable market information, inadequate past experience, and high volatility of the market conditions. For the purpose of decision-making, the uncertainty is classified as:

(a) complete ignorance, and
(b) partial ignorance.

In case of *complete ignorance*, investment decisions are taken by the investor using their own judgement or using any of the rational criteria. What criterion he chooses depends on his attitude towards risk. The investor’s attitude towards risk may be that of

(i) a risk averter,
(ii) a risk neutral or
(iii) a risk seeker or risk lover.

In case of *partial ignorance*, on the other hand, there is some knowledge about the future market conditions; some information can be obtained from the experts in the field, and some probability estimates can be made. The available information may be incomplete and unreliable. Under this condition, the decision-makers use their subjective judgement to assign an *a priori probability* to the outcome or the pay-off of each possible action such that the *sum of such probability distribution is always equal to one*. This is called *subjective probability distribution*. 
2.3 BUSINESS DECISION RULES

Let us go through the various economic concepts which assist in making business decision rules.

2.3.1 Concepts of Opportunity Cost

The opportunity cost is essentially opportunity lost because of scarcity of resources. The concept of opportunity cost is related to the alternative uses of scarce resources. As noted earlier, resources, both natural and man-made, are scarce in relation to demand for them to satisfy the ever-growing human needs. Resources, though scarce, have alternative uses. The scarcity and the alternative uses of the resources give rise to the concept of opportunity cost.

In the context of a business firm, resources available to a business unit—be it an individual firm, a joint stock corporation or a multinational—are limited. But the limited resources available to a firm can be put to alternative uses.

The difference between actual earning and its opportunity cost is called economic gain or economic profit. The concept of opportunity cost assumes a great significance where economic gain is neither insignificant nor very large because then it requires a careful evaluation of the two alternative options.

The applicability of the opportunity cost concept is not limited to decisions on the use of financial resources. The concept can be applied to all other kinds of issues involved in business decisions, especially where there are at least two alternative options involving costs and benefits. For example, suppose a firm has to take a decision on whether to fire an efficient labour officer (for treating labour unkindly) in settlement of a dispute with the labour union or to allow the matter to be taken to the labour court. If the firm decides to fire the labour officer, then the loss of an efficient and reliable labour officer is the opportunity cost of buying peace with the labour union. If the firm decides to retain the labour officer, come whatever may, then the cost of prolonged litigation, the cost arising out of a possible labour strike and the consequent reduction in output are the opportunity costs of retaining the labour officer. Given the two options, the firm will have to evaluate the cost and benefit of each option and take a decision accordingly.

2.3.2 Production Possibility Curve

Societies cannot have all that they want because resources are scarce and technology is given. In reality, however, both human and non-human resources available to a country keep increasing over time and technology becoming more and more efficient and productive. Availability of human resources increases due to a natural process of increase in population, and non-human resources (especially capital goods and raw materials) increase due to creative nature of human beings. Non-human resources have been increasing due to human efforts to create more and better of capital goods, to discover new kinds and sources of raw materials,
and to create a new and more efficient technique of production. Such factors bring about a change in production possibilities and production possibilities frontier of an economy.

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

The Production Possibilities Frontier

Production possibilities refer to the alternative combinations of goods and services that a society is capable of producing with its given resources and state of technology. With reference to our model specified above, production possibilities are the alternative combinations of maximum food and clothing that the country can produce by making full use of its labour and capital, given the technology. For example, let us suppose that, given the availability of labour, capital and technology, the alternative production possibilities open to the country are given in Table 2.1. These production possibilities given in Table 2.1 can be presented in the form of a diagram as shown in Fig. 2.1. In this diagram, vertical axis measures food production and horizontal axis measures production of clothing. By graphing the alternative production possibilities given in Table 2.1, we locate points A, B, C, and E shown in Fig. 2.1. A number of intermediate points can be located between any two of these points. By joining these points, we get a curve PF. This curve is called production possibilities frontier (PPF). The production possibilities frontier (PF) shows all the alternative combinations of two goods (food and clothing) that can be produced by making full use of all the available resources (labour and capital), given the state of technology. Each point on the PPF shows a different combination of two goods. For example, shows that if the country chooses point P on the production possibilities frontier, PF, it can produce 8 thousand tons of food and no clothing.

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

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

**Fig 2.1 The Production Possibilities Frontier**

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

**2.3.3 Incremental Concepts**

The concept of ‘incremental’ value is similar to the concept of ‘marginal’ value but with a difference. While ‘marginal’ concept is basically a theoretical concept, ‘incremental’ is a practical concept. Marginal principle can be applied only when change in quantity or value, e.g., $MC$ and $MR$, can be calculated precisely in unit terms. In general, however, firms find it difficult to estimate $MC$ and $MR$ as defined conceptually. The reason is that most business firms produce and sell their products in bulk, not in terms of units unless, of course, it is the case of production and sale of such large-unit goods as aeroplanes, ships, large buildings, turbines, etc. Where
production and sale activities are carried out in bulk and where both fixed and variable costs are subject to change, business managers use the incremental principle in their business decisions.

The incremental principle is applied to business decisions which involve bulk production and a large increase in total cost and total revenue. Such an increase in total cost and total revenue is called "incremental cost" and "incremental revenue" respectively, related to "incremental output".

Let us first explain the concept of incremental cost. Conceptually, incremental costs can be defined as the costs that arise due to a business decision. Thus, an increase in the total cost of production due to a business decision is incremental cost.

The incremental revenue, on the other hand, is the increase in revenue due to a business decision, e.g., a decision to increase production and sale of the firm’s product. When a business decision is successfully implemented, it does result in a significant increase in its total revenue. The increase in the total revenue resulting from a business decision is called incremental revenue.

Incremental Reasoning in Business Decision

The use of the incremental concept in business decisions is called incremental reasoning. The incremental reasoning is used for accepting or rejecting a business proposition or option.

It may be added at the end, by way of comparison, that the marginal concept (especially when defined and measured by calculus) is used in economic analysis where a high degree of precision is involved, whereas the incremental concept is used where large values, especially of cost and revenue are involved. Besides, incremental concept and reasoning are used in business decisions more frequently than the marginal concept. There are at least two reasons for this. First, marginal concept used in business analysis is generally associated with one (marginal) unit of output produced or sold whereas most business decisions involve large quantities and values. Second, the precise calculation of marginal change (defined in terms of the first derivative of a function) is neither practicable nor necessary in real life business decisions. However, marginal concept is of great significance in theoretical analysis.

2.3.4 Cardinal and Ordinal Approaches to Consumer Behaviour

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

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

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

On the other hand, from a consumer’s point of view, utility is a post-consumption phenomenon as one derives satisfaction from a commodity only when one consumes or uses it. Utility in the sense of satisfaction is a ‘subjective’ or ‘relative’ concept because (i) a commodity need not be useful for all—cigarettes do not have any utility for non-smokers, and meat has no utility for strict vegetarians; (ii) utility of a commodity varies from person to person and from time to time; and (iii) a commodity need not have the same utility for the same consumer at different points of times, at different levels of consumption and for different moods of a consumer. In consumer analysis, only the ‘subjective’ concept of utility is used.

Having explained the concept of utility, we now turn to some quantitative concepts related to utility used in utility analysis, viz. total utility and marginal utility.

Cardinal Utility

Some early psychological experiments on an individual’s responses to various stimuli led neo-classical economists to believe that utility is measurable and cardinally quantifiable. According to neo-classical economists, utility can be measured in terms of money. That is, utility of a unit of a commodity for a person is equal to the amount of money he is willing to pay for it. This belief gave rise to the concept of cardinal utility. It implies that utility can be assigned a cardinal number like 1, 2, 3, etc. Neo-classical economists built up the theory of consumption on the assumption that utility is cardinaly measurable. They coined and used a term ‘util’ meaning ‘units of utility’. In their measure of utility, they assumed (i) that one ‘util’ equals one unit of money, and (ii) that utility of money remains constant.
It has, however, been realized over time that absolute or cardinal measurement of utility is not possible. Difficulties in measuring utility have proved to be insurmountable.

**Ordinal Utility**

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

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

**Two Approaches to the Consumer Demand Analysis**

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

(i) **Cardinal Utility Approach**, attributed to Alfred Marshall and his followers, is also called the Neo-classical Approach.

(ii) **Ordinal Utility Approach**, pioneered by J.R. Hicks, a Nobel laureate, and R.G.D. Allen, is also called the Indifference Curve Analysis.

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

**2.3.5 Time Value of Money**

All business decisions are taken with a certain time perspective. The time perspective refers to the duration of time period extending from the relevant past and foreseeable future taken in view while taking a business decision. Relevant past refers to the period of past experience and trends which are relevant for business decisions with long-run implications. All business decisions do not have
the same time perspective. Some have short-run outcome or pay-off and, therefore, involve short-run time perspective. For example, a decision to buy explosive materials for manufacturing crackers involves short-run demand prospects. Similarly, a decision regarding building inventories of finished product involves a short-run time perspective. There are, however, a large number of business decisions which have long-run repercussions, e.g., investment in plant, building, machinery, land, spending on labour welfare activities, expansion of the scale of production, introduction of a new product, advertisement, bribing a government officer and investment abroad. The decision about such business issues, for example, the decision regarding the introduction of a new product may not be profitable in the short-run but may prove very profitable in the long-run.

The business decision-makers must assess and determine the time perspective of business propositions well in advance and make decisions accordingly. Determination of time perspective is of great significance especially where projections are involved. The decision-makers must decide on an appropriate future period for projecting the value of a variable. Otherwise, projections may prove meaningless from analysis point of view and decisions based thereon may result in poor pay-offs. For example, in a business decision regarding the establishment of a Management Institute, projecting a short-run demand and taking a short-run time perspective will be unwise, and in buying explosive materials for manufacturing crackers for Deepawali, a long-run time perspective is unwise.

### Check Your Progress

1. Define opportunity cost.
2. Name the two approaches used to analyse consumer behaviour.

### 2.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Opportunity cost is a term in economics that refers to the value of what you have to give up in order to choose something else.
2. The two approaches used to analyse consumer behavior are the following:
   - Cardinal utility approach
   - Ordinal utility approach

### 2.5 SUMMARY

- Managerial decisions are taken at different levels of sophistication. While some business decisions require only ‘rule-of-thumb’ technique, some others involve the use of sophisticated techniques.
Certainty is the state of perfect knowledge about the market conditions. In the state of certainty, there is only one rate of return on the investment and that rate is known to the investors.

From the business decision-making point of view, risk refers to a situation in which a business decision is expected to yield more than one outcome and the probability of each outcome is known to the decision makers or it can be reliably estimated.

Uncertainty refers to a situation in which there is more than one outcome of a business decision and the probability of no outcome is known nor can it be reliably estimated.

The scarcity and the alternative uses of the resources give rise to the concept of opportunity cost.

The difference between actual earning and its opportunity cost is called economic gain or economic profit. The concept of opportunity cost assumes a great significance where economic gain is neither insignificant nor very large because then it requires a careful evaluation of the two alternative options.

Production possibilities refer to the alternative combinations of goods and services that a society is capable of producing with its given resources and state of technology.

The incremental revenue is the increase in revenue due to a business decision, e.g., a decision to increase production and sale of the firm’s product.

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

From consumer’s angle, utility is the psychological feeling of satisfaction, pleasure, happiness or well-being, which a consumer derives from the consumption, possession or the use of a commodity.

The cardinal and ordinal concepts of utility arise out of question whether “utility is measurable”.

All business decisions are taken with a certain time perspective. The time perspective refers to the duration of time period extending from the relevant past and foreseeable future taken in view while taking a business decision.

The business decision-makers must assess and determine the time perspective of business propositions well in advance and make decisions accordingly.

2.6 KEY WORDS

- **Risk**: It means a low probability of an expected outcome.
• **Uncertainty**: It refers to a situation in which there is more than one outcome of a business decision and the probability of no outcome is known nor can it be reliably estimated.

• **Incremental costs**: These are costs that arise due to a business decision.

### 2.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

**Short Answer Questions**

1. What is the significance of opportunity cost?
2. Define economic profit.
3. Write a short note on the time value of money.

**Long Answer Questions**

1. What is the role of managerial economics in decision making?
2. Explain the production possibilities frontier with the help of an example.
3. Describe the two approaches used to analyse consumer behaviour.

### 2.8 FURTHER READINGS


**Websites**

https://blogs.economictimes.indiatimes.com/et-commentary/heres-how-indias-widening-income-distribution-can-be-redressed/

UNIT 3 CONSUMER BEHAVIOUR

3.0 INTRODUCTION

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

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

These questions are answered by the analysis of consumer behaviour. The study of consumer behaviour gives rise to the Theory of Consumer Demand. In this unit, you will study about the concept of marginalism, equilibrium and...
equi-marginalism and their role in business decision-making, utility analysis, cardinal and ordinal utility approaches and the properties of an indifference curve.

### 3.1 OBJECTIVES

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

- Define marginalism and equilibrium
- State the equi-marginal principle
- Discuss total and marginal utility
- Explain the Marshallian approach
- Describe the properties of indifference curve

### 3.2 MARGINALISM, EQUILIBRIUM AND EQUI-MARGINALISM AND THEIR ROLE IN BUSINESS DECISION MAKING

The concept of ‘marginal’ value is widely used in economic analysis, for example *marginal utility in consumer analysis, marginal productivity in production analysis and marginal revenue and marginal cost in pricing theory*. Marginality concept assumes special significance where maximization or minimization problem is involved, for example, maximization of a consumer’s utility, maximization of a firm’s profit, minimization of cost, etc. The term ‘marginal’ refers to the change (increase or decrease) in total quantity or value due to a one-unit change in its determinant. For example, given the factor prices, the total cost of production of a commodity depends on the number of units produced. In this case, ‘marginal cost’ ($MC$) can be defined as the rise in total cost as a result of producing one additional unit of a commodity. The marginal cost ($MC$) can be worked out as follows.

$$
MC = MC_n - MC_{n-1}
$$

where $MC_n = \text{total cost of producing } n \text{ units}$ and $MC_{n-1} = \text{total cost of producing } n-1 \text{ units}$.

For example, suppose total cost ($TC$) of producing 100 units of a commodity is $`2500$. When production is increased to 101 units $TC$ increases to $`2550$. In this case, $TC_n = `2550$, $TC_{n-1} = `2500$ (where $n = 101$ and $n-1 = 100$).

Then

$$
MC = TC_n - TC_{n-1} = `2550 - `2500 = `50
$$

Similarly, $MR$ can be defined as the change in $TR$ due to the sale of one additional unit of a product. It can also be defined as:

$$
Marginal \text{ revenue } (MR) = TR_n - TR_{n-1},
$$
where \( TR_n \) = total revenue from the sale of \( n \) units and \( TR_{n-1} \) = total revenue from the sale of \( n-1 \) units.

Alternatively, if \( TC \) and \( TR \) are given in the form of functions, then \( MC \) and \( MR \) are defined as the \textit{first derivatives} of the \( TC \) and \( TR \) functions, respectively. Suppose \( TC \) and \( TR \) functions are given as

\[
TC = f(Q) \quad \text{and} \quad TR = f(Q)
\]

Then the first derivative is calculated through calculus as follows.

\[
MC = \frac{\partial TC}{\partial Q} \quad \text{and} \quad MR = \frac{\partial TR}{\partial Q}
\]

If \( TC \) and \( TR \) are given in the form of curves, then \( MC \) and \( MR \) are defined in terms of the \textit{slopes} of the \( TR \) and \( TC \) curves. Marginal value derived from a function in the form of its first derivative is not the same as defined in terms of a unit. More on this aspect of the ‘marginal’ concept will follow in a subsequent section.

The decision rule: Suppose a profit-maximizing firm is faced with a problem—how much to produce so that profit is maximized. One simple decision rule under the marginal principle is that a business activity (production and sale) must be carried out so long as its \( MR > MC \). As regards profit maximizing output, economists use the marginality principle to set a necessary condition for profit maximizing output. The necessary condition for profit maximizing output is that \( MC \) must be equal to \( MR \). That is, profit is maximum where

\[
MR = MC
\]

In simple words, the profit of a firm is maximized at that level of output and sale where the cost of producing one additional unit equals the revenue from the sale of that unit of output.

The application of the marginal principle for profit maximization has certain serious \textit{limitations} which must be borne in mind.

One, it can be applied only where the management has the \( TC \) and \( TR \) data for each and every unit of output or where the management is fully aware of the cost of producing one additional unit and the price expected to be received from the sale of that unit.

Two, the concept of ‘marginal’ value, when used in cost analysis, reduces the value of \( MC \) to the change in \textit{variable cost} only. Therefore, marginal analysis can be applied to a situation in which only the variable cost changes.

3.2.1 Equi-marginal Principles

The equi-marginal principle was originally associated with consumption theory and the law is called ‘the law of equi-marginal utility’. The law of equi-marginal
utility states that a utility maximizing consumer distributes his consumption expenditure between various goods and services he/she consumes in such a way that the marginal utility derived from each unit of expenditure on various goods and service is the same. This pattern of distribution of consumption expenditure maximizes a consumer’s total utility.

The law of equi-marginal principle was over time applied by business managers to allocation of resources between their alternative uses with a view to maximizing profit in case a firm carries out more than one business activity. This principle suggests that available resources (inputs) should be so allocated between the alternative options that the marginal productivity gains (MP) from the various activities are equalized. For example, suppose a firm has a total capital worth `100 million which it has the option of spending on three projects, A, B and C. Each of these projects requires a unit expenditure of `10 million. Suppose also that the marginal productivity schedule of each unit of expenditure on the three projects is given as shown in Table 3.1.

<table>
<thead>
<tr>
<th>Units of Expenditure (‘10 million)</th>
<th>Marginal Productivity (MP) (‘ in million)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Project A</td>
</tr>
<tr>
<td>1st</td>
<td>50</td>
</tr>
<tr>
<td>2nd</td>
<td>45</td>
</tr>
<tr>
<td>3rd</td>
<td>35</td>
</tr>
<tr>
<td>4th</td>
<td>20</td>
</tr>
<tr>
<td>5th</td>
<td>10</td>
</tr>
</tbody>
</table>

Note: Subscripts 1, 2, 3… indicate the order of the unit of expenditure on Projects A, B and C.

Going by the equi-marginal principle, the firm will allocate its total resources (‘100 million) among the projects A, B and C in such a way that marginal product of each project is the same, i.e., $MP_A = MP_B = MP_C$. It can be seen from Table 3.1, that, going by this rule, the firm will spend 1st, 2nd, 7th, and 10th unit of finance on Project A, 3rd, 5th, and 8th unit on Project B, and 4th, 6th and 9th unit on Project C. In all, it puts 4 units of its finances in Project A, 3 units each in Projects B and C. In other words, of the total finances of `100 million, a profit maximizing firm would invest `40 million in Project A, `30 million each in Projects B and C. This pattern of investment maximizes the firm’s productivity gains. No other pattern of investment will ensure this objective.

Now the equi-marginal principle can be formally stated. It suggests that a profit (gain) maximizing firm allocates its resources in a proportion such that

\[ MP_A = MP_B = MP_C = \ldots = MP_N \]
If cost of project \((COP)\) varies from project to project, then resources are so allocated that \(MP\) per unit of \(COP\) is the same. That is, resources are allocated in such proportions that

\[
\frac{MP_1}{COP_1} = \frac{MP_2}{COP_2} = \frac{MP_3}{COP_3} = \ldots = \frac{MP_n}{COP_n}
\]

The equi-marginal principle can be applied only where (i) firms have limited investible resources, (ii) resources have alternative uses, and (iii) the investment in various alternative uses is subject to diminishing marginal productivity or returns.

### 3.3 UTILITY ANALYSIS

Let us go through the various economic terms associated with utility analysis.

#### 3.3.1 Total and Marginal Utility

Assuming that utility is measurable and additive, total utility may be defined as the sum of the utility derived by a consumer from the various units of a good or service he consumes at a point or over a period of time. Suppose a consumer consumes four units of a commodity, \(X\), at a time and derives utility from the successive units of consumption as \(u_1, u_2, u_3\), and \(u_4\). His total utility \((U_X)\) from commodity \(X\) can be measured then as follows.

\[
U_X = u_1 + u_2 + u_3 + u_4
\]

If a consumer consumes a large number of commodities, say, \(n\) number of commodities, his total utility, \(TU\), is the sum of the total utility derived from each commodity. For instance, if the consumption goods are \(X, Y\) and \(Z\) and their total respective utilities are \(U_X, U_Y\) and \(U_Z\), then

\[
TU = U_X + U_Y + U_Z
\]

**Marginal Utility**

Marginal utility is another very important concept used in economic analysis. Marginal utility can be defined as the utility derived from the marginal or one additional unit consumed. It may also be defined as the addition to the total utility resulting from the consumption of one additional unit. Marginal Utility \((MU)\) thus refers to the change in the Total Utility \((i.e., \Delta TU)\) obtained from the consumption of an additional unit of a commodity, say \(X\). It may be expressed as

\[
MU = \frac{\Delta TU}{\Delta Q}
\]

where \(TU\) = total utility, and \(\Delta Q\) = change in quantity consumed by one unit.

Another way of expressing marginal utility \((MU)\), when the number of units consumed is \(n\), can be as follows.

\[
MU\ of\ \text{\(n\)th\ unit} = TU_n - TU_{n-1}
\]
3.4 CARDINAL UTILITY APPROACH OR THE MARSHALLIAN APPROACH

In the preceding unit, you were briefly introduced to the concept of cardinal utility and ordinal utility approaches to consumer behaviour. In this section, you will study the cardinal utility approach also known as the Marshallian approach in detail with the help of tables and figures.

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

(i) how does a consumer decide the optimum quantity of a commodity that he or she chooses to consume, i.e., how does a consumer attain his/her equilibrium in respect to each commodity?

(ii) how does he or she allocate his/her disposable income between various commodities of consumption so that his/her total utility is maximized?

The theory of consumer behaviour based on the cardinal utility approach seeks to answer the above questions on the basis of the following assumptions.

Assumptions

(i) **Maximization of satisfaction:** Every rational consumer intends to maximize his/her satisfaction from his/her given money income.

(ii) **Rationality:** It is assumed that the consumer is a rational being in the sense that he or she satisfies his/her wants in the order of their preference. That is, he or she buys that commodity first which is expected to yield the highest utility and the last which is likely to the least utility.

(iii) **Limited money income:** The consumer has a limited money income to spend on the goods and services he or she chooses to consume. Limitedness of income, along with utility maximization objective makes the choice between goods inevitable.

(iv) **Utility is cardinally measurable:** The cardinalists have assumed that utility is cardinally measurable and that utility of one unit of a commodity equals the units of money a consumer is prepared to pay for it, i.e., 1 util = 1 unit of money.

(v) **Diminishing marginal utility:** The consumption of a commodity is subject to the law of diminishing marginal utility. That is, it is assumed that the utility gained from the successive units of a commodity consumed decreases as a person consumes them. This is an axiom of the theory of consumer behaviour.
(vi) **Constant marginal utility of money**: The cardinal utility approach assumes that marginal utility of money remains constant whatever the level of a consumer’s income.

(vii) **Utility is additive**: Cardinalities utility approach assumes the cardinality measurable utility is additive. It implies that utility derived from each unit of commodity and from different commodities can be added numerically. For example, suppose a consumer consumes three units of a commodity and \( MU \) derived from each unit is given as \( U_1, U_2, \) and \( U_3 \). In that case,

\[
\text{Total Utility (TU)} = U_1 + U_2 + U_3.
\]

### 3.4.1 Law of Diminishing Marginal Utility

The law of diminishing marginal utility is one of the fundamental laws of economics pertaining to the analysis of consumer behaviour. This law states that as the quantity consumed of a commodity goes on increasing, the utility derived from each successive unit consumed goes on decreasing, consumption of all other commodities remaining constant. In simple words, when a person continues to consume more and more units of a commodity at a point of time, e.g., chocolates, the utility that he derives from each successive chocolate goes on diminishing. The law of diminishing marginal utility is illustrated below numerically.

To illustrate the law of diminishing marginal utility, let us assume that a consumer consumes only one commodity \( X \), and that utility is measurable in quantitative terms. Let us also suppose that total and marginal utility schedules of commodity \( X \) are given as in Table 3.2. The law of diminishing marginal utility is illustrated numerically in Table 3.2 and graphically in Fig. 3.1.

<table>
<thead>
<tr>
<th>Units of commodity ( X )</th>
<th>Total utility ( (TU) )</th>
<th>Marginal utility ( (MU) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>60</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>65</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>60</td>
<td>– 5</td>
</tr>
<tr>
<td>6</td>
<td>45</td>
<td>– 15</td>
</tr>
</tbody>
</table>

As shown in Table 3.2, with the increase in the number of units of commodity \( X \) consumed per unit of time, \( TU \) increases but at a diminishing rate. The diminishing \( MU \) is shown in the last column of Table 3.2. Figure 3.1 illustrates graphically the law of diminishing \( MU \). The rate of increase in \( TU \) as a result of increase in the number of units consumed is shown by the \( MU \) curve in Figure 3.1. The downward sloping \( MU \) curve shows that marginal utility goes on decreasing as consumption increases. At 4 units consumed, the \( TU \) reaches its maximum level, the point of saturation marked by point \( M \). Beyond this point, \( MU \) becomes negative and
TU begins to decline. The downward sloping MU curve illustrates the law of diminishing marginal utility.

**Fig. 3.1 Total and Diminishing Marginal Utility of Commodity X**

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

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

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

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

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

**Fourth,** the mental condition of the consumer must remain normal during the period of consumption. A person drinking whisky may feel a greater pleasure with successive pegs because of change in his mental status due to intoxication.

Given these conditions, the law of diminishing marginal utility holds universally. In some cases, e.g., accumulation of money, collection of hobby items like stamps, old coins, rare paintings and books, melodious songs, etc., the marginal utility may initially increase rather than decrease. But eventually it does decrease. As a matter of fact, the law of diminishing marginal utility operates generally under normal conditions.
It concluded that as price decreases, demand increases. This inverse relationship between demand and price gives rise to demand curve.

### 3.5 ORDINAL UTILITY APPROACH OR THE HICKS APPROACH

In the preceding section, we have discussed the consumer behaviour based on *cardinal utility approach*. In this section, we proceed to discuss consumer behaviour on the basis of *ordinal utility approach*. The cardinal utility approach is based on the assumption that utility is cardinally or numerically measurable while ordinal utility approach is based on the assumption that utility is measurable only ordinally, not cardinally or numerically. In other words, utility can be expressed only comparatively. For example, a consumer consuming two goods, X and Y, can express his feeling of utility as utility of X is ‘higher than’ or ‘lower than’ that of Y. Although the concept of ‘ordinal utility’ was evolved much earlier, the analysis of consumer behaviour based on ordinal utility concept was developed by J.R. Hicks and R.G.D. Allen in 1934. The *ordinal utility approach* is also known as Hicks-Allen Approach. However, the basic theory of consumer behaviour was developed by J. R. Hicks. In order to analyse consumer behaviour, Hicks used a new tool of analysis called *Indifference curve*. In this section, we will discuss consumer behaviour following the ordinal utility approach. We will first explain the ‘indifference curve’, the tool of analysis and then proceed to analyse consumer behaviour through the indifference curve.

**Assumptions of Ordinal Utility Theory**

1. **Rationality**: The consumer is assumed to be a rational being. Rationality means that a consumer aims at maximizing his total satisfaction given his income and prices of the goods and services he consumes and his decisions are consistent with this objective.

2. **Ordinal utility**: Indifference curve analysis is based on the assumption that utility is only ordinally expressible. That is, the consumer is only able to express the order of his preference for different baskets of goods.

3. **Transitivity and consistency of choice**: Consumer’s choices are assumed to be transitive. *Transitivity* of choice means that if a consumer prefers $A$ to $B$ and $B$ to $C$, he must prefer $A$ to $C$. Or, if he treats $A = B$ and $B = C$, he must treat $A = C$. *Consistency* of choice means transitivity of choice over a period of time. For example, it means that if a consumer prefers $A$ to $B$ in one period, he does not prefer $B$ to $A$ in another period or even treat them as equal.

4. **Nonsatiety**: It is also assumed that the consumer has not reached the point of saturation in case of any commodity. Therefore, a consumer always prefers a larger quantity of all the goods.
5. **Diminishing marginal rate of substitution**: The consumer consumes substitutable goods and can substitute one good for another. The marginal rate of substitution is the rate at which a consumer is willing to substitute one commodity (X) for another (Y) so that his total satisfaction remains the same. The marginal rate of substitution is given as $\Delta Y / \Delta X$. The ordinal utility approach assumes that $\Delta Y / \Delta X$ goes on decreasing when a consumer continues to substitute X for Y.

### 3.5.1 Indifference Curve Analysis

An indifference curve may be defined as the *locus of points each representing a different combination of two substitute goods, which yield the same utility or level of satisfaction to the consumer*. Therefore, he is indifferent between any two combinations of two goods when it comes to making a choice between them. Such a situation arises because he consumes a large number of goods and services and often finds that one commodity can be substituted for another. The consumer can, therefore, substitute one commodity for another, and can make various combinations of two substitute goods which give him the same level of satisfaction. Since each combination yields the same level of satisfaction, he would be indifferent between the combinations when he has to make a choice. When such combinations are plotted graphically, it produces a curve. This curve is called the indifference curve. An indifference curve is also called iso-utility curve or equal utility curve.

For example, let us suppose that a consumer consumes two goods, X and Y, and he makes five combinations $a$, $b$, $c$, $d$ and $e$ of the two substitute commodities, X and Y, as presented in Table 3.3. All these combinations yield the same level of satisfaction.

**Table 3.3 Indifference Schedule of Commodities X and Y**

<table>
<thead>
<tr>
<th>Combination</th>
<th>Units of Commodity Y</th>
<th>Units of Commodity X</th>
<th>Total Utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>25</td>
<td>3</td>
<td>$U$</td>
</tr>
<tr>
<td>$b$</td>
<td>15</td>
<td>5</td>
<td>$U$</td>
</tr>
<tr>
<td>$c$</td>
<td>8</td>
<td>9</td>
<td>$U$</td>
</tr>
<tr>
<td>$d$</td>
<td>4</td>
<td>17</td>
<td>$U$</td>
</tr>
<tr>
<td>$e$</td>
<td>2</td>
<td>30</td>
<td>$U$</td>
</tr>
</tbody>
</table>

Table 3.3 is an indifference schedule—a schedule of various combinations of two goods—X and Y—yielding the same level of utility. Therefore, the consumer is indifferent between the combinations. The last column of the table shows an undefined utility ($U$) derived from each combination of X and Y. The combinations $a$, $b$, $c$, $d$ and $e$ given in Table 3.3 are plotted and joined by a smooth curve (as shown in Fig. 3.2). The resulting curve is known as *indifference curve*. On this curve, one can locate many other points between any two points showing different combinations of X and Y which yield the same level of satisfaction. Therefore, the
consumer is indifferent between the combinations which may be located on the indifferent curve.

**Indifference Map:** We have drawn a single indifference curve in Fig. 3.2 on the basis of the indifference schedule given in Table 3.3. The combinations of the two commodities, X and Y, given in the indifference schedule or those indicated by the indifference curve are by no means the only combinations of the two commodities. The consumer may make many other combinations with less of one or both of the goods—each combination yielding the same level of satisfaction but less than the level of satisfaction indicated by the indifference curve $IC$ in Fig. 3.2. As such, an indifference curve below the one given in Fig. 3.2 can be drawn, say, through points $f, g$ and $h$. Similarly, the consumer may make many other combinations with more of one or both the goods—each combination yielding the same satisfaction but greater than the satisfaction indicated by $IC$. Thus, another indifference curve can be drawn above $IC$, say, through points $j, k$ and $l$. This exercise may be repeated as many times as one wants, each time generating a new indifference curve.

Fig 3.2 Indifference Curve

In fact, the space between X and Y axes is known as the *indifference plane* or *commodity space*. This plane is full of finite points and each point on the plane indicates a different combination of goods X and Y. Intuitively, it is always possible to locate any two or more points on the indifference plane indicating different combinations of goods X and Y yielding the same level of satisfaction. It is thus possible to draw a number of indifference curves without intersecting or being tangent to one another as shown in Fig. 3.3. The set of indifference curves $IC_1, IC_2, IC_3$ and $IC_4$ drawn in this manner make the *indifference map*. In fact, an indifference map may contain any number of indifference curves, ranked in the order of consumer’s preferences.
Marginal Rate of Substitution (MRS)

An indifference curve is formed by substituting one good for another. The rate at which one good is substituted for another is called Marginal Rate of Substitution (MRS). In other words, the MRS is the rate at which one commodity can be substituted for another, the level of satisfaction remaining the same. The MRS between two commodities X and Y, may be defined as the quantity of X which is required to replace one unit of Y (or quantity of Y required to replace one unit of X) in the combination of the two goods so that the total utility remains the same. This implies that the utility of X (or Y) given up is equal to the utility of additional units of Y (or X) added to the combination. The MRS is expressed as $\frac{\Delta Y}{\Delta X}$, moving down the curve.

MRS goes on Diminishing: The basic postulate of ordinal utility theory is that if a consumer goes on substituting one good for another, the MRS goes on decreasing. It means that the quantity of a commodity that a consumer is willing to sacrifice for an additional unit of another goes on decreasing when he goes on substituting one commodity for another. The diminishing marginal rate of substitution causes the indifference curves to be convex to the origin.

Why Does MRS Diminish? The MRS diminishes along the IC curve because, in most cases, no two goods are perfect substitutes for one another. In case any two goods are perfect substitutes, the indifference curve will be a straight line having a negative slope showing constant MRS. Since goods are not perfect substitutes, the subjective value attached to the additional quantity (i.e., subjective MU) of a commodity decreases fast in relation to the other commodity whose total quantity is decreasing. Therefore, when the quantity of one commodity (X) increases and that of the other (Y) decreases, the subjective MU of Y increases and that of X decreases. Therefore, the consumer becomes increasingly unwilling...
to sacrifice more units of $Y$ for one unit of $X$. But, if he is required to sacrifice additional units of $Y$, he will demand increasing units of $X$ to maintain the level of his satisfaction. As a result, the $MRS$ decreases.

### Properties of Indifference Curve

Indifference curves have the following four basic properties:

1. Indifference curves slope downward to right;
2. Indifference curves of imperfect substitutes are convex to the origin;
3. Indifference curves do not intersect nor are they tangent to one another;
4. Upper indifference curves indicate a higher level of satisfaction.

These properties of indifference curves, in fact, reveal the consumer’s behaviour, his choices and preferences. They are, therefore, very important in the modern theory of consumer behaviour. Let us now look at their implications.

#### 1. Indifference Curves Slope Downward to Right.

In the words of Hicks, “... so long as each commodity has a positive marginal utility, the indifference curve must slope downward to the right”, as shown in Fig. 3.3. The negative slope of an indifference curve implies two requisits: (a) that the two commodities can be substituted for each other; and (b) that if the quantity of one commodity decreases, quantity of the other commodity must so increase that the consumer stays at the same level of satisfaction. The conditions make the indifference curve slope downward to the right. In case two goods are perfect substitutes, then the change in the combination of two goods produces an indifference line, not a curve.

#### 2. Indifference Curves are Convex to Origin.

Indifference curves are not only negatively sloped, but are also convex to the origin as shown in Fig 3.3. The convexity of the indifference curves is caused by two factors:

(i) the two commodities are imperfect substitutes for one another, and

(ii) the marginal rate of substitution ($MRS$) between the two goods decreases as a consumer moves along an indifference curve.

The postulate of diminishing $MRS$, is based on the fact that if a consumer substitutes one commodity ($X$) for another ($Y$), his willingness to sacrifice more units of $Y$ for one additional unit of $X$ decreases, as quantity of $Y$ decreases. There are two reasons for diminishing $MRS$: (i) two commodities are not, in general, perfect substitutes for one another, and (ii) $MU$ of a commodity increases as its quantity decreases and vice versa, and, therefore, more and more units of the other commodity are needed to keep the total utility constant.

#### 3. Indifference Curves Neither Intersect Nor are Tangent with One Another.

If two indifference curves intersect or are tangent with one another, it will yield two impossible conclusions: (i) that two equal combinations of two goods yield two different levels of satisfaction, and (ii) that two different combinations—one being larger than the other—yield the same level of satisfaction. Such conditions
are impossible if the consumer’s subjective valuation of a commodity is greater than zero.

4. Upper Indifference Curves Represent a Higher Level of Satisfaction than the Lower Ones. An indifference curve placed above and to the right of another represents a higher level of satisfaction than the lower one. The reason is that an upper indifference curve contains all along its length a larger quantity of one or both the goods than the lower indifference curve. And a larger basket of commodities is supposed to yield a greater satisfaction than the smaller one, provided MU of goods is greater than zero.

Budgetary Constraints on Consumer’s Choice: Limited Income and Prices

Given the indifference map, a utility maximizing consumer would like to reach the highest possible indifference curve on his indifference map. But the consumer has two strong constraints: (i) he has a limited income, and (ii) he has to pay a price for the goods. Given the prices, the limitedness of income acts as a constraint on how high a consumer can ride on his indifference map. This is known as budgetary constraint. In a two-commodity model, the budgetary constraint may be expressed through a budget equation as

\[ P_x \cdot Q_x + P_y \cdot Q_y = M \]

where \( P_x \) and \( P_y \) are prices of goods \( X \) and \( Y \) respectively; \( Q_x \) and \( Q_y \) are their respective quantities; and \( M \) is the consumer’s money income.

The budget equation states that the total expenditure of the consumer on goods \( X \) and \( Y \) cannot exceed his total income, \( M \). The quantities of \( X \) and \( Y \) that a consumer can buy, given his income (\( M \)) and prices, \( P_x \) and \( P_y \), can be easily obtained from the budget equation, as shown below.

\[ Q_x = \frac{M}{P_x} - \frac{P_y}{P_x} \cdot Q_y, \]
\[ Q_y = \frac{M}{P_y} - \frac{P_x}{P_y} \cdot Q_x. \]

Now, \( Q_x \) or \( Q_y \) may be alternatively assigned any positive numerical value and the corresponding values of \( Q_y \) and \( Q_x \) may be obtained. When the values of \( Q_x \) and \( Q_y \) are plotted on the \( X \) and \( Y \) axes, we get a line with a negative slope, which is called the budget line or price line.

The budget line shows the alternative options of commodity combinations available to the consumer given his income and the prices of \( X \) and \( Y \).

Shifts in the Budget Line

The budget line is drawn on the basis of the budget equation, given as \( M = P_x \cdot Q_x + P_y \cdot Q_y \). Any change in the parameters of the budget equation, \( \text{viz.}, M, P_x \) and \( P_y \), make the budget line shift upward or downward or swivel left or right and up or down. If consumer’s income (\( M \)) increases, prices remaining constant, the budget line...
line shifts upwards remaining parallel to the original budget line. Suppose the original budget line is given by line \( AB \) in Fig. 3.4. If \( M \) increases (prices remaining the same), the budget line \( AB \) will shift to \( CD \). And, if \( M \) decreases by the same amount, the budget line will shift backward to its original position \( AB \). Income remaining the same, if prices change, the budget line changes its position and slope. For example, if \( M \) and \( P_x \) remain constant and \( P_y \) decreases to half then the budget line will be \( AF \). Similarly, \( M \) and \( P_y \) remaining constant, if \( P_x \) increases, the budget line shifts to \( EB \).

**Slope of the Budget Line**

Another important aspect of the budget line that matters in determining a consumer’s equilibrium is its *slope*. The slope of the budget line \((AB)\) in Fig. 3.4, is given as

\[
\frac{\Delta Q_y}{\Delta Q_x} = \frac{OA}{OB}
\]

Since \( OA = M/P_y \) (at \( X = 0 \)) and \( OB = M/P_x \) (at \( Y = 0 \)), the slope of the budget line \( AB \) in Fig. 3.4 may be rewritten as

\[
\frac{OA}{OB} = \frac{M/P_y}{M/P_x} = \frac{P_x}{P_y}
\]

**Fig. 3.4 Shift in the Budget Space**

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

**Check Your Progress**

1. Define marginal cost.
2. Who developed the ordinal utility approach?
3. What is an indifference curve?
3.6 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Marginal cost (MC) can be defined as the rise in total cost as a result of producing one additional unit of a commodity.

2. The ordinal utility approach was developed by J.R. Hicks and R.G.D. Allen in 1934. The ordinal utility approach is also known as Hicks-Allen Approach.

3. An indifference curve may be defined as the locus of points each representing a different combination of two substitute goods, which yield the same utility or level of satisfaction to the consumer.

3.7 SUMMARY

- Consumer demand is the basis of all productive activities. Just as 'necessity is the mother of invention', demand is the mother of production.
- The concept of 'marginal' value is widely used in economic analysis, for example marginal utility in consumer analysis, marginal productivity in production analysis and marginal revenue and marginal cost in pricing theory.
- The law of equi-marginal principle was over time applied by business managers to allocation of resources between their alternative uses with a view to maximizing profit in case a firm carries out more than one business activity.
- Marginal utility can be defined as the utility derived from the marginal or one additional unit consumed. It may also be defined as the addition to the total utility resulting from the consumption of one additional unit.
- The central theme of the consumption theory – be it based on ordinal utility or cardinal utility approach – is the utility maximizing behaviour of the consumer.
- The law of diminishing marginal utility states that as the quantity consumed of a commodity goes on increasing, the utility derived from each successive unit consumed goes on decreasing, consumption of all other commodities remaining constant.
- An indifference curve may be defined as the locus of points each representing a different combination of two substitute goods, which yield the same utility or level of satisfaction to the consumer.
- An indifference curve is formed by substituting one good for another. The rate at which one good is substituted for another is called Marginal Rate of Substitution (MRS).
3.8 KEY WORDS

- Marginal Rate of Substitution (MRS): It is defined as the rate at which one good is substituted for another.
- Budget line: It shows all the combinations of two commodities that a consumer can afford at given market prices and within the particular income level.

3.9 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short Answer Questions

1. State the limitations of the application of the marginal principle for profit maximization.
2. Write a short note on the cardinal utility approach or the Marshallian approach.
3. What are the assumptions of the ordinal utility approach?

Long Answer Questions

1. Explain the equi-marginal principle.
2. Discuss the law of diminishing marginal utility.
3. Describe the properties of an indifference curve.

3.10 FURTHER READINGS


Websites

https://blogs.economictimes.indiatimes.com/et-commentary/heres-how-indias-widening-income-distribution-can-be-redressed/
UNIT 4  DEMAND ANALYSIS

Structure
4.0 Introduction
4.1 Unit Objectives
4.2 Meaning, Functions and Applications of Demand Analysis
  4.2.1 Law of Demand
  4.2.2 Determinants of Demand
  4.2.3 Demand Estimation and Forecasting
4.3 Elasticity of Demand: Types and Measures
  4.3.1 Price Elasticity of Demand
  4.3.2 Cross Elasticity of Demand
  4.3.3 Income Elasticity of Demand
  4.3.4 Advertisement or Promotional Elasticity
4.4 Answers to Check Your Progress Questions
4.5 Summary
4.6 Key Words
4.7 Self Assessment Questions and Exercises
4.8 Further Readings

4.0 INTRODUCTION

Demand is an economic term alluding to a consumer’s desire and willingness to pay a price for a particular good/service. Keeping all other factors constant, an increase in the price of a good or service will decrease demand and vice versa. This unit will introduce you to the concept of demand which includes the determinants of demand, law of demand, demand estimation and demand forecasting, application of demand in analysis, types and measures of elasticity of demand their role in business decisions.

4.1 UNIT OBJECTIVES

After going through this unit, you will be able to:
- Discuss the function and application of demand analysis
- Define the law of demand
- List the determinants of demand
- Explain demand estimation and forecasting
- Describe the types and measures of elasticity of demand
4.2 MEANING, FUNCTIONS AND APPLICATIONS OF DEMAND ANALYSIS

The accomplishment of any business enterprise generally depends on sales and sales is governed by market demand behaviour. Market demand analysis is one of the fundamental necessities for the existence of any business enterprise. Analysis of market demand for a product/service is essential for the management in order to take decisions regarding production, cost allocation, product pricing, advertising, inventory holdings and other factors. How much the business enterprise must work hard to produce largely depends upon the demand for its product. If demand fails to fulfill the number of items produced, then the two must be balanced by creating a new demand through better advertisement and publicity. In case, the future demand for the product is likely to be more, the more the inventories that the firm should hold. If the demand for the product is high, a higher price can be charged, with other things remaining the same.

Demand analysis has two significant applications in managerial decision-making:

- Forecasting Demand
- Manipulating Demand

Market demand analysis assists the manager to take decisions regarding:

(a) sales forecasting with a sound basis and greater accuracy; (b) guidelines for demand manipulation through advertising and sales promotion programmes;
(c) production planning and product improvement; (d) pricing policy;
(e) determination of sales quotas and performance appraisal of personnel in the sales department; and (f) size of market for a given product and matching market share.

4.2.1 Law of Demand

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

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

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

The law of demand can be illustrated through a demand schedule. A demand schedule is a series of quantities which consumers would like to buy per unit of time at different prices. To illustrate the law of demand, an imaginary demand schedule for tea is given in Table 4.1. It shows seven alternative prices and the corresponding quantities (number of cups of tea) demand per day. Each price has a unique quantity demanded, associated with it. As price per cup of tea decreases, daily demand for tea increases, in accordance with the law of demand.

**The Demand Curve**

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

<table>
<thead>
<tr>
<th>Price per cap of tea (Rs.)</th>
<th>No. of cups of tea demand per consumer per day</th>
<th>Symbols representing per price-quantity combination</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>2</td>
<td>A</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>B</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>C</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>D</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>E</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>F</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>G</td>
</tr>
</tbody>
</table>

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

Why Demand Curve Slopes Downward to the Right

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

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

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

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

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

$$MU_m = P_c = MU_c$$

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

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

(a) **Expectations regarding future price.** When consumers expect a continuous increase in the price of a durable commodity, they buy more of it despite increase in its price. They do so with a view to avoiding the pinch of still higher price in future. Similarly, when consumers anticipate a considerable decrease in the price in future, they postpone their purchases and wait for the price to fall to the expected level rather than buy the commodity when its price initially falls. Such decisions of the consumers are contrary to the law of demand.

(b) **Status goods.** The law does not apply to the commodities which serve as a ‘status symbol’, enhance social prestige or display wealth and richness, e.g., gold, precious stones, rare paintings and antiques, etc. Rich people buy such goods mainly because their prices are high.

(c) **Giffen goods.** An exception to this law is also the classic case of Giffen goods named after Robert Giffen (1837–1910).

4.2.2 Determinants of Demand

Market demand for a product depends on a number of factors, called **determinants of demand.** The knowledge of the determinants of market demand for a product and the nature of relationship between the demand and its determinants proves very helpful in analyzing and estimating demand for the product. It may be noted at the very outset that a host of factors determine the market demand for a product.

In general, however, following are the factors that determine the market demand for a product:

1. **Price of the Product:** The price of a product is one of the most important determinants of its demand in the long-run and the only determinant in the short-run. The price of a product and its quantity demanded are inversely related. The law of demand states that the quantity demanded of a product which its consumers/users would like to buy per unit of time, increases when its price falls and decreases when its price increases, other factors remaining constant. The assumption ‘other factors remaining constant’ implies that factors other than price remain constant, particularly income of the consumers, prices of the substitutes and complementary goods, consumer’s taste and preferences, and number of consumers, remain unchanged. Other factors remaining constant, price is the main determinant of market demand especially in short run.
2. Price of the Related Goods: The demand for a commodity is also affected by the changes in the prices of its related goods. Related goods are classified under two categories: (i) substitutes, and (ii) complementary goods.

(i) Substitutes. Two commodities are deemed to be substitutes for one another if they satisfy the same want and change in the price of one changes the demand for its substitute in the same direction. That is, increase in the price of a good increases demand for its substitutes.

(ii) Complements. A commodity is considered to be a complement for another when it complements the use of the other. In case of complements, the use of the two goods goes together. For example, petrol is a complement to cars, butter and jam to bread, milk and sugar to tea and coffee, electricity to computer, chair to table, etc. In economic sense, two goods are termed as complementary to one another if an increase in the price of one causes a decrease in demand for the other. By definition, there is an inverse relation between the demand for a good and the price of its complement.

3. Consumer’s Income: Income is the basic determinant of quantity of a product demanded as it determines the purchasing power of the consumer. That is why people with higher current disposable incomes spend a larger amount on consumer goods and services than those with lower income. Income-demand relationship is of a more varied nature than that between demand and its other determinants. While other determinants of demand, e.g., product’s own price and the price of its substitutes are more significant in the short-run, income as a determinant of demand is equally important in both short-run and long-run.

The relationship between the demand for a commodity, say X, and the household income (Y), assuming all other factors to remain constant, is expressed by a demand function such as

\[ D_x = f(Y), \Delta D_x / \Delta Y > 0 \]

Before we proceed to discuss income-demand relationships, it will be useful to note that consumer goods of different nature have different relationships with incomes of different categories of consumers. The managers need, therefore, to be fully aware of the goods they are dealing with and their relationship with the income of consumers, particularly in regard to the assessment of both existing and prospective demand for a product.

For the purpose of income-demand analysis, consumer goods and services may be grouped under four broad categories, viz. (a) essential consumer goods, (b) inferior goods, (c) normal goods, and (d) luxury or prestige goods. Let us now look into the relationship between income and the different goods. This relationship is presented through Engel curves.

(a) Essential consumer goods (ECG). The goods and services in this category are called ‘basic needs’ and are consumed by all persons of a society, e.g., foodgrains, salt, vegetable oils, matches, cooking fuel, minimum clothing and
housing. Quantity demanded of this category of goods increases with increase in consumer’s income but only up to a certain limit, even though the total expenditure may increase in accordance with the quality of goods consumed, other factors remaining the same.

(b) Inferior goods (IG). Inferior and superior goods are widely known to both consumers and sellers. For instance, every consumer knows that millet is inferior to wheat and rice; bidi (indigenous cigarette) is inferior to cigarette, cotton clothes are inferior to silk clothes, kerosene is inferior to cooking gas; bike is inferior to car; non-AC car is inferior to AC-car, and so on and so forth. In economic sense, however, a commodity is deemed to be inferior if its demand decreases with the increase in consumer’s income beyond a certain level of income.

c) Normal goods (NG). Technically, normal goods are those that are demanded in increasing quantities as consumers’ income rises. Clothing, house, furniture, and automobiles are some of the important examples of this category of goods. Demand for normal goods increases rapidly with the increase in the consumer’s income but slows down with further increases in income.

d) Luxury and prestige goods (LG). What is and what is not a luxury good is a matter of consumer’s perception of the need for a commodity. Conceptually, however, all such goods that add to the pleasure and prestige of the consumer without enhancing his earning capacity or efficiency fall in the category of luxury goods. Demand for such goods arises beyond a certain level of consumer’s income, i.e., consumption of luxury goods at a certain level high level of income. Producers of such items, while assessing the demand for their product, should consider the income change in the richer section of the society, and not merely the per capita income.

4. Consumer’s Taste and Preference: Consumer’s taste and preference play an important role in determining the demand for a product. Taste and preference generally depend on life-style, social customs, religious values attached to a commodity, habit of the people, the general levels of living of the society, and age and sex of the consumers. Change in these factors changes consumers’ taste and preferences. As a result, consumers reduce or give up the consumption of some goods and add new ones to their consumption pattern.

This piece of information is useful for the manufacturers of goods and services subject to frequent changes in fashion and style, at least in two ways: (i) they can make quick profits by designing new models of their product and popularising them through advertisement, and (ii) they can plan production better and can even avoid over-production if they keep an eye on the changing fashions.

5. Advertisement Expenditure: Products are advertised with the objective of promoting sales of the product. Advertisement helps in increasing demand for the product in at least four ways: (a) by informing the potential consumers about the availability of the product; (b) by showing its superiority over the rival product; (c) by influencing consumers’ choice against the rival products; and (d) by setting
new fashions and changing tastes. The impact of such effects shifts the demand upward to the right. In other words, other factors remaining the same, as expenditure on advertisement increases, volume of sales increases to an extent.

6. Consumers’ Expectations: Consumers’ expectations regarding the future prices, income, and supply position of goods, etc. play an important role in determining the demand for goods and services in the short-run. If consumers expect a high rise in the price of a storable commodity, they would buy more of it at its high current price with a view to avoiding the pinch of a high price rise in future. On the contrary, if consumers expect a fall in the price of certain goods, they postpone their purchase of such goods with a view to taking advantage of lower prices in future, mainly in the case of non-essential goods. This behaviour of consumers reduces the current demand for goods whose prices are expected to decrease in the future.

7. Demonstration and Snob Effect: When new commodities or new models of existing ones appear in the market, rich people buy them first. But once new commodities are in vogue, many households buy them not because they have a genuine need for them but because their neighbours have bought these goods. The purchases made by the latter category of the buyers arise out of such feelings as jealousy, competition and equality in the peer group, social inferiority and the desire to raise their social status. Purchases made on account of these factors are the result of what economists call ‘Demonstration effect’ or the ‘Bandwagon effect’. These factors have a positive effect on demand. On the contrary, when a commodity becomes the thing of common use, some people, mostly rich, decrease or give up the consumption of such goods. This is known as the ‘Snob effect’. It has a negative effect on the demand for the related goods.

8. Consumer-Credit Facility: Availability of credit to the consumers from the sellers, banks, relations and friends, or from other sources, enduces the consumers to buy more than what they would buy in the absence of credit facility. That is why consumers who can borrow more can consume more than those who cannot borrow. Credit facility mostly affects the demand for durable goods, particularly those which require bulk payment at the time of purchase.

9. Population of the Country: The total domestic demand for a product of mass consumption depends also on the size of the population. Given the price, per capita income, tastes and preferences etc., the larger the population, the larger the demand for a product. With an increase (or decrease) in the size of population and with the employment percentage remaining the same, demand for the product tends to increase (or decrease). The global perception that India offers the largest market in the world is based on the fact that she has the second largest population – albeit with a low purchasing power – in the world.

10. Distribution of National Income: The level of national income is the basic determinant of the market demand for a product—the higher the national income, the higher the demand for all normal goods and services. Apart from its level, the
distribution pattern of national income is also an important determinant of the overall demand for a product. If national income is unevenly distributed, i.e., if a majority of the population belongs to the lower income groups, market demand for essential goods, including inferior ones, will be the largest whereas the demand for other goods will be relatively lower.

4.2.3 Demand Estimation and Forecasting

The business world is characterized by risk and uncertainty and, therefore, most business decisions are made under the condition of risk and uncertainty. One way to reduce the adverse effects of risk and uncertainty is to acquire knowledge about the future demand prospects for the product. The information regarding the future demand for the product is obtained by demand forecasting. Demand forecasting is predicting the future demand for firm’s product. A reliable forecast of the future demand for the product helps a great deal in the following areas of business management.

- Determining the production target,
- Planning and scheduling production,
- Acquiring inputs (labour, raw material and capital),
- Making provision for finances,
- Formulating pricing strategy,
- Planning advertisement.

Steps in Demand Forecasting

The demand can be forecast reliably only when forecast is made systematically and scientifically and when it is fairly reliable. The following steps are generally taken to make systematic demand forecasting.

(i) Specifying the objective
(ii) Determining the time perspective
(iii) Making choice of method for demand forecasting
(iv) Collection of data and data adjustment
(v) Estimation and interpretation of results

Methods of Demand Forecasting: An Overview

There are various methods of demand forecasting. As pointed out above, the choice of method for forecasting demand depends on the purpose and kind of demand forecasting and availability of required data. The various methods of demand forecasting are listed below.

1. Survey Methods
   (i) Consumer Survey – direct interview
   (ii) Opinion Poll Methods
2. Statistical Methods
   (i) Trend Projection
   (ii) Barometric Methods
   (iii) Econometric Methods

   All these methods have different kinds of sub-methods.

4.3 ELASTICITY OF DEMAND: TYPES AND MEASURES

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

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

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

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

\[
E_d = \frac{\Delta Q}{\Delta Y} = \frac{\Delta Q}{\Delta Y} \times \frac{Y}{Q}
\]

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

Let us now discuss the elasticity of demand with reference to its different determinants.

4.3.1 Price Elasticity of Demand

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

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

In technical terms, the measure of elasticity of demand is called *elasticity coefficient* measured by the following formula:

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

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

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

\[
E_d = \frac{\Delta Q}{\Delta Y} = \frac{\Delta Q}{\Delta Y} \times \frac{Y}{Q}
\]

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

Let us now discuss the elasticity of demand with reference to its different determinants.

4.3.1 Price Elasticity of Demand

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

\[
e_p = \frac{\text{Percentage change in quantity demanded}}{\text{Percentage change in price}}
\]
A general formula for calculating coefficient of price elasticity, as given in
Eq. (4.1), is given as follows:

\[ e_p = \frac{\Delta Q}{{Q}} \frac{\Delta P}{{P}} = \frac{\Delta Q}{{Q}} \frac{P}{{\Delta P}} \]

where \( Q \) = original quantity demanded, \( P \) = original price, \( \Delta Q \) = change in quantity

demanded and \( \Delta P \) = change in price.

It is important to note here that a minus sign (–) is generally inserted in the

formula before the fraction in order to make the elasticity coefficient a non-negative

value.

The price elasticity can be measured between any two points on a demand
curve (called arc elasticity) or at a point (called point elasticity).

Measuring Price Elasticity from a Demand Function

The price elasticity of demand for a product can be measured directly from the
demand function. In this section, we describe the method of measuring price
elasticity of demand for a product from the demand function—both linear and
non-linear. It may be noted here that if a demand function is given, arc elasticity

can be measured simply by assuming two prices and working out \( \Delta P \) and \( \Delta Q \). We

will, therefore, confine ourselves here to point elasticity of demand with respect to
price.

Measuring Price Elasticity from a Linear Demand Function

Suppose that a linear demand function is given as

\[ Q = 100 - 5P \]

Given the demand function, point elasticity can be measured for any price.

For example, suppose one has to measure elasticity at \( P = 10 \). Point elasticity is

measured as

\[ e_p = \frac{\Delta Q}{{Q}} \frac{P}{{\Delta P}} \]

The term \( \Delta Q/\Delta P \) in the elasticity formula gives the slope of the demand
curve. The slope of the demand curve can be found by differentiating the demand
function. Thus

\[ \frac{\Delta Q}{\Delta P} = \frac{\delta(100-5P)}{\delta P} = -5 \]

Having obtained the slope of the demand curve as \( \Delta Q/\Delta P = -5 \), \( e_p \) at \( P = 10 \) can be calculated as follows. At price \( P = 10 \), \( Q = 100 - 5(10) = 50 \). By

substituting these values into the elasticity formula,
Measuring Price Elasticity from a Non-linear Demand Function

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

\[ Q = aP^b \]

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

\[ e_p = \frac{\frac{\partial Q}{\partial P}}{\frac{Q}{P}} \]  

(4.3)

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

\[ \text{slope of demand curve} = \frac{\partial Q}{\partial P} = -b \frac{Q}{P} \]  

(4.4)

By substituting Eq. (4.4) in Eq. (4.3), \( e_p \) can be expressed as

\[ e_p = -b \left( \frac{P}{Q} \right) \]  

(4.5)

Since \( Q = aP^b \), by substitution, we get

\[ e_p = -b \frac{aP^{b-1}}{aP} = -b \]  

(4.6)

Equation (4.6) shows that when a demand function is of a multiplicative or power form, price elasticity coefficient equals the power of the variable \( P \). This means that price elasticity in the case of a multiplicative demand function remains constant all along the demand curve regardless of a change in price.

Determinants of Price Elasticity of Demand

We have noted above that price-elasticity of demand for a product may vary between zero and infinity. However, price-elasticity of demand, at a given price, varies from product to product depending on the following factors.
1. **Availability of Substitutes.** The higher the degree of closeness of the substitutes, the greater the elasticity of demand for the commodity. On the other hand, sugar and salt do not have close substitutes and hence their price-elasticity is lower.

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

3. **Weightage in the Total Consumption.** If proportion of income spent on a commodity is large, its demand will be more elastic. On the contrary, if the proportion of income spent on a commodity is small, its demand is less price-elastic.

4. **Time Factor in Adjustment of Consumption Pattern.** Price-elasticity of demand depends also on the time consumers need to adjust their consumption pattern to a new price: the longer the time available, the greater the price-elasticity.

5. **Range of Commodity Use.** The range of uses of a commodity also influences the price-elasticity of its demand. The wider the range of the uses of a product, the higher the elasticity of demand for the decrease in price.

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

**Application of Price Elasticity**

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

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

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

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

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

$$MR = MC$$  \hspace{1cm} (4.7)

As has been shown

$$MR = P [1 + (1/e_p)]$$  \hspace{1cm} (4.8)

By substituting Eq. (4.8) for $MR$ in Eq. (4.7), the profit maximizing condition can be expressed as

$$P [1 + (1/e_p)] = MC$$  \hspace{1cm} (4.9)

The optimal price can be worked out from Eq. (4.9) as follows.

$$P = MC / [1 + (1/e_p)]$$  \hspace{1cm} (4.10)

Given the Eq. (4.10), if point elasticity of demand curve and marginal cost ($MC$) of a firm are known, the optimal price can be easily determined. For example, suppose point elasticity of demand at point on demand curve is estimated as $e_p = -2$ and firm’s $MC = 50$. In that case, optimal price can be worked out as follows.

$$P = 50 / [1 + (1/-2)]$$

$$= 100$$

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

Some Problems on Price Elasticity of Demand

Problem: 1

A cinema charges ₹8 per ticket for evening screenings and sells 250 tickets a night on average. They estimate that the price elasticity of demand for tickets is $e_p = -1.6$.

Calculate the expected number of tickets sold if they reduce the ticket price to ₹7.
**Demand Analysis**

**NOTES**

**Self-Instructional Material**

---

**Answer:**

Price elasticity of demand = \( \frac{\% \text{ change in Qty Demanded}}{\% \text{ change in Price}} \)

If Price elasticity of demand = (-) 1.6 then ticket sales will rise by 1.6 \( \times \) 12.5\% = 20\%

20\% of 250 is 50 extra tickets

Expected tickets sold rises to 300

**Problem: 2**

A mall raises the price of car parking from ₹ 3 per day to ₹ 5 per day and finds that usage of car parks contracts from 1,200 cars a day to 900 cars per day.

Calculate the price elasticity of demand for this price change and calculate whether total revenue from the car park rises or falls.

**Answer:**

% change in price = (+) 66.7\%

% change in demand = (-) 25\%

Price elasticity of demand = \(-\frac{25}{66.7}\) = 0.375 (i.e. demand is price inelastic)

**Total revenue:**

@ ₹ 3 per day – revenue = ₹3 \( \times \) 1,200 = ₹3,600

@ ₹ 5 per day – revenue = ₹5 \( \times \) 900 = ₹4,500

Revenue rises when Price elasticity of demand <1 and a business raises their average selling price.

**4.3.2 Cross Elasticity of Demand**

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

\[
\varepsilon_{t,c} = \frac{\text{Percentage change in demand for tea (Q)}}{\text{Percentage change in price of coffee (P)}}
\]

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

\[
\frac{P}{Q} \frac{\Delta O}{\Delta P}
\]

... (4.11)

Similarly, cross-elasticity of demand for coffee with respect to change in the price of tea is measured as follows.
The same formula is used to measure the cross-elasticity of demand for a good with respect to a change in the price of its complementary goods. Electricity to electrical gadgets, petrol to automobiles, butter to bread, sugar to tea and coffee, are the examples of complementary goods.

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

**Uses of Cross Elasticity**

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

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

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

### 4.3.3 Income Elasticity of Demand

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

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

$$ e_y = \frac{\% \text{ change in demand for product } X}{\% \text{ change in consumer’s income } Y} = \frac{\Delta X}{X} / \frac{\Delta Y}{Y} $$

$$ e_y = \frac{Y}{X} \frac{\Delta X}{\Delta Y} $$

... (4.13)
(where \( X' = \text{quantity of } X \text{ demanded} \); \( Y = \text{disposable income} \); \( \Delta X' = \text{change in quantity of } X \text{ demanded} \); and \( \Delta Y = \text{change in income} \))

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

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

Nature of Commodity and Income-Elasticity

For all normal goods, income-elasticity is positive though the degree of elasticity varies in accordance with the nature of commodities. Consumer goods of the three categories, viz., necessities, comforts and luxuries have different elasticities. The general pattern of income-elasticities of different goods for increase in income and their effect on sales are given in Table 4.2. As Table 4.2 shows, income elasticity of essential goods in less than 1. It is so because of Engel’s law. Income elasticity of ‘comforts’ equals 1. And, in case of luxury goods, \( e_y > 1 \).

<table>
<thead>
<tr>
<th>Consumer goods</th>
<th>Co-efficient of income-elasticity</th>
<th>Effect on sales with change in income</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Essential goods</td>
<td>Less than one ( (e_y &lt; 1) )</td>
<td>Less than proportionate change in sale</td>
</tr>
<tr>
<td>2. Comforts</td>
<td>Almost equal to unity ( (e_y \approx 1) )</td>
<td>Almost proportionate change in sale</td>
</tr>
<tr>
<td>3. Luxuries</td>
<td>Greater than unity ( (e_y &gt; 1) )</td>
<td>More than proportionate increase in sale</td>
</tr>
</tbody>
</table>

Uses of Income-Elasticity in Business Decisions

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

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

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

4.3.4 Advertisement or Promotional Elasticity

The expenditure on advertisement and on other sales-promotion activities does help in promoting sales, but not at the same degree at all levels of the total sales and total ad-expenditure. The concept of advertisement elasticity is useful in determining the optimum level of advertisement expenditure. The concept of advertisement elasticity assumes a greater significance in deciding on advertisement expenditure, particularly when there is competitive advertising by the rival firms. Advertisement elasticity \( e_A \) of sales is measured as

\[
e_A = \frac{\% \text{ change in sales}}{\% \text{ change in Ad-expenditure}}
\]

\[
e_A = \frac{\Delta S / S}{\Delta A / A} = \frac{A}{\Delta A / \Delta S} S
\]

where \( S = \) sales; \( \Delta S = \) increase in sales; \( A = \) initial advertisement cost, and \( \Delta A = \) additional expenditure on advertisement.

Suppose, for example, a company increases its advertising expenditure from \( \text{Rs} 10 \) million to \( \text{Rs} 12 \) million, and as a result, its sales increase from 5,000 units to 6,000 units. In this case \( \Delta A = 12 \) million – 10 million = \( \text{Rs} 2 \) million, and \( \Delta S = 6,000 – 5,000 = 1000 \) units. By substituting these values in ad-elasticity formula (4.14), we get

\[
e_A = \frac{1000}{2} \times \frac{10}{5000} = 1
\]
It means that a one per cent increase in ad-expenditure increases sales by 1 per cent.

**Interpretation of Advertisement Elasticity:** The advertisement elasticity of sales varies between \( e_A = 0 \) and \( e_A = \infty \) depending on the nature of the product, the level of market supplied, the trend in consumers’ income, the competitive strength of the competitors, etc. The interpretation of some measures of advertising elasticity is given below.

<table>
<thead>
<tr>
<th>Elasticities</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( e_A = 0 )</td>
<td>Sales do not respond to the advertisement expenditure.</td>
</tr>
<tr>
<td>( e_A &gt; 0 ) but ( &lt; 1 )</td>
<td>Increase in total sales is less than proportionate to the increase in advertisement expenditure.</td>
</tr>
<tr>
<td>( e_A = 1 )</td>
<td>Sales increase in proportion to the increase in advertisement expenditure.</td>
</tr>
<tr>
<td>( e_A &gt; 1 )</td>
<td>Sales increase at a higher rate than the rate of increase of advertisement expenditure.</td>
</tr>
</tbody>
</table>

**Determinants of Advertisement Elasticity**

Some important factors that determine the level of ad-elasticity are the following.

(i) The level of total sales
(ii) Advertisement by rival firms
(iii) Cumulative effect of past advertisement
(iv) Other factors

**Check Your Progress**

1. What is a demand curve?
2. State the factors that determine the market demand for a product.

**4.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS**

1. A demand curve is a locus of points showing various alternative price-quantity combinations.
2. The factors that determine the market demand for a product are the following:
   - Price of the product,
   - Price of the related goods—substitutes, complements and supplements,
   - Level of consumers’ income,
   - Consumers’ taste and preferences,
   - Advertisement of the product,
• Consumers’ expectations about future price and supply position,
• Demonstration effect and ‘bandwagon effect’,
• Consumer-credit facility,
• Population of the country (for the goods of mass consumption),
• Distribution pattern of national income, etc.

4.5 SUMMARY

• The term ‘demand’ refers to the quantity demanded of a commodity per unit of time at a given price. It implies also a desire backed by ability and willingness to pay.
• The relationship between price and demand is expressed by the law of demand. The law of demand states that quantity of a product demanded per unit of time increases when its price falls, and decreases when its price increases, other factors remaining constant.
• The law of demand can also be presented through a curve called demand curve. A demand curve is a locus of points showing various alternative price-quantity combinations.
• Diminishing marginal utility is also responsible for increase in demand for a commodity when its price falls. When a person buys a commodity, he exchanges his money income with the commodity in order to maximize is satisfaction.
• Market demand for a product depends on a number of factors, called determinants of demand. The knowledge of the determinants of market demand for a product and the nature of relationship between the demand and its determinants proves very helpful in analyzing and estimating demand for the product.
• The information regarding the future demand for the product is obtained by demand forecasting. Demand forecasting is predicting the future demand for firm’s product.
• There are various methods of demand forecasting. The choice of method for forecasting demand depends on the purpose and kind of demand forecasting and availability of required data.
• In general terms, the elasticity of demand is defined as the degree of responsiveness of demand for a product to change in its determinants. The measure of the degree of responsiveness of demand to change in its determinants gives the measure of the extent of relationship between the demand for a product and any of its determinants. There are various sub-types including price, income, cross and advertising elasticity of demand.
4.6 KEY WORDS

- **Income effect:** The increase in demand on account of increase in real income is known as income effect.
- **Elasticity of demand:** It is defined as the degree of responsiveness of demand for a product to change in its determinants.

4.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short Answer Questions

1. State the law of demand.
2. What are the exceptions to the law of demand?
3. Write a short note on diminishing marginal utility.
4. Define demand forecasting.
5. Briefly mention the methods of demand forecasting.

Long Answer Questions

1. Discuss the application of demand analysis.
2. Explain the application of price elasticity of demand.
3. Explain income elasticity of demand.
4. What is cross elasticity of demand? Mention its uses.

4.8 FURTHER READINGS


Websites


## UNIT 5  SUPPLY ANALYSIS

### Structure
- **5.0 Introduction**
- **5.1 Objectives**
- **5.2 The Law of Supply and Determinants of Supply**
  - 5.2.1 Elasticity of Supply: Measures and Significance
  - 5.2.2 Derivations of Market Demand
  - 5.2.3 Demand Estimation and Forecasting
- **5.3 Demand Supply Equilibrium**
- **5.4 Giffen Paradox**
- **5.5 Answers to Check Your Progress Questions**
- **5.6 Summary**
- **5.7 Key Words**
- **5.8 Self Assessment Questions and Exercises**
- **5.9 Further Readings**

### 5.0 INTRODUCTION

Supply and demand are economic terms which occupy a significant place in the analysis of market economy. In the previous unit, you studied about demand analysis and this unit will introduce you supply analysis. This unit will introduce you to the law of supply, the measures of elasticities of supply, derivation of market demand, demand supply equilibrium and finally, the Giffen paradox.

### 5.1 OBJECTIVES

After going through this unit, you will be able to:
- State the law of supply
- Discuss the measures of elasticities of supply
- Explain the derivation of market demand
- Describe the demand supply equilibrium
- Define the Giffen paradox
Supply Analysis

5.2 THE LAW OF SUPPLY AND DETERMINANTS OF SUPPLY

In the previous unit, we discussed the law of demand. In this section, we turn to the other side of the market and to discuss the law of supply, specifically the law of market supply. Market supply means the quantity of a commodity which all its producers or sellers offer to sell at a given price, per unit of time. Market supply, like market demand, is the sum of supplies of a commodity made by all individual firms.

The law of supply can be stated as the supply of a product increases with the increase in its price and decreases with decrease in its price, other things remaining constant. It implies that the supply of a commodity and its price are positively related. This relationship holds under the assumption that “other things remaining the same”. “Other things” include cost of production, change in technology, price of related goods (substitutes and complements), and weather and climate in case of agricultural products.

The Supply Factors

In reality, the supply of a commodity depends on its price, cost of production and production technology. In other words, supply of a product \( X \) is the function of its price \( P_x \), cost of production \( C_x \) and technology of its production \( T_x \). In its functional form

\[
Q_x = f(P_x, C_x, T_x)
\]

In simple theory of supply, however, the law of supply is expressed generally in terms of price-quantity relationship supply function is expressed as

\[
Q_x = f(P_x) \quad ...(5.1)
\]

The Supply Schedule

The supply schedule is a tabular presentation of the supply function. In fact, a supply schedule is a table showing alternative prices of a commodity and the corresponding quantity that suppliers are willing to offer for sale. Table 5.1 presents a hypothetical supply schedule of shirts per month.

<table>
<thead>
<tr>
<th>Price (in `)</th>
<th>100</th>
<th>200</th>
<th>300</th>
<th>400</th>
<th>600</th>
<th>800</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supply</td>
<td>10</td>
<td>40</td>
<td>55</td>
<td>70</td>
<td>75</td>
<td>80</td>
</tr>
</tbody>
</table>

Table 5.1 Supply Schedule for Shirts
It can be seen from Table 5.1 that at price \( \$100 \) per shirt, only 10 thousand shirts would be supplied per month. When price increases to \( \$200 \), suppliers offer 40 thousand shirts for sale. And, when price rises to \( \$400 \), supply rises to 70 thousand shirts, and so on.

The Supply Curve

A supply curve is a graphical depiction of the supply schedule. The supply curve \( SS' \) given in Fig. 5.1 has been obtained by plotting the data in Table 5.1. The points \( S, P, Q, R, T \) and \( S' \) show the price-quantity combinations on the supply curve \( SS' \). This supply curve depicts the law of supply. The upward slope of the supply curve indicates the rise in the supply of shirts with the rise in its price and fall in the supply with fall in its price.

![Supply Curve of Shirts](image)

Shift in the Supply Curve

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

(a) Change in Input Prices. When input prices go down, the use of inputs increases or more inputs can be used at a given total cost. As a result, product supply increases and the supply curve shifts to the right, as shown by the supply curve \( S'S' \). in Fig. 5.2. Similarly, when input increase, product supply curve shifts leftward to \( S'S' \).
(b) **Technological Progress.** Technological changes that reduce cost of production or increase factor efficiency increase the product supply. For instance, introduction of high yielding varieties of paddy and new techniques of cultivation increased per acre yield of rice in India in the 1970s. Such changes make the supply curve shift to the right.

(c) **Price of Product Substitutes.** In many kinds of production activities, it is possible to produce a substitute product. For example, a refrigerator company can also produce ACs; Tatas famous for the production of trucks produce also cars; Maruti Udyog can produce trucks, and so on. Fall in the price of one of the product substitutes may lead to the rise in the supply of other due to capacity utilization for profit maximisation.

![Fig. 5.2 Shift in the Supply Curve](image)

(d) **Level of Competition and Size of the Industry.** The supply of a commodity depends also on whether an industry is monopolized or competitive. If a monopolized industry is made competitive, the total supply increases. Besides, if size of an industry increases because new firms join the industry, the total supply will increase and supply curve will shift rightward.

(e) **Government Policy.** When government imposes restrictions on production, e.g., import quota on inputs, excise taxation, etc., production tends to fall. Such restrictions make supply curve to shift leftward.

(f) **Non-Economic Factors.** The factors like war, drought, flood, communal riots, epidemics, etc. also affect adversely the supply of commodities.

**The Supply Function**

The law of supply states only the nature of relationship between the price and the quantity supplied. A supply function quantifies this relationship. The supply function is, in fact, a statement which states the relationship between the quantity supplied of a commodity (as a dependent variable) and its determinants (as independent variables). A supply function is written in its general form as
\[ Q_s = xP_s \] …(5.2)

where \( Q_s \) = quantity supplied of commodity \( X \), \( x = \Delta Q / \Delta P \), and \( P_s \) = price.

Given the supply function (5.2), a supply schedule can be generated by substituting numerical values for \( P_s \) and assuming \( x = 10 \). For example, if \( P_s = 2 \), \( Q_s = 10 \times 2 = 20 \) and if \( P_s = 5 \), \( Q_s = 10 \times 5 = 50 \) and so on. By plotting the supply schedule, a supply curve can be obtained. (For procedure, refer to the section on demand function).

### 5.2.1 Elasticity of Supply: Measures and Significance

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

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

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

\[ e_p = \frac{\Delta Q/Q}{\Delta P/P} \]

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

### Determinants of the Price Elasticity of Supply

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

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

It is important to note here that short and long periods are not fixed in terms of days, months or years. It varies depending on the nature of the product. For example, for the supply of perishable commodities like milk and fish in a city, a week’s time may be a short period. For agricultural products, 6 months may be a short period. But in regard to the local supply of petroleum products in India, a period of five years or even more may be regarded as a short period.
Law of Diminishing Returns. The other factor that determines the elasticity of supply is the Law of Diminishing Returns. We will discuss this law later in detail when we take up the laws of production. Here, it suffices to say that if the law of diminishing returns come in force at an early level of production, cost increases rapidly. As a result, supply tends to becomes less and less elastic.

5.2.2 Derivations of Market Demand

Market demand can be defined as the sum of individual demands for a product at a price per unit of time. We may recall that the quantity demanded of a commodity by an individual per unit of time, at a given price, is known as ‘individual demand’ for that commodity. The aggregate of individual demands for a product is called market demand for the product. In other words, the sum of quantity demand by all the consumers/users of a commodity per unit of time at a given price, all other things remaining the same, is called ‘market demand’ for that product.

For example, suppose there are only three consumers—A, B and C—of a commodity X which has a fixed price. Consumers A, B and C consume 100 units, 200 units and 300 units respectively, of commodity X monthly. Therefore, the monthly market demand for commodity X equals 100 + 200 + 300 = 600 units.

Derivation of Market Demand Curve

If individual demand schedules or individual demand functions are known, the market demand schedule and market demand curve can easily be derived. The market demand curve can be derived by adding up (i) the individual demand schedules, and (ii) the individual demand functions. In this section, we illustrate the derivation of market demand curve by using these two methods.

Derivation of Market Demand Curve from Individual Demand Schedules

Suppose again that there are three consumers (A, B and C) of a commodity X and their monthly demand schedules for the commodity are given in Table 5.2. The table shows the quantity demanded of commodity X individually by the three consumers at different prices of commodity X. The last column shows the market demand, i.e., the sum of individual demands for commodity X. The market demand shows the total quantity of commodity X demanded per month by the three consumers at different prices.

Given the individual and market demand schedules, the market demand curve can be obtained by plotting the market demand against the respective prices. This is illustrated in Fig. 5.3 by the curve $D_{MR}$. 

NOTES
Derivation of Market Demand Curve from Individual Demand Functions

Market demand curve can also be drawn by summing up the individual demand functions. Suppose individual demand functions of consumers A, B and C for commodity X are given as follows.

- **A’s demand function**: \( D_A = 100 - 10P_x \)
- **B’s demand function**: \( D_B = 75 - 7.5P_x \)
- **C’s demand function**: \( D_C = 50 - 5P_x \)

Given the individual demand functions, the market demand function can be obtained by adding them up. Thus,
Market Demand function:

\[ D_M = (100 - 10P) + (75 - 7.5P) + (50 - 5P) \]

\[ D_M = 225 - 22.5P \]

This market demand function can be converted into a market demand schedule by assigning numerical values to \( P \). The market demand curve can then be drawn by plotting the demand schedule.

5.2.3 Demand Estimation and Forecasting

We have already discussed this concept in Unit 4, let us revise the concept here briefly. Demand forecasting and estimation is used by business enterprises to obtain valuable information about the markets in which they operate and the markets in which they plan to launch their products. Forecasting and estimation are terms whose usage is synonymous—that basically mean predicting what will happen in the future. If business enterprises do not use demand forecasting and estimation, then they might enter markets which are not worthy of the products produced by them.

The fundamental objective of demand forecasting and estimation is to analyse the business’s potential demand. This will facilitate managers to make accurate decisions about pricing, business growth and market potential. We can understand this with the help of an example. For instance, if the market demand for pasta is extremely high in a city but there are few competitors, then the business managers can price pasta slightly higher than if the demand was less. Established businesses use the tool of demand forecasting and estimation if they wish to venture into a new market. If the demand for their product is currently low, but will increase in the future, they will wait to enter the market.

5.3 DEMAND SUPPLY EQUILIBRIUM

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

The Concept of Equilibrium

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

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

**Determination of Equilibrium Price**

The equilibrium price in a free market is determined by the market forces of demand and supply. In order to analyse how equilibrium price is determined, we have to analyse the process through which market forces bring the suppliers’ plan in balance with the buyers’ plan. For this purpose, let us use our example of demand and supply schedules for shirts. Suppose that the market demand and supply schedules for shirts are given as shown in Table 5.3.

<table>
<thead>
<tr>
<th>Price per Shirt (₹)</th>
<th>Demand ('000 shirts)</th>
<th>Supply ('000 shirts)</th>
<th>Market Position</th>
<th>Effect on Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>80</td>
<td>10</td>
<td>Shortage</td>
<td>Rise</td>
</tr>
<tr>
<td>200</td>
<td>55</td>
<td>28</td>
<td>Shortage</td>
<td>Rise</td>
</tr>
<tr>
<td>300</td>
<td>40</td>
<td>40</td>
<td>Equilibrium</td>
<td>Stable</td>
</tr>
<tr>
<td>400</td>
<td>28</td>
<td>30</td>
<td>Surplus</td>
<td>Fall</td>
</tr>
<tr>
<td>500</td>
<td>20</td>
<td>55</td>
<td>Surplus</td>
<td>Fall</td>
</tr>
<tr>
<td>600</td>
<td>15</td>
<td>60</td>
<td>Surplus</td>
<td>Fall</td>
</tr>
</tbody>
</table>

Table 5.3 places the demand and supply schedules side by side against the same price schedule. As the table shows, there is only one price of shirts (₹ 300) at which the market is in equilibrium, i.e., the quantity demanded and the quantity supplied are equal at 40 thousand shirts. At all other prices, the shirt market is in disequilibrium as either demand exceeds supply or supply exceeds demand. At all prices below ₹ 300, demand exceeds supply showing shortage of shirts in the market. Likewise, at all prices above ₹ 300 supply exceeds demand showing surplus supply.

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

**Demand and Supply Adjustment**

Let the price be initially set at ₹ 100. At this price, the quantity demanded exceeds the quantity supplied by 70 thousands shirts. The shortage gives sellers an
Supply Analysis

NOTES

Self-Instructional Material

opportunity to raise the price. Increase in price enhances the profit margin. This induces firms to produce more in order to maximize their profits. This trend will continue till price rises to \( 300 \). As Table 3.4 shows, at price \( 300 \), the buyers are willing to buy 40 thousands shirts. This is exactly the number of shirts that the sellers would like to sell at this price. At this price, there is neither shortage nor surplus of shirts in the market. This price is therefore the equilibrium price.

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

Graphical Presentation

The determination of equilibrium price is illustrated graphically in Fig. 5.4. The demand curve \( DD' \) and the supply curve \( SS' \) have been obtained by plotting the demand and supply schedules, respectively, on the same price and quantity axes.

As Fig. 5.4 shows, demand and supply curves for shirts intersect at point \( E \) determining the equilibrium price at \( 300 \). At this price, the quantity demanded (40 thousand shirts) equals the quantity supplied. Thus, the equilibrium price is \( 300 \) and equilibrium quantity is 40 thousand shirts. The equilibrium condition is not fulfilled at any other point on the demand and supply curves.
Supply Analysis

Algebra of Demand-Supply Equilibrium

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

Let the demand function for commodity X be given as

\[ Q_d = 150 - 5P_x \]

and supply function as

\[ Q_s = 10P_x \]

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

\[ Q_s = Q_d \]  

...(5.3)

By substituting supply and demand functions into Eq. 5.3, we get

\[ 10P_x = 150 - 5P_x \]  

...(5.4)

Given the Eq. (5.4) the equilibrium price can be worked out as follows:

\[ 10P_x = 150 - 5P_x \]  

...(5.5)

or

\[ 10P_x + 5P_x = 150 \]

\[ 15P_x = 150 \]

\[ P_x = 10 \]

Thus, at equilibrium, \( P = 10 \). That is, the equilibrium price is 10.

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

The algebraic determination of equilibrium price and quantity can also be demonstrated graphically. It is shown in Fig. 5.5. The demand curve \( DD' \) has been drawn using the demand function \( Q_d = 150 - 5P_x \) and the supply curve \( SS' \) using the supply function \( Q_s = 10P_x \).

The demand and the supply curves intersect at point P. A perpendicular drawn from point P to the quantity axis determines the equilibrium quantity at 100 units and a line drawn from point P to the price axis determines the equilibrium price at 10. At this price, the quantity demanded equals the quantity supplied at 100 units (Fig. 5.5).
5.4 GIFFEN PARADOX

Giffen paradox was introduced by Marshall as an exception to the law of demand. A Giffen good is defined as an inferior good whose demand increases when its price increases. There are several inferior commodities (much cheaper than its superior substitutes) consumed by the poor households as an essential commodity. If the price of such goods increases (price of its substitute remaining constant), its demand increases instead of decreasing because, in case of a Giffen good, income effect of a price rise is greater than its substitution effect. The reason is, when price of an inferior good increases, income remaining the same, poor people cut the consumption of the superior substitutes so that they are able to buy sufficient quantity of the inferior good to meet their basic need. For instance, let us suppose that the monthly minimum consumption of foodgrains by a poor household is 20 kg of bajra (an inferior good) and 10 kg of wheat (a superior good). Suppose also that bajra sells at ₹5 per kg and wheat at ₹10 per kg and that the household spends its total income of ₹200 on these items. Now, if price of bajra increases to ₹6 per kg, the household will be forced to reduce the consumption of wheat by 5 kg and increase that of bajra by the same quantity in order to meet its minimum monthly consumption requirement of 30 kg of food-grains, its expenditure on foodgrains remaining the same. The consumer substitutes bajra for wheat because he can in no other way meet his basic minimum needs. Obviously, the household’s demand for bajra increases from 20 kg to 25 kg per month despite the increase in its price.
Check Your Progress
1. What is a supply schedule?
2. Define a free market.

5.5 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. The supply schedule is a tabular presentation of the supply function. In fact, a supply schedule is a table showing alternative prices of a commodity and the corresponding quantity that suppliers are willing to offer for sale.
2. A free market is one in which market forces of demand and supply are free to take their own course: there is no outside control on price, demand and supply.

5.6 SUMMARY

- The law of supply can be stated as the supply of a product increases with the increase in its price and decreases with decrease in its price, other things remaining constant.
- The supply schedule is a tabular presentation of the supply function. In fact, a supply schedule is a table showing alternative prices of a commodity and the corresponding quantity that suppliers are willing to offer for sale.
- Technological changes that reduce cost of production or increase factor efficiency increase the product supply.
- The law of supply states only the nature of relationship between the price and the quantity supplied. A supply function quantifies this relationship.
- Like the law of demand, the law of supply states only the nature of relationship between the change in the price of a commodity and the quantity supplied thereof.
- Market demand can be defined as the sum of individual demands for a product at a price per unit of time.
- In a general sense, the term equilibrium means the “state of rest”. It indicates the condition where forces working in opposite direction are in balance.
- In a free market, disequilibrium itself creates the condition for equilibrium.
- Giffen paradox was introduced by Marshall as an exception to the law of demand. A Giffen good is defined as an inferior good whose demand increases when its price increases.
5.7 KEY WORDS

- **Price elasticity of supply**: It is the measure of responsiveness of the quantity supplied of a good to the changes in its market price.
- **Market demand**: It can be defined as the sum of individual demands for a product at a price per unit of time.
- **Giffen good**: It is defined as an inferior good whose demand increases when its price increases.

5.8 SELF ASSESSMENT QUESTIONS AND EXERCISES

**Short Answer Questions**

1. State the law of supply.
2. Write a short note on the supply function.
3. What are the determinants of price elasticity of supply?

**Long Answer Questions**

1. Explain the supply curve with the help of diagrams.
2. Discuss the derivation of market demand curve.
3. Describe the determination of equilibrium price.

5.9 FURTHER READINGS


**Websites**

https://blogs.economictimes.indiatimes.com/et-commentary/heres-how-indias-widening-income-distribution-can-be-redressed/

UNIT 6  PRODUCTION FUNCTIONS

Structure
6.0 Introduction
6.1 Objectives
6.2 Production Functions: Short-Run and Long-Run Production Function
6.3 Theory of Production and Managerial Uses of Production Function
   6.3.1 Short-Run Laws of Production
   6.3.2 Isoquants
   6.3.3 Long-Run Laws of Production
6.4 Cobb-Douglas and Other Production Functions: Empirical Estimation
6.5 Answers to Check Your Progress Questions
6.6 Summary
6.7 Key Words
6.8 Self Assessment Questions and Exercises
6.9 Further Readings

6.0 INTRODUCTION

Once business firms take decision on ‘what to produce’ the major issues that arise are ‘how to produce’ and ‘how much to produce’. These issues arise because achieving optimum efficiency in production and minimizing cost for a given production is one of the prime concerns of the business managers. In fact, the very survival of a firm in a competitive market depends on their ability to produce at a competitive cost. Therefore, managers of business firms endeavour to minimize the production cost of a given output or, in other words, maximize the output from a given quantity of inputs. In their effort to minimize the cost of production, the fundamental questions that managers are faced with are:

(i) How can production be optimized with given resources?
(ii) How does output respond to change in quantity of inputs?
(iii) How does technology matter in reducing the cost of production?
(iv) How can the least-cost combination of inputs be achieved?
(v) Given the technology, what happens to the rate of return when more plants are added to the firm?

The theory of production provides a theoretical answer to these questions through abstract models built under hypothetical conditions. The production theory may therefore not provide solutions to the real life problems. But it does guide to provide tools and techniques to analyse the real-life production conditions and to find solutions to the practical business problems.
This unit presents a detailed discussion on the theory of production and its managerial uses. In addition, the unit also deals with the application of isoquants, long-run and short-run laws of production, Cobb-Douglas function and other production functions.

6.1 OBJECTIVES

After going through this unit, you will be able to:
- Explain the short-run and long-run production function
- Describe the theory of production
- Identify the uses of production function
- Define isoquants
- State Cobb-Douglas and other production functions

6.2 PRODUCTION FUNCTIONS: SHORT-RUN AND LONG-RUN PRODUCTION FUNCTION

This section presents a brief discussion on the basic concepts and terminology used in the analysis of the theory of production.

Meaning of Production

In general sense of the term, ‘production’ means transforming inputs (labour, capital, raw materials, time, etc.) into an output with value added. This concept of production is however limited to only ‘manufacturing’. In economic sense, the term ‘production’ means a process by which resources (men, material, time, etc.) are transformed into a different and more useful commodity or service. In other words, a process by which men, material, capital and time are converted into value added products is called production.

Input and Output

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

The term ‘inputs’ needs some more explanations. Production process requires a wide variety of inputs, depending on the nature of product. But, economists classified inputs as (i) labour, (ii) capital, (iii) land, (iv) raw materials and (v) entrepreneurship. Technology and time are also treated as inputs in the modern concept of production.
**Fixed and Variable Inputs**

Inputs are classified as

(i) *fixed inputs*, and

(ii) *variable inputs*.

Fixed and variable inputs are defined in economic sense and also in technical sense. The two concepts of input are explained below in economic as well as in technical sense.

*Fixed Inputs* In economic sense, a fixed input is one whose supply is inelastic in the short-run. Therefore, all of its users together cannot buy more of it in the short-run. In technical sense, a fixed factor is one that remains fixed (or constant) for a certain level of output.

*Variable Input* A variable input is defined as one whose supply in the short-run is elastic, e.g., labour and raw material, etc. All the users of such factors can employ a larger quantity in the short-run as well as in the long-run. Technically, a variable input is one that changes with the change in output. In the long-run, all inputs are variable.

**Short-run and Long-run**

Production of a good involves time. The reference to time period involved in production process is another important concept used in production analysis. The two reference periods are short-run and long-run. The short-run refers to a period of time in which the supply of certain inputs (e.g., plant, building, machinery, etc.) is fixed or is inelastic and are used in a fixed quantity. In the short-run, therefore, production of a commodity can be increased by increasing the use of only variable inputs like labour and raw materials.

On the other hand, long-run refers to a period of time in which the supply of all the inputs is elastic, but not enough to permit a change in technology. That is, in the long-run, all the inputs are variable. Therefore, in the long-run, production of a commodity can be increased by employing more of both variable and fixed inputs.

It is important to note that ‘short-run’ and ‘long-run’ are economists’ jargon. They do not refer to any specific time period. While in some industries short-run may be a matter of few weeks or few months, in some others (e.g., in housing, shipping, flying, electricity and power industries), it may mean three or more years.

The economists use another term, i.e., *very long-run* which refers to a period in which the technology of production is also subject to change or can be improved. In the very long-run, the production function also changes. The technological advances result in a larger output from a given quantity of inputs per unit of time.
Production Function

Production function is a mathematical presentation of input-output relationship. More specifically, a production function states the technological relationship between inputs and output in the form of an equation, a table or a graph. In its general form, it specifies the inputs required for the production of a commodity or service. In its specific form, it states the extent of quantitative relationships between inputs and output. Besides, the production function represents the technology of a firm or of an industry. For example, suppose production of a product, say X, depends on labour (L) and capital (K), then production function is expressed in equation form as:

\[ Q = f(L, K) \]

A real-life production function is generally very complex. It includes a wide range of inputs, viz., (i) land and building; (ii) labour including manual labour, engineering staff and production manager, (iii) capital, (iv) raw material, (v) time, and (vi) technology. All these variables enter the actual production function of a firm. The long-run production function is generally expressed as:

\[ Q = f(LB, \ L, \ K, \ M, \ T, \ t) \]

where \( LB = \) land and building, \( L = \) labour, \( K = \) capital, \( M = \) raw materials, \( T = \) technology and \( t = \) time.

The economists have however reduced the number of variable inputs used in a production function to only two, viz., labour (L) and capital (K), for the sake of convenience and simplicity in the analysis of input-output relations. It has logical reasoning also. A production function with two variable inputs, K and L, is expressed as:

\[ Q = f(K, L) \]

The reasons for excluding other inputs are following.

Land and building (LB), as inputs, are constant for the economy as a whole, and hence they do not enter into the aggregate production function. However, land and building are not a constant variable for an individual firm or industry. In the case of individual firms, land and building are lumped with ‘capital’.

In case of ‘raw materials’ it has been observed that this input ‘bears a constant relation to output at all levels of production’. For example, cloth size bears a constant relation to the number of garments. Similarly, for a given size of a house, the quantity of bricks, cement, steel, etc., remains constant, irrespective of the number of houses constructed. To consider another example, in car manufacturing of a particular brand or size, the quantity of steel, number of the engine, and number of tyres and tubes are fixed per car. Since in case of ‘raw materials’ the relationship between inputs and output is fixed, the output can be easily estimated given the quantity of inputs. Therefore, raw materials are left out of production function.
So is the case, generally, with time and space. Also, technology (T) of production remains constant over a period of time. That is why, in most production functions, only labour and capital are included.

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

To illustrate the algebraic form of production function, let us suppose that a coal mining firm employs only two inputs—capital (K) and labour (L)—in its coal production activity. So the coal output depends on the number of labour and capital units employed to produce coal. As such, the general form of its production function may be expressed symbolically as

\[ Q = f(K, L) \] …(6.1)

where \( Q \) = the quantity of coal produced per time unit; \( K \) = capital; and \( L \) = labour.

The production function (6.1) implies that quantity of coal produced depends on the quantity of capital (K) and labour (L) employed to produce coal. Increasing coal production will require increasing \( K \) and \( L \). Whether the firm can increase both \( K \) and \( L \) or only \( L \) depends on the time period it takes into account for increasing production, i.e., whether the firm considers a short-run or a long-run.

By definition, as noted above, short-run is a period in which supply of capital is inelastic. In the short-run, therefore, the firm can increase coal production by increasing only labour since the supply of capital in the short run is fixed. Long-run is a period during which supply of both labour and capital becomes elastic, i.e., it increases over time. In the long-run, therefore, the firm can employ more of both capital and labour. Accordingly, there are two kinds of production functions:

(i) Short-run production function; and
(ii) Long-run production function.

The two kinds of production functions are described here briefly.

(i) Short-run Production Function. A short-run production function is a single variable function. The single variable factor is labour (L), capital (K) remaining constant, expressed as

\[ Q = f(L) \text{, where } K \text{ indicates constant capital.} \] …(6.2)

An estimated short-run production function may take any of the following forms depending on input-output relationship.

(a) Linear function: \( Q = a + bL \).
(b) Quadratic function: \( Q = a + bL - cL^2 \).
(c) Cubic function: \( a + bL + cL^2 - dL^3 \); or
(d) Power function: \( Q = aL^b \) (\( b > 1 \)).
(ii) Long-run Production Function: In the long-term production function, both $K$ and $L$ are treated as variable factors and the function takes the following form.

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

As mentioned above, a production function can be expressed in the form of an equation, a graph or a table, though each of these forms can be converted into its other forms. We illustrate here how a production function in the form of an equation can be converted into its tabular form. Consider, for example, the Cobb-Douglas production function—the most famous and widely used production function—given in the form of an equation as

$$Q = AK^aL^b$$

(where $K = \text{Capital}$, $L = \text{Labour}$, and $A$, $a$ and $b$ are parameters, and $b = 1 - a$)

Production function (6.3) gives the general form of Cobb-Douglas production function. The numerical values of parameters $A$, $a$ and $b$, can be estimated by using actual factory data on production, capital and labour.

### 6.3 THEORY OF PRODUCTION AND MANAGERIAL USES OF PRODUCTION FUNCTION

Let us now discuss the significant aspects of the theory of production.

#### 6.3.1 Short-Run Laws of Production

The laws of production state the relationship between output and input. In the short-run, input-output relations are studied with one variable input (labour), other inputs (especially, capital) held constant. The laws of production under these conditions are called the ‘Laws of Variable Proportions’ or the ‘Laws of Returns to a Variable Input’. In this section, we explain the ‘laws of returns to a variable input’.

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

*The Law of Diminishing Returns*. The law of diminishing returns states that when more and more units of a variable input are used with a given quantity of fixed inputs, the total output may initially increase at increasing rate and then may be at a constant rate, but it will eventually increase at diminishing rates. That is, the marginal increase in total output decreases eventually when additional units of a variable factor are used, given quantity of fixed factors.

**Assumptions.** The law of diminishing returns is based on the following assumptions:

(i) labour is the only variable input, capital remaining constant;

(ii) labour is homogeneous;
(iii) the state of technology is given; and
(iv) input prices are given.

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

\[ Q = f(L), \text{ } K \text{ constant} \]

Now let us suppose that labour and production data of the coal mining firm are given as in cols. (1) and (2) of Table 6.2. When regression technique is applied to estimate the production function, it produces an empirical production function as given below.

\[ Q = -L^3 + 15L^2 + 10L \quad \ldots (6.4) \]

Given the production function (6.4), we may assign different numerical values to \( L \) in the function and work out a series of \( Q \), i.e., the quantity of coal that can be produced with different number of workers. For example, if \( L = 5 \), then by substitution, we get

\[ Q = -5^3 + 15 \times 5^2 + 10 \times 5 = -125 + 375 + 50 = 300 \]

What we need now is to analyse the input-output relationship with the objective of taking production decision. To this end, we need is to work out marginal productivity of labour \( MP_L^1 \) to find the trend in the contribution of the marginal labour and average productivity of labour \( AP_L \) to find the average contribution of labour.

**Marginal Productivity of Labour \( MP_L^1 \)** can be obtained by differentiating the production function (6.4). Thus,

\[ MP_L^1 = \frac{\partial Q}{\partial L} = -3L^2 + 30L + 10 \quad \ldots (6.5) \]

By substituting numerical value for labour \( (L) \) in Eq. (6.5), \( MP_L^1 \) can be obtained at different levels of labour employment. However, this method can be used only where labour is perfectly divisible and \( \partial L \to 0 \). Since, in our example, each unit of \( L = 1 \), calculus method cannot be used.

Alternatively, where labour can be increased by a minimum of one unit, \( MP_L^1 \) can be obtained as

\[ MP_L^1 = TP_L - TP_{L-1} \]

The \( MP_L^1 \) worked out by this method is presented in Col. 3 of Table 6.1.
Average Productivity of Labour (APL) can be obtained by dividing the production function (6.4) by $L$. Thus,

$$\text{AP}_L = \frac{-L^2 + 15L + 10}{L} \quad \ldots (6.6)$$

Now $\text{AP}_L$ can be obtained by substituting the numerical value for $L$ in Eq. (6.6). $\text{AP}_L$ obtained by this method is given in Col. 4 of Table 6.2.

**Table 6.1 Three Stages of Production**

<table>
<thead>
<tr>
<th>No. of Workers (N)</th>
<th>Total Product (TP$_L$) (tonnes)</th>
<th>Marginal Product* (MP$_L$)</th>
<th>Average Product (AP$_L$)</th>
<th>Stages of Production (based on MP$_L$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24</td>
<td>24</td>
<td>1</td>
<td>I</td>
</tr>
<tr>
<td>2</td>
<td>72</td>
<td>48</td>
<td>36</td>
<td>Increasing</td>
</tr>
<tr>
<td>3</td>
<td>138</td>
<td>66</td>
<td>46</td>
<td>returns</td>
</tr>
<tr>
<td>4</td>
<td>216</td>
<td>78</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>300</td>
<td>84</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>384</td>
<td>84</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>462</td>
<td>78</td>
<td>66</td>
<td>II</td>
</tr>
<tr>
<td>8</td>
<td>528</td>
<td>66</td>
<td>66</td>
<td>Diminishing</td>
</tr>
<tr>
<td>9</td>
<td>576</td>
<td>48</td>
<td>64</td>
<td>returns</td>
</tr>
<tr>
<td>10</td>
<td>600</td>
<td>24</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>594</td>
<td>-</td>
<td>64</td>
<td>III</td>
</tr>
<tr>
<td>12</td>
<td>552</td>
<td>-</td>
<td>42</td>
<td>Negative returns</td>
</tr>
</tbody>
</table>

*MPS = TP$_L$ - TP$_{L-1}$, MP$_L$ calculated by differential method will be different from that given in Col. 3.

The information contained in Table 6.1 is presented graphically in panels (a) and (b) of Fig. 6.1. Panel (a) of Fig. 6.1 presents the total product curve (TP$_L$) and panel (b) presents marginal product (MP$_L$) and average product (AP$_L$) curves. The TP$_L$ schedule demonstrates the laws of returns to variable input ($L$). As the curve TP$_L$ shows, the total output increases at an increasing rate till the employment of the 5th worker, as indicated by the increasing slope of the TP$_L$ curve. (See also Col. 3 of the table). Employment of the 6th worker contributes as much as the 5th worker. Note that beyond the employment of the 6th worker, although TP$_L$ continues to increase (until the 10th worker), the rate of increase in TP$_L$ (i.e., MP$_L$) begins to fall. This shows the operation of the law of diminishing returns.
The three stages in production. Table 6.1 and Fig. 6.1 present the three usual stages in the application of the laws of returns to variable input.

In **Stage I**, $TP_L$ increases at increasing rate. This is indicated by the rising $MP_L$ till the employment of the 5th and 6th workers. Given the production function (6.4), the 6th worker produces as much as the 5th worker. The output from the 5th and the 6th workers represents an intermediate stage of constant returns to the variable factor, labour.

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

To conclude, the law of diminishing returns can be stated as follows. Given the employment of the fixed factor (capital), when more and more workers are employed, the return from the additional worker may initially increase but will eventually decrease.
Factors Behind the Laws of Returns. As shown in Fig. 6.1, the marginal productivity of labour ($MP_L$) increases in Stage I, whereas it decreases in Stage II. In other words, in Stage I, Law of Increasing Returns to variable input is in operation and in Stage II, the law of Diminishing Returns is in application. The reasons which underly the application of the laws of returns in Stages I and II may be described as follows.

One of the important factors causing increasing returns to a variable factor is the **indivisibility of fixed factor** (capital). The size of a capital is given and, given the technology, it requires an optimum number of labour. If labour is less than optimum number, capital cannot be divided to suit the number of workers. Therefore, if labour is less than its optimum number, capital remains underutilized. Let us suppose that optimum capital-labour combination is 1:6. If capital is indivisible and less than 6 workers are employed, then capital would remain underutilized. When more and more workers are added, utilization of capital increases. As a result, productivity of capital increases. This gives increasing returns to variable input, labour.

The second and the most important reason for increase in labour productivity is the **division of labour**, i.e., assigning works to labour according to their skill. This becomes possible with the employment of additional labour. Division of labour according to their skill and specialization increases productivity of labour.

Thus increase in productivity of both the factors, capital and labour, increases the output at increasing rate.

Once the optimum capital-labour ratio is reached, employment of additional labour amounts to underutilization of labour. So the productivity of marginal labour decreases. Also, with increasing number of workers, capital remaining the same, capital-labour ratio goes on decreasing. As a result, productivity of labour begins to decline. This marks the beginning of the second stage—the stage of decreasing returns.

Application of the Law of Diminishing Returns

The law of diminishing returns to the variable input is an empirical law, frequently observed in various production activities. This law, however, may not apply uniformly to all kinds of productive activities since it is not as true as the law of gravitation. In some productive activities, the law of diminishing returns may operate quickly, in some its operation may take a little longer time and in some others, it may not appear at all. This law has been found to operate in agricultural production more regularly than in industrial production. The reason is, in agriculture, natural factors play a predominant role whereas man-made factors play the major role in industrial production. Despite the limitations of the law, if increasing units of an input are applied to the fixed factors, the marginal returns to the variable input decrease eventually.
The Law of Diminishing Returns and Business Decisions. The law of diminishing returns as presented graphically has a relevance to the business decisions. The graph can help in identifying the rational and irrational stages of operations. It can also tell the business managers the number of workers (or other variable inputs) to apply to a given fixed input so that, given all other factors, output is maximum. As Fig. 6.1 exhibits, capital is presumably underutilized in Stage I. So a firm operating in Stage I is required to increase labour, and a firm operating in Stage III is required to reduce labour, with a view to maximizing its total production. From the firm’s point of view, setting an output target in Stages I and III is irrational because setting output target in stage I means underutilization of capital and setting output target in Stage III means accepting negative productivity of labour. The only meaningful and rational stage from the firm’s point of view is Stage II in which the firm can find answer to the question ‘how many workers to employ’ to maximize production, given the labour productivity. Figure 6.1 shows that the firm should employ a minimum of 7 workers and a maximum of 10 workers even if labour is available free of cost. This means that the firm has a limited choice—ranging from 7 to 10 workers. How many workers to employ against the fixed capital and how much to produce can be answered, only when the price of labour, i.e., wage rate, and product price are known. This question is answered below.

Optimum Employment of Labour

It may be recalled from Fig. 6.1 that an output maximizing coal-mining firm would like to employ 10 workers since at this level of employment, the output is maximum. The firm can, however, employ 10 workers only if workers are available free of cost. But labour is not available free of cost—the firm is required to pay wages to the workers. The objective of the firm to maximize profit, not output. Therefore, the question arises as to how many workers would the firm employ—10 or less or more than 10—to maximize its profit. A simple answer to this question is that the number of workers to be employed depends on the output that maximizes the firm’s profit, given the product price and the wage rate. This point can be proved as follows.

Profit is maximum where

\[ MC = MR \]

The same rule can be applied to determine the optimum level of employment, with some modification, of course. In our example here, since labour is the only variable input, marginal cost \( (MC) \) equals marginal wages \( (MW) \), i.e., \( MC = MW \).

As regards \( MR \), in case of factor employment, the concept of Marginal Revenue Productivity (MRP) is used. The marginal revenue productivity is the value of product resulting from the marginal unit of variable input (labour). In specific
terms, marginal revenue productivity (MRP) equals marginal physical productivity (MP_L) of labour multiplied by the price (P) of the product, i.e.,

$$\text{MRP} = \text{MP}_L \times P$$

### 6.3.2 Isoquants

In the preceding section, we have discussed the short-term laws of production, i.e., technological relationship between inputs and output assuming labour to be the only variable input, capital held constant. In this section, we proceed to discuss the long-term laws of production, i.e., the nature of relationship between inputs and output under the condition that both the inputs, capital and labour, are variable factors. In the long-run, supply of both the inputs is supposed to be elastic and, therefore, firms can use larger quantities of both labour and capital. With larger employment of capital labour, the scale of production increases. The nature of changing relationship between changing scale of inputs and output is referred to the laws of returns to scale. The laws of returns to scale are generally explained through the production function and isoquant curve technique. The most common and simple tool of analysis is isoquant curve technique. We will, therefore, first introduce and elaborate on this tool of analysis. The laws of returns to scale will then be explained through isoquant curve technique. The laws of returns to scale through production function will be explained in the next section.

**Isoquant: The Tool of Analyses**

The term ‘isoquant’ has been derived from the Greek word *iso* meaning ‘equal’ and Latin word *quantus* meaning ‘quantity’. The ‘isoquant curve’ is, therefore, also known as ‘Equal Product Curve’ or ‘Production Indifference Curve’. An isoquant curve can be defined as the locus of points representing various combinations of two inputs—capital and labour—yielding the same output. An ‘isoquant curve’ is analogous to an ‘indifference curve’, with two points of distinction: (a) an indifference curve is made of two consumer goods while an isoquant curve is constructed of two producer goods (labour and capital), and (b) an indifference curve represents an unmeasured level of satisfaction whereas an isoquant represents the quantity of output of a commodity.

Isoquant curves are drawn on the basis of the following assumptions:

(i) there are only two inputs, viz., labour (L) and capital (K), to produce a commodity X;
(ii) both L and K and product X are perfectly divisible;
(iii) the two inputs—L and K—can substitute each other but at a diminishing rate as they are imperfect substitutes; and
(iv) the technology of production is given.
Given these assumptions, it is technically possible to produce a given quantity of commodity $X$ with various combinations of capital and labour. The factor combinations are so formed that the substitution of one factor for the other leaves the output unaffected. This technological fact is presented through an isoquant curve ($IQ_1 = 100$) in Fig. 6.2. The curve $IQ_1$ all along its length represents a fixed quantity, 100 units of product $X$. This quantity of output can be produced with a number of labour-capital combinations. For example, points $A$, $B$, $C$, and $D$ on the isoquant $IQ_1$ show four different combinations of inputs, $K$ and $L$, as given in Table 6.2, all yielding the same output—100 units. Note that movement from $A$ to $D$ indicates decreasing quantity of $K$ and increasing number of $L$. This implies substitution of labour for capital such that all the input combinations yield the same quantity of commodity $X$, i.e., $IQ_1 = 100$.

**Table 6.2 Capital Labour Combinations and Output**

<table>
<thead>
<tr>
<th>Points</th>
<th>Input Combinations</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>$OK_1 + OL_1$</td>
<td>= 100</td>
</tr>
<tr>
<td>$B$</td>
<td>$OK_1 + OL_2$</td>
<td>= 100</td>
</tr>
<tr>
<td>$C$</td>
<td>$OK_1 + OL_3$</td>
<td>= 100</td>
</tr>
<tr>
<td>$D$</td>
<td>$OK_1 + OL_4$</td>
<td>= 100</td>
</tr>
</tbody>
</table>

**Marginal Rate of Technical Substitution (MRTS)**

An isoquant curve as explained above, shows various combinations of two inputs—labour and capital—that can technically produce the same quantity of a commodity. The movement along the isoquants means substituting one input for the other—capital for labour or labour for capital. Given the technology, the inputs are so
NOTES

substituted for one another that output remains the same. The rate at which one input is substituted for the other is called the marginal rate of technical substitution (MRTS).

The MRTS is a very important concept used in determining the shape and properties of isoquants and also in analyzing the production with two variable inputs. Therefore, let us look at the concept of the MRTS in detail before we proceed to discuss the properties of isoquants.

In simple words, MRTS is the rate at which a marginal unit of labour can be substituted for a marginal unit of capital or other way round, total output remaining the same. The MRTS gives the slope of the isoquant at different levels of input combinations.

The MRTS is measured as follows.

\[
MRTS = \frac{K_c - K_p}{L_c - L_p} = \frac{\Delta K}{\Delta L} = \text{Slope of the isoquant} \quad \ldots \(6.7\)
\]

where subscript \(C\) means current and \(p\) means previous.

It is important to note that the MRTS determined the shape of the isoquants. If two inputs are perfect substitutes, the isoquant takes the form of straight line sloping downward to the right and in case of imperfect substitutes—the most general case—the isoquant takes a curvilinear form convex to origin, as shown in Fig. 6.2.

6.3.3 Long-Run Laws of Production

Laws of Returns to Scale

Having introduced the isoquants—the basic tool of analysis—we now return to the theory of long-run production, specifically the laws of returns to scale. The laws of returns to scale explain the nature of change in output in response to a proportional and simultaneous change in inputs. Increasing inputs proportionately and simultaneously is, in fact, an expansion of the scale of production.

When a firm expands its scale of production, i.e., it increases both the inputs in a certain proportion, then there are three technical possibilities of increase in production:

(i) Total output may increase more than proportionately;
(ii) Total output may increase proportionately; and
(iii) Total output may increase less than proportionately.

Accordingly, there are three kinds of laws of returns to scale:

(i) The law of increasing returns to scale;
(ii) The law of constant returns to scale, and
(iii) The law of diminishing returns to scale.
So far as the sequence of the laws of ‘returns to scale’ is concerned, the law of increasing returns to scale is followed by the law of constant returns and then by the law of diminishing returns to scale. This is the most common sequence of the laws.

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

**Law of Increasing Returns to Scale**

When inputs, $K$ and $L$, are increased at a certain proportion and output increases more than proportionately, it exhibits the law of increasing returns to scale. For example, if quantities of both the inputs, $K$ and $L$, are successively doubled and the resultant output is more than doubled, then the law of returns to scale is said to be in operation. The increasing returns to scale is illustrated in Fig. 6.3. In Fig. 6.3, lines $OB$ and $OC$ represent the expansion path. The movement from point $a$ to $b$ on the line $OB$ means doubling the inputs. It can be seen in Fig. 6.3 that input-combination increases from $1K + 1L$ to $2K + 2L$. As a result of doubling the inputs, output is more than doubled: it increases from 10 to 25 units, i.e., an increase of more than double. Similarly, the movement from point $b$ to point $c$ indicates 50% increase in inputs as a result of which the output increases from 25 units to 50 units, i.e., by 100%. Clearly, output increases more than the proportionate increase in inputs. This kind of relationship between the inputs and output exemplifies the law of increasing returns to scale.

![Fig. 6.3 Increasing Returns to Scale](image)

**The Factors Behind Increasing Returns to Scale**

The factors that lead to increasing returns to scale are known as internal economics of scale. There are at least three plausible factors causing increasing returns to scale.
(i) **Technical and managerial indivisibilities.** Certain inputs, particularly mechanical equipments and managers, used in the process of production are available in a given size. Such inputs cannot be divided into parts to suit small scale of production. For example, half a turbine cannot be used and one-third or a part of a composite harvester and earth-movers cannot be used to produce the relevant product.

In regard to managerial manpower, if scale of production is small, managers remain under-employed because half of a production manager cannot be employed, if part-time employment is not acceptable to the manager. Because of indivisibility of machinery and managers, given the state of technology, they have to be employed in a minimum quantity even if scale of production is much less than the capacity output. Therefore, when scale of production is expanded by increasing all the inputs, the productivity of indivisible factors increases exponentially because of technological advantage. This results in increasing returns to scale.

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

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

**Law of Constant Returns to Scale**

When the increase in output is proportionate to the increase in inputs, it exhibits constant returns to scale. For example, if quantities of both the inputs, \( K \) and \( L \), are doubled and output is also doubled, then the returns to scale are said to be constant. Constant returns to scale are illustrated in Fig. 6.4. The lines \( OA \) and \( OB \) are ‘expansion paths’ indicating two hypothetical techniques of production with optimum capital-labour ratio. The isoquants marked \( Q = 10, Q = 20 \) and \( Q = 30 \) indicate the three different levels of output. In the figure, the movement from points \( a \) to \( b \) indicates doubling both the inputs—increasing capital from \( 1K \) to \( 2K \) and labour from \( 1L \) to \( 2L \). When inputs are doubled, output is also doubled, i.e., output increases from 10 to 20.
Similarly, the movement from point $b$ to $c$ indicates a 50 per cent increase in both labour and capital. This increase in inputs results in an increase of output from 20 to 30 units, i.e., a 50 per cent increase in output. In simple words, a 50 per cent increase in inputs leads to a 50 per cent increase in output. This relationship between a proportionate change in inputs and the same proportional change in outputs may be summed up as follows.

\[
\begin{align*}
1K + 1L & \Rightarrow 10 \\
2K + 2L & \Rightarrow 20 \\
3K + 3L & \Rightarrow 30
\end{align*}
\]

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

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

**Decreasing Returns to Scale**

The firms are faced with decreasing returns to scale when a certain proportionate increase in inputs, $K$ and $L$, leads to a less than proportionate increase in output. For example, when inputs are doubled and output is less than doubled, then decreasing returns to scale is in operation. The decreasing returns to scale is illustrated in Fig. 6.5. As the figure shows, when the inputs $K$ and $L$ are doubled,
i.e., when capital-labour combination is increased from $1K + L$ to $2K + 2L$, the output increases from 10 to 18 units. This means that when capital and labour are increased by 100 per cent, output increases by only 80 per cent. That is, increasing output is less than the proportionate increase in inputs. Similarly, movement from point $b$ to $c$ indicates a 50 per cent increase in the inputs. But, the output increases by only 33.3 per cent. This exhibits decreasing returns to scale.

**Fig. 6.5 Decreasing Return to Scale**

**Causes of Diminishing Return to Scale**

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

**Laws of Returns to Scale through Production Function**

The laws of returns to scale may be explained more precisely through a production function. Let us assume a production function involving two variable inputs ($K$ and $L$) and one commodity $X$. The production function may then be expressed as

$$Q_x = f(K, L) \quad \cdots (6.8)$$

where $Q_x$ denotes the quantity of commodity $X$.

Let us also assume that the production function is homogeneous. A production function is said to be homogeneous when all the inputs are increased in the same
production and the proportion can be factored out. And, if all the inputs are increased by a certain proportion (say, \( k \)) and output increases in the same proportion (\( k \)), then production is said to be homogeneous of degree 1. This kind of production function may be expressed as follows.

\[
kQ_i = f(kK, kL) \quad \text{...(6.9)}
\]

or

\[
k = k(K, L)
\]

A *homogeneous production function* of degree 1, as given in Eq. (6.9), implies *constant returns to scale*. Eq. (6.9) shows that increase in inputs, \( K \) and \( L \), by a multiple of \( k \), increases output, \( Q \), by the same multiple (\( k \)). This means constant returns to scale.

The constant returns to scale may not be applicable at all the levels of increase in inputs. Increasing inputs \( K \) and \( L \) in the same proportion may result in increasing or diminishing returns to scale. In other words, it is quite likely that if all the inputs are increased by a certain proportion, output may increase more or less than proportionately. For example, if all the inputs are doubled, the output may increase by less than or more than double. Then the production function may be expressed as

\[
hQ = f(kK, kL) \quad \text{...(6.10)}
\]

where \( h \) denotes \( h \)-times increase in \( Q \), as a result of \( k \)-times increase in inputs, \( K \) and \( L \). The proportion \( h \) may be greater than \( k \), equal to \( k \), or less than \( k \). Accordingly, it reveals the three laws of returns to scale:

(i) If \( h = k \), production function reveals constant returns to scale.
(ii) If \( h > k \), it reveals increasing returns to scale.
(iii) If \( h < k \), it reveals decreasing returns to scale.

6.4 COBB-DOUGLAS AND OTHER PRODUCTION FUNCTIONS: EMPIRICAL ESTIMATION

Cobb-Douglas Production Function—The Multiplicative Power Function

In the preceding section, the laws of returns to scale have been illustrated through a normal and a power function. In this section, we show the application of the most widely used production functions—the multiplicative *power function*. The most popular production function of this category is *Cobb-Douglas Production Function* of the form

\[
Q = AK^a L^b \quad \text{...(6.11)}
\]

where \( A \) is a positive constant; \( a \) and \( b \) are positive fractions; and \( b = 1 - a \).

The Cobb-Douglas production function is often used in its following form.

\[
Q = AK^a L^{1-a} \quad \text{...(6.12)}
\]
Properties of Cobb-Douglas Production Function. A power function of this kind has several important properties.

First, the multiplicative form of the power function can be changed into its log-linear form as

\[ \log Q = \log A + a \log K + b \log L \]  \hspace{1cm} (6.13)

In its logarithmic form, the function becomes simple to handle and can be empirically estimated using linear regression analysis.

Secondly, power functions are homogeneous and the degree of homogeneity is given by the sum of the exponents \( a \) and \( b \). If \( a + b = 1 \), then the production function is homogeneous of degree 1 and implies constant returns to scale.

Thirdly, parameters \( a \) and \( b \) represent the elasticity coefficient of output for inputs \( K \) and \( L \), respectively. The output elasticity coefficient \( (\varepsilon) \) in respect of capital may be defined as proportional change in output as a result of a given change in \( K \), keeping \( L \) constant. Thus,

\[ \varepsilon_k = \frac{\partial Q}{\partial K} \frac{Q}{Q} = \frac{\partial Q}{\partial K} \frac{K}{Q} \]  \hspace{1cm} (6.14)

By differentiating the production function \( Q = AK^aL^b \) with respect to \( K \) and substituting the result in Eq. (6.14), we can find the elasticity coefficient. We know that

\[ \frac{\partial Q}{\partial K} = a AK^{a-1} L^b \]

By substituting the values for \( Q \) and \( Q/K \), we get

\[ \varepsilon_k = a AK^{a-1} L^b \left( \frac{K}{AK^a L^b} \right) = a \]  \hspace{1cm} (6.15)

Thus, output-elasticity coefficient for \( K \) is \( 'a' \). The same procedure may be adopted to show that \( b \) is the elasticity coefficient of output for \( L \).

Fourthly, constants \( a \) and \( b \) represent the relative share of inputs, \( K \) and \( L \), in total output \( Q \). The share of \( K \) in \( Q \) is given by

\[ \frac{\partial Q}{\partial K} K \]

Similarly, the share of \( L \) in \( Q \) is given by

\[ \frac{\partial Q}{\partial L} L \]

The relative share of \( K \) in \( Q \) an e obtained as

\[ \frac{\partial Q}{\partial K} L = \frac{1}{Q} a AK^{a-1} L^b K = a \]

Similarly, it can be shown that \( b \) gives the relative share of \( L \) in \( Q \).
Finally, Cobb-Douglas production function in its general form, \( Q = K^\alpha L^{1-\alpha} \) implies that at zero cost, there will be zero production.

Some Input-Output Relationships

Some of the concepts used in production analysis can be easily derived from the Cobb-Douglas production function as shown below.

(i) Average Product (AP) of \( L \) and \( K \):
\[
AP_L = A \left( \frac{K}{L} \right)^{1-\alpha} \\
AP_K = A \left( \frac{L}{K} \right)^\alpha
\]

(ii) Marginal Product of \( L \) and \( K \):
\[
MP_L = aA \left( \frac{K}{L} \right)^{\alpha} = a \left( \frac{Q}{L} \right) \\
MP_K = (a-1)A \left( \frac{L}{K} \right)^{\alpha} = (1-a)\frac{Q}{K}
\]

(iii) Marginal Rate of Technical Substitution:
\[
MRTS_{L,K} = \frac{MP_L}{MP_K} = \left[ \frac{\alpha}{1-a} \right] \frac{K}{L}
\]

CES Production Function

In addition to the Cobb-Douglas production function, there are several other forms of production function, viz., ‘constant elasticity substitution’ (CES), ‘variable elasticity of substitution’ (VES), Leontief-type, and linear-type. Of these, the constant elasticity substitution (CES) production function is more widely used, apart from Cobb-Douglas production function. We will, therefore, discuss the CES production function briefly.

The CES production function is expressed as
\[
Q = A[\alpha K^\beta + (1-\alpha)L^\beta]^{-\beta} \quad \text{... (6.16)}
\]

or
\[
Q = A[\alpha L^\beta + (1-\alpha)K^\beta]^{-\beta} \quad \text{... (6.17)}
\]

where \( L = \text{labour} \), \( K = \text{capital} \), and \( A \), \( \alpha \) and \( \beta \) are the three parameters.

An important property of the CES production function is that it is homogeneous of degree 1. This can be proved by increasing both the inputs, \( K \) and \( L \), by a constant factor and finding the final outcome. Let us suppose that inputs \( K \) and \( L \) are increased by a constant factor \( m \). Then the production function given in Eq. (6.16) can be written as follows.
\[
Q’ = A[(\alpha mK)^\beta + (1-\alpha)(mL)^\beta]^{-\beta} \\
= A[m^\beta \{\alpha K^\beta + (1-\alpha)L^\beta\}]^{-\beta} \\
= (m^\beta)^{-\beta} \cdot A[\alpha K^\beta + (1-\alpha)L^\beta]^{-\beta} \quad \text{... (6.17)}
\]
As given in Eq. (6.16) the term $A[\alpha K^\beta + (1 - \alpha) L^\beta]^{\frac{1}{\beta}} = Q$. By substitution, therefore, we get

$$Q' = mQ$$

Thus, the CES production function is homogeneous of degree 1.

Given the production function (6.16), the marginal product of capital ($K$) can be obtained as

$$\frac{\delta Q}{\delta K} = \alpha \frac{Q^{\beta+1}}{K}$$

and of labour ($L$) as

$$\frac{\delta Q}{\delta K} = (1 - \alpha) \frac{Q^{\beta+1}}{L}$$

The rate of technical substitution (RTS) can be obtained as

$$RTS = \frac{\alpha}{1 - \alpha} \frac{L^{\beta+1}}{K}$$

**Merits of CES Production Function.** CES production function has certain advantages over the other functions:

(i) it is a more general form of production function;

(ii) it can be used to analyze all types of returns to scale, and

(iii) it removes many of the problems involved in the Cobb-Douglas production function.

**Limitations.** The CES production function has, however, its own limitations. Some economists claim that it is not a general form of production function as it does not stand the empirical test. In other words, it is difficult to fit this function to empirical data. Also, Uzawa finds that it is difficult to generalize this function to $n$-number of factors. Besides, in this production function, parameter $\beta$ combines the effects of two factors, $K$ and $L$. When there is technological change, given the scale of production, homogeneity parameter $\beta$ may be affected by both the inputs. This production function does not provide a measure to separate the effects on the productivity of inputs.

---

**Check Your Progress**

1. What is a variable input?
2. State the assumptions on which the law of diminishing returns is based.
6.5 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. A variable input is defined as one whose supply in the short-run is elastic, e.g., labour and raw material, etc. All the users of such factors can employ a larger quantity in the short-run as well as in the long-run.

2. The law of diminishing returns is based on the following assumptions:
   (i) labour is the only variable input, capital remaining constant;
   (ii) labour is homogeneous;
   (iii) the state of technology is given; and
   (iv) input prices are given.

6.6 SUMMARY

- Once business firms take decision on ‘what to produce’ the major issues that arise are ‘how to produce’ and ‘how much to produce’. These issues arise because achieving optimum efficiency in production and minimizing cost for a given production is one of the prime concerns of the business managers.
- In general sense of the term, ‘production’ means transforming inputs (labour, capital, raw materials, time, etc.) into an output with value added. This concept of production is however limited to only ‘manufacturing’.
- In economic sense, production process may take a variety of forms other than manufacturing.
- Inputs are classified as
  - fixed inputs, and
  - variable inputs.
- Production function is a mathematical presentation of input-output relationship. More specifically, a production function states the technological relationship between inputs and output in the form of an equation, a table or a graph. In its general form, it specifies the inputs required for the production of a commodity or service.
- There are two kinds of production functions:
  - Short-run production function; and
  - Long-run production function
- The marginal revenue productivity is the value of product resulting from the marginal unit of variable input (labour).
• The term ‘isoquant’ has been derived from the Greek word *iso* meaning ‘equal’ and Latin word *quantus* meaning ‘quantity’. The ‘isoquant curve’ is, therefore, also known as ‘Equal Product Curve’ or ‘Production Indifference Curve’.

• In simple words, MRTS is the rate at which a marginal unit of labour can be substituted for a marginal unit of capital or other way round, total output remaining the same.

• The factors that lead to increasing returns to scale are known as *internal economics of scale*. There are at least three plausible factors causing increasing returns to scale.

• The decreasing returns to scale are attributed to the *diseconomies of scale*. The economists find that the most important factor causing diminishing returns to scale is ‘the diminishing return to management’, i.e., managerial diseconomies.

• In addition to the Cobb-Douglas production function, there are several other forms of production function, viz., ‘constant elasticity substitution’ (CES), ‘variable elasticity of substitution’ (VES), Leontief-type, and linear-type.

### 6.7 KEY WORDS

• An isoquant curve: It can be defined as the locus of points representing various combinations of two inputs—capital and labour—yielding the same output.

• Marginal rate of technical substitution (MRTS): It is the rate at which a marginal unit of labour can be substituted for a marginal unit of capital or other way round, total output remaining the same.

### 6.8 SELF ASSESSMENT QUESTIONS AND EXERCISES

**Short Answer Questions**

1. What is the significance of the production function in economics?
2. Write a short note on the short-run and long-run production function.
3. Briefly mention the use of isoquant curves.
4. What is the Cobb-Douglas production function?

**Long Answer Questions**

1. Describe the theory of production and its managerial uses.
2. Discuss the application of the law of diminishing returns.
3. Explain the three kinds of laws of returns to scale.

6.9 FURTHER READINGS


Websites

https://blogs.economictimes.indiatimes.com/et-commentary/heres-how-indias-widening-income-distribution-can-be-redressed/

UNIT 7  FORMS OF MARKETS

7.0  INTRODUCTION

Knowing input–output relations and output-cost relations is not the objective of business firms. Profit maximization is the objective of business firms, as assumed in traditional theory of firm. Profit is maximized when \( TR - TC \) is maximized. Since \( TR = P \times Q \), where \( P \) = price, and \( Q \) = quantity sold, given the cost-output relation, the main task for profit maximization is to find the profit maximizing price. This takes us to the theory of price determination. The theory or model of price determination developed by economic pundits is based on the market conditions. Market conditions are determined by the number of firms in an industry and the nature and degree of competition among the firms. The number of firms and degree of inter-firm competition determine the market structure of the industry. The economists have proposed different kinds of theoretical models for price determination on the basis of the nature and the characteristics of the market. Market conditions provide the playing field of the business firms.

In this unit, you will study about the meaning and characteristics of market structure, price determination by demand and supply functions and pricing and output decisions under monopoly, monopolistic competition and oligopoly.

7.1  UNIT OBJECTIVES

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

- Identify the characteristics of the market structure
• Evaluate price determination by demand and supply functions
• Discuss pricing and output decisions under monopoly, monopolistic competition and oligopoly

7.2 MEANING AND CHARACTERISTICS OF MARKETS

In general usage, the term ‘market’ means a place or locality where goods and services are bought and sold, e.g., Connaught Place market, Chandani Chowk market, Karol Bagh Market, etc. The term ‘market’ is also used with reference to a product, e.g., wheat market, vegetable market (Sabji mandi), gold market, stock market, etc.

In economics, the word ‘market’ is used in a rather abstract sense. The market means a system by which sellers and buyers of a commodity interact to settle its price and the quantity to be bought and sold. According to Samuelson and Nordhaus, “A market is a mechanism by which buyers and sellers interact to determine the price and quantity of a good or service”. Market for a commodity consists of the buyers and sellers who interact to settle its price and quantity to be transacted. The sellers and buyers may be individuals, firms, factories, dealers and agents.

Some important aspects of the market concept are the following ones:

(i) A market need not be situated in a particular place or locality. The geographical area of a market depends on how far and wide are the buyers and sellers scattered. It may be as small as a fish market in a corner of a city or as large as the entire world, e.g., the global markets for arms, cars, electronic goods, aeroplanes, computers, oil, medicines, etc.

(ii) Buyers and sellers need not come into personal contact with each other. The transactions can be carried out through postal services or telecommunication system—telephone, fax, agents, or e-mail etc. People do buy many goods and services without ever meeting the supplier.

(iii) The word ‘market’ may refer to a commodity or service (e.g., fruit market, car market, share market, money market, paper market, labour market, etc.) or to a geographical area, Bombay market, Indian market, Asian market, or the world market.

(iv) The economists distinguish between markets also on the basis of (a) nature of goods and services, e.g., factor market, input market and output market; (b) number of firms and degree of competition, e.g., perfectly competitive market (very large number of firms), monopolistic market (many firms with differentiated products); oligopolistic market, and so on.
The Market System: An Overview

Market system refers to the process by which buyers and sellers of a product interact to settle the price of the product and carry out the sale-purchase transactions. The market system works on the basis of a basic market principle. The market principle is based on the fundamental laws of demand and supply. Buyers create the demand and sellers create the supply of the product. Demand and supply work as two opposite market forces. The working of the market system is governed by the fundamental laws of demand and supply. The laws of demand and supply play a crucial role in determining the price of a product and the size of the market—total demand and total supply. A clear understanding of how markets work is essential for business decisions on production and future planning. As noted above, market system works on the basis of two market forces—demand and supply. The demand and supply sides work on the laws of demand and supply, respectively.

7.2.1 Market Equilibrium and Changes in Market Equilibrium

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

The Concept of Market Equilibrium

In physical sense, the term equilibrium means the ‘state of rest’. In general sense, it means that forces working in opposite directions are in balance. In the context of market analysis, equilibrium refers to a state of market in which quantity demanded of a commodity equals the quantity supplied of the commodity. The equality of demand and supply produces an equilibrium price. The equilibrium price is the price at which quantity demanded of a commodity equals its quantity supplied. That is, at equilibrium price, demand and supply are in equilibrium. Equilibrium price is also called market-clearing price. Market is cleared in the sense that there is no unsold stock and no unsupplied demand.

Determination of Market Price

Equilibrium price of a commodity in a free market is determined by the market forces of demand and supply. In order to analyse how equilibrium price is determined, we need to integrate the demand and supply curves. For this purpose, let us use our earlier example of shirts. Let us suppose that the weekly market demand and supply schedules for shirts in Delhi are given as shown in Table 7.1.
### Table 7.1 Weekly Demand and Supply Schedules for Shirts

<table>
<thead>
<tr>
<th>Price per Shirt (( \text{₹} ))</th>
<th>Demand (( \text{₀₀₀ shirts} ))</th>
<th>Supply (( \text{₀₀₀ shirts} ))</th>
<th>Market Position</th>
<th>Effect on Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>80</td>
<td>10</td>
<td>Shortage</td>
<td>Rise</td>
</tr>
<tr>
<td>200</td>
<td>55</td>
<td>28</td>
<td>Shortage</td>
<td>Rise</td>
</tr>
<tr>
<td>300</td>
<td>40</td>
<td>40</td>
<td>Equilibrium</td>
<td>Stable</td>
</tr>
<tr>
<td>400</td>
<td>28</td>
<td>50</td>
<td>Surplus</td>
<td>Fall</td>
</tr>
<tr>
<td>500</td>
<td>20</td>
<td>55</td>
<td>Surplus</td>
<td>Fall</td>
</tr>
<tr>
<td>600</td>
<td>15</td>
<td>60</td>
<td>Surplus</td>
<td>Fall</td>
</tr>
</tbody>
</table>

As the table shows, there is only one price of shirts (\( \text{₹} 300 \)) at which quantity demanded per week equals the quantity supplied at 40 thousand shirts. It means that the shirt market in Delhi is in equilibrium at price \( \text{₹} 300 \). At all other prices, the shirt market is in disequilibrium—the state of imbalance between supply and demand. When market is in the state of disequilibrium, either demand exceeds supply or supply exceeds demand. As the table shows, at all prices below \( ₹300 \), demand exceeds supply showing shortage of shirts in the market. Likewise, at all prices above \( ₹300 \), supply exceeds demand showing excess supply.

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

**Market Mechanism: How the Market Brings about Balance**

*Market mechanism* is a process of interaction between the market forces of demand and supply to determine the equilibrium price. To understand how it works, let the price of shirts be initially set at \( ₹100 \). At this price, the number of shirts demanded (80,000) exceeds the quantity supplied (10,000) by 70 thousand shirts. The shortage gives sellers an opportunity to raise the price and it prepares buyers to accept and pay a higher price. As a result, price goes up. Increase in price enhances the profit margin. This induces firms to produce and sell more in order to maximize their profits. This trend continues till price rises to \( ₹300 \). As Table 7.1 shows, at price \( ₹300 \), the buyers are willing to buy 40 thousand shirts. This is exactly the number of shirts that sellers would like to sell at this price. At this price, there is neither shortage nor excess supply of shirts in the market. Therefore, \( ₹300 \) is the equilibrium price. The market is, therefore, in equilibrium.
Similarly, at all prices above ₹300, supply exceeds demand showing excess supply of shirts in the market. The excess supply forces the competing sellers to cut down the price in order to clear their unsold stock. Some firms find low price unprofitable and go out of the market and some cut down their production. Therefore, supply of shirts goes down. On the other hand, fall in price invites more customers. This process continues until price of shirts falls to ₹300. At this price, demand and supply are in balance and market is in equilibrium. Therefore, price at ₹300 per shirt is equilibrium price.

**Graphical Illustration of Price Determination.** The determination of equilibrium price is illustrated graphically in Fig. 7.1. The demand curve $DD'$ and the supply curve $SS'$ have been obtained by plotting the demand and supply schedules, (Table 7.1) on the price and quantity axes.

As Fig. 7.1 shows, demand and supply curves intersect at point $E$ determining the equilibrium price at ₹300. At this price, the quantity demanded (40 thousand shirts) equals the quantity supplied. Thus, the equilibrium price is ₹300 and equilibrium quantity is 40 thousand shirts. The equilibrium condition is not fulfilled at any other point on the demand and supply curves. Therefore, if price is set at any point other than ₹300, there would be either excess supply or shortage of shirts in the market.

![Fig. 7.1 Equilibrium of Demand and Supply: Price Determination](image-url)
far exceeds the demand. The excess supply equals 45 thousand shirts. The suppliers would, therefore, lower down the price gradually in order to get rid of the unsold stock and cut down the supply simultaneously. On the other hand, when price falls, the law of demand comes into operation and demand for shirts increases too. In this process, the supply-demand gap is reduced. This process continues until price reaches \( \$300 \) at point \( E \), the point of equilibrium where demand and supply are equal at 40 thousand shirts. At this price, the market is in equilibrium and there is no inherent force at work which can disturb it.

Likewise, if price is initially set at \( \$100 \), the buyers would be willing to buy 80 thousand shirts whereas suppliers would be willing to supply only 10 thousand shirts. Thus, there would be a shortage of 70 thousand shirts in the market as shown by the distance \( JK \) in Fig. 7.1. Due to shortage the buyers would be willing to pay a higher price. This will lead to increase in price which will encourage the suppliers to increase their supply. This process of adjustment will continue as long as demand exceeds supply. When price rises to \( \$300 \), the market reaches its equilibrium.

**Price Determination by Demand and Supply Functions**

In the previous section, we have illustrated graphically how equilibrium of demand and supply is determined at the point of intersection of the demand and supply curves. If demand and supply functions are known, the equilibrium quantity and equilibrium price can also be determined numerically. Suppose demand function for a commodity \( X \) is given as follows.

\[
Q_d = 150 - 5P_x
\]

and the supply function as

\[
Q_s = 10P_x
\]

We know that market equilibrium takes place when quantity supplied equals quantity demanded, i.e., at equilibrium, \( Q_s = Q_d \). By substituting supply and demand functions for \( Q_s \) and \( Q_d \), respectively, we get equilibrium price as follows.

\[
10P_x = 150 - 5P_x
\]

At equilibrium price \( P_x = 10 \), the quantities supplied and demanded are in equilibrium. Equilibrium demand and supply can be obtained by substituting 10 for \( x \) in demand and supply functions as shown below.

\[
10P_x = 150 - 5P_x
\]

\[
10(10) = 150 - 5(10) = 100.
\]
Fig. 7.2 Determination of Equilibrium Price and Quantity

The algebraic determination of equilibrium price and quantity is illustrated graphically in Fig. 7.2. The demand curve $DD'$ has been drawn by using the demand function $Q_d = 150 - 5P$, and the supply curve $SS'$ by using the supply function $Q_s = 10P$. As the figure shows, demand and supply curves intersect at point $P$. A perpendicular drawn from point $P$ to the quantity axis determines the equilibrium quantity at 100 units and a line drawn perpendicular from point $P$ to the price axis determines the equilibrium price at $10$. At this price, the quantity demanded equals the quantity supplied and hence the market is in equilibrium.

Shift in Demand and Supply Curves and Market Equilibrium

When demand increases or decreases price remaining constant, the demand curve shifts upwards or downwards, respectively, it is called shift in demand curve. Similarly, price remaining constant when supply increases or decreases, supply curve shifts rightward or leftward, respectively, it is called shift in supply curve. In this section, we show how market equilibrium is effected by the shift in demand and supply curves.

Shift in Demand Curve

Whenever there is a shift in the demand and/or supply curve, there is also a shift in the equilibrium point. The effect of shift in the demand curve on the equilibrium is shown in Fig. 7.3. Suppose that the initial demand curve is given by the curve $DD'$ and supply curve by $SS'$. The demand and supply curves intersect at point $P$. The equilibrium price is determined at $PQ$ and equilibrium quantity at $OQ$. Let the demand curve now shift from its position $DD'$ to $DD''$, supply curve remaining the same. The demand curve $DD''$ intersects the supply curve $SS'$ at point $M$. Thus, shift in the demand curve causes a shift in the equilibrium from point $P$ to point $M$. At the new market equilibrium, quantity demanded and supplied increases from
$OQ$ to $ON$ and price increases from $PQ$ to $MN$. Note that, the supply curve remaining the same, a rightward shift in the demand curve results in a higher equilibrium price and quantity.

**Shift in Supply Curve**

Figure 7.4 shows the effect of shift in the supply curve on the equilibrium. Suppose that the demand curve is given as $DD'$ and the initial supply curve as $SS'$. The curves $DD'$ and $SS'$ intersect at point $P$, determining equilibrium price at $PQ$ and equilibrium quantity at $OQ$. Let the supply curve now shift from its position $SS'$ to $SS^\prime_2$, demand curve remaining unchanged. The new supply curve $SS^\prime_2$ intersects the demand curve at point $M$. Thus, a new equilibrium is struck at point $M$ where equilibrium price is $MN$ and equilibrium quantity is $ON$. Note that a rightward shift in the supply curve, demand curve remaining the same, causes equilibrium price to fall and output to increase.

**Simultaneous Shift in Demand and Supply Curves**

We have seen above how a rightward shift in the demand curve causes a *rise* in market price (Fig. 7.3), and how a rightward shift in the supply curve (Fig. 7.4) causes a *fall* in the market price. Let us now look at the effect of simultaneous and parallel shifts in demand and supply curves on the equilibrium price and output. The effect of a simultaneous and parallel rightward shift in demand and supply curves on the equilibrium price and output depends on how big or small is the relative shift in demand and supply curves. The *simultaneous and parallel* shift in demand and supply curves in different measures and its effect on equilibrium price and output are illustrated in parts (a) and (b) of Fig. 7.5.

Part (a) of Fig. 7.5 shows that if the shift in the supply curve is bigger than that in the demand curve, then equilibrium price decreases and output increases. For example, suppose that initial demand and supply curves are given by $DD_1$ and $SS_1$, respectively, intersecting at point $E_1$ and determining equilibrium price at $P_1$ and output at $Q_1$. Let the demand curve shift to $DD_2$ and supply curve from $SS_1$ to $SS_2$, intersecting at point $E_2$. Note that shift in the supply curve is bigger than that in...
the demand curve. As a result, equilibrium price falls to $P_3$ and output increases to $Q_3$. But, if demand and supply curves shift in equal measure, as shown by $DD_2$ and $SS_2$, equilibrium price remains unchanged though output increases to $Q_2$.

Similarly, part (b) of the figure shows the effect of a bigger shift in the demand curve on the equilibrium price and output. It can be seen in the figure that the shift in demand curve from $DD_1$ to $DD_2$ is bigger than the shift in the supply curve from $SS_1$ to $SS_2$. In this case, both equilibrium price and output increase.

### 7.2.2 Market Equilibrium: Practical Importance

The following constitutes the practical importance of market equilibrium:

- It facilitates in anticipating specific results of changes in the economic factors.
- It assists in deriving realistic solutions for practical economic problems.
- It helps in defining the equilibrium position of the entire economy.
- It assists in understanding the functioning of the working economic system.
- It provides the foundation for modern welfare economics.
- It is helpful in comprehending the market problems and various pricing processes.

### 7.3 PRICING FUNCTIONS UNDER DIFFERENT MARKET STRUCTURES

Let us study the pricing functions related with the various market structures.

#### 7.3.1 Market Structures and Pricing Decisions

Market structure is extremely complex and so is the system of price determination. This has lead to formation of different theories of price determination under different
kinds of market conditions. In this section, we discuss briefly the market structure – the market morphology.

**Market structure** refers to the number of firms in an industry and the degree of competition among the firms. The categorization of market structure is based on the following factors.

- Number of firms—the sellers
- Degree and nature of competition
- Level of product differentiation
- Possibility of entry and exit of the firms

The number of sellers of a product in a market determines the nature and degree of competition in the market. The nature and degree of competition make the structure of the market. Depending on the number of sellers and the degree of competition, the market structure is broadly classified as given in Table 7.2.

<table>
<thead>
<tr>
<th>Market structure degree of</th>
<th>No. of firms and industry production differentiation</th>
<th>Nature of over price where prevalent</th>
<th>Control of marketing</th>
<th>Method of</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Perfect Competition</td>
<td>Large no. of firms with financial markets and some homogeneous products</td>
<td>Farm products</td>
<td>None</td>
<td>Market exchange or auction</td>
</tr>
<tr>
<td>(a) Monopolistic competition</td>
<td>Many firms with real or perceived product differentiation</td>
<td>Manufacturing: tea, toothpaste, TV sets, shoes, refrigerators, etc.</td>
<td>Some</td>
<td>Competitive advertising, quality rivalry</td>
</tr>
<tr>
<td>(b) Oligopoly</td>
<td>Little or no product differentiation</td>
<td>Aluminium, steel, cigarettes, cars, passenger cars, etc.</td>
<td>Some</td>
<td>Competitive, advertising, quality rivalry</td>
</tr>
<tr>
<td>(c) Monopoly</td>
<td>A single producer, without close substitute</td>
<td>Public utilities: Telephones, Electricity, etc., regulated</td>
<td>Considerable but usually large</td>
<td>Promotional advertising if supply</td>
</tr>
</tbody>
</table>


**Market Structure and Pricing Decisions**

The market structure determines a firm’s power to fix the price of its product a great deal. The degree of competition determines a firm’s degree of freedom in determining the price of its product. The degree of freedom implies the extent to which a firm is free or independent of the rival firms in taking its own pricing decisions.
decisions. Depending on the market structure, the degree of competition varies between zero and one. And, a firm’s discretion or the degree of freedom in setting the price for its product varies between one and none in the reverse order of the degree of competition. As a matter of rule, the higher the degree of competition, the lower the firm’s degree of freedom in pricing decision and control over the price of its own product and vice versa. Let us now see how the degree of competition affects pricing decisions in different kinds of market structures.

Under perfect competition, a large number of firms compete against each other for selling their product. Therefore, the degree of competition under perfect competition is close to one, i.e., the market is highly competitive. Consequently, firm’s discretion in determining the price of its product is close to none. In fact, in perfectly competitive market, price is determined by the market forces of demand and supply and a firm has to accept the price determined by the market forces. If a firm uses its discretion to fix the price of its product above or below its market level, it loses its revenue and profit in either case. For, if it fixes the price of its product above the ruling price, it will not be able to sell its product, and if it cuts the price down below its market level, it will not be able to cover its average cost. In a perfectly competitive market, therefore, firms have little or no choice in respect to price determination.

As the number of firms decreases, the degree of competition decreases. And, as a result, firm’s control over the price and its discretion in pricing decision increases. For example, under monopolistic competition, where degree of competition is high but less than one, the firms have some discretion in setting the price of their products. Under monopolistic competition, the degree of freedom depends largely on the number of firms and the level of product differentiation. Where product differentiation is real, firm’s discretion and control over the price is fairly high and where product differentiation is nominal or only notional, firm’s pricing decision is highly constrained by the prices of the rival products.

When the number of firms is few, the market takes the form of an oligopoly. Under oligopoly, the degree of competition is quite low, lower than that under monopolistic competition. The firms, therefore, have a good deal of control over the price of their products and can exercise their discretion in pricing decisions, especially where product differentiation is prominent. However, the fewness of the firms gives them an opportunity to form a cartel or to make some settlement among themselves for fixation of price and non-price competition.

In case of a monopoly, the degree of competition is close to nil. An uncontrolled monopoly firm has full freedom to determine the price of its product. A monopoly, in the true sense of the term, is free to fix any price for its product, of course, under certain constraints, viz., (i) the objective of the firm, and (ii) demand conditions.

The theory of pricing provides different models of how price is determined in different kinds of market structures. The characteristics of different kinds of
market structures and price determination in each type of market has been discussed in a theoretical framework in the subsequent section since most markets in modern times are either digopolistic or monopolistically competitive. It implies that most firms have some monopoly power.

Check Your Progress
1. What do you understand by the term ‘market clearing price’?
2. What are the factors which determine the categorization of market structure?

7.4 PRICE AND OUTPUT DECISIONS UNDER DIFFERENT COMPETITIVE CONDITIONS

Let us study the pricing and output decisions under different competitive conditions namely, monopoly, monopolistic competition and oligopoly.

7.4.1 Monopoly

Let us begin with the theory of price and output determination in another kind of rare market – monopoly. In modern times, most monopolies are established or governed by the government. Price in government monopolies and government governed monopolies are determined under social considerations, not under market conditions. In this section, however, we discuss price and output determination under the conditions of private monopolies. Let us first look at definition of monopoly and factors that lead to emergence of private monopolies.

Definition of Monopoly

The term pure monopoly means an absolute power of a firm to produce and sell a product that has no close substitute. In other words, a monopolized market is one in which there is only one seller of a product having no close substitute. The cross elasticity of demand for a monopoly product is either zero or negative. A monopolized industry is a single-firm industry. Firm and industry are identical in a monopoly setting. In a monopolized industry, equilibrium of the monopoly firm signifies the equilibrium of the industry.

However, the precise definition of monopoly has been a matter of opinion and purpose. For instance, in the opinion of Joel Deal, a noted authority on managerial economics, a monopoly market is one in which ‘a product of lasting distinctiveness is sold. The monopolized product has distinct physical properties recognized by its buyers and the distinctiveness lasts over many years.’ Such a definition is of practical importance if one recognizes the fact that most of the commodities have their substitutes varying in degree and it is entirely for the consumers/users to distinguish between them and to accept or reject a commodity as a substitute. Another concept of pure monopoly has been offered by E.H. Chamberlin who envisages monopoly as the control of all goods and services by...
Monopoly Pricing and Output Decision

Short-Run Analysis

The theory of pricing and output determination under monopoly is based on profit maximization hypothesis, given the revenue and cost conditions. Although cost conditions, i.e., AC and MC curves, in a competitive and monopoly market are generally identical, revenue conditions, i.e., AR and MR curves, are different under monopoly. Unlike a competitive firm, a monopoly firm faces a downward sloping demand curve. The reason is a monopolist has the option and power to reduce the price and sell more and to raise the price and sell less. Therefore, given the price-demand relationship, demand curve under monopoly is a typical downward sloping demand curve.

In case of a demand curve is sloping downward, marginal revenue (MR) curve lies below the AR curve and, technically, the slope of the MR curve is twice that of AR curve.

The short-run revenue and cost conditions faced by a monopoly firm are presented in Fig. 7.6. Firm’s average and marginal revenue curves are shown by the AR and MR curves, respectively, and its short-run average and marginal cost curves are shown by SAC and SMC curves, respectively. The price and output decision rule for profit maximizing monopoly is the same as for a firm in the competitive industry.

As noted earlier, profit is maximized at the level of output at which $MC = MR$. Given the profit maximization condition, a profit maximizing monopoly firm
chooses a price-output combination at which \( MR = SMC \). Given the monopoly firm’s cost and revenue curves in Fig. 7.6, its \( MR \) and \( SMC \) intersect at point \( N \). An ordinate drawn from point \( N \) to \( X \)-axis, determines the profit maximizing output at \( OQ \). The ordinate \( NQ \) extended to the demand curve (\( AR = D \)) gives the profit maximizing price \( PQ \). It means that given the demand curve, the output \( OQ \) can be sold per time unit at only one price, i.e., \( PQ (= OP) \). Thus, the determination of equilibrium output simultaneously determines the equilibrium price for the monopoly firm. Once price is fixed, the unit and total profits are also simultaneously determined. Hence, the monopoly firm is in a state of equilibrium.

At output \( OQ \) and price \( PQ \), the monopoly firm maximizes its unit and total profits. Its per unit monopoly or economic profit (i.e., \( AR – SAC \)) equals \( PQ – MQ = PM \). Its total profit, \( \pi = OQ \times PM \). Since \( OQ = P_m, \pi = P \times PM = \text{area} \, OQ \times PM \), as shown by the shaded rectangle. Since cost and revenue conditions are not expected to change in the short-run, the equilibrium of the monopoly firm will remain stable.

**Determination of Monopoly Price and Output: Algebraic Solution**

The determination of price and output by a monopoly firm in the short-run is illustrated above graphically (see Fig. 7.6). Here, we present an algebraic solution to the problem of determination of equilibrium price and output under monopoly.

Suppose demand and total cost functions for a monopoly firm are given as follows.

Demand function : 
\[ Q = 100 – 0.2 \, P \] …(7.1.1)

Price function : 
\[ P = 500 – 5Q \] …(7.1.2)

Cost function : 
\[ TC = 50 + 20Q + Q^2 \] …(7.2)

Given the demand and cost functions, the problem before the monopoly firm is to find the profit maximizing output and price. The problem can be solved as follows.

We know that profit is maximum at output at which \( MR = MC \). So the first step is to find \( MR \) and \( MC \) from the demand function and cost function, respectively. We have noted earlier that \( MR \) and \( MC \) are the first derivation of \( TR \) and \( TC \) functions, respectively. \( TC \) function is given, but \( TR \) function is not. So, let us find \( TR \) function first. We know that

\[ TR = PQ \]

Since \( P = 500 – 5Q \), by substitution, we get

\[ TR = (500 – 5Q) \times Q \]

\[ TR = 500Q – 5Q^2 \] …(7.3)

Given the \( TR \) function (7.3), \( MR \) can be obtained by differentiating the \( TR \) function.
Forms of Markets

\[ MR = \frac{\partial TR}{\partial Q} = 500 - 10Q \]

Likewise, \( MC \) can be obtained by differentiating the \( TC \) function (7.2).

\[ MC = \frac{\partial TR}{\partial Q} = 20 + 2Q \]

Now that \( MR \) and \( MC \) function are known, profit maximizing output can be easily obtained. Recall that profit is maximum where \( MR = MC \). As given above,

\[ MR = 500 - 10Q \]

and

\[ MC = 20 + 2Q \]

By equalizing the \( MR \) and \( MC \) functions, we get profit maximizing output \( (Q) \) as follows.

\[ MR = MC \]

\[ 500 - 10Q = 20 + 2Q \]

\[ 480 = 12Q \]

\[ Q = 40 \]

The output \( Q = 40 \) is the profit maximizing output.

Now profit maximizing price can be obtained by substituting 40 for \( Q \) in the price function (7.1.2). Thus, \( P = 500 - 5(40) = 300 \). Profit maximizing price is `300.

Given the \( TR \) and \( TC \) functions, the total profit \( (\pi) \) can be obtained as follows.

\[ \pi = TR - TC \]

By substitution, we get

\[ \pi = 500Q - 5Q^2 - (50 + 20Q + Q^2) \]

\[ = 500Q - 5Q^2 - 50 - 20Q - Q^2 \]

By substituting profit maximizing output (40) for \( Q \), we get

\[ \pi = 500(40) - 5(40)(40) - 50 - 20(40) - (40 \times 40) \]

\[ = 20,000 - 8,000 - 50 - 800 - 1600 = 9,550 \]

Thus, total maximum profit comes to `9,550.

Does a Monopoly Firm Always Earn Abnormal Profit?

There is no certainty that a monopoly firm will always earn an economic or abnormal profit. Whether a monopoly firm earns abnormal profit or normal profit or incurs loss depends on:

(i) its cost and revenue conditions;
Forms of Markets

NOTES

Self-Instructional Material

(ii) threat from potential competitors; and
(iii) government policy in respect of monopoly.

If a monopoly firm operates at the level of output where \( MR = MC \), its profit depends on the relative levels of \( AR \) and \( AC \). Given the level of output, there are three possibilities.

(i) if \( AR > AC \), there is abnormal profit for the firm,
(ii) if \( AR = AC \), the firm earns only normal profit, and
(iii) if \( AR < AC \), though only a theoretical possibility, the firm makes losses.

Long-Run analysis

The decision rules regarding optimal output and pricing in the long-run are the same as in the short-run. In the long-run, however, a monopolist gets an opportunity to expand the scale of its production with a view to enhancing its long-run profits. The expansion of the plant size may, however, be subject to such conditions as

\( a \) size of the market,
\( b \) expected economic profit, and
\( c \) risk of inviting legal restrictions. Let us assume, for the time being, that none of these conditions limits the expansion of a monopoly firm and discuss the price and output determination in the long-run.

The equilibrium of monopoly firm and its price and output determination in the long-run is shown in Fig. 7.7. The \( AR \) and \( MR \) curves show the market demand and marginal revenue conditions faced by the monopoly firm. The \( LAC \) and \( LMC \) show the long-run cost conditions. It can be seen in Fig. 7.7, that monopoly’s \( LMC \) and \( MR \) intersect at point \( P \) determining profit maximizing output at \( OQ_2 \). Given the \( AR \) curve, the price at which the total output \( OQ_2 \) can be sold is \( P_2Q_2 \). Thus, in the long-run, equilibrium output will be \( OQ_2 \) and price \( P_2Q_2 \). This output-price combination maximizes monopolist’s long-run profit. The total long-run monopoly profit is shown by the rectangle \( LMSP_2 \).

Fig. 7.7 Equilibrium of Monopoly in the Long-run
Comparison of Short-run and Long-run Equilibrium

It can be seen in Fig 7.7 that compared to short-run equilibrium, the monopolist produces a larger output and charges a lower price and makes a larger monopoly profit in the long-run. In the short-run, monopoly’s equilibrium is determined at point $A$, the point at which $SMC_1$ intersects the $MR$ curve. Thus, monopoly’s short-run equilibrium output is $OQ_1$, which is less than long-run output $OQ_2$. But the short-run equilibrium price $P_1Q_1$ is higher than the long-run equilibrium price $P_2Q_2$. The total short-run monopoly profit is shown by the rectangle $JP_1TK$ which is much smaller than the total long-run profit $LP_2SM$. This, however, is not necessary: it all depends on the short-run and long-run cost and revenue conditions.

It may be noted at the end that if there are barriers to entry, the monopoly firm may not reach the optimal scale of production ($OQ_2$) in the long-run, nor can it make full utilization of its existing capacity. The firm’s decision regarding plant expansion and full utilization of its capacity depends solely on the market conditions.

7.4.2 Monopolistic Competition

By definition, monopolistic competition refers to a market structure in which a large number of sellers sell differentiated products, which are close substitutes for one another. Incidentally, a close substitute is one whose cross-elasticity is close to unity or greater. Monopolistic competition combines the basic elements of both perfect competition and monopoly.

The element of monopoly in monopolistic competition arises from the fact that each firm has an absolute right to produce and sell a branded or patented product. Other firms are prevented by laws from producing and selling a branded product of other firms. This gives a firm monopoly power over production, pricing and sale of its own-branded product. For example, consider toilet soap industry. There are a number of brand names available in the market, e.g., Lux, Liril, Palmolive, Fairglow, Pears, Rexona, Lifebuoy, Carmel, Godrej, Cinthol, Ponds, Dove, Dettol and so on. Each of these branded toilet soaps is produced and sold by a company having monopoly power over the product. Similarly, Maruti Udyog Limited has monopoly power for producing and selling cars under the brand name Maruti. No other car manufacturing company can produce and sell cars under this brand name, so is the case with all other car manufacturing companies.

The element of competition arises from the fact that each generic branded product has several close substitutes and firms selling branded products of the same generic category have to compete for the market share. Considering again our example of toilet soaps, all the companies producing and selling branded toilet soaps are in intensive competition for capturing the largest possible market share. One index of the competition between them is the amount that they spend advertising their product. These features of the toilet soap industry make it monopolistically competitive. Toothpaste industry with a number of branded product names (e.g., Binaca, Colgate, Close-up, Pepsodent, Forhans, Cibaca, Neem, Meswak, Signal,
Promise, Prestige and so on) is another example of monopolistic competition. So is the case with major industrial products in India, e.g., electrical tubes and bulbs, TV sets, refrigerators, air conditioners, personal computers, textile goods, tea, coffee, cigarettes, soft drinks, cold creams, shampoos, detergents, shaving blades, shaving cream, hair oils, hair dyes, shoes, wrist watches, steel, cement, mobile phones and so on.

Some of the industries looking monopolistically competitive may be oligopolistic in which there are only a few sellers selling differentiated or homogenous products. The question as to what makes a market monopolistically competitive or oligopolistic will be taken up in the next chapter which deals with oligopoly markets. Let us now look at the general characteristics of monopolistic competitions.

**Price and output Decisions in the Short-Run**

The price and output determination models of monopolistic competition was developed by Chamberlin as explained here.

Although monopolistic competition is characteristically close to perfect competition, pricing and output decisions under this kind of market are similar to those under monopoly. The reason is that a firm under monopolistic competition faces a downward sloping demand curve. This kind of demand curve is the result of (i) a strong preference of a section of consumers for the product and (ii) the quasi-monopoly of the seller over the supply. The strong preference or brand loyalty of the consumers gives the seller an opportunity to raise the price and yet retain some customers. Besides, since each product is a substitute for the other, the firms can attract the consumers of other products by lowering their prices.

![Fig. 7.8 Price-Output Determination under Monopolistic Competition](image)

The short-term pricing and output determination under monopolistic competition is illustrated in Fig. 7.8. The AR and MR curves and SMC and SAC
Price and Output Determination in the Long-Run

The mechanism of price and output determination in the long-run under monopolistic competition is illustrated graphically in Fig. 7.9. To begin the analysis, let us suppose that, at some point of time, firm’s revenue curves are given as $AR$ and $MR$, and long-run cost curves as $LAC$ and $LMC$. As the figure shows, $MR$ and $LMC$ intersect at point $M$ determining the equilibrium output at $OQ_2$ and price at $P_2Q_2$. At price $P_2Q_2$, the firms make a supernormal or economic profit of $P_2T$ per unit of output. This situation is similar to short-run equilibrium.

Let us now see what happens in the long run. The supernormal profit brings about two important changes in a monopolistically competitive market in the long run.

**First**, the supernormal profit attracts new firms to the industry. As a result, the existing firms lose a part of their market share to new firms. Consequently, their demand curve shifts downward to the left. This kind of change in the demand
curve is shown in Fig. 7.9 by the shift in AR curve from AR₁ to AR₂ and the MR curve from MR₁ to MR₂.

Second, the increasing number of firms intensifies the price competition between them. Price competition increases because losing firms try to regain or retain their market share by cutting down the price of their product. And, new firms in order to penetrate the market set comparatively low prices for their product. The price competition increases the slope of the firms’ demand curve or, in other words, it makes the demand curve more elastic. Note that AR₂ has a greater slope than AR₁ and MR₂ has a greater slope than MR₁.

The ultimate picture of price and output determination under monopolistic competition is shown at point P₁ in Fig. 7.9. As the figure shows, LMC intersects MR₂ at point N where firm’s long-run equilibrium output is determined at OQ₁ and price at P₁Q₁. Note that price at P₁Q₁ equals the LAC at the point of tangency. It means that under monopolistic competition, firms make only normal profit in the long-run. Once all the firms reach this stage, there is no attraction (i.e., super normal profit) for the new firms to enter the industry, nor is there any reason for the existing firms to quit the industry. This signifies the long-run equilibrium of the industry.

7.4.3 Oligopoly

Oligopoly is defined as a market structure in which there are a few sellers selling homogeneous or differentiated products. In case oligopoly firms sell a homogeneous product, it is called pure or homogeneous oligopoly. For example, industries producing bread, cement, steel, petrol, cooking gas, chemicals, aluminium and sugar are industries characterized by homogeneous oligopoly. And, where firms of an oligopoly industry sell differentiated products, it is called differentiated or heterogeneous oligopoly. Automobiles, television sets, soaps and detergents, refrigerators, soft drinks, computers, cigarettes, etc. are some examples of industries characterized by differentiated or heterogeneous oligopoly.

Be it pure or differentiated, “Oligopoly is the most prevalent form of market organization in the manufacturing sector of the industrial nations…” In non-industrial nations like India also, a majority of big and small industries have acquired the features of oligopoly market. The market share of 4 to 10 firms in 84 big and small industries of India is given in Table 7.3.

<table>
<thead>
<tr>
<th>Market share (%)</th>
<th>No. of industries</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – 24.9</td>
<td>8</td>
</tr>
<tr>
<td>25 – 49.9</td>
<td>11</td>
</tr>
<tr>
<td>50 – 74.9</td>
<td>15</td>
</tr>
<tr>
<td>75 – 100</td>
<td>50</td>
</tr>
<tr>
<td>Total</td>
<td>84</td>
</tr>
</tbody>
</table>
Oligopoly Models: An Overview

As already mentioned, under oligopolistic conditions, rival firms indulge in an intricate pattern of actions, reactions and counter-actions showing a variety of behavioural patterns. As Baumol puts it, "Under [these] circumstances, a very wide variety of behaviour pattern becomes possible. Rivals may decide to get together and cooperate in the pursuit of their objectives,… or, at the other extreme, may try to fight each other to the death. Even if they enter an agreement, it may last or it may break down." The economists have, therefore, found it extremely difficult to make a systematic analysis of price and output determination under oligopoly. This has, however, not deterred the economists from their efforts to build theoretical models to analyse the behaviour of oligopoly firms.

In accordance with the wide variety of behavioural patterns, the economists have developed a variety of analytical models based on different behavioural assumptions. The widely known oligopoly models include Cournot’s duopoly model (1838), Bertrand’s leadership model (1880), Edgeworth’s duopoly model (1897), Stackelberg’s model (1933), Sweezy’s kinked demand curve model (1939), Neumann and Morgenstern Game Theory model (1944) and Baumol’s sales maximization model (1959). None of these models, however, provides a universally acceptable analysis of oligopoly, though these models do provide an insight into oligopolistic behaviour.

In this section, we discuss some selected oligopoly models with the purpose of showing the behaviour of oligopoly firms and working of the oligopolistic markets. The analytical models discussed here are selected on the basis of how price and output are determined under price competition among oligopoly firms and cartel system. Specifically, we discuss here the following oligopoly models.

1. Cournot’s Model of Oligopoly

Augustin Cournot, a French economist, was the first to develop a formal oligopoly model in 1838 in the form of a duopoly model. Cournot developed his model with the example of two firms, each owning a well of mineral water and water being produced at zero cost. To illustrate his model, Cournot made the following assumptions.

(a) There are two firms, each owning an artesian mineral water well;
(b) Both the firms operate their wells at zero cost;
(c) Both of them face a demand curve with constant negative slope;
(d) Each seller acts on the assumption that his competitor will not react to his decision to change his output and price. This is Cournot’s behavioural assumption.
On the basis of this model, Cournot has concluded that each seller ultimately supplies one-third of the market and both the firms charge the same price. And, one-third of the market remains unsupplied.

Cournot’s duopoly model is illustrated in Fig. 7.10. The demand curve for mineral water is given by the AR curve and their MR by the MR curve. To begin with, let us suppose that firm A is the only seller of mineral water in the market. By assumption, its $MC = 0$. Following the profit maximizing rule, it sells quantity $OQ_A$ where its $MC = 0 = MR$, at price $OP_A$. Its total profit is $OP_AOQ_A$. This is the maximum profit seller A can make given the demand curve.

![Fig. 7.10 Price and Output Determination under Duopoly: Cournot’s Model](image)

Now let another firm B enter the market. The market open to B is the market unsupplied by A. This market equals $QM$ which is half of the total market. That is, B can sell its product in the remaining half of the market. B assumes that A will not change its price and output because A is making maximum profit. In other words, B assumes that A will continue to sell $OQ_A$ at prices $OP_A$. Thus, the market available to firm B is $QM$ and the relevant part of the demand curve is $PM$. The $MR$ curve corresponding to B’s demand curve $PM$ is given by $PN$. B’s MR curve, $PN$, it bisects $QM$ at point $N$ where $QN = NM$. In order to maximize its revenue, B sells $QN$ at price $OP_B = PN$. Its total revenue is maximum at $QRPN$ which equals its total profit. Note that B supplies only $QN = (1/2)/2 = 1/4$ of the market.

With the entry of firm B, price falls to $OP_B$. Price falls because A’s customers will switch over to firm B. Therefore, firm A is forced to reduce its price. Faced with this situation, firm A adjusts its price and output to the changed conditions. Firm A assumes that firm B will not change its output $QN$ and price $OP_B$ as it is making maximum profit. Accordingly, firm A assumes that firm B will continue to supply $QN = 1/4$ of the market. Thus, firm A assumes that it has $3/4 (= 1 - 1/4)$ of the market available to it. To maximize its profit, firm A supplies 1/2 of the remaining $3/4$ of the market, i.e., $1/2 \times 3/4 = 3/8$ of the market. It is noteworthy that A’s market share has fallen from 1/2 to 3/8.
Now it is firm B’s turn to react. Following Cournot’s assumption, firm B assumes that firm A will continue to supply only 3/8 of the market and the rest of the market is open to him, which equals $1 - 3/8 = 5/8$. To maximize his profit under the new conditions, firm B supplies half of the remaining market, i.e., firm B supplies $1/2 \times 5/8 = 5/16$ of the market. It is now for firm A to reappraise the situation and adjust his price and output accordingly.

This process of action and reaction continues in successive periods. In the process, firm A continues to lose his market share and firm B continues to gain. Eventually, a situation is reached when their market shares equal 1/3 each. Any further attempt to adjust output produces the same result. The firms, therefore, reach their equilibrium position where each one supplies one-third of the market and both charge the same price and one-third of the market remains unsupplied.

2. Sweezy’s Model of Oligopoly: Kinked-Demand Curve Model

The concept and form of the kinked-demand curve was first developed and used by Chamberlin in his theory of monopolistic competition. Later, Hall and Hitch used kinked-demand curve to explain rigidity of prices in oligopolistic market. But, neither Chamberlin nor Hall and Hitch used kinked demand curve as a tool of analysis in their respective theories. Paul M. Sweezy used the kinked-demand curve to establish that price remains stable in oligopolistic market. Sweezy’s model is described below.

The kinked-demand curve model developed by Paul M. Sweezy has features common to most oligopoly pricing models. This is the best-known model to explain the behaviour of the oligopolistic firms. It must, however, be noted at the outset that kinked-demand curve analysis does not deal with price and output determination. It only seeks to establish that once a price-quantity combination is determined, an oligopoly firm does not find it profitable to change its price even when there is a considerable change in the cost of production and change in demand for the product.

The logic behind this proposition is as follows. An oligopoly firm believes that if it reduces the price of its product, the rival firms would follow and neutralize the expected gain from price reduction. But, if it raises the price, the firms would either maintain their prices or even go for price cutting, so that the price-raising firm loses a part of its market to the rival firms. This behaviour is true of all the firms. The oligopoly firms would, therefore, find it more desirable to maintain the prevailing price and output. Let us now see how Sweezy has proved this point of view by using kinked demand curve technique.

Sweezy’s Kinked-Demand Curve Model: In order to analyse the effects of possible reactions of the rival firms on the demand for the product of the firm initiating the change in price, let us make the following assumptions.

1. There are four oligopoly firms—A, B, C and D;
2. Market demand is divided between the firms and all the firms face a uniform demand curve;

3. All the firms are in equilibrium at a point of time, all maximizing their profit.

Sweezy's kinked-demand curve model is presented in Fig. 7.11 on the basis of these assumptions. Let us suppose that the individual demand curve is given by the curve $MN$ and the price is initially fixed at $PQ$. Given the price at $PQ$, let firm A take the lead in changing its price. Let us now examine the effect of various kinds of possible reactions of the rival firms on demand for A's product.

![Fig. 7.11 Kinked-Demand curve analysis](image)

- **Reaction (i).** When firm A increases or decreases its price, the rival firms follow the suit. Then, firm A finds itself moving along the demand curve $MN$. It does not gain or lose.

- **Reaction (ii).** If rival firms do not react to price changes made by the firm A, its demand curve will be $DD'$.

- **Reaction (iii).** When firm A raises its price and rival firms do not follow, then firm A loses a part of its market share to the rival firms. Then the relevant demand curve for firm A is $DP$. But, when firm A decreases its price, rival firms react by cutting down their own prices by an equal amount of even more. This counter move by the rival firms prevents firm A from taking any advantage of price cut. Therefore, firm A is forced to move along the $PN$ part of the demand curve.

Thus, the two relevant segments of demand curve for firm A (below point $P$) is $PN$ and $PD$ beyond point $P$. When the two relevant segments of the two
Forms of Markets

Consider now the relationship between $AR$ and $MR$. We know that slope of the $AR$ curve is twice the slope by the $MR$ curve. The $MR$ curve drawn on the basis of this relationship, will take a shape as shown of $DJKL$ in Fig. 7.11. It is discontinuous between points $J$ and $K$, at output $OQ$. The segment $DJ$ of the $MR$ curve corresponds to $DP$ segment of the demand curve and $KL$ segment of $MR$ curve corresponds to $PN$ segment of the demand curve.

Now let us see why price gets stabilized at $PQ$ even when there is change in cost curves. Suppose $MC$ curve of the firms is given as $MC$, which intersects $MR$ at point $K$. Point $K$ satisfies the necessary condition for profit maximization ($MR = MC$). Therefore, oligopoly firms are in equilibrium at output $OQ$ and they are making maximum profit at price $PQ$. Now, if $MC$ curve shifts upwards to $MC_n$ or to any level between points $J$ and $K$, their profit would not be affected because profit maximization condition remains undistributed. Therefore, they have no motivation for increasing or decreasing their price. It is always beneficial for them to stick to the price $PQ$ and output $OQ$. Thus, both price and output are stable. This is how price gets stabilized under oligopoly. The oligopoly firms would go for changing their price and output only if $MC$ rises beyond point $J$ or falls below point $K$ (in Fig. 7.11).

To conclude, the basic point that Sweezy’s model establishes is the rigidity of price in oligopolistic market. As shown in Fig. 7.11, once price is determined in oligopoly, it remains rigidly fixed because profit maximizing firms’ have no incentive to change price even if $MC$ curve shifts upwards or downwards within the range between points $J$ and $K$. Therefore, price remains constant over a period of time.

3. Collusive Models of Oligopoly

The oligopoly models discussed in the previous sections are based on the assumption that the oligopoly firms act independently; they are in competition with one another; and there is no collusion or any kind of agreement between the firms. The oligopoly models of this category are called non-collusive models. In reality, however, oligopoly firms are found to be in some kind of collusion or agreement—open or secret, explicit or tacit, written or unwritten, and legal or illegal—with one another for at least three major reasons.

First, collusion eliminates or reduces the degree of competition between the firms and gives them some monopolistic powers in their price and output decisions.

Secondly, collusion reduces the degree of uncertainty surrounding the oligopoly firms and ensure profit maximization.

Thirdly, collusion creates some kind of barriers to the entry of new firms.
The models that deal with the collusive oligopolies are called **collusive oligopoly models**. Collusion between firms may take many forms depending on their relative strength and objective of collusion, and on whether collusion is legal or illegal. There are, however, two major forms of collusion between the oligopoly firms: (i) cartel, i.e., firms’ association and (ii) price leadership agreements.

On the basis of these forms of collusion between oligopoly firms, collusive oligopoly models are classified as:

1. Cartel models, and
2. Price leadership models.

In this section, we will discuss these two types of oligopoly collusive models.

### 4. Price and Output Determination under Cartel Models

**Cartel Models of Collusion.** A cartel is a formal collusion of the oligopoly firms in an industry with a purpose. A general purpose of cartels is to centralize certain areas of managerial decisions and functions of individual firms of the industry, with a view to promoting common benefits. Under cartel system, “the firms jointly establish a cartel organization to make price and output decisions, to establish production quotas for each firm and to supervise market activation of the firms in the industry”. Cartels are formed with the purpose of (i) eliminating uncertainty in business, and (ii) restraining harmful interfirm competition.

Cartels may be in the form of open or secret collusion. Whether open or secret, cartel agreements are explicit and formal, in the sense, that agreements are enforceable on member firms not observing the Cartel Rules or dishonouring the agreements. Cartels are, therefore, regarded as the perfect form of collusion. Cartels and cartel-type agreements between the firms in manufacturing and trade are illegal in most countries. Yet, cartels in the broader sense of the term exist in the form of trade associations, professional organizations and the like.

Although a cartel performs a variety of functions for its members, the two important functions of cartels are following:

1. Fixing price for joint profit maximization, and
2. Market sharing between its members.

### 5. Price Leadership Models

**Price leadership is an informal position of a firm in most oligopolistic industries.** Price leadership may emerge spontaneously due to technical reasons or out of a tacit or explicit agreement between the firms to assign a leadership role to one of them. The spontaneous price leadership may be the result of such technical reasons as size, efficiency, economies of scale or firm’s ability to forecast market conditions accurately or a combination of these factors. The most typical case of price leadership is the leading role played by the dominant firm, i.e., the largest
firm in the industry. The dominant firm takes lead in determining the price and in making price changes and the smaller ones follow.

Sometimes, price leadership is barometric by a firm whose decisions and actions serve as a barometer for the decisions and actions taken by other firms. In the barometric price leadership one of the firms, not necessarily the dominant one, takes the lead generally in announcing a change in price, particularly when such a change is due but is not effected due to uncertainty in the market.

The price leadership is found under both product homogeneity and product differentiation. There may be, however, price differences commensurate with product differentiation. Price differentials may also exist on account of cost differentials.

Another important aspect of price leadership is that it often serves as a means to price discipline and price stabilization. Achievement of this objective establishes an ‘effective price leadership’. Such a leadership can, however, exist and work effectively only under the following conditions.

(i) There is a small number of firms in oligopoly;
(ii) Entry to the industry is restricted;
(iii) Products are, by and large, homogeneous;
(iv) Demand for industry is inelastic or has very low elasticity, and
(v) Firms have almost similar cost curves.

Price Leadership Models: Given the conditions for effective working of price leadership, the economists have envisaged the following kinds of price leadership models.

1. Price leadership by low cost firm,
2. Price leadership by dominant firm, and
3. Price leadership by barometric firm.

Check Your Progress

3. Where does the element of monopoly in monopolistic competition arises from?
4. Define oligopoly.
5. What is barometric price leadership?

7.5 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Market-clearing price is also known as the equilibrium price. It implies that the market is cleared in the sense that there is no unsold stock and no unsupplied demand.
2. The factors which determine the categorization of market structure are the following:
   - Number of firms—the sellers
   - Degree and nature of competition
   - Level of product differentiation
   - Possibility of entry and exit of the firms

3. The element of monopoly in monopolistic competition arises from the fact that each firm has an absolute right to produce and sell a branded or patented product.

4. Oligopoly is defined as a market structure in which there are a few sellers selling homogeneous or differentiated products.

5. In the barometric price leadership one of the firms, not necessarily the dominant one, takes the lead generally in announcing a change in price, particularly when such a change is die but is not affected due to uncertainty in the market.

7.6 SUMMARY

- In economics, the word ‘market’ is used in a rather abstract sense. The market means a system by which sellers and buyers of a commodity interact to settle its price and the quantity to be bought and sold.
- Market system refers to the process by which buyers and sellers of a product interact to settle the price of the product and carry out the sale-purchase transactions. The market system works on the basis of a basic market principle.
- In the context of market analysis, equilibrium refers to a state of market in which quantity demanded of a commodity equals the quantity supplied of the commodity. The equality of demand and supply produces an equilibrium price.
- Equilibrium price of a commodity in a free market is determined by the market forces of demand and supply. In order to analyze how equilibrium price is determined, we need to integrate the demand and supply curves.
- Market mechanism is a process of interaction between the market forces of demand and supply to determine the equilibrium price.
- Market structure refers to the number of firms in an industry and the degree of competition among the firms.
- The market structure determines a firm’s power to fix the price of its product a great deal. The degree of competition determines a firm’s degree of freedom in determining the price of its product.
• The term pure monopoly means an absolute power of a firm to produce and sell a product that has no close substitute. In other words, a monopolized market is one in which there is only one seller of a product having no close substitute.

• The decision rules regarding optimal output and pricing in the long-run are the same as in the short-run. In the long-run, however, a monopolist gets an opportunity to expand the scale of its production with a view to enhancing its long-run profits.

• By definition, monopolistic competition refers to a market structure in which a large number of sellers sell differentiated products, which are close substitutes for one another.

• Oligopoly is defined as a market structure in which there are a few sellers selling homogeneous or differentiated products. In case oligopoly firms sell a homogeneous product, it is called pure or homogeneous oligopoly.

7.7 KEY WORDS

• Market mechanism: It is a process of interaction between the market forces of demand and supply to determine the equilibrium price.

• Market structure: It refers to the number of firms in an industry and the degree of competition among the firms.

• Monopoly: It means an absolute power of a firm to produce and sell a product that has no close substitute.

7.8 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short Answer Questions
1. Mention the various aspects of the market concept.
2. How is the market price determined?
3. State the practical importance of market equilibrium.
4. Write a short note on the types of market structures.

Long-Answer Questions
1. Discuss shift in demand and supply curves.
2. Explain pricing and output decisions under monopoly with the help of diagrams.
3. Evaluate the pricing and output decisions under monopolistic competition.
4. Describe the various oligopoly models.
7.9 FURTHER READINGS


Websites

https://blogs.economictimes.indiatimes.com/et-commentary/heres-how-indias-widening-income-distribution-can-be-redressed/

Several mathematicians and economists have made numerous attempts to find a reasonable explanation to strategic behaviour of firms. The first important contribution to this field was made by a mathematician John von Neumann and an economist Oskar Morgenstern in 1944. Their contribution was in the form of game theory. The game theory brings out the strategy used by the oligopoly firms to determine the best possible action to maximize their predetermined objective. Although many other economists have contributed to game theory, Martin Shubik is regarded as the ‘most prominent proponent of the game theory approach’. A more recent and in-depth work on the game theory and its application to economics and management problems can be found in the work of Prajit K. Dutta.

In this unit, you will study about the strategic behaviour of the firms, the application of the game theory to oligopolistic strategy, prisoner’s dilemma, Nash equilibrium and its implications.

8.1 UNIT OBJECTIVES

After going through this unit, you will be able to:
- Analyse the strategic behaviour of the firms
- Discuss the game theory
- Explain the prisoner’s dilemma
8.2 STRATEGIC BEHAVIOUR OF THE FIRMS: PRICE AND NON-PRICE COMPETITION

We have already discussed the non-price competition strategies in relation to monopolistic competition and oligopoly in Unit 7. Let us revisit some of the basic concepts before moving on to learning about game theory and its application. Price competition is the competition that lays focus on the issue of price as more and attempts to stay ahead of the the price of the competitors existing in the market.

Non-price competition is the competition that lays emphasis on the factors other than the price of the product. The main objective is to differentiate from the products that are available as alternatives in the market. Non-price competition can be compared and contrasted with price competition, that is, where a business enterprise tries to make a distinction of its product or service from competing products on the basis of a low price.

Business enterprises engage in non-price competition, despite the additional costs involved because it is generally more profitable than selling for a lower price and evades the hazard of a price war. Although a non-price competition strategy can be utilized by any business enterprise, yet it is most commonly used under oligopolies and monopolistic competition because these firms can be extremely competitive.

The business enterprise can also distinguish its product offering on the basis of quality of service, extensive distribution, customer focus, or any other sustainable competitive advantage other than price. The objective is to try to persuade consumers that they should purchase these products, not only because they are cheaper but also as they are better than those made by competitors in one way or the other. Non-price competition characteristically involves promotional expenditures (such as advertising, selling staff, the locations convenience, sales promotions, coupons, special orders, or free gifts), marketing research, new product development, and brand management costs. Business enterprises will certainly engage in non-price competition because it is generally more profitable than selling for a lower price and evades the hazard of a price war. For instance, branded goods often sell more items than do their local counterparts, due to promotion and publicity despite being more expensive. Non-price competition may also promote innovation as business enterprise try to distinguish their product.

A non-price competition strategy can be used by any business enterprise. However, it is most common among oligopolies and monopolistic competition because these firms can be extremely competitive.
Before we proceed to discuss the game theory, it is helpful to understand the meaning and purpose of the game theory and some basic terms and tools used in the analysis and application of game theory.

1. **The Game Theory**: In all kinds of games there are two teams. In all games, the objective of the players of each team is to win the game. To win the game, players make their play-strategy and take action in anticipation of possible reactions of the opposite team and plan their own counter action. This concept of strategic play has been applied by von Neumann and Morgenstern to strategic play of oligopoly firms. As a player, each firm formulates its strategic play and estimates its effects on its objective, called **pay-off**. The pay-off may be **positive**, **negative** or **zero-sum** for a firm taking strategic action. Accordingly, if a strategic action taken by a firm may yield some gains to the firm and counteraction by the rival firm neutralises the gain, it is a **zero-sum game**. If both the firms—action-taking firm and rival firms—gain from the strategic action taken by a firm, it is a **positive-sum game**. And, if both the firms, action-taking firm and rival firms, make losses from the strategic action taken by a firm, it is a **negative-sum game**.

2. **Interdependence**: The game theory has been formulated on the basis of a realistic assumption of **interdependence** of oligopoly firms. It implies that decision-making of the firms under oligopoly is **interdependent**. That is, while taking a business decision – be it price determination, advertising, introduction of a new product or brand, setting-up a new production unit, or any other issue – oligopoly firms take into account the possible action and reaction of the rival firms. This kind of behaviour of the oligopoly firms shows their **interdependence**.

3. **Strategy**: The term ‘strategy’ means the course of action to be taken by the oligopoly firms with the purpose of gaining most from an action under the condition of unknown reaction of the rival firms. For example, suppose there are two firms – A and B. Firm A plans to cut down the price of its product. But Firm A is not sure of the possible reaction of the rival Firm B. There are two possible reactions of the rival Firm B: (i) it may cut down its own price, and (ii) it may not cut down the price. The price cutting firm A will assess its gain and losses under these conditions and chose the best option. This is the strategy of the price cutting firm.

4. **Pay-off Matrix**: The ‘pay-off matrix’ is tabular recording of gains and losses of a firm taking an action under different kinds of anticipated reactions of the rival firms. Recall the above example of two firms, A and B. Firm A estimates its gains and losses in terms of increase in the sales of its product under the following conditions:

   (i) Estimated increase in its sales when Firm A does not reduce its price;
Notes

5. Dominant Strategy: As noted above, strategy means the course of action planned by an oligopoly firm with the purpose of gaining most from its action. The pay-off of the strategy may be high or it may be low depending on the counteraction taken by the rival firm. A strategic action that yields the best outcome whatever the reaction of the rival firms is called dominant strategy. In the context of game theory, dominant strategy can be defined as the strategy that gives the best payoff no matter what counteraction is taken by the rival firm.

Having looked at the ‘basics’ of the game theory, we proceed now to discuss the game theory and its application to business decision-making. In game theory, the decision-making problem of the oligopoly firms is best exemplified by, what game theorists call, the prisoners' dilemma. We begin our discussion with the concept of prisoners' dilemma – an example of dilemma faced by the oligopoly firms in decision-making.

8.3.1 Prisoner's Dilemma

The nature of the decision-making problems faced by the oligopoly firms is exemplified in game theory by prisoner's dilemma. To illustrate prisoners' dilemma, let us suppose that two persons, A and B, are partners in illegal activities. They are arrested under the suspicion of being involved in cricket match-fixing. They are lodged in separate jails with no possibility of communication between themselves. They are interrogated by CBI officials under the following conditions disclosed to each of them in isolation.

1. If you confess your involvement in match fixing, you will get a 5-year imprisonment.
2. If you deny your involvement and your partner denies too, you will be set free for lack of evidence.
3. If you confess and your partner does not confess and you turn approver, then you get 2-year imprisonment and the other person will get 10-year imprisonment.

Given these conditions, each suspect has two options open to him: (i) to confess or (ii) not to confess. Now, both A and B face a dilemma on how to decide whether or not to confess. While taking a decision, both have a common
objective, i.e., to minimize the period of imprisonment. Given this objective, the option is quite simple that both of them deny their involvement in match-fixing. But, there is no certainty that if one denies his involvement, the other will also deny—the other one may confess and turn approver. With this uncertainty, the dilemma in making a choice still remains. For example, if \( A \) denies his involvement, and \( B \) confesses and turns approver (settles for a 2-year imprisonment), then \( A \) gets a 10-year jail term. So is the case with \( B \). If they both confess, then they get a 5-year jail term each. Then what to do? That is the dilemma. The nature of their problem of decision-making is illustrated in Table 8.1 in the form of a ‘pay-off matrix’. The pay-off matrix shows the pay-offs of their different options in terms of the number of years in jail.

### Table 8.1 Prisoners’ Dilemma: The Pay-off Matrix

<table>
<thead>
<tr>
<th>A’s Options</th>
<th>B’s Options</th>
<th>Confess</th>
<th>Deny</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confess</td>
<td>A</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Deny</td>
<td>A</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

Given the conditions, it is quite likely that both the suspects may opt for ‘confession’, because neither \( A \) knows what \( B \) will do, nor \( B \) knows what \( A \) will do. When they both confess, each one gets a 5-year jail term. This is the second best option. For his decision to confess, \( A \) might formulate his strategy in the following manner. He is supposed to reason out the case in this way: If I confess (though I am innocent), I will get a maximum of 5 years’ imprisonment. But, if I deny (which I must) and \( B \) confesses and turns approver then I will get 10 years’ imprisonment. That will be the worst scenario. It is quite likely that suspect \( B \) also reasons out their case in the same manner, even if he too is innocent. If they both confess, they would get jail-term for 5 years and would avoid 10 years’ imprisonment, the maximum possible jail sentence under the law. This is the best they could achieve under the given conditions.

**Relevance of Prisoners’ Dilemma to Oligopoly**

The prisoners’ dilemma illustrates the nature of problems oligopoly firms are confronted with in the formulation of their business strategy with respect to such problems as strategic advertising, price cutting or cheating the cartel if there is one. Look at the nature of problems an oligopoly firm is confronted with when it plans to increase its \( \text{advertisement} \) expenditure (ad-expenditure for short). The basic issue is whether or not to increase the ad-expenditure. If the answer is ‘do not increase’, then the following questions arise. Will the rival firms increase ad-expenditure or will they not? If they do, what will be the consequences for the firm under consideration? And, if the answer is ‘increase’, then the following questions arise. What will be the reaction of the rival firms? Will they increase or will they not
increase their ad-expenditure? What will be the pay-off if they do not and what if they do? If the rival firms do increase their advertising, what will be the pay-off to the firm? Will the firm be a net gainer or a net loser? The firm planning to increase ad-spending will have to find the answer to these queries under the conditions of uncertainty. To find a reasonable answer, the firm will have to anticipate actions, reactions and counter-actions by the rival firms and chalk out its own strategy. It is in case of such problems that the case of prisoners’ dilemma becomes an illustrative example.

8.3.2 Application of Game Theory to Oligopolistic Strategy

Let us now apply the game theory to our example of ‘whether or not to increase ad-expenditure’, assuming that there are only two firms, A and B, i.e., the case of a duopoly. We know that in all games, the players have to anticipate the moves of the opposite player(s) and formulate their own strategy to counter them. To apply the game theory to the case of ‘whether or not to increase ad-expenditure’, the firm needs to know or anticipate the following two kinds of reactions of the rival firm and their pay-offs.

(i) The counter moves by the rival firm in response to increase in ad-expenditure by this firm, and
(ii) The pay-offs of this strategy under two conditions:
(a) when the rival firm does not react, and
(b) the rival firm does make a counter move by increasing its ad-expenditure.

In order to find solution to its problem, the firm anticipates the possible reactions of the rival firms and estimates their possible outcomes. The firm will then take decision on the best possible strategy for playing the game and achieving its objective of, say, increasing sales and capturing a larger share of the market. The best possible strategy in game theory is called the ‘dominant strategy’. A dominant strategy is one that gives optimum pay-off, no matter what the opponent does. Thus, the basic objective of applying the game theory is to arrive at the dominant strategy.

Suppose that the possible outcomes of the ad-game under the alternative moves are given in the pay-off matrix presented in Table 8.2.

**Table 8.2 Pay-off Matrix of the Ad-Game**

<table>
<thead>
<tr>
<th></th>
<th>Increase Ad</th>
<th>Don’t increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase Ad</td>
<td>A 20 B 10</td>
<td>A 30 B 0</td>
</tr>
<tr>
<td>Don’t increase</td>
<td>A 10 B 15</td>
<td>A 15 B 5</td>
</tr>
</tbody>
</table>
As the matrix shows, if Firm A decides to increase its ad-expenditure, and Firm B counters A’s move by increasing its own ad-expenditure, A’s sales go up by $20 million and those of Firm B by $10 million. And, if Firm A increases its advertisement and B does not, then A’s sales increase by $30 million and there are no sales gain for Firm B. One can similarly find the pay-offs of the strategy ‘Don’t increase’ in case of both firms. As shown in Table 8.2, if Firm A does not increase its Ad-spending and Firm B does increase its Ad-spending, then A’s sales increases by $10 million and of B by $15 million.

Given the pay-off matrix, the question arises: What strategy should Firm A choose to optimize its gain from extra ad-expenditure, irrespective of counter-action by the rival Firm B. It is clear from the pay-off matrix that Firm A will choose the strategy of increasing the ad-expenditure because, no matter what Firm B does, its sales increase by at least $20 million. This is, therefore, the dominant strategy for Firm A. A better situation could be that when Firm A increases its expenditure on advertisement, Firm B does not. In that case, sales of Firm A could increase by $30 million and sales of Firm B do not increase. But there is a greater possibility that Firm B will go for counter-advertising in anticipation of losing a part of its market to Firm A in future. Therefore, a strategy based on the assumption that Firm B will not increase its ad-expenditure involves a great deal of uncertainty. Under these conditions, the first option gives the dominant strategy for Firm A.

8.3.3 Nash Equilibrium and Implications

In the preceding section, we have used a very simple example to illustrate the application of game theory to an oligopolistic market setting, with the following simplifying assumptions.

(i) The strategy formulation is a one-time affair,

(ii) Only one firm initiates the competitive warfare and other firms only react to action taken by one firm, and

(iii) The there exists a dominant strategy—a strategy which gives an optimum solution.

The real-life situation is, however, much more complex. There is a continuous one-to-one and tit-for-tat kind of warfare. Actions, reactions and counter-actions are regular phenomena. Under these conditions, a dominant strategy is often non-existent. To analyze this kind of situation, John Nash, an American mathematician, developed a technique, which is known by his name as Nash equilibrium. Nash equilibrium technique seeks to establish that each firm does the best it can, given the strategy of its competitors and a Nash equilibrium is one in which none of the players can improve their pay-off given the strategy of the other players. In case of our example, Nash equilibrium can be defined as one in which none of the firms can increase its pay-off (sales) given the strategy of the rival firm.
The Nash equilibrium can be illustrated by making some modifications in the pay-off matrix given in Table 8.2. Now we assume that action and counter-action between Firms A and B is a regular phenomenon and the pay-off matrix that appears finally is given in Table 8.3. The only change in the modified pay-off matrix is that if neither Firm A nor Firm B increases its ad-expenditure, then pay-offs change from (15, 5) to (25, 5).

### Table 8.3 Nash Equilibrium: Pay-off Matrix of the Ad-Game

**(Increase in sales in million *)

<table>
<thead>
<tr>
<th></th>
<th>Increase AD</th>
<th>Don’t increase</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A’s Options</strong></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Increase AD</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>Don’t increase</td>
<td>10</td>
<td>25</td>
</tr>
</tbody>
</table>

It can be seen from the pay-off matrix (Table 8.3) that Firm A no longer has a **dominant strategy**. Its optimum decision depends now on what Firm B does. If Firm B increases its ad-expenditure, Firm A has no option but to increase its advertisement expenditure. And, if Firm A reinforces its advertisement expenditure, Firm B will have to follow suit. On the other hand, if Firm B does not increase its ad-expenditure, Firm A does the best by increasing its ad-expenditure. Under these conditions, the conclusion that both the firms arrive at is to increase ad-expenditure if the other firm does so, and ‘don’t increase’, if the competitor ‘does not increase’. In the ultimate analysis, however, both the firms will decide to increase the ad-expenditure. The reason is that if none of the firms increases its ad-outlay, Firm A gains more in terms of increase in its sales (*25 million*) and the gain of Firm B is much less (*5 million*) only. And, if Firm B increases advertisement expenditure, its sales increase by *10 million*. Therefore, Firm B would do best to increase its ad-expenditure. In that case, Firm A will have no option but to do likewise. Thus, the **final conclusion** that emerges is that both the firms will go for advertisement war. In that case, each firm finds that it is doing the best given what the rival firm is doing. This is the Nash equilibrium.

However, there are situations in which there can be more than one Nash equilibrium. For example, if we change the pay-off in the south-east corner from (25, 5) to (22, 8); each firm may find it worthless to wage advertisement war and may settle for ‘don’t increase’ situation. Thus, there are two possible Nash equilibria.

### Check Your Progress

1. Why do businesses engage in non-price competition?
2. What is a pay-off matrix?
3. What is a zero-sum game?
4. State the basic objective of applying game theory to oligopolistic strategy.
8.4 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Business enterprises engage in non-price competition, despite the additional costs involved because it is generally more profitable than selling for a lower price and evades the hazard of a price war.

2. The ‘pay-off matrix’ is tabular recording of gains and losses of a firm taking an action under different kinds of anticipated reactions of the rival firms.

3. If a strategic action taken by a firm may yield some gains to the firm and counteraction by the rival firm neutralizes the gain, it is a zero-sum game.

4. The basic objective of applying game theory is to arrive at the dominant strategy.

8.5 SUMMARY

- Price competition is the competition that lays focus on the issue of price as more and attempts to stay ahead of the price of the competitors existing in the market.

- Non-price competition is the competition that lays emphasis on the factors other than the price of the product. The main objective is to differentiate from the products that are available as alternatives in the market.

- Business enterprises engage in non-price competition, despite the additional costs involved because it is generally more profitable than selling for a lower price and evades the hazard of a price war.

- The pay-off may be positive, negative or zero-sum for a firm taking strategic action. Accordingly, if a strategic action taken by a firm may yield some gains to the firm and counteraction by the rival firm neutralizes the gain, it is a zero-sum game.

- The game theory has been formulated on the basis of a realistic assumption of interdependence of oligopoly firms. It implies that decision-making of the firms under oligopoly is interdependent.

- The ‘pay-off matrix’ is tabular recording of gains and losses of a firm taking an action under different kinds of anticipated reactions of the rival firms.

- The prisoners’ dilemma illustrates the nature of problems oligopoly firms are confronted with in the formulation of their business strategy with respect to such problems as strategic advertising, price cutting or cheating the cartel if there is one.
• The Nash equilibrium technique seeks to establish that each firm does the best it can, given the strategy of its competitors and a Nash equilibrium is one in which none of the players can improve their pay-off given the strategy of the other players.

8.6 KEY WORDS

• Strategy: It is a general plan or set of plans intended to achieve something, especially over a long period.
• Duopoly: It is the most basic form of oligopoly, a market dominated by a small number of companies.

8.7 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short Answer Questions
1. What is a dominant strategy in the context of game theory?
2. Give an example to illustrate the prisoner’s dilemma.
3. Write a short note on the relevance of prisoner’s dilemma to oligopoly.

Long Answer Questions
1. Analyse the strategic behaviour of the firms.
2. Discuss the application of game theory to oligopolistic strategy.
3. Give examples to illustrate the use of the Nash equilibrium technique.

8.8 FURTHER READINGS

Strategic Behaviour of the Firms and Game Theory

Websites
https://blogs.economictimes.indiatimes.com/et-commentary/heres-how-indias-widening-income-distribution-can-be-redressed/
UNIT 9 COST AND RETURN

Structure
9.0 Introduction
9.1 Objectives
9.2 Relevant Costs for Decision-Making
9.3 Traditional Theory of Cost: Cost Function and Cost Output Relationship
   9.3.1 Short-Run Cost-Output Relations
   9.3.2 Long-Run Cost-Output Relations
   9.3.3 Economies and Diseconomies of Scale
9.4 Modern Theory of Cost
9.5 Cost Control and Cost reduction
9.6 Cost Behaviour and Business Decision
9.7 Answers to Check Your Progress Questions
9.8 Summary
9.9 Key Words
9.10 Self Assessment Questions and Exercises
9.11 Further Readings

9.0 INTRODUCTION

The theory of cost provides conditions for minimizing the cost of production. In addition, analysis of cost of production is very important in almost all kinds of business decisions, especially those related to the weak points of production management; determining the output level for cost minimization; determining the price of the product and dealers’ margin; and estimating and projecting the cost of business operation. Cost functions help in determining the relationship between outputs and costs. In this unit, we will study the relevant costs for decision-making and the traditional and modern theories of cost.

9.1 OBJECTIVES

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

- Discuss the relevant costs for decision-making
- Explain the traditional theory and modern theories of cost
- Describe the cost function and cost output relationship
- Discuss the concept of cost control, cost reduction, cost behaviour and business decisions
9.2 RELAVANT COSTS FOR DECISION-MAKING

The cost concepts that are relevant to business operations and decisions can be grouped on the basis of their nature and purpose under two overlapping categories: (i) cost concepts used for accounting purposes, and (ii) analytical cost concepts used in economic analysis of business activities. We will discuss some important concepts of the two categories. It is important to note here that this classification of cost concepts is only a matter of analytical convenience.

**Accounting Cost Concepts**

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

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

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

3. **Actual or Explicit Costs and Implicit or Imputed Costs**: The Actual or Explicit costs are those which are actually incurred by the firm in payment for labour, material, plant, building, machinery, equipment, travelling and transport, advertisement, etc. The total money expenses, recorded in the books of accounts are, for all practical purposes, the actual costs. Actual cost comes under the accounting cost concept.
In contrast to explicit costs, there are certain other costs that do not take the form of cash outlays, nor do they appear in the accounting system. Such costs are known as implicit or imputed costs. Opportunity cost is an important example of implicit cost. For example, suppose an entrepreneur does not utilize his services in his own business and works as a manager in some other firm on a salary basis. If he sets up his own business, he foregoes his salary as manager. This loss of salary is the opportunity cost of income from his own business. This is an implicit cost of his own business. Thus, implicit wages, rent, and implicit interest are the wages, rent and interest that an owner’s labour, building and capital respectively, can earn from their second best use.

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

4. Out-of-Pocket and Book Costs: The items of expenditure that involve cash payments or cash transfers, both recurring and non-recurring, are known as out-of-pocket costs. All the explicit costs (e.g., wages, rent, interest, cost of materials and maintenance, transport expenditure, electricity and telephone expenses, etc.) fall in this category. On the contrary, there are certain actual business costs that do not involve cash payments, but a provision is therefore made in the books of account and they are taken into account while finalizing the profit and loss accounts. Such expenses are known as book costs. In a way, these are payments made by a firm to itself. Depreciation allowances and unpaid interest on the owner’s own funds are the example of book costs.

Analytical Cost Concepts

The analytical cost concepts refers to the different cost concepts that are used in analysing the cost-output relationship with increase in inputs and output and also the cost concepts that figure in analysing the effect of expansion of production on the society as a whole.

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

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

2. Total, Average and Marginal Costs. Total cost (TC) refers to the total outlays of money expenditure, both explicit and implicit, on the resources used to
produce a given level of output. It includes both fixed and variable costs. The total cost for a given output is measured as

\[ TC = \text{Total fixed cost} + \text{Total variable cost} \]

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

\[ AC = \frac{TC}{Q} \]

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

\[ MC = \frac{dTC}{dQ} \]

Total, average and marginal cost concepts are used in the economic analysis of firm’s production and in pricing decisions. These cost concepts are discussed in further detail in the following section.

3. **Short-Run and Long-Run Costs.** Short-run and long-run cost concepts are related to variable and fixed costs, respectively, and often figure in economic analysis cost-output relationship.

- **Short-run** refers to the time period during which scale of production remains unchanged. The costs incurred in the short-run are called short-run costs. It includes both the variable and the fixed costs. From analytical point of view, short-run costs are those that vary with the variation in output in short-run, the size of the firm remaining the same. Therefore, short-run costs are treated as **variable costs**.

- **Long-run** costs, on the other hand, are those that are incurred to increase the scale of production in the long-run. The costs that are incurred on the fixed factors like plant, building, machinery, etc., are known as long-run costs. It is important to note that the running cost and depreciation of the capital assets are included in the short-run or variable costs.

Furthermore, long-run costs are by implication the costs that are incurred in the long-run. In the long run, however, even the fixed costs become variable costs as the size of the firm or scale of production increases. Broadly speaking, ‘the short-run costs are those associated with variables in the utilization of fixed plant or other facilities whereas long-run costs are associated with the changes in the size and kind of plant.’

4. **Incremental Costs and Sunk Costs.** Conceptually, **incremental costs** are closely related to the concept of marginal cost but with a relatively wider connotation. While marginal cost refers to the cost of the marginal unit (generally one unit) of output, incremental cost refers to the total additional cost associated with the decisions to expand the output or to add a new variety of product, etc.
5. Historical and Replacement Costs. Historical cost refers to the cost incurred in past on the acquisition of productive assets, e.g. land, building, machinery, etc., whereas replacement cost refers to the expenditure made for replacing an old asset. These concepts owe their significance to the unstable nature of input prices. Stable prices over time, other things given, keep historical and replacement costs on par with each other. Instability in asset prices makes the two costs differ from each other.

As regards their application, historical cost of assets is used for accounting purposes, in the assessment of the net worth of the firm whereas replacement cost figures in business decisions regarding the renovation of the plant.

Private costs are those which are actually incurred or provided for by an individual or a firm on the purchase of goods and services from the market. For a firm, all the actual costs, both explicit and implicit, are private costs. Private costs are internalized costs that are incorporated in the firm’s total cost of production.

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

Check Your Progress
1. Define explicit costs.
2. Under which category does the running cost and depreciation of the capital assets come?

9.3 TRADITIONAL THEORY OF COST: COST FUNCTION AND COST OUTPUT RELATIONSHIP

The theory of cost deals with how cost of production changes with change in output. In other words, the cost theory deals with cost-output relations. The basic principle of the cost-output relationship is that the total cost increases with increase in output. This simple statement of an observed fact is of little theoretical and practical importance. What is important from a theoretical and managerial decision...
Cost and Return

NOTES

Self-Instructional Material

point of view is the rate of increase in total cost with increase in output and the
direction of change in the average cost (AC) and the marginal cost (MC). The
direction of change in AC and MC—whether AC and MC decrease or increase
or remain constant—depends on the nature of the cost function. A cost function is
a symbolic statement of the technological relationship between the cost and output.
The general form of the cost function is written as

\[ TC = f(Q), \quad \Delta TC/\Delta Q > 0 \quad \ldots (9.1) \]

The actual form of cost function depends on whether the time framework
chosen for cost analysis is short-run or long-run. It is important to recall here that
some costs remain constant in the short-run while all costs are variable in the long-
run. Thus, depending on whether cost analysis pertains to short-run or to long-
run, there are two kinds of cost functions: (i) short-run cost functions, and
(ii) long-run cost functions. Accordingly, the cost output relations are analyzed in
short-run and long-run framework.

9.3.1 Short-Run Cost-Output Relations

In this section, we will analyse the cost-output relations in the short-run. The long-
run cost output relations discussed in the following section. Before we discuss the
short-run cost-output relations, let us first look at the cost concepts and the
components used to analyze the short-run cost-output relations.

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

\[ TC = TFC + TVC \quad \ldots (9.2) \]

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

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

\[ AC = \frac{TC}{Q} = \frac{TFC + TVC}{Q} = \frac{TFC}{Q} + \frac{TVC}{Q} = AFC + AVC \]

Thus,

\[ AFC = \frac{TFC}{Q} \quad \text{and} \quad AVC = \frac{TVC}{Q} \]

and

\[ AC = AFC + AVC \quad \ldots (9.3) \]

Marginal cost (MC) can be defined as the change in the total cost due to
change in the total output by one unit, i.e.,
In case $TC$ is expressed in functional form, $MC$ is measured as the first derivative of cost function, i.e., $\frac{\partial TC}{\partial Q}$.

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

**Short-Run Cost Functions and Cost Curves**

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

1. **Linear Cost Function.** When total cost increases at a constant rate with increase in production, it produces a linear cost function. A linear cost function takes the following form.

   \[ TC = a + bQ \quad \ldots(9.5) \]

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

   Given the cost function [Eq. (9.5)] $AC$ and $MC$ can be obtained as follows.

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

   and

   \[ MC = \frac{\partial TC}{\partial Q} = b \]

   Note that since '$b$' is a constant coefficient, $MC$ remains constant, throughout in case of a linear cost function.

   Note that, in case of a linear cost function, while $MC$ remains constant, $AC$ continues to decline with the increase in output. This is so simply because of the logic of the linear cost function.

2. **Quadratic Cost Function.** When $TC$ increases at increasing rate with constant increase in output ($Q$), the $TC$ data produces a quadratic cost function expressed as

   \[ TC = a + bQ + Q^2 \quad \ldots(9.6) \]
where $a$ and $b$ are constants and $TC$ and $Q$ are total cost and total output, respectively.

Given the cost function (9.6), $AC$ and $MC$ can be obtained as follows.

$$AC = \frac{TC}{Q} \cdot \frac{a + bQ + Q^2}{Q} + b + Q \quad \ldots(9.7)$$

$$MC = \frac{\partial TC}{\partial Q} = b + 2Q \quad \ldots(9.8)$$

The cost curves that emerge from the cost function in this case are that while fixed cost remains constant, $TVC$ is increasing at an increasing rate. The rising $TVC$ sets the trend in the total cost ($TC$). Note that $MC$ and $AVC$ are rising at a constant rate whereas $AC$ declines till a certain point and then begins to increase.

3. Cubic Cost Function When $TC$ increases first at decreasing rate and then of increasing rate with increase in production, the $TC$ data produces a cubic cost function. A cubic cost function is of the form

$$TC = a + bQ - cQ^2 + Q^3 \quad \ldots(9.9)$$

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

$TFC$ remains fixed for the whole range of output, and hence, takes the form of a horizontal line—$TFC$. The $TVC$ curve shows two different trends with increase in output. The total variable cost first increases at a decreasing rate and then at an increasing rate with the increase in the output. The rate of increase can be obtained from the slope of $TVC$ curve. The two patterns of change in the $TVC$ curve stems directly from the law of increasing and diminishing returns to the variable inputs. So long as the law of increasing returns is in operation, $TVC$ increases at decreasing rate. And, when the law of diminishing returns comes into operation output increases at decreasing rate causing $TVC$ to increase at increasing rate.

**Average Fixed Cost (AFC):** As already mentioned, the costs that remain fixed for a certain level of output make the total fixed cost in the short-run. We know that

$$AFC = \frac{TFC}{Q} \quad \ldots(9.10)$$

**Average Variable Cost (AVC).** As defined above,

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

**Average Cost (AC)** The average cost ($AC$) is defined as $AC = \frac{TC}{Q}$. 
Law of Diminishing Returns and the Cost Curves

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

In the initial stage of production, both AFC and AVC are declining because of internal economies. Since $\text{AC} = \text{AFC} + \text{AVC}, \text{AC}$ is also declining. This shows the operation of the law of increasing returns to the variable input in the initial stage of production. But beyond a certain level of output while AFC continues to fall, AVC starts increasing because of a faster increase in the TVC. Consequently, the rate of fall in AC decreases. Beyond this level of output, AC starts increasing which shows that the law of diminishing returns comes into operation.

A downward trend in the MC shows increasing marginal productivity of the variable input due mainly to internal economies resulting from increase in production. Similarly, an upward trend in the MC shows increase in TVC, on the one hand, and decreasing marginal productivity of the variable input, on the other.

Some Important Relationships between Different Measures of Cost

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

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

(b) When AFC falls but AVC increases, change in AC depends on the rate of change in AFC and AVC.
   
   (i) If decrease in AFC > increase in AVC, then AC falls,
   (ii) If decrease in AFC = increase in AVC, AC remains constant and
   (iii) If decrease in AFC < increase in AVC, then AC increases.

(c) AC and MC are related in following ways.
   
   (i) When MC falls, AC follows, over a certain range of output. When MC is falling, the rate of fall in MC is greater than that of AC, because while MC is attributed to a single marginal unit, AC is distributed over the entire output. Therefore, AC decreases at a lower rate than MC.
   (ii) Similarly, when MC increases, AC also increases but at a lower rate for the reason given in (i). There is, however, a range of output...
over which the relationship does not exist. Compare the behaviour of
MC and AC over the range of output from 6 units to 10 units (see
Fig.9.5). Over this range of output, MC begins to increase while AC
continues to decrease. The reason for this can be seen in Table 9.1:
when MC starts increasing, it increases at a relatively lower rate that is
sufficient only to reduce the rate of decrease in AC—not sufficient to
push the AC up.

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

Output Optimization in the Short-Run

In this section, we show the application of the same technique to cost-minimising
output.

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

\[ TC = 200 + 5Q + 2Q^2 \]  

As noted earlier, the level of output is optimized at the level of production at
which MC = AC. In other words, at optimum level of output, AC = MC. Given
the cost function in Eq. (9.11),

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

and

\[ MC = \frac{\partial TC}{\partial Q} = 5 + 4Q \]  

By equating AC and MC equations, i.e., Eqs. (9.12) and (9.13), respectively,
and solving them for Q, we get the optimum level of output. Since at equilibrium,
AC = MC

\[ \frac{200}{Q} + 5 + 2Q = 5 + 4Q = 2Q \]

\[ 2Q^2 = 200 \]

\[ Q = 10 \]

Thus, given the cost function (9.20), the optimum output is 10.
9.3.2 Long-Run Cost-Output Relations

In the preceding section, we have discussed the short-run theory of cost. It tells us how cost changes when production is increased by increasing the variable input (labour), fixed input (capital) remaining constant. In this section, we discuss the long-run theory of cost. In the context of the production theory, long run refers to a period in which firms can use more of both the inputs – labour and capital – to increase their production. The long-run theory of cost deals with the long-run cost-output relationship. In other words, long-run theory of cost states the nature of relationship between output and cost with increase in scale of production.

To understand the long-run-cost-output relations and to derive long-run cost curves, it will be helpful to imagine that a long-run is composed of a series of short-run production decisions. As a corollary of this, long-run cost curve is composed of a series of short-run cost curves. With this perception of long-run-cost-out relationship, we may now show the derivation of the long-run cost curves and study their relationship with output.

Long-run Total Cost Curve (LTC)

In order to draw the long-run total cost curve, let us begin with a short-run situation. Suppose that a firm having only one plant has its short-run total cost curve as given by $STC_1$, in panel (a) of Fig. 9.1. Let us now suppose that the firm decides to add two more plants over time, one after the other. As a result, two more short-run total cost curves are added to $STC_1$, in the manner shown by $STC_2$ and $STC_3$ in Fig. 9.1 (a). The $LTC$ can now be drawn through the minimum points of $STC_1$, $STC_2$, and $STC_3$, as shown by the $LTC$ curve corresponding to each $STC$.

Long-run Average Cost Curve (LAC)

Like $LTC$, long-run average cost curve ($LAC$) is derived by combining the short-run average cost curves ($SAC$). Note that there is one $SAC$ associated with each $STC$. The curve $SAC_1$ in panel (b) of Fig. 9.1 corresponds to $STC_1$ in panel (a). Similarly, $SAC_2$ and $SAC_3$, in panel (b) correspond to $STC_2$ and $STC_3$ in panel (a), respectively. Thus, the firm has a series of $SAC$ curves, each having a bottom point showing the minimum $SAC$. For instance, $C_1, Q_1$ is minimum $AC$ when the firm has only one plant. The $AC$ decreases to $C_2, Q_2$ when the second plant is added and then rises to $C_3, Q_3$ after the addition of the third plant. The $LAC$ curve can be drawn through the $SAC_1$, $SAC_2$, and $SAC_3$, as shown in Fig. 9.1 (b). The $LAC$ curve is also known as the ‘Envelope Curve’ or ‘Planning Curve’ as it serves as a guide to the entrepreneur in his plans to expand production.
The SAC curves can be derived from the data given in the STC schedule, from STC function or straightaway from the LTC curve. Similarly, LAC curve can be derived from LTC schedule, LTC function or from LTC curve.

The relationship between LTC and output, and between LAC and output can now be easily derived. It is obvious from the LTC that the long-run cost-output relationship is similar to the short-run cost-output relation. With the subsequent increases in the output, LTC first increases at a decreasing rate, and then at an increasing rate. As a result, LAC initially decreases until the optimum utilization of the second plant and then it begins to increase. These cost-output relations follow the "law of returns to scale". When the scale of the firm expands, LAC, i.e., unit cost of production, initially decreases, but ultimately increases as shown in Fig. 9.1 (b). The decrease in unit cost is attributed to the internal and external economies of scale and the eventual increase in cost, to the internal and external diseconomies of scale. The economies and diseconomies of scale are discussed in the following section.
Long-run Marginal Cost Curve (LMC)

The long-run marginal cost curve (LMC) is derived from the short-run marginal cost curves (SMCs). The derivation of LMC is illustrated in Fig. 9.2 in which SACs, SMCs, and LAC are the same as in Fig. 9.1 (b). To derive the LMC, consider the points of tangency between SACs and the LAC, i.e., points A, B and C. In the long-run production planning, these points determine the minimum LAC at the different levels of production. Each of these outputs has an SMC. For example, if we draw a perpendicular from point A, it intersects SMC at point M determining SMC at output Q1. The same process can be repeated for points B and C to find out SMC at outputs Q2 and Q3. Note that points B and C determine SMC at BQ2 and CQ3, respectively. A curve drawn through points M, B and N, as shown by the LMC, represents the behaviour of the marginal cost in the long-run. This curve is known as the long-run marginal cost curve, LMC. It shows the trends in the marginal cost in response to the changes in the scale of production.

Some important inferences may be drawn from Fig. 9.1. The LMC must be equal to SMC for the output at which the corresponding SAC is tangent to the LAC. At the point of tangency, LAC = SAC. Another important point to be noted is that LMC intersects LAC when the latter is at its minimum, i.e., point B. It indicates that there is one and only one short-run plant size whose minimum SAC coincides with the minimum LAC. This point is B where

\[ SAC_1 = SMC_1 = LAC = LMC \]

Optimum Plant Size and Long-Run Cost Curves

Conceptually, the optimum size of a firm is one which ensures the most efficient utilization of resources. Practically, the optimum size of the firm is one that minimizes the LAC. Given the state of technology over time, theoretically there is technically a unique size of the firm and level of output associated with the least-cost concept. In Fig. 9.2, the optimum size of the firm consists of two plants represented by

---

**Fig. 9.2** Derivation of LMC

---
SAC₁ and SAC₂. The two plants together produce \(OQ_2\) units of a product at minimum long-run average cost (LAC) of \(BQ_2\). The downtrend in the LAC indicates that until output reaches the level of \(OQ_2\), the firm is of less than optimal size. Similarly, expansion of the firm beyond production capacity \(OQ_2\), causes a rise in SMC and, therefore, in LAC. It follows that given the technology, a firm aiming to minimize its average cost over time must choose a plant that gives minimum LAC where \(SAC = SMC = LAC = LMC\). This size of plant assures the most efficient utilization of the resources. Any change in output level—increase or decrease—will make the firm enter the area of inoptimality.

9.3.3 Economies and Diseconomies of Scale

While optimization of output in the long run is an important concern of business firms, cost minimization is an equally import decision area. Cost of production depends not only on internal factor—the productivity of inputs—but also on many external factors—the factors that arise out of the firm. Since we are concerned in this unit with the theory of cost, in this section we give a detailed analysis of internal and external economies and diseconomies of scale and how they determine the trend in cost of production. To begin with, let us have look at the trend of long-run average cost curve (LAC).

As shown in Fig. 9.1(b), LAC decreases with the expansion of production scale up to \(OQ_2\) and then it begins to rise. Decrease in LAC is caused by the economies of scale and increase in LAC is caused by diseconomies of scale. Economies of scale result in cost saving and diseconomies lead to rise in cost. Economies and diseconomies of scale determine also the returns to scale. Increasing returns to scale operate till economies of scale are greater than the diseconomies of scale, and returns to scale decrease when diseconomies are greater than the economies of scale. When economies and diseconomies are in balance, returns to scale are constant. In this section, we briefly discuss the various kinds of economies and diseconomies of scale and their effect on cost of production.

Economies of Scale

The economies of scale are classified as

(a) Internal or Real Economies, and
(b) External or Pecuniary Economies.

A. Internal Economies

Internal economies, also called ‘real economies’, are those that arise within the firm with addition of new production plants. This means that internal economies are available exclusively to the expanding firm. Internal economies may be classified under the following categories,

(i) Economies in production;
(ii) Economies in marketing;
(iii) Managerial economies, and
(iv) Economies in transport and storage.

(i) Economies in Production: Economies in production arise from two sources:
(a) technological advantages, and (b) advantages of division of labour based on specialization and skill of labour.

(ii) Economies in Purchase of Inputs: Economies in input purchases arise from the large-scale purchase of raw materials and other material inputs and large-scale selling of the firm’s own products. As to economies in the purchase of inputs, the large-size firms normally make bulk purchases of their inputs. The large scale purchase entitles the firm for certain discounts in input prices and other concessions that are not available on small purchases. As such, the growing firms gain economies on the cost of their material inputs.

(iii) Managerial Economies: Managerial economies arise from (a) specialization in managerial activities, i.e., the use of specialized managerial personnel, and (b) systemization of managerial functions. For a large-size firm, it becomes possible to divide its management into specialized departments under specialized personnel, such as production manager, sales manager, HR manager, financial manager, etc. The management of different departments by specialized managers increases the efficiency of management at all the levels of management because of the decentralization of decision-making. It increases production, given the cost. Large-scale firms have the opportunity to use advanced techniques of communication, telephones and telex machines, computers, and their own means of transport. All these lead to quick decision-making, help in saving valuable time of the management and, thereby, improve the managerial efficiency. For these reasons, managerial cost increases less than proportionately with the increase in production scale up to a certain level, of course.

(iv) Economies in Transport and Storage: Economies in transportation and storage costs arise from fuller utilization of transport and storage facilities. Transportation costs are incurred both on production and sales sides. Similarly, storage costs are incurred on both raw materials and finished products. The large-size firms may acquire their own means of transport and they can, thereby, reduce the unit cost of transportation, at least to the extent of profit margin of the transport companies. Besides, own transport facility prevents delays in transporting goods. Some large-scale firms have their own railway tracks from the nearest railway point to the factory, and thus they reduce the cost of transporting goods in and out. For example, Bombay Port Trust has its own railway tracks, oil companies have their own fleet of tankers. Similarly, large-scale firms can create their own godowns in various centres of product distribution and can save on cost of storage.

B. External or Pecuniary Economies of Scale

External economies are those that arise outside the firm and accrue to the expanding firms. External economies appear in the form of money saving on inputs, called
pecuniary economies. Pecuniary economies accrue to the large-size firms in the form of discounts and concessions on (i) large scale purchase of raw material, (ii) large scale acquisition of external finance, particularly from the commercial banks; (iii) massive advertisement campaigns; (iv) large scale hiring of means of transport and warehouses, etc. These benefits are available to all the firms of an industry but large scale firms benefit more than small firms.

Diseconomies of Scale

The economics of scale have their own limits, i.e., scale economies exist only up to a certain level of production scale. The expansion of scale of production beyond that limit creates condition for diseconomies of scale. Diseconomies of scale are disadvantages that arise due to the expansion of production scale beyond its optimum level and lead to rise in the cost of production. Like economies, diseconomies may be internal and external. Let us describe the nature of internal and external diseconomies in some detail.

1. Internal Diseconomies. Internal diseconomies are those that are exclusive and internal to a firm as they arise within the firm. Like everything else, economies of scale have a limit too. This limit is reached when the advantages of division of labour and managerial staff have been fully exploited; excess capacity of plant, warehouses, transport and communication systems, etc., is fully used; and economy in advertisement cost tapers off. Although some economies may still exist, diseconomies begin to outweigh the economies and the costs begin to rise.

Managerial Inefficiency. Diseconomies begin to appear first at the management level. Managerial inefficiencies arise, among other things, from the expansion of scale itself. With fast expansion of the production scale, personal contacts and communications between (i) owners and managers, (ii) managers and labour, and (iii) between the managers of different departments or sections get rapidly reduced. The lack of fast or quick communication causes delays in decision-making affecting production adversely.

Secondly, close control and supervision is replaced by remote control management. With the increase in managerial personnel, decision-making becomes complex and delays in decision-making become inevitable.

Thirdly, implementation of whatever decisions are taken is delayed due to coordination problem in large scale organisations.

Finally, with the expansion of the scale of production, management is professionalized beyond a point. As a result, the owner’s objective function of profit maximization is gradually replaced by managers’ utility function, like job security and high salary, standard or reasonable profit target, satisfying functions. All these lead to laxity in management and, hence to a rise in the cost of production.

Labour Inefficiency: Increasing number of labour leads to a loss of control over labour management. This affects labour productivity adversely. Besides, increase
in the number of workers encourages labor union activities that cause loss of output per unit of time and hence, rise in the cost of production.

2. External Diseconomies. External diseconomies are the disadvantages that arise outside the firm, especially in the input markets, due to natural constraints, specially in agriculture and extractive industries. With the expansion of the firm, particularly when all the firms of the industry are expanding, the discounts and concessions that are available on bulk purchases of inputs and concessional finance come to an end. More than that, increasing demand for inputs puts pressure on the input markets and input prices begin to rise causing a rise in the cost of production. These are pecuniary diseconomies.

On the production side, the law of diminishing returns to scale come into force due to excessive use of fixed factors, more so in agriculture and extractive industries. For example, excessive use of cultivable land turns it into barren land; pumping out water on a large scale for irrigation causes the water table to go down resulting in rise in cost of irrigation; extraction of minerals on a large scale exhausts the mineral deposits on upper levels and mining further deep causes rise in cost of production; extensive fishing reduces the availability of fish and the catch, even when fishing boats and nets are increased. These kinds of diseconomies make the LAC move upward.

9.4 MODERN THEORY OF COST

Up till now have learnt about the traditional theory of costs, in this section we will learn the modern theory of cost. Unlike the traditional theory of cost, the modern theory does not agree to the U-shape of the cost curve. As per the modern theory, the short run cost curve has a saucer-type shape and the long-run average cost curve is either L-shaped or inverse J-shaped. This is because of the existence of built-in reserve capacity which imparts flexibility and enables the plant to produce larger output without adding to the costs.

Short-Run Cost Curves

The SAVC, SMC and the AFC has a saucer shape. Further the AVC is saucer-shaped because the reserve capacity is maintained by the firm and the AFC is downward sloping even as the output expands. The SAC curve has a U-shape even in the modern version.

Long-Run Cost Curves

The long-run average cost curve is not U-shaped in the modern theory, in fact it is L-shaped of inverse J-Shaped. This can be understood through the returns to scale. The traditional theory assumed a U shaped Long-run cost curve because of the assumption that after the optimal level of output, the diseconomies of scale overtake the economies of scale. As per the modern theory, the long run average costs includes basically production and managerial costs of which the average
production costs continue to fall even at large scales while the managerial costs per unit of output may rise only gradually and at large scales of output.

9.5 COST CONTROL AND COST REDUCTION

The primary aim of business firms is to produce more products at the least possible cost. Profit maximization becomes essential in such conditions. The cost accounting systems have gained prominence because the success of the firms depends on their ability to reduce costs while selling more products. Cost control, essentially, deals with the cost-effective ways of operations. This majorly involves the avoidance or elimination of wastages in the current operations.

The primary aim of business firms is to produce more products at the least possible cost. Profit maximization becomes essential in such conditions. The cost accounting systems have gained prominence because the success of the firms depends on their ability to reduce costs while selling more products. Cost control, essentially, deals with the cost-effective ways of operations. This majorly involves the avoidance or elimination of wastages in the current operations.

The advantages of cost control are:

- The profit margins increase along with it becoming the top competition for others
- It is the necessary condition for ensuring efficient productivity
- The competitive prices will ensure that the sales are stable and employment of the workforce
- It reduces the costs involved in producing products thereby reducing its overall costs
- Without cost control, it is very much possible that the firm’s profits start to decrease even with an increase in the volume of sales

There are mainly two types of standards of cost control and reduction:

- External: This includes the company adjudging its performance in comparison to its competitions. This is done through measuring cost ratios.
- Internal: These are standards used by the companies to identify places where costs can be reduced. Some of the tools under these standards are:
  - Budgetary control: This is nothing but a plan of action for a given period of time in numerical terms. The organizational activities are not only forecasting but carried forward based on the anticipated revenues and expenses.
  - Standard costing: This refers to the assigning and use of certain set standard costs based on which actual costs are computed. The differences in such comparisons are then corrected to minimize wastage.
9.6 COST BEHAVIOUR AND BUSINESS DECISION

Cost behaviour primarily refers to the phenomenon of studying the pattern in which costs change in relation to volume and activity. The categorization of cost behaviour is done on the basis of:

- Costs which change with the activity or volume
- Costs which remain constant throughout
- Costs which fall somewhere in the middle

As you can understand, these types of costs are total costs, fixed costs and variable costs. We have already seen how these affect output and production earlier in the unit, now let’s see why studying costs are important in business decisions.

The correct forecasting or estimation of costs is crucial as it assists the managers ensure that their planning is accurate. Based on this planning, organizing activities are undertaken which ultimately reflects on the business performance. Several researchers stress on the importance of studying cost behaviour as it is crucial to making financial decisions in the present, as well as planning and decision-making in the future. Further, the significance of cost behaviour to business decision can be understood by the fact that any half-baked knowledge will result in the company incurring serious losses. Proper study of cost behaviour is helpful in ensuring the operating margins are kept under control and that the wastages are identified and eliminated.

Check Your Progress

3. What is quadratic cost function?
4. What does the operation of the law of increasing returns to variable inputs in the initial stage of production means in terms of costs?
5. Name the curve whose shape remains a U-type in the modern theory of cost.

9.7 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Explicit costs are those which area actually incurred by the firm in payment for labour, material, plant, building, machinery, equipment, travelling and transport, etc.
2. The running cost and depreciation of the capital assets are included in the short-run or variable costs.
3. When the total cost increase at increasing rate with constant increase in output, the TC data produces a quadratic cost function.

4. In the initial stages of production, both AFC and AVC are declining because of internal economies. Since AC=AFC+AVC, AC is also declining. This shows the operation of the law of increasing returns to the variable input in the initial stage of production.

5. In the modern theory of cost, the SAC curve has the same U-shape.

### 9.8 SUMMARY

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

- Accounting cost concepts include opportunity cost and actual cost; business costs and full costs; explicit and implicit cost; and out-of-pocket and book costs.

- Analytical cost concepts are: fixed and variable costs; total, average and marginal costs; short-run and long-run costs; incremental costs and sunk costs; and historical and replacement costs.

- The theory of cost deals with how costs of production changes with change in output. The basic principle of the cost-output relationship is that the total cost increases with increase in output.

- The short-run total cost is composed of two major elements (i) total fixed cost and, (ii) total variable cost. That is, in the short-run, TC=TFC+TVC.

- Long-run cost curve is composed of a series of short-run curves.

- Decrease in LAC is caused by diseconomies of scale. Economies of scale result in cost saving and diseconomies of scale determine also the returns to scale. Diseconomies of scale are disadvantages that arises due to the expansion of production scale beyond its optimum level and lead to rise in the cost of production.

- Unlike the traditional theory of cost, the modern theory does not agree to the U-shape of the cost curve. As per the modern theory, the short-run cost curve has a saucer-type shape and the long-run average cost curve is either L-shaped or inverse J-shaped. This is because of the existence of built-in-reserve capacity which imparts flexibility and enables the plant to produce larger output without adding to the costs.

- The primary aim of business firms is to produce more products at the least possible cost. Profit maximization becomes essential in such conditions. The cost accounting systems have gained prominence because the success
of the firms depends on their ability to reduce costs while selling more products. Cost control, essentially, deals with the cost-effective ways of operations. This majorly involves the avoidance or elimination of wastages in the current operations.

- Cost behaviour primarily refers to the phenomenon of studying the pattern in which costs change in relation to volume and activity. The correct forecasting or estimation of costs is crucial as it assists the managers ensure that their planning is accurate. Based on this planning, organizing activities are undertaken which ultimately reflects on the business performance.

### KEY WORDS

- **Analytical cost concepts:** It refers to the different cost concepts that are used in analysing the cost-output relationship with increase in inputs and output and also the cost concepts that figure in analysing the effect of expansion of production on the society as a whole.
- **Short-run:** It refers to the time period during which scale of production remains unchanged.
- **Long-run:** It refers to those costs that are incurred to increase the scale of production in the long-run.
- **Total cost:** It is defined as the actual cost that are incurred to produce a given quantity of output.
- **Marginal cost:** It is defined as the change in the total cost due to change in the total output by one unit.

### SELF ASSESSMENT QUESTIONS AND EXERCISES

**Short Answer Questions**

1. Differentiate between business costs and full costs.
2. What are incremental costs?
3. Explain social costs.
4. Write a short note on cost control and reduction.
5. What is the modern theory of cost?

**Long Answer Questions**

1. Explain the accounting and analytical cost concepts relevant to decision-making.
2. Describe the cost curves produced by linear, quadratic and cubic cost functions.

3. Mention some of the important relationships between difference measures of cost.

4. Examine long-run cost-output relations.

5. Discuss the economies and diseconomies of scale.

9.11 FURTHER READINGS


Websites

https://blogs.economictimes.indiatimes.com/et-commentary/heres-how-indias-widening-income-distribution-can-be-redressed/

UNIT 10 PENETRATION AND SKIMMING STRATEGIES

Structure
10.0 Introduction
10.1 Objectives
10.2 Concept of Profit
   10.2.1 Types of Profit
10.3 Theory of Profit
   10.3.1 Dynamic Theory of Profit
   10.3.2 Innovation Theory of Profit
   10.3.3 Uncertainty – Bearing Theory of Profit
   10.3.4 Risk Bearing Theory of Profit
10.4 Cost Volume Profit Analysis
   10.4.1 Total Contribution Margin
   10.4.2 Basic Assumptions of Cost Volume Profit Analysis
10.5 Pricing of a New Product
10.6 Answers To Check Your Progress
10.7 Summary
10.8 Key Words
10.9 Self Assessment Questions and Exercises
10.10 Further Readings

10.0 INTRODUCTION

In ordinary language, profit is the surplus of income over expenses of production according to a businessman. Profit is the fourth component of factor pricing. It is the amount left with the entrepreneur after he has made all other payments rent to land, wage to labour and interest on capital. Therefore profit can be called as remuneration to the entrepreneur for his service in the production process. No doubt profits are associated with entrepreneur and functions. In this unit, you will study about the concept of profit, types and theories of profit and cost volume profit analysis.

10.1 OBJECTIVES

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

- Interpret the concept of profit
- Name the types of profit
- Discuss the theories of profit
10.2 CONCEPT OF PROFIT

In general, profit is the difference between cost and the price of the commodity. In more technical terms, profit implies a positive residual of an entrepreneur after deducting the total cost from the total revenue. It is worth mentioning that profits are defined as residual income left after the payment of the contractual rewards to other factors of production. The entrepreneur while engaging other factors of production enters into contract with them. He, thus, pays wages to the workers, rent on the land employed, and interest on the loans taken at the rate already fixed by contracts. In fact, the entrepreneur makes payment to these factors much in advance of the realization of the value of the output produced after sale of the product. What is left after paying the contractual rewards of other factors employed is the profit of entrepreneur. Thus, profits are non-contractual income, therefore, they may be positive or negative, whereas the contractual income of other factors such as wage, rent, interest are always positive and never negative.

The term profit has been differently defined by different economists. Some definitions are as follows:

- According to Ulmer: “Profit may be defined the difference between firm’s total revenue and its total cost.”
- According to Pearce: “Profit is the difference between the revenue generated from the sale of output and the full opportunity cost of factors used in the production of that output.”
- According to Prof. J.B. Clark, “Profit arises due to disequilibrium of demand and supply.”
- According to Prof. Hawley, “Profit of an enterprises is not reward of management or co-ordination but at the risk and responsibilities that he undertakes.”
- F.H. Knight— “Profit is the reward for risk bearing but for uncertainty bearing.”

10.2.1 Types of Profit

Let us now study the types of profit.

1. Gross Profit

Gross profit refers to that part of the income of a businessman which is available to him after all payment to the contractually hired factors and other current...
obligations like taxes and depreciation charges. In other words, the difference between total revenue of entrepreneur and total explicit costs is called gross profit. Thus,

\[
\text{Gross profit} = \text{Total Revenue} - \text{Explicit cost}
\]

For instance, the total revenue of business man is ₹ 25,000. He pays ₹ 6000 as wage, ₹ 5000 as interest and ₹ 2000 by the way of rent. Therefore, his total explicit cost will be ₹ 13000 (6000 + 5000 + 2000).

Gross profit includes many constituents. The main constituents of gross profit are depreciation and replacement of capital, monopoly gains, and rewards for other factors supplied by the entrepreneur and so on.

2. Net Profit

Pure profit is the amount that is available to the entrepreneur for assuming risk inseparable from all business in anticipation of demand. In simple words, pure profit can be calculated by deducting the implicit costs comprising rent of entrepreneur’s land, interest on entrepreneur’s capital, wage for entrepreneur’s labour and so forth from gross profit. Symbolically, we can write it as: \( \text{NP} = \text{GP} - \text{IC} \).

\[
\text{Net profit} = \text{Gross profit} - \text{Implicit Cost}
\]

\[
\text{Pure/Net Profit} = \text{Total Revenue} - \text{Explicit costs} - \text{Implicit Cost}
\]

Or

\[
\text{Net profit} = \text{Total Revenue} - \text{Total Cost}
\]

Difference between Net Profit and Gross Profit

The difference between gross profit and net profit is based on certain arguments which are given as the following:

- **Wider**: Gross profit is wider as compared to net profit. In fact, net profit is a part of gross profit.
- **Costs**: Gross profit includes the explicit cost while net profit includes implicit costs.
- **Perfect competition**: Under the condition of perfect competition, net profit can be obtained but a part of net profit is earned.
- **Loss**: Actual loss can be calculated from gross profit, but in the case of net profits this not possible.
- **Long period**: In a long period, net profit is not available but only a part of gross profit is earned.
10.3 THEORY OF PROFIT

Profit is the reward for the enterprise. No doubt profit is associated with the entrepreneur and his function but the economists from time to time have expressed diverse and conflicting views about the nature, origin and role of profit. Till today, there is no complete agreement among economists about the nature and origin of profits.

Some economists have held the view that the function of the entrepreneur is to organize and coordinate other factors of production. According to them, the entrepreneur earns profit for his performing function. As per this view, enterprise is a special type of labour and profit is a special form of wage. Others have described the entrepreneur as performing the joint and inseparable functions of responsibility (ultimate risk bearing) and control (decision making). The entrepreneur earns profit because he takes a risk of incurring losses when his price and output policies prove to be incorrect in view of the future business movements. Thus, Hawley ascribed it to the entrepreneur’s risk-taking ability. Schumpeter has assigned to the entrepreneur the role of an innovator and profits as a reward for introducing innovations. F.H. Knight has emphasised uncertainty in the economy as a factor which gives rise to profits and bearing of uncertainty is the task assigned to the entrepreneur. Thus, profit is an income which arises out of change, uncertainty and friction inherent in a dynamic world.

Some economists have described profit as a non-functional income. The early classical economists regarded profit as accruing to the capitalists who supplied capital and owned the business. They did not distinguish between interest and profit. J.M. Keynes expressed the view that profits resulted from the favourable movements of general price level. J. Robinson, E.H. Chamberlin and M. Kalecki have associated profit with the imperfect competition and monopoly. According to them, the greater the degree of imperfection or in other words the greater the degree of monopoly power, the greater the profits made by the entrepreneur.

This profit has been associated by Knight with uncertainty, by Schumpeter with innovations, by Hawley with risk-bearing, and by Joan Robinson and Chamberlin with the degree of monopoly power. As a matter of fact, profits arise from all these sources.

"Profit may come to units as a result of monopoly or monopsony, as a reward for innovation, as reward for correct estimate of uncertain factors, either particular to the industry or general to the whole economy."
10.3.1 Dynamic Theory of Profit

Professor J.B. Clark propounded his dynamic theory of profit in 1900. According to him, profit is defined as the difference between price and cost of production of a commodity. He argued that in a stationary state where no changes in conditions of demand and supply are taking place, the prices paid to the factors on the basis of their marginal productivity would exhaust the total value of production and profit would accrue to the entrepreneur. Profits accrue when selling prices of the goods exceed their cost of production. If there is no change either in the condition of demand or in the condition of supply, competitive equilibrium will persist and, therefore, no profits will be earned by the entrepreneur. On the contrary, if due to change in either demand or supply, price exceeds cost of production, profit will arise. If due to these changes, price falls below the cost of production, negative profits, that is, losses will accrue to the entrepreneur. It is evident that changes disturb the equilibrium and thereby give rise to profits. In other words, profits arise due to disequilibrium caused by the changes in demand and supply conditions.

It should be noted that due to disequilibrium profits arise from unanticipated changes in demand or cost conditions. If these changes could be foreseen in advance, then suitable adjustments can be made according to the anticipated changes so that forces of competition do not drive profits to zero.

Now the question is; what changes occur in the economy and give rise to profits? J.B. Clark mentioned five changes that occur in the dynamic economy and which give rise to profits. These five changes are mentioned as the following:

- Increase in population.
- Changes in the quantity and quality of human wants
- Changes in methods or techniques of production
- Changes in the amount of capital
- Changes in the forms of business organization

These changes are constantly taking place and bring about divergence between price and cost and thereby give rise to profits—positive and negative. If the demand for a commodity increases due to increase in population or increase in the income of the people or due to the increase in consumer’s preference for the commodity, the price of the commodity will rise, and if the cost remains the same, profits would accrue to the entrepreneurs producing the commodity. On the other hand, cost of production may go down as a result of the adoption of a new technique of production, or as a result of low cost of the raw material, and if prices remain constant or do not fall to the same extent, then profit would arise.
Apart from the five changes, there are other changes also which take place in the economy. All the changes which take place and as a result of which profits arise in a dynamic economy may be classified into two types— (i) innovations and (ii) exogenous changes.

(i) Innovation represents changes which are introduced by individual entrepreneurs themselves. The entrepreneur earns large profits from introducing innovations such as a new product, a new and cheaper method of production, a new method of marketing the product, a new way of advertisement. The innovational changes may either reduce the cost or increase the demand for the product and thereby bring profits into existence. New innovations are being continuously introduced by entrepreneurs and profits continue to arise out of them.

(ii) Exogenous changes refer to those changes which are external to the firms or industries in an economy. These changes affect all firms in an industry or sometimes all the industries in the economy. Examples of exogenous changes include outbreak of wars, periods of inflation and sometimes business depression and falling prices, changes in the monetary and fiscal policies of government affecting favourably or unfavourably, changes in technology of production, changes in tastes and preferences of the consumers, changes in income and spending habits of the people, changes in the availability of substitute products, alteration in the legislative and legal environment affecting the industries, and changes in preferences between income and leisure. All these changes affect the cost or demand of the products and give rise to profits— positive and negative as the case may be. For instance, during war when prices of goods mount up, and costs lag behind, the entrepreneurs make a lot of profits. Similarly, when inflation takes place due to the increase demand for goods caused by rising incomes, increasing population and expansion in the money supply, huge profits accrue to the firms.

10.3.2 Innovation Theory of Profit
The innovation theory of profit is propounded by Professor J.A. Schumpeter. It has been held by J.A. Schumpeter that the main function of the entrepreneur is to introduce innovation in the economy and profits are the reward which he earns for performing his functions. According to him, “Profits arise due to the introduction of innovations in the production process or the sale of output produced.” Innovations refer to those measures which either curtail the cost of production or lead to an increase in demand of the product. In the first case, innovations are included in the form of introduction of a new machinery, new and cheaper technique or process of production, exploitation of new
source of raw material, a new and better method of organising the firm and so forth. In the second case, innovations are those which change the demand or utility function thereby increasing demand. In the second case, innovations are included in the form of introduction of a new product, a new variety or design of the product, a new and superior method of advertising, discovery of new markets and so forth.

If an innovation proves successful, that is, if it achieves its aim of either reducing the cost of production or enhancing the demand for the product, it will give rise to profit. Profits emerge due to successful innovation whereby either the cost of production falls below the prevailing price of the product or the entrepreneur is able to sell more units at a better price than before; consequently, the entrepreneur earns profit very wisely.

It is here worth mentioning that profits brought about by a particular innovation are only temporary in nature and tend to fade as the competitors of the industry might imitate that innovation. Hence, as long as the entrepreneur enjoys a monopoly position in the market, he will earn profits. However, when his innovative measure is imitated by the competitors then his profits will start declining gradually. Now, the entrepreneur will further strive to incorporate some new innovations so that he may remain ahead of others. In a competitive and progressive economy, the entrepreneur always continues to introduce new innovations and, thus, earning profits.

**Criticism:** The innovation theory of profit is subject to the following limitations:

- Schumpeter has not given importance to the element of uncertainty. As long as there is no uncertainty, profits will be equal to wages.
- Schumpeter has overseen the risk taking ability of the entrepreneur.
- Schumpeter calls economic profits a temporary surplus. In the long run, innovations get reduced and so there will be no profits.

**10.3.3 Uncertainty – Bearing Theory of Profit**

Professor F.H. Knight in his classic work, “Risk, uncertainty and Profit” propounded the theory of profit. According to him, “Profit is the reward not for risk bearing but for uncertainty bearing.” He accepts that the dynamic changes and risk bearing are all important elements of profit. As the future is uncertain and unpredictable, any entrepreneur who probes in the darkness does so only with a view to earn some profit.

Knight has greatly developed the theory of profits based on uncertainty. He has distinguished between risk and uncertainty on the one hand and predictable and unpredictable changes on the other hand. According to him, dynamic changes give rise to profits.
If there were no changes or if the changes were foreseen and predictable, there would have been no uncertainty about future and no profits. If the future conditions become completely predictable, then competition would certainly adjust things to the ideal state where all prices would equal costs and profit would not emerge. It is the divergence of the actual conditions from those which are expected and on the basis of which business arrangements have been made that give rise to uncertainty and profit.

Entrepreneurs have to make estimates of the future conditions regarding demand for the product and other factors which affect price and cost. In view of their estimates and anticipations, they make contract with the suppliers of factors of production in advance at fixed rates of remuneration. They realize the value of the output produced and sold in the markets. It is to be noted that between the time of contracts and the sale of the output, many changes may take place which can upset anticipation for good or worse and thereby give rise to profits—positive and negative.

If the conditions prevailing at the time of the sale of output could be known and predicted when the entrepreneurs enter into contractual relationships with the factors of production about their rates of remuneration, there would have been no uncertainty and, therefore, no profits. Thus, uncertainty leads to complete ignorance about the future conditions of demand and supply.

We, thus, see that profit is a residual and non-contractual income which accrues to the entrepreneurs because of the fact of uncertainty. It is the entrepreneur who bears uncertainty and earns profit as a reward for that bearing that hazard.

Now the question is what are changes causing uncertainty? There are two types of changes which take place and are responsible for uncertainty. The first type of change refers to the innovations (for example, introduction of a new product or a new cheaper method of production). The second type is changes in the taste and fashions of the people, changes in the government policies and laws especially taxation, wages and labour policies and laws, movement of prices as a result of inflation and depression. All these changes cause uncertainty and bring profits either positive or negative into existence.

Due to the changes there are continuously taking place in the economy, the entrepreneur has to face many risks. Professor F.H. Knight has drawn a distinction between insurable risk and non-insurable risk. The entrepreneur faces risk of fire, theft, accident and from other factors which may cause him huge losses. But these kinds of risks—fire, theft and accident can be insured against by paying a fixed premium. Insurance premium is included in the cost of production. Thus, no uncertainty arises due to insurable risks as far as the individual entrepreneur is
concerned and, therefore, they cannot give rise to profit.

Only those risks can be insured the probability of whose occurrence can be calculated. But there are many risks which cannot be insured and, therefore, they have to be borne by the entrepreneurs. These non-insurable risks relate to the outcomes of the price output decisions to be taken by the entrepreneurs.

**Criticism**

The main shortcomings of the theory are mentioned as the following:

- Knight’s theory is based on unrealistic assumptions. It considers that uncertainty is the root cause for restricting the supply of the entrepreneur. There are so many other factors, i.e., lack of knowledge and others, which are equally responsible for the same.
- Knight’s theory is criticised on the ground that it is a vague theory of profit determination. In the present times, MNCs have assumed such great importance that they have succeeded in bringing an element of certainty in their profit.
- Knight’s theory of profit in incomplete. This theory gives no place to other functions of the entrepreneur namely, management, coordination and so forth.
- In joint stock companies, the entrepreneurial functions are divided among the shareholders, managers and directors. The theory is not clear as to who bears uncertainty.

**10.3.4 Risk Bearing Theory of Profit**

The risk theory of profit was propounded by F.B. Hawley who regarded risk taking ability as the main function of the entrepreneur. Profit is the residual income which the entrepreneur receives because he assumes risks. The entrepreneur exposes his business to risk and receives in turn a reward in the form of profits because the task of risk-taking is irksome. Profit is ‘an excess of payment above the actuarial value of the risk. No entrepreneur will be willing to undertake risks if he gets only the normal return. Therefore, the reward for risk taking must be higher than the actual value of the risk. In Hawley’s words, ‘The profit of an undertaking, of the residue of the product after the claims of land, capital and labour (furnished by others or by the undertaker himself) are satisfied, is not the reward of management or coordination, but of risks and responsibilities that the undertaken subjects himself to.’

According to Hawley, the entrepreneur can avoid certain risk for a fixed payment to the insurance company. But he cannot get rid of all risks by means of insurance, for if he is able to do so, he would cease to be an entrepreneur and
Penetration and Skimming Strategies

NOTES

would earn only wages of management and no profit.

However, when the entrepreneur transfers his risk to the insurance company, he abdicates his risk-taking function to the latter which receives the profit. The reward of the insurance company is not the premium it receives, but the difference between that premium and the loss it eventually suffers. So profit is the reward of risk taking, especially of 'wisely selected' risks. It is to be noted that not all entrepreneurs are capable of taking risks. Those who remain in business are able to earn an excess of payment above the actuarial value of the risk.

Its criticisms

Like other theories, the risk theory of profit has also been criticised on the following grounds:

- **Meaning of risk is unclear**: Hawley does not clarify the meaning of risk.
- **Profits due to the entrepreneurial ability**: Risk taking is not only an entrepreneurial function which leads to the emergence of profits. Profits also accrue due to the organizational and coordinating ability of the entrepreneurs.
- **Profits are the reward of avoiding risks**: According to Carver, those entrepreneurs who are able to avoid risks earn profits. Hence, profits arise not because risks are undertaken but because they are avoided by able entrepreneurs.
- **Amount of profit not related to size of risk**: The quantum of profit is any way not related to the size of the risk undertaken. If it were so, every entrepreneur would involve himself into huge risks in order to earn larger profits.
- **Incomplete theory**: There is little empirical evidence to prove that entrepreneurs earn more in risky enterprises. Hawley's risk theory is also an incomplete theory profit.

10.4 COST VOLUME PROFIT ANALYSIS

The success of a business as measured in term of profit depends upon its sales. The volume of sales must be sufficient to cover all costs and allow a satisfactory margin for net income. When the proportion of the fixed costs in a business becomes large in relation to total costs, then volume becomes an extremely important factor in achieving profitability. For example a business with only variable costs would be able to report net income at any level of sales as long as price exceeds the
variable cost rates. However, a business with only fixed costs cannot accrue profit until the contribution from sales is equal to the amount of fixed expenses. Therefore, a minimum level of sales is absolutely essential in a business that incurs fixed expenses.

It is essential to know that change in volume can have a profound impact on the profits of a business. Cost volume profit analysis has been developed as a management tool to enable analysis of the following variables:

(i) Price
(ii) Quantity
(iii) Variable cost
(iv) Fixed cost

The local point of the cost-volume profit analysis is on the effect that changes in volume have on fixed and variable costs. Volume may be regarded as the number of units sold. Typically, the theory of cost-volume-profit analysis is explained in term of units. However, using units as the measure of volume for completing break-even point or target income point requires that the business sells only a single product. Since all business from a practical viewpoint sell multiple products, the practical use of cost volume profit analysis requires that volume should be measured in term of sales.

Cost volume profit analysis for a single product business

The primary objective of the management in using cost volume profit analysis is to determine the target income point and not the break-even point. The break-even point is simply the quantity of sales that achieve net zero income. It is that level of sales where total sales equals total expenses. The break-even point may be computed as follows:

\[
\text{B.E.P. (in units)} = \frac{F}{S - V}
\]

Where
- \( F \) = Fixed cost
- \( S \) = Selling price (per unit)
- \( V \) = Variable cost (per unit)

Let \( F = 20,000 \), \( S = 10 \), \( V = 6 \)

\[
\text{B.E.P. in rupees} = \frac{20,000}{10 - 6} = 20,000 \times \frac{4}{4} = 50,000 \text{ Units}
\]
Cost volume profit analysis for multiple products

For a business selling multiple products, it is necessary to use the monetary value of sales as the measure of volume. The cost volume profit analysis for a multiple product business can be calculated as follows:

\[ S = \frac{I + F}{1 - V} \]

Where,  
- \( S \) = Sales Volume  
- \( F \) = Fixed Cost  
- \( V \) = Variable cost (percentage)

This equation may be used to complete the value level of sales required to attain a desired level of income. For example, assume that a company's variable cost percentage is 80 and fixed cost is 10,000. Furthermore, assume that management has set a profit goal of 50,000. What must be the value of volume of sales in order to attain the 50,000 profit objective?

\[
\frac{I + F}{1 - V} = \frac{50,000 + 10,000}{1 - 0.8} = \frac{60,000}{0.2} = 3,00,000
\]

Graphical Illustration of Cost-Volume Profit Analysis

Since the fundamental relationship of cost volume profit analysis is basically mathematical in nature, the elements of cost volume profit analysis can be illustrated graphically. The general procedure is to plot the revenue and cost functions on the same graph. In order to illustrate the cost volume profit graphically, the following data has been assumed:

- \( P = 10 \); \( V = 6 \); \( F = 20,000 \)

For purpose of preparing the graph assume different levels increasing activity by increments of 1,000 units.

We can explain cost-volume profit analysis with the help of Figure 10.1.
Fig. 10.1 Cost Volume Profit Graph

Fig. 10.1 represents the complete cost volume profit graph. The graph can be easily used to show the net income and net loss. In the graph, volume is shown on the X axis and cost/sale is shown on the Y axis. E is the break-even point at this point TR equal to TC, it is the point of zero profit. This point is also known as no profit, no loss or profit beginning point. Any sale below this point shows loss and above this point shows profit. Thus, the area to the right of point E is the profit zone.

The contribution of margin concept

Use of cost volume profit analysis requires understanding the concept of contribution margin. Contribution margin is simply stated as total sales less total variable costs. Contribution = Sales – Variable Cost

Let = 80,000 – 48,000

= ₹ 32,000

The contribution margin rate and the contribution margin percentage are often called contribution margin ratio. A ratio may be expressed on a unit basis or a percentage basis. The concept of contribution margin provides a rather unique way of interpreting the activity of a business. At the start of the operating period, a business with fixed expenses would show a loss. At zero sales, the loss would be
equal to the total fixed expenses (cost). As each unit of product is sold, the loss is gradually reduced by the contribution margin of each unit sold. No profit can be reported until total contribution equals total fixed cost. After break-even point, each unit sold contributes to net income, an amount equal to the contribution margin per unit of product.

10.4.1 Total contribution margin

It is important to understand that the term ‘contribution’ means a contribution first to fixed costs. As previously mentioned, there can be profit in a business until total contribution equals total fixed cost/expenses. When this occurs, the business has reached the break-even point (B.E.) point. Break-even point is that point where the quantity of sales cause total contribution margin to be exactly equal to total fixed expenses.

Contribution margin per unit of product

The use of cost volume profit analysis as a decision-making tool also requires understanding the concept of contribution margin per unit of product. Contribution margin is illustrated in Figure 10.2.

![Contribution Margin Graph](image)

TVC stands for total variable cost. TFC stands for total fixed cost. TFC is added to TVC in order to arrive at TC (Total cost). The difference between TR and TVC is the contribution margin.

Margin Safety

This type of breakeven analysis can be used to calculate the level of sales which must be attained to avoid loss or to calculate the Margin Safety (MS). MS is the
difference between the firm’s actual level of sales and sales at the BEP point. It is expressed as the following:

\[ MS = \text{Actual sales revenue} - \text{B.E. Sales} \]

However, firms calculate the MS in terms of ratio as

\[ \text{MS Ratio} = \frac{MS}{\text{Actual Sales}} \]

The MS is an indicator of the strength of a firm. If margin is large, it shows that the firm can make profit even if it has to face difficulties. On the other hand, if the margin is small, a small reduction in sales can lead to loss. MS is nil at the BE Point because actual sales volume is equal to the cost.

### 10.4.2 Basic Assumptions of Cost Volume Profit Analysis

Cost volume profit analysis may be used as a profit planning tool and thereby it is useful in decision making. However, effective use of cost volume profit analysis for planning purposes requires understanding of certain basic assumptions. Unless the following assumptions are substantially met, any attempt to use cost volume profit analysis in a real world situation may prove to be inaccurate and misleading.

The basic assumptions underlying the cost volume profit analysis can be summarized as follows:

- Within a relevant range of volume, the variables namely, price, quantity, fixed costs and variable costs are subject to managerial control.
- Price and the variable cost rate are constant within the relevant range of activity.
- This assumption simply means that variable costs and revenue are assumed to vary early with changes in volume. Stated differently, changes in volume have no effect on price, the variable cost rate and fixed costs.
- In a company selling multiple products, the sales mix ratio remains constant with changes in total sales.
- In a company that uses absorption costing unless sales equals production, there exists no unique break-even point. When direct costing is used, no problems arise where production varies from sales. In direct costing, fixed manufacturing overhead is treated as a period charge.

Cost volume profit analysis is a powerful analytical tool. It can be effectively used in different kinds of decisions. Cost volume profit analysis is based on the theory of cost behaviour and as such it is imperative that the management has a good understanding of the cost behaviour.
10.5 PRICING OF A NEW PRODUCT

A new product may simply be either another brand name added to the existing ones or an altogether new product. Pricing a new brand for which there are several substitutes available in the market is not as big a problem as pricing a new product for which close substitutes are not available. In case of the former (where there are several substitutes available in the market), the market provides adequate information regarding cost, demand and availability of market and so forth. In this case, pricing depends on the nature of the market. In pricing a new product without close substitutes, however, problems arise due to lack of information; hence, there is some degree of uncertainty.

Thus, the pricing policy in respect of new products depends on whether or not close substitutes are available. Depending on whether or not close substitutes are available in pricing a new product, generally two kinds of pricing policies are available: (i) Penetration Price Policy and (ii) Skimming Price Policy.

1. Penetration Price Policy: In contrast to the skimming price policy, the penetration price policy involves a reserve strategy. This pricing policy is generally adopted in the case of new products for which substitutes are available. This policy requires fixing a lower initial price designed to penetrate the market as quickly as possible and is intended to maximize the profits in the long run. Therefore, the firm pursuing the penetration price policy sets a low price of the product in the initial stage. As the product slowly becomes popular in the market, its price is gradually increased. The success of the penetration price policy requires the existence of the following conditions:

(i) The short run demand for the product should have elasticity greater than unity $e_p > 1$. It helps in capturing the market at lower prices.
(ii) Economics of large scale production are available to the firm with the increase in sales. Otherwise, increase in production would result in increase in costs which might reduce the competitiveness of the price.
(iii) The product should have a high cross elasticity in relation to rival products for the initial lower price to be effective.
(iv) The potential market for the product is fairly large and has a good deal of future prospects.
(v) The product by nature should be such that it can be easily accepted and adopted by the consumers.

2. Skimming Price Policy: The skimming price policy is adopted where close substitutes of a new product are not available. This pricing strategy is intended to skim the cream of the market, i.e., consumer’s surplus, by setting a high initial
Penetration and Skimming Strategies

NOTES

Self-Instructional Material

185

price, three or four times the ex-factory price, and a subsequent lowering of prices in a series of reduction especially in case of consumer durables. The initial high price would generally be accompanied by heavy sales promoting expenditure. This policy succeeds due to the following reasons:

(i) In the initial stage of the introduction of products, demand is relatively inelastic because of consumers design for distinctiveness by the consumption of a new product.

(ii) Cross elasticity is usually very low for lack of a close substitute.

(iii) Step by step price cut help skimming consumers surplus available at the lower segment of demand curve.

(iv) High initial prices are helpful in recouping the development costs.

The post skimming strategy includes the decisions regarding the time and size of price reduction. The appropriate occasion for price reduction is the time of saturation of the top level demand or when a strong competition is apprehended. As regards the rate of price reduction, when the product is on its way to losing its distinctiveness, the price cut should be appropriately larger. But if the product has retained its exclusiveness, a series of small price reductions would be more appropriate.

The choice between the two strategic price policies depends on the following aspects:

(i) The rate of market growth

(ii) The rate of erosion of distinctiveness, and

(iii) The cost structure of the producers

If the rate of the market growth is slow for such reasons as lack of information, consumers’ hesitation and other factors then the penetration price policy would be unsuitable. The reason is that a low price will not imply sale of more number of units. If the pioneer product is likely to lose its distinctiveness at a faster rate then the skimming price policy would be unsuitable. It should be followed when the lead time i.e., the period of distinctiveness is fairly long. If its cost structure shows increasing return overtime; penetration pricing policy would be more suitable, since it enables the producer to reduce the cost and prevents potential competitors from entering the market in the short run.

Government Control over Pricing

General price controls—controls on prices of many goods—are often imposed when the public becomes alarmed that inflation is out of control.

In India, price control has been one of the major instruments used by the government for accomplishing economic goals and for the implementation of Five
Year Plans.

Why is price control required?

The price mechanism in a free enterprise economy leads to an optimum allocation of resources subject to certain assumptions.

The imposition of the price control mechanism in the context of the Indian economy needs to consider the following aspects:

(i) It is vital to analyse the choice of products within the framework of economic justification for imposing price controls on them.

(ii) It is necessary to evaluate the impact of price control mechanism on the flow of investment resources especially on those products that are subject to price control.

Check Your Progress

1. State one difference between gross profit and net profit.
2. Mention the five changes which according to J.B. Clark give rise to profits in a dynamic economy.
3. Who propounded the innovation theory of profit?

10.6 ANSWERS TO CHECK YOUR PROGRESS

1. Actual loss can be calculated from gross profit, but in the case of net profits this is not possible.
2. The five changes which according to J.B. Clark give rise to profits in a dynamic economy are the following:
   - Increase in population.
   - Changes in the quantity and quality of human wants
   - Changes in methods or techniques of production
   - Changes in the amount of capital
   - Changes in the forms of business organization
3. The innovation theory of profit is propounded by Professor J.A. Schumpeter.

10.7 SUMMARY

- In ordinary language, profit is the surplus of income over expenses of production according to a businessman. Profit is the fourth component of factor pricing.
In general, profit is the difference between coast and the price of the commodity. In more technical terms, profit implies a positive residual of an entrepreneur after deducting the total cost from the total revenue.

- Profits are non-contractual income, therefore, they may be positive or negative, whereas the contractual income of other factors such as wage, rent, interest are always positive and never negative.

- Gross profit refers to that part of the income of a businessman which is available to him after all payment to the contractually hired factors and other current obligations like taxes and depreciation charges.

- Pure profit is the amount that is available to the entrepreneur for assuming risk inseparable from all business in anticipation of demand.

- Profit is the reward for the enterprise. No doubt profit is associated with the entrepreneur and his function but the economists from time to time have expressed diverse and conflicting views about the nature, origin and role of profit.

- J.M. Keynes expressed the view that profits resulted from the favourable movements of general price level. J. Robinson, E.H. Chamberlin and M. Kalecki have associated profit with the imperfect competition and monopoly.

- Professor J.B. Clark propounded his dynamic theory of profit in 1900. According to him, profit is defined as the difference between price and cost of production of a commodity.

- It should be noted that due to disequilibrium profits arise from unanticipated changes in demand or cost conditions. If the changes could be foreseen in advance, then suitable adjustments can be made according to the anticipated changes so that forces of competition do not drive profits to zero.

- The innovation theory of profit is propounded by Professor J.A. Schumpeter. It has been held by J.A. Schumpeter that the main function of the entrepreneur is to introduce innovation in the economy and profits are the reward which he earns for performing his functions.

- Professor F.H. Knight in his classic work, “Risk, uncertainty and Profit” propounded the theory of profit. According to him, “Profit is the reward not for risk bearing but for uncertainty bearing.” He accepts that the dynamic changes and risk bearing are all important elements of profit. As the future is uncertain and unpredictable, any entrepreneur who probes in the darkness does so only with a view to earn some profit.

- The risk theory of profit was propounded by F.B. Hawley who regarded risk taking ability as the main function of the entrepreneur. Profit is the residual income which the entrepreneur receives because he assumes risks.
• The success of a business as measured in terms of profit depends upon its sales. The volume of sales must be sufficient to cover all costs and allow a satisfactory margin for net income.

• The primary objective of the management in using cost-volume-profit analysis is to determine the target income point and not the break-even point. The break-even point is simply the quantity of sales that achieve net zero income.

• Break-even point is that point where the quantity of sales cause total contribution margin to be exactly equal to total fixed expenses.

• Cost-volume-profit analysis may be used as a profit planning tool and thereby it is useful in decision making. However, effective use of cost-volume-profit analysis for planning purposes requires understanding of certain basic assumptions.

• A new product may simply be either another brand name added to the existing ones or an altogether new product. Pricing a new brand for which there are several substitutes available in the market is not as big a problem as pricing a new product for which close substitutes are not available.

• Depending on whether or not close substitutes are available in pricing a new product, generally two kinds of pricing policies are available: (i) Penetration Price Policy and (ii) Skimming Price Policy.

### 10.8 KEY WORDS

- **Gross profit**: It is the difference between total revenue of entrepreneur and total explicit costs.

- **Break-even point**: It is the production level where total revenues equals total expenses.

### 10.9 SELF ASSESSMENT QUESTIONS AND EXERCISES

**Short Answer Questions**

1. Define profit.
2. Write a short note on the types of profit.
3. What are the criticisms raised against the uncertainty-bearing theory of profit?
4. Mention the major tenets of the risk theory of profit.
5. What is margin safety? How is it calculated?
Long Answer Questions

1. Discuss the Innovation theory propounded by J.A. Schumpeter.
2. Explain the Dynamic theory of profit.
3. Analyse the use of Cost volume profit analysis as a decision-making tool.
4. Describe the pricing policies adopted for launching new products in the market.

10.10 FURTHER READINGS


Websites

https://blogs.economictimes.indiatimes.com/et-commentary/heres-how-indias-widening-income-distribution-can-be-redressed/

UNIT 11 PROFIT AND INVESTMENT ANALYSIS

NOTES

Structure
11.0 Introduction
11.1 Objectives
11.2 Meaning and Measurement of Profit
11.3 Theories of Pricing
  11.3.1 Cost-Plus Pricing
  11.3.2 Multiple Product Pricing
  11.3.3 Pricing in the Life-Cycle of a Product
  11.3.4 Transfer Pricing
  11.3.5 Peak Load Pricing
11.4 Profit Planning and Forecasting
  11.4.1 Profit and Wealth Maximization
  11.4.2 Cost Volume Profit Analysis
11.5 Investment Analysis and Evaluation
  11.5.1 Internal Rate of Return (IRR)
  11.5.2 Net Present Value (NPV)
  11.5.3 Adjusted Present Value (APV)
11.6 Answers to Check Your Progress Questions
11.7 Summary
11.8 Key Words
11.9 Self Assessment Questions and Exercises
11.10 Further Readings

11.0 INTRODUCTION

The first and foremost responsibility of business managers is to achieve the objective of the firm they manage. Therefore, all their managerial decisions, policies, strategy and functions are directed towards the attainment of the objective of the firm. A question that arises here is: What is the objective of business firms? The quick answer to this question is ‘making maximum profit, of course!’ However, economists do not agree with this answer. According to Baumol, a Nobel laureate in Economics and an authority on business economics, ‘It is most frequently assumed in economic analysis that the firm is trying to maximize its total profit. However, there is no reason to believe that all business firms pursue the same objective’. In fact, it is common experience when interviewing executives to find that they will agree to every plausible goal about which they are asked. The say they want to maximize sales and also to maximize profits; that they wish ... to minimize cost; and so on’. In this unit, we will concentrate on the profit maximizing objective of the business firms. We will also explain how firms formulate their strategic pricing policy and determine the price of their product. Finally, we will also discuss the investment related tools which are employed by companies for achieving good returns.
11.1 OBJECTIVES

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

- Describe the meaning and measurement of profit
- Discuss the different theories of pricing
- Explain the profit and wealth maximization objective of firms
- Recall some tools of investment analysis and evaluation

11.2 MEANING AND MEASUREMENT OF PROFIT

Profit means different things to different people. “The word ‘profit’ has different meaning to businessmen, accountants, tax collectors, workers and economists and it is often used in a loose polemical sense that buries its real significance…” In a general sense, ‘profit’ is regarded as income accruing to entrepreneurs, in the same sense as wages accrue to the labour; rent accrues to the owners of rentable assets; and interest accrues to the moneylenders. To a layman, profit means all income that flows to the investors. To an accountant, ‘profit’ means the excess of revenue over all paid-out costs including both manufacturing and overhead expenses, also known as ‘accounting profit’. It is more or less the same as ‘net profit’. For all practical purposes, profit means accounting profit plus non-allowable expenses. Profit figures published by the business firms are profits conforming to accounting concept of profit. Economist’s concept of profit is of ‘pure profit’, also called ‘economic profit’ or ‘just profit’. Pure profit is a return over and above the opportunity cost, i.e., the income which a businessman might expect from the second best alternative use of his resources. The accounting and economic concepts of profit are discussed below in detail.

Accounting Profit vs. Economic Profit

The two important concepts of profit that figure in business decisions are ‘economic profit’ and ‘accounting profit’. From conceptual clarity point of view, it is useful to understand the difference between the two concepts of profit. As already mentioned, in accounting sense, profit is surplus of total revenue over and above all paid-out costs, including both manufacturing and overhead expenses. Accounting profit may be calculated as follows.

\[
\text{Accounting profit} = TR - (W + R + I + M)
\]

where \(TR\) = total revenue, \(W\) = wages and salaries, \(R\) = rent, \(I\) = interest, and \(M\) = cost of materials.

Obviously, while calculating accounting profit, only explicit or book costs, i.e., the costs recorded in the books of accounts, are considered.
It should also be noted that economic or pure profit makes provision also for (a) insurable risks, (b) depreciation, and (c) necessary minimum payment to shareholders to prevent them from withdrawing their capital. Pure profit may thus be defined as ‘a residual left after all contractual costs have been met, including the transfer costs of management, insurable risks, depreciation and payments to shareholders sufficient to maintain investment at its current level’. Thus,

\[ \text{Pure profit} = \text{Total revenue} - (\text{explicit costs} + \text{implicit costs}) \]

Alternatively, pure profit may be defined as follows.

\[ \text{Pure profit} = \text{Accounting profit} - (\text{opportunity cost} + \text{unauthorized payments, e.g., bribes}) \]

Pure profit so defined may not be necessarily positive for a single firm in a single year—it may be negative, since it may not be possible to decide beforehand the best way of using the resources. Besides, in economics, pure profit is considered to be a short-term phenomenon—it does not exist in the long run, especially under perfectly competitive market conditions.

11.3 THEORIES OF PRICING

Up till now we have discussed the conventional theories of price determination under the conditions of different market structures, under the postulate that firms’ objective is to maximize profit. The alternative theories of firm offered by the next generation of economists have built their theories assuming different objectives of firms. A section of economists has built game theory and have shown its application to business decision-making. As noted in previous units, all these theories and strategic models have their own deficiencies and problems in application because of increasing complexity of the business world. The other economists have recognized the complexity of business world and have explained how firms formulate their strategic pricing policy and determine the price of their product.

We begin our discussion by explaining the cost-plus pricing method of product pricing.
11.3.1 Cost-Plus Pricing

Cost-plus pricing is also known as ‘mark-up pricing’, ‘average cost pricing’ and ‘full cost pricing’. The cost-plus pricing is the most common method of pricing used by the manufacturing firms. The general practice under this method is to add a ‘fair’ percentage of profit margin to the average variable cost (AVC). The formula for setting the price is given as

\[ P = AVC + AVC(m) \]  \hspace{1cm} \text{... (11.1)}

where \( AVC \) = average variable cost, and \( m \) = mark-up percentage, and \( AVC(m) = \) gross profit margin (GPM).

The mark-up percentage \( (m) \) is fixed so as to cover average fixed cost \( (AFC) \) and a net profit margin \( (NPM) \). Thus,

\[ AVC(m) = AFC + NPM \]  \hspace{1cm} \text{... (11.2)}

The general procedure followed by the firms for arriving at \( AVC \) and price fixation may be summarized as follows.

The first step in price fixation is to estimate the average variable cost. For this, the firm has to ascertain the volume of its output for a given period of time, usually one accounting or fiscal year. To ascertain the output, the firm uses figures of its ‘planned’ or ‘budgeted’ output or takes into account its normal level of production. If the firm is in a position to compute its optimum level of output or the capacity output, the same is used as standard output in computing the average cost.

The next step is to compute the total variable cost (TVC) of the ‘standard output.’ The TVC includes direct cost, i.e., the cost of labour and raw material, and other variable costs e.g., electricity and transportation cost, etc. These costs added together give the total variable cost. The ‘Average Variable Cost’ (AVC) is then obtained by dividing the total variable cost (TVC) by the ‘standard output’ (\( Q \)), i.e.,

\[ AVC = \frac{\text{TVC}}{Q} \]

After \( AVC \) is obtained, a ‘mark-up’ in the form of some percentage of \( AVC \) is added to it as profit margin and the price is fixed. While determining the mark-up, firms always take into account ‘what the market will bear’ and the degree of competition in the market.

11.3.2 Multiple Product Pricing

Most microeconomic models of price determination are based on the assumption that a firm produces a single, homogeneous product. In actual practice, however, production of a single homogeneous product by a firm is an exception rather than a rule. Almost all firms have more than one product in their line of production. Even the most specialized firms produce a commodity in multiple models, styles
and sizes, each so much differentiated from the other that every model or size of
the product may be considered a different product. For example, the various
models of refrigerators, TV sets, cell phones, computers and car models etc.
produced by the same company may be treated as different products for at least
pricing purpose. The various models are so differentiated that consumers view
them as different products and, in some cases, as close substitutes for each other.
It is for this reason that each model or product has different \( AR \) and \( MR \) curves
and that one product of the firm competes against the other product. The pricing
under these conditions is known as multi-product pricing or product-line pricing.

The major problem in pricing multiple products is that each product has a
separate demand curve. But, since all the products are produced under one
establishment by interchangeable production facilities, they have only one joint
and one inseparable marginal cost curve (MC). That is, while revenue curves, \( AR \)
and \( MR \), are separate for each product, cost curves, \( AC \) and \( MC \), are inseparable.
Therefore, the marginal rule of pricing cannot be applied straightaway to fix the
price of each product separately. The problem, however, has been provided with
a solution by E.W. Clemens. The solution is similar to the technique employed to
illustrate third degree price discrimination under profit maximization assumption.
As a discriminating monopoly tries to maximize its revenue in all its markets, so
does a multi-product firm in respect of each of its products.

11.3.3 Pricing in the Life-Cycle of a Product

The life-cycle of a product is generally divided into five stages: (i) Introduction or
initial stage, (ii) Growth, (iii) Maturity, (iv) Saturation, and (v) Decline. The
introduction phase is the period taken to introduce the product to the market.
The total sale during this period is limited to the quantity put on the market for trial
with considerable advertisement. The sales during this period remain almost
constant. Growth is the stage, after a successful trial, during which the product
gains popularity among the consumers and sales increase at an increasing rate as
a result of cumulative effect of advertisement over the initial stage. Maturity is the
stage in which sales continue to increase but at a lower rate and the total sale
eventually becomes constant. During the saturation period the total sale
saturates—there is no considerable increase or decrease in the sales. After the
saturation stage, comes the stage of decline in which total sales begin to decline
for such reasons as (i) increase in the availability of substitutes, and (ii) the loss of
distinctiveness of the product.

The pricing strategy varies from stage to stage over the life-cycle of a
product, depending on the market conditions. From the pricing strategy point of
view, growth and maturity stages may be treated likewise.

11.3.4 Transfer Pricing

Large size firms often divide their production process into different product divisions
or their subsidiaries. Also, growing firms add new divisions or departments to the
existing ones. The firms then transfer some of their production activities to other divisions. The goods and services produced by the new divisions are used by the parent organization. In other words, the parent division buys the product of its subsidiaries. Such firms face the problem of determining an appropriate price for the product transferred from one division or subsidiary to the parent body. This problem becomes much more difficult when each division has a separate profit function to maximize. Pricing of intra-firm 'transfer product' is referred to as 'transfer pricing'. One of the most systematic treatments of the transfer pricing technique has been provided by Hirshleifer. We will discuss here briefly his technique of transfer pricing.

To begin with, let us suppose that a refrigerator company established a decade ago used to produce and sell refrigerators fitted with compressors bought from a compressor manufacturing company. Now the refrigerator company decides to set up its own subsidiary to manufacture compressors. Now the problem for the company is how to price the product of its subsidiary under the following conditions.

(i) Both parent and subsidiary companies have their own profit functions to maximize, and
(ii) The refrigerator company has the option of using all the compressors produced by its subsidiary and/or to sell the compressors in a competitive market and its demand is given by a straight horizontal line.

Given these conditions of the model, transfer pricing is discussed under two conditions.

(i) The parent company uses the entire output of its subsidiary and there is no external market for the compressors, and
(ii) There does exist a competitive market for the compressor and refrigerator company sells also in the open market.

### 11.3.5 Peak Load Pricing

There are certain non-storable products, e.g., electricity, telephones, transport and security services, etc., which are demanded in varying measures during the day as well as night. For example, consumption of electricity reaches its peak in daytime. It is called 'peak-load' time. It reaches its bottom in the night. This is called 'off-peak' time. Electricity consumption peaks in daytime because all business establishments, offices and factories come into operation. Electricity consumption decreases during nights because most business establishments are closed and household consumption falls to its basic minimum. In Delhi, demand for electricity peaks during summers due to use of ACs and coolers, and it declines to its minimum level during winters. Similarly, consumption of telephone services is at its peak at daytime and at its bottom at nights. Another example of 'peak' and 'off-peak' demand is of railway and air services. During festivals, summer holidays, 'Pooja' vacations, etc., the demand for railway and air travel services rises to its peak.
A technical feature of such products is that they cannot be stored. Therefore, their production has to be increased in order to meet the ‘peak-load’ demand and reduced to ‘off-peak’ level when demand decreases. Had they been storable, the excess production in ‘off-peak’ period could be stored and supplied during the ‘peak-load’ period. But this cannot be done. Besides, given the installed capacity, their production can be increased but at an increasing marginal cost (MC).

Check Your Progress
1. Define pure profit.
2. Mention some of the other names for cost-plus pricing.
3. Why does the total sales begin to decline in the last stage of the life-cycle of a product?

11.4 PROFIT PLANNING AND FORECASTING

Profit forecasting is essential for profit planning in any organization. The projective of earnings in the future after taking in to account different factors like costing policies, pricing policies, depreciation policies, etc., is known as profit forecasting. This profit estimation is advantageous to the firm for making different projections related to costs, profits and sales volume.

Some researches have suggested different approaches to profit forecasting. As per Joel Deal, profit forecasting can be done through the following three approaches: (1) break-even analysis; (2) spot projection which includes projection for the specified future period; and environmental analysis which is analysis related to the factors in the economic environment which are beyond the scope of the enterprises.

Several elements of business have an impact on the profits of the business firms. These elements are known to have some type of pattern or the other. Elements like rate of output, material costs, wages, efficiency work in a pattern at the company, industry and national levels. They are interdependent and have a bearing on each other. If these patterns are studied in a proper manner, these can be used to improving the chances of the firms to earn profits and in fact help them find the relation between specific profit making variables. Profit planning and forecasting aims to decode this functional relationship so as to work for the profit maximization of the firms.

11.4.1 Profit and Wealth Maximization

The conventional economic theory assumes profit maximization as the only objective of business firms—profit measured as TR–TC. Profit maximization as the objective of business firms has a long history in economic literature. It forms the basis of conventional price theory. Profit maximization is regarded as the most
reasonable and analytically the most ‘productive’ business objective. The strength
of this assumption lies in the fact that this assumption ‘has never been unambiguously
disproved’. Besides, profit maximization assumption has a greater predictive power. It
helps in predicting the behaviour of business firms in the real world and also the
behaviour of price and output under different market conditions. No other
hypothesis explains and predicts the behaviour of firms better than the profit
maximization assumption. Nevertheless, the profit maximization has been
questioned strongly by some modern economists. This created a controversy on
objectives of business firms. The controversy has been discussed below in detail.

Controversy on Profit Maximization: Theory vs. Practice

As discussed above, the conventional theory of firm assumes profit maximization
as the sole objective of the business firms. Some modern economists, however,
refute the profit maximization assumption because, in their opinion, it is practically
non-achievable. Their own findings reveal that business firms, especially big
corporations, pursue several other objectives, rather than profit maximization.
However, some modern economists have strongly defended the profit maximization
objective. This has created a controversy on the profit maximization objective of
the business firms. In this section, we discuss briefly the arguments against and for
profit maximization objective.

Arguments against Profit-Maximization Objective

(i) The first argument against the profit maximization objective is based on
the dichotomy between the ownership and management of business firms.
It is argued that, in modern times, due to rapid growth of large business
corporations, management of business firms has got separated from the
ownership. The separation of management from ownership gives managers
an opportunity and also the discretion to set firm’s goals other than profit
maximization. The researches conducted by the economists reveal that, in
practice, business managers pursue such objectives as: (a) maximization
of sales revenue, (b) maximization of the value of the firm, i.e., the net
worth of the firm, (c) maximization of managerial utility function,
(d) maximization of firm’s growth rate, (e) making a target profit,
(f) retaining and increasing market share, and so on.

(ii) Another argument against profit maximization objective is that traditional
theory of firm assumes managers to have full and perfect knowledge of
market conditions and of the possible future development in business
environment of firm. The firm is thus supposed to be fully aware of its demand
and cost conditions in both short and long runs. Briefly speaking under
profit maximization objective, a complete certainty about the market
conditions is assumed. Some modern economists question the validity of
this assumption. They argue that the firms do not possess the perfect
knowledge of their costs, revenue and future business environment. They operate in the world of uncertainty. Most price and output decisions are based on probabilities.

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

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

The Defence of Profit Maximization

The arguments against profit-maximization objectives have been strongly rejected by other economists. They argued strongly that pricing theory does have relevance to the actual pricing policy of the business firms. A section of economists has strongly defended the profit maximization objective and ‘marginal principle’ of pricing and output decisions.

The conventional economic theorists defend the profit maximization hypothesis on the following grounds.

1. **Profit is indispensable for firm’s survival.** The survival of all the profit-oriented firms in the long run depends on their ability to make a reasonable profit depending on the business conditions and the level of competition. What profit is reasonable may be a matter of opinion. But, making profit is a necessary condition for the survival of the firm. Once the firms begin to make profit, they try to maximize it.
2. **Achieving other objectives depends on firm’s ability to make profit.**

Many other objectives of business firms have been cited in economic literature, e.g., maximization of managerial utility function, maximization of long-run growth, maximization of sales revenue, satisfying all the concerned parties, increasing and retaining market share, etc. The achievement of such alternative objectives depends wholly or at least partly on the primary objective of making profit.

3. **Evidence against profit maximization objective is not conclusive.**

Profit maximization is a time-honoured objective of business firms. Although this objective has been questioned by many researchers, some economists have argued that the evidence against it is not conclusive or unambiguous.

4. **Profit maximization objective has a greater predicting power.**

Compared to other business objectives, profit maximization objective has been found to provide a much more powerful basis for predicting certain aspects of firms’ behaviour. As Friedman has argued, the validity of the profit maximization objective cannot be judged by *a priori* logic or by asking business executives, as some economists have done. In his opinion, ultimate test of its validity lies in its ability to predict the business behaviour and the business trends.

5. **Profit is a more reliable measure of a firm’s efficiency.**

Though not perfect, profit is the most quick and reliable measure of the efficiency of a firm. It is also the source of internal finance. Profit as a source of internal finance assumes a much greater significance when financial market is highly volatile. The recent trend shows a growing dependence on the internal finance in the industrially advanced countries. In fact, *in developed countries, internal sources of finance contribute more than three-fourths of the total finance.*

6. **Finally,** according to Milton Friedman, whatever one may say about firms’ motivations, if one judges their motivations by their managerial acts, profit maximization appears to be a *more valid business objective.*

11.4.2 Cost Volume Profit Analysis

The theory of cost guides both the analysts and the firms to find the nature of change in cost of production with change in output. Also, it provides conditions for determining cost minimizing output. However, despite its high analytical value, cost theory does not provide answer to certain basic questions pertaining to business decisions of new firms. Some such questions are: (i) given the product price and cost, what is the minimum output to recover the total cost? (ii) what is the output at which total cost (TC) and total revenue (TR) break-even? and (iii) if anticipated TR and TC functions produce non-linear curves, what output would maximize the
profit? The economists have developed a new technique, in addition to cost analysis, that can be used to find answer to these questions. The technique is known as **Break-Even analysis** known also as **Profit Contribution Analysis** and **Cost-Volume Profit Analysis**.

### Meaning of Break-even Analysis

The break-even analysis is an important analytical technique used to study the relationship between the total costs, total revenue and total profit and loss over the whole range of stipulated output. The break-even analysis is a technique of having a preview of profit prospects and a tool of profit planning. It integrates the cost and revenue estimates to ascertain the profits and losses associated with different levels of output.

The relationship between cost and output and between price and output may be linear or non-linear in nature.

### 11.5 INVESTMENT ANALYSIS AND EVALUATION

Investment is an activity of spending resources (money, labour and time) on creating assets that can generate income over a long period of time or which enhances the returns on the existing assets. In a broader sense of the term, investments that generate returns over a number of years can be classified under following categories.

1. **Investment in Financial Assets** including bank deposits, deposits with companies, contribution to provident fund (in excess of compulsory deduction), shares and debentures, government bonds and treasury bills, purchase of NSC, buying units, personal lending, etc.
2. **Investment in Physical Assets** including purchase of land, building, machinery, plants, etc.
3. **Investment in Human Capital** including expenditure on skill formation through education and training that increases productivity and earning capacity of a person.
4. **Miscellaneous Investment** including expenditure on replacement of depreciated and obsolete machinery, product diversification, R&D, installation of safety measures for employees, pollution control for public health and safety, and meeting legal requirements.

**Investment Decisions under Certainty**

In this section, we discuss the methods of investment decisions under the condition of certainty. In the context of investment, **certainty** means that investors have complete knowledge about the market conditions, especially the investment opportunities, cost of capital and the expected returns on the investment.
Specifically, certainty means that there is only one outcome of investment and it is known and assured to the investor. For example, money put in fixed deposit in a bank yields an assured return with full certainty.

In this section, we discuss the various criteria for measuring a certain return from a planned investment. Of the several criteria applied for evaluating the profitability of the various kinds of projects, the three most commonly used criteria under certainty are following:

(i) Pay-back (or pay-out) period;
(ii) Net discounted present value, and
(iii) Internal rate of return or marginal efficiency of capital.

These criteria are equally applicable to a variety of investment decisions regarding new investments and those pertaining to replacement, scrapping, and widening or deeping of capital. Incidentally, from analysis point of view, there is no structural difference between decisions on new investment and those on replacement.

11.5.1 Internal Rate of Return (IRR)

The Internal Rate of Return (IRR) is also called Marginal Efficiency of Investment (MEI), Internal Rate of Project (IRP) and Break-even Rate (BER). The concept of IRR can be illustrated with an example. For example, if a one-year project costing ₹100 million yields ₹120 million at the end the year, then its internal rate of return (r) can be obtained as follows.

\[
\frac{120 \text{ million}}{(1 + r)} = \frac{100 \text{ million}}{(1 + r)100} = 120
\]

and

\[ r = 0.20 \]

The IRR of this project is 0.20 or 20 per cent. No other value of r can equate the NPV of the project with its cost.

The IRR or MEI is defined as ‘the rate of interest or return which renders the discounted present value of its expected future marginal yields exactly equal to the investment cost of project’. In other words, ‘IRR is the rate of return (r) at which the discounted present value of receipts and expenditures are equal’. The IRR of a project yielding a stream of returns over n years and involving different investment costs can be obtained by using the formula given in Eq. (11.3).

\[
\sum_{j=1}^{n} \frac{R_j}{(1 + r)^j} = \sum_{j=1}^{n} \frac{C_j}{(1 + r)^j}
\]

or

\[
\sum_{j=1}^{n} \frac{R_j}{(1 + r)^j} - \sum_{j=1}^{n} \frac{C_j}{(1 + r)^j} = 0
\]
The Concept of Present Value: The Time Value of Money. The concept of present value of money is very well reflected in the proverb 'a bird in the hand is worth two in the bush'. In general, money received today is valued more than money receivable tomorrow. Cash in hand is valued more because it gives (i) liquidity and (ii) an opportunity to invest it and earn return (interest) on it. This is called the time value of money. The concept of the time value of money is very often applied to investment decisions. Generally, there is a time-lag between investment and its returns. When an investment is made today, it begins to yield returns at some future date. The time gap between the investment and the first return from the investment is called 'time lag'. During the time lag, the investor loses interest on the expected incomes. This implies that a rupee received today is worth more than a rupee receivable at some future date. Or conversely, a rupee expected one year hence is worth less than a rupee today. In the context of the time value of money, the present value of a future income is lower than its value if received today.

The concept of present value of money can be explained through an example. Suppose that a sum of \( \text{`100} \) held in cash today is deposited in a bank at 10 per cent rate of interest. After one year, \( \text{`100} \) today will increase to \( \text{`110} \). The amount (principal + interest) is worked out as follows.

\[
\text{Amount} = 100 + 100 \times \left( \frac{10}{100} \right) = 100 + 10 = 110
\]

It follows that \( \text{`110} \) expected one year hence is worth only \( \text{`100} \) today. This means that \( \text{`100} \) is the present value of \( \text{`110} \) to be earned after a period of one year at the interest rate of 10 per cent. The present value (PV) of \( \text{`110} \) can be obtained as follows.

\[
\text{PV of } \text{`110} = \frac{110}{(1 + 0.1)} = 100
\]

The present value of a future income may thus be defined as its value discounted at the current rate of interest. Alternatively, the present value of an amount expected at a future date is the amount of money which must be invested today to get that amount in future.

The Formula for Computing Present Value. Suppose that an amount \( X_0 \) is invested for a period of one year at a compounding interest rate. At the end of the year, the total receipt, say \( X_1 \) can be expressed as
\[ X_1 = (X_0 + rX_0) = X_0(1 + r) \] …(11.5)

Equation (11.5) shows that \( X_0 \) increases at the rate of \( (1 + r) \) to take the value \( X_1 \) after one year. It implies that if \( X_0 \) is discounted at the same rate of interest, it gives its present value (PV). The formula for computing the present value is given below.

\[
PV \text{ of } X_{n-1} = \frac{X_n}{(1 + r)} = X_0 \left( \frac{1}{1 + r} \right) \] \quad ...(11.6)

In Eq. (11.6), \( 1/(1 + r) \) is the discount rate for one year. Given the rate of interest (i.e., the numerical value for \( r \)), any income receivable after one year can be discounted to its present value. For example, the present value of an income of \`500 expected after one year at 10 per cent interest per annum (where \( r = 0.10 \)), can be calculated as

\[
PV = 500 \left( \frac{1}{1 + 0.10} \right) = 454.55
\]

It means that, at 10 per cent interest rate, the present value of `500 expected after one year is 454.55. Given the formula in Eq. (11.6), the discount rate (\( d_1 \)) for an income receivable after 2 years will be \( 1/(1 + r)^2 \), and for an income receivable after 3 years, \( d_2 = 1/(1 + r)^3 \) and so on. The formula for discount rate for the \( n \)th year is given as

\[
d_n = \frac{1}{(1 + r)^n} \] \quad ...(11.7)

The present value (PV) formula of an amount receivable in the \( n \)th year is given as

\[
PV = X_n \left[ \frac{1}{(1 + r)^n} \right] \] \quad ...(11.8)

or

\[
PV = \frac{X_n}{(1 + r)^n} \]

**Present Value of an Income Stream.** The formula for calculating the total present value (TPV) of a stream of annual return (\( R \)) over \( n \) year is given as

\[
TPV = \frac{R}{(1 + r)} + \frac{R}{(1 + r)^2} + \frac{R}{(1 + r)^3} + \cdots + \frac{R}{(1 + r)^n} \] \quad ...(11.9a)

or

\[
TPV = \sum_{i=1}^{n} \frac{R}{(1 + r)^i} \]

or

\[
TPV = \frac{R}{(1 + r)^{n+1}} \] \quad ...(11.9b)

**Net Present Value and Investment Decision.** Having noted the concept of present value (PV) and the method of calculating PV of a future income, let us now see how investment decisions are taken on the basis of present value. In fact,
Profit and Investment Analysis

NOTES

present value (PV) adjusted for the cost of investment provides the basis of investment decisions. The PV adjusted for its cost is called ‘Net Present Value’. The investment decision—accepting or rejecting a project—is taken on the basis of net present value. The net present value (NPV) may be defined as the difference between the present value (PV) of an income stream and the cost of investment (C), i.e.,

\[ NV = PV - C \]

or

\[ NV = \sum_{j=1}^{n} \frac{R_j}{(1 + r)^j} - C \]  \hspace{1cm} (11.10)

where \( C \) is the total cost of investment without any recurring expenditure.

The investment decision rules can be specified as follows.

(i) If \( NPV > 0 \), the project is acceptable;

(ii) If \( NPV = 0 \), the project is accepted or rejected on non-economic considerations;

(iii) If \( NPV < 0 \), the project is rejected.

If investment is a recurring expenditure, the total present cost (TPC) for \( n \) years can be calculated in the same manner as present value of an income stream is calculated, i.e.,

\[ TPC = \sum_{j=1}^{n} \frac{C_j}{(1 + r)^j} \]  \hspace{1cm} (11.11)

And then,

\[ NPV = \sum_{j=1}^{n} \frac{R_j - C_j}{(1 + r)^j} \]

\[ = \sum_{j=1}^{n} \frac{R_j - C_j}{(1 + r)^j} \]  \hspace{1cm} (11.12)

The investment decision rule in this case is the same as given above. If the NPV is positive (i.e., \( NPV > 0 \)), the project is profitable and acceptable. The firm can borrow any amount at the existing interest rate (\( r \)) and invest in it. When a choice between two projects has to be made, the one with higher NPV would be chosen.

11.5.3 Adjusted Present Value (APV)

The Adjusted Present Value can be defined as the Net Present Value (NPV) or investment adjusted for the interest and tax advantages of leveraging debt provided that equity is the only source of financing.

Amongst the many ways in which a company finances its projects, a company bring into use the unleveraged shareholders’ equity alone as an investment source for the projects. In such situations, the unleveraged cash flows from the shareholder’s equity is utilized to repay associated debts. The company in such circumstances becomes entitled to significant tax deductions on the interest
component of these payments. The tax deductions received here is advantageous for the companies to the limit of project’s ultimate profitability, as they increase the project’s bottom line. Adjusted present values is used in such situations to analyse a project’s profitability. This measure reflects the project or investment’s NPV adjusted for the tax benefits from interest obligations on outstanding debts associated with the project or investment.

Any project which is using the un-leveraged equity cash flows as a source of finance and is receiving tax deductions for it can be assessed for profitability using APV measures.

Check Your Progress

4. List some of the objectives of business firms which are dependent wholly or at least partly on the primary objective of making profit.

5. What is the Net Present Value (NPV)?

11.6 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Pure profit is a return over and above the opportunity cost, i.e., the income which a businessman might expect from the second best alternative use of his resources.

2. Cost-plus pricing is also known as ‘mark-up pricing’, ‘average cost pricing’ and ‘full cost pricing’.

3. The total sales begin to decline in the last stage of the life-cycle of a product for such reasons as (i) increase in the availability of substitutes, and (ii) the loss of distinctiveness of the product.

4. Some of the objectives of business firms which are dependent wholly or at least partly on the primary objective of making profit are maximization of managerial utility function, maximization.

5. The net present value (NPV) is defined as the difference between the present value (PV) of an income stream and the cost of investment.

11.7 SUMMARY

- Pure profit is a return over and above the opportunity cost, i.e., the income which a businessman might expect from the second best alternative use of his resources.

- The accounting profit may be calculated as follows: Accounting profit = TR - (W + R + I + M), while calculating accounting profit, only explicit or book
Costs, i.e., the costs recorded in the book of accounts are considered. Economic profit also takes into account the implicit and imputed costs.

- Cost-plus pricing is also known as ‘mark-up pricing’, ‘average cost pricing’ and ‘full cost pricing’. The cost-plus pricing is the most common method of pricing used by the manufacturing firms. The general practice under this method is to add a ‘fair’ percentage of profit margin to the average variable cost (AVC).

- Most microeconomic models of price determination are based on the assumption that a firm produces a single, homogeneous product. In actual practice, however, production of a single homogeneous product by a firm is an exception rather than a rule. It is for this reason that each model or product has different $AR$ and $MR$ curves and that one product of the firm competes against the other product. The pricing under these conditions is known as multi-product pricing or product-line pricing.

- The life-cycle of a product is generally divided into five stages: (i) Introduction or initial stage, (ii) Growth, (iii) Maturity, (iv) Saturation, and (v) Decline. The pricing strategy varies from stage to stage over the life-cycle of a product, depending on the market conditions.

- Pricing of intra-firm ‘transfer product’ is referred to as ‘transfer pricing’.

- There are certain non-storable products, e.g., electricity, telephones, transport and security services, etc., which are demanded in varying measures during the day as well as night. The pricing of such products is known as peak load pricing.

- Profit forecasting means projection of future earnings after considering all the factors affecting the size of business profits, such as firm’s pricing policies, costing policies, depreciation policy, and so on.

- Profit maximization is regarded as the most reasonable and analytically the most ‘productive’ business objective.

- The break-even analysis is an important analytical technique used to study the relationship between the total costs, total revenue and total profit and loss over the whole range of stipulated output.

- The IRR or MEI is defined as the rate of interest or return which renders the discounted present value of its expected future marginal yields exactly equal to the investment cost of project.

- The concept of present value of an amount expected at a future date is the amount of money which must be invested today to get that amount in future.

- Adjusted present value is the net present value or investment adjusted for the interest and tax advantages of leveraging debt provided that equity is the only source of financing.
11.8 KEY WORDS

- **Pure Profit**: It is a return over and above the opportunity cost.
- **Cost-Plus Pricing**: It is a most common method of pricing used by the manufacturing firms which includes adding a fair percentage of profit margin to the average variable cost.
- **Break-Even Analysis**: It refers to the technique which integrates the cost and revenue estimates to ascertain the profits and losses associated with different levels of output.
- **Investment**: It is an activity of spending resources on creating assets that can generate income over a long period of time or which enhances the returns on the existing assets.

11.9 SELF ASSESSMENT QUESTIONS AND EXERCISES

**Short Answer Questions**

1. Differentiate between accounting and economic profit.
2. What is the general procedure followed by the firms for arriving at AVC and price fixation?
3. List the five stages of a life-cycle of a product.
4. What is cost volume profit analysis?

**Long Answer Questions**

1. Describe the different theory of pricing.
2. Discuss the arguments for and against profit maximization objective.
3. Explain the tools of NPV, APV and IRR.

11.10 FURTHER READINGS


NOTES

Websites

https://blogs.economictimes.indiatimes.com/et-commentary/heres-how-indias-widening-income-distribution-can-be-redressed/

UNIT 12 MACRO-ECONOMIC FACTORS

<table>
<thead>
<tr>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.0 Introduction</td>
</tr>
<tr>
<td>12.1 Objectives</td>
</tr>
<tr>
<td>12.2 Meaning and Nature of Macro-Economic Factors</td>
</tr>
<tr>
<td>12.2.1 Importance of Macro-Economic Factors</td>
</tr>
<tr>
<td>12.3 Economic Growth and Development</td>
</tr>
<tr>
<td>12.4 Business Cycle: Phases and Business Decisions</td>
</tr>
<tr>
<td>12.5 Inflation</td>
</tr>
<tr>
<td>12.5.1 Factors Causing Inflation</td>
</tr>
<tr>
<td>12.5.2 Factors Causing Deflation</td>
</tr>
<tr>
<td>12.5.3 Control Measures</td>
</tr>
<tr>
<td>12.6 Balance of Payment Trend and Its Implications on Managerial Decisions</td>
</tr>
<tr>
<td>12.7 Answers to Check Your Progress Questions</td>
</tr>
<tr>
<td>12.8 Summary</td>
</tr>
<tr>
<td>12.9 Key Words</td>
</tr>
<tr>
<td>12.10 Self Assessment Questions and Exercises</td>
</tr>
<tr>
<td>12.11 Further Readings</td>
</tr>
</tbody>
</table>

12.0 INTRODUCTION

The subject matter of managerial microeconomics consists of theories of demand and supply, theory of production, cost-output relationships, market mechanism, theory of price determination under different kinds of market structure, pricing principles and prices, and capital budgeting and investment decisions. Managerial decisions on these issues, except on consumers’ demand, fall within the purview of the decision powers of the managers. Most micro level managerial decisions are made generally with short-run perspective assuming that the business environment of the economy would continue to remain the same and changes therein, if any, are predictable and can be taken into account while making business decisions.

However, managerial decision-making, especially in case of large and medium scale business corporations, does not remain always confined to short-run and internal microeconomic issues. In case of large and medium scale business firms, managerial decisions are taken with a long-run perspective of economic conditions and business environment of the country.

Changing economic conditions and business environment have significant impact on the business decisions. As a result, business decisions on internal matters are likely to be affected by the deteriorating economic environment of the country. Macroeconomics is the study of the factors that determine the economic
environment of the country and business prospects. Therefore, macroeconomics constitutes a subject matter of managerial economics.

In this unit, you will study about the meaning, nature and importance of macro-economic factors, economic growth and development, business cycle, factors causing inflation and deflation and finally, balance of payment trend and its implications on managerial decisions.

12.1 OBJECTIVES

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

- Discuss the meaning, nature and importance of macro-economic factors
- Analyse economic growth and development
- Explain the phases of business cycles
- Identify the factors causing inflation and deflation
- Describe balance of payment trend and its implication on managerial decisions

12.2 MEANING AND NATURE OF MACRO-ECONOMIC FACTORS

Macroeconomics is essentially the study of the behaviour and performance of the economy as a whole. More importantly, it studies the relationship and interaction between the ‘factors and forces’ that determine the level and growth of national output and employment, general price level, and the balance of payments positions of an economy. This definition too should be treated only as a working definition of macroeconomics.

To comprehend better the subject matter of macroeconomics, let us look at the kinds of questions that macroeconomics seeks to answer.

- What determines the levels of economic activities, national income, total output, the general price level, and the overall employment in a country?
- How is the equilibrium level of national income determined?
- How product and monetary markets interact to bring about a general equilibrium in the economy?
- What causes fluctuations in the national output and employment?
- What determines the general level of prices in a country?
- What determines the level of foreign trade and trade balance?
- What causes disequilibrium in the balance of payments of a country?
How do the monetary and fiscal policies of the government affect the economy?

What economic policies can steer the economy on the path of growth?

These are some major theoretical questions that macroeconomics seeks to answer.

Besides, there are various areas and kinds of business decisions which have a long-run perspective and long-run financial implications, e.g., business decisions related to setting up a new production unit like Reliance planning to set up a new power plant in Uttar Pradesh; innovation and introduction of a new product like Tatas planning to launch a low-price car; adopting a new production technology, setting up a new industry or retail chains or shopping malls; going for franchise of a multinational company; and so on. All such business decisions have to be taken in view of the long-term perspective of business environment of the country. Business environment of a country is constituted of economic, social and political environment of the country. Therefore, business decision-makers have to take into account the economic, social and political environment of the country. For example, establishment of Special Economic Zones (SEZs) and the problem of land acquisition for SEZs has created both social and political problems—social problem because farmers lose the source of their livelihood and this often leads to their resistance and loss of human lives as it happened in Singur and Nandigram of West Bengal; political problem because this issue has to be resolved at the political level. The social and political issues fall outside the purview of the macroeconomic aspects of managerial decisions. However, macroeconomic aspects of business environment are very much a part of managerial economics. The need for the study of economic aspects of business environment extends the scope of managerial economics to include also the study of macroeconomics.

Nature of Macroeconomics

The nature of a subject is determined on the basis of whether the subject has a purely positive (theoretical) or normative (policy) orientation or both. A positive science has theoretical orientation in the sense that it aims at formulation of theories and laws. And a science having normative orientation aims at setting norms for finding solution to practical problems and provides policy guidelines. Macroeconomics has both positive (theoretical) and policy orientations.

Macroeconomics as a positive science has theoretical orientation. As a positive science, it uses macroeconomic models to explain the behaviour of macroeconomic variables—national income, aggregate consumption, savings and investment, price level, and employment. This results in formulation of macroeconomic theories. An important aspect of macroeconomic theories is that they provide a framework and analytical models to analyze the macroeconomic phenomena. The theories of income determination, consumption, investment,
employment, price-level determination, product and money market equilibrium, exchange rate and balance-of-payments constitute the main body of the macroeconomic theories.

Macroeconomics as a policy or normative science provides guidelines for the formulation of economic policies. And also, it studies the working of macroeconomic policies of the government and investigates into the effects of public policies on the economy. It provides a framework for policy evaluation and guidelines for the formulation of appropriate macroeconomic policies. In fact, the ultimate aim of macroeconomic studies is to formulate macroeconomic policies for macroeconomic management. It is perhaps for this reason that some economists consider macroeconomics as a purely policy science. According to some economists, “Macroeconomics is first and foremost a policy science”. In fact, it analyzes the various macro aspects of the economy with the aim of providing the base and logic for policy formulation with the objective of controlling and guiding the economy on the path of stable growth of output and employment. Thus, macroeconomics as a policy science provides a framework and instruments for managing the economy and guiding it on the path of growth and stability.

12.2.1 Importance of Macro-Economic Factors

The importance of macroeconomics can be viewed from two angles: (i) how it contributes to government’s economic policy formulation for the management of the economy, and (ii) how it guides business managers for taking appropriate business decisions and business planning. It has been proved over time that macroeconomics has contributed a great deal to both these areas. In this section, we discuss briefly how macroeconomics contributes to these two areas.

Application of Macroeconomic Management

In modern times, most countries are confronted with three major macroeconomic problems: (i) problems in achieving a sustainable growth rate, (ii) problem of unemployment, and (iii) problem of inflation. Given these problems, the basic economic functions of the government are (i) ensuring a sustainable economic growth rate; (ii) ensuring full employment, (iii) ensuring price stability. Macroeconomic problems of the country have to be resolved effectively as the economic fate of the country and its people depends on its economic strength. Macroeconomics contributes a great deal in formulating economic policies to achieve these ends.

Both developed and developing countries are constantly confronted with some or the other kind of macroeconomic problems, e.g., recession and depression, stagflation, balance-of-payment deficits, outflow of capital, mounting debt burden or a country falling into debt trap, and so on. These problems have to be solved if eventual economic collapse like the Great Depression
of 1930s has to be averted. Even if economic catastrophes of this magnitude do not take place, macroeconomic problems like recession, unemployment, inflation and increasing external debt burden need to be solved because they have serious socio-political implications for the country in general and the government in particular.

12.3 ECONOMIC GROWTH AND DEVELOPMENT

The overall economic trend in a country is the most important element of its business environment. A growing economy with a considerably high and sustained growth rate in its GDP over a long period of time provides a promising business prospect and builds business confidence.

Meaning of Economic Growth

In general, economic growth means continuous increase in GDP or GNP over a period of time at a reasonably high rate. In real sense of the term, however, economic growth means a sustained increase in per capita national output or net national product over a long period of time. It implies that the rate of increase in total output must be greater than the rate of population growth. Economic growth implies a considerable and sustained increase in per capita income with or without increase in population.

Another qualification of economic growth is that the national output should be composed of such goods and services which satisfy the maximum wants of the maximum number of people. Besides, for economic growth to be genuine, the increase in output must be sustained over a long period of time. A short-run increase followed by a similar decrease in the output does not mean economic growth. Also seasonal, occasional and cyclical increases in output do not satisfy the conditions of sustained economic growth.

Determinants of Economic Growth

There are five most important determinants of economic growth of a country, viz.,

(i) Human resources and its quality,
(ii) Natural resources of the country,
(iii) The rate of capital formation,
(iv) Technological development, and
(v) Political and Social environment.

The first four factors are considered as the “four wheels” of economic growth. The social and political factors make the fifth factor. Let us now see how these factors contribute to economic growth of a country.
1. Human Resource and its Quality

Human resource of a country is the most crucial factor in its economic growth. Human resource is comprised of the available labour force and its quality. Quality of labour force depends on the level of its education, training, skills, and its inventive and innovative abilities. Quantity and quality of manpower are both equally important. However, an excess supply of unskilled labour force, as is the case in most LDCs including India, is of little consequence. On the other hand, scarcity of skilled labour in the US is proving a serious constraint to its economic growth. The labour force along with its skill is the source of all goods and services.

Apart from quantity and quality, an appropriate combination of labour with different skills is also very important in making optimum use of human resources.

2. Natural Resources

Natural resources of a country include the area of usable land, and resources on the land surface and underground. Land surface resources include sources of natural water (rivers and lakes), forests, landscape, etc. Underground resources include oil and natural gas and minerals. Favourable climatic and environmental conditions add to the natural resources endowments of a country. The countries with rich natural resource endowments have a much larger growth potential than those lacking natural resources.

However, natural resources are passive factors of growth. The exploitation and use of natural resources depends on the quality of manpower, availability of capital and technology. The countries endowed with rich natural resources and a highly skilled and motivated manpower can do miracles in economic growth.

3. Capital Formation

Capital is defined as man-made means of production. In practical sense of the term, capital includes machinery, plant and building, means of transport and communication, electricity plants, and social overheads like roads, railways, schools, colleges, hospitals, etc. Creating or acquiring man-made means of production is known as capital formation or capital accumulation. Capital formation enhances the availability of capital per worker. A high capital/labour ratio enhances the productivity of labour. In other words, with a high rate of capital formation, a larger quantity of goods and services are produced per unit of time. This means a high growth rate.

Capital formation requires saving men and material resources from their use in consumer goods and transforming them into producer goods. In economic jargon, capital formation means sacrificing current consumption and saving incomes to be invested in capital goods (machinery, plant, building and equipment etc.). In general, the countries with a high rate of saving and investment have a higher rate of economic growth. Also, as the rate of saving and investment increases, the rate of economic growth increases too.
4. Technological Development

Technology used in production is the fourth vital determinant of economic growth. Technology refers to scientific methods and techniques of production. In effect, technology means the nature and kind of machinery and technical equipment used with a given amount of labour. Capital-labour ratio is a broad measure of technology. Technological development means improving the technique of production through research and innovations. It results in a larger output from a given number of labour, materials and time. Invention of steam engines and railways, telephones and wireless, electricity, airplanes and computers are a few examples of technological developments over the past two centuries. Historical evidence shows that countries which achieved technological development at a rapid pace have made big strides in the field of economic growth. Economically backward countries which are unable to make technological development on their own have to import technology from other countries. Countries using inefficient technology have evidently grown at a slower pace.

5. Social and Political Factors

Social and political systems, organizations, institutions, social values, etc., also play an important role in the development process of an economy. Social factors like customs, traditions, beliefs, institutions, social (communal) harmony, and attitude towards the material life and well-being, determine, to a considerable extent, the pace of economic growth. A society of illiterate and ignorant people living with superstitions and unscientific beliefs resists modern ways of life and rational organization of society. Such a society finds it very difficult to achieve a high growth rate.

The form of government and its economic roles and policies matter to a great extent in determining the level and the rate of economic growth of a country. A government that plays a promotional role, provides adequate and efficient industrial infrastructure, builds an efficient system of public utilities (health and education system), invests in industries in which private investment is inadequate and removes weaknesses of the market system—helps economic growth. On the other hand, a government that throttles business and, thereby, activities through its restrictive economic policies, controls and regulations—as did the Indian government prior to the 1991 reforms through its licence, permit, quota raj—encourages inefficiency and malallocation of resources and restrains economic growth.

12.4 BUSINESS CYCLE: PHASES AND BUSINESS DECISIONS

Business cycles, i.e., repeated phases of economic booms and slumps are essentially a perpetual feature of the economic environment of a country. Business cycles influence business prospects tremendously and set the trend for future business.
Macro-Economic Factors

The period of prosperity opens up new and larger opportunities for investment, employment and production, and thereby promotes business. On the contrary, the period of depression reduces business opportunities. A profit maximizing entrepreneur must, therefore, analyze and take in view the economic environment of the period prior to making decisions, particularly those pertaining to forward planning.

Phases of Business Cycles

Business cycles refer to the periodic booms and slumps in economic activities, generally compared to ‘ebb and flow’. The ups and downs in the economy are reflected by the fluctuations in aggregate economic magnitudes, including total production, investment, employment, prices, wages, bank credits, etc. The upward and downward movements in these magnitudes show different phases of business cycles. Basically, there are only two phases in a cycle, viz., prosperity and depression. However, considering the intermediate stages between prosperity and depression, the various phases of trade cycle are enumerated as follows:

(i) Expansion of economic activities,
(ii) Peak of boom or prosperity,
(iii) Recession, the downtrend,
(iv) Trough, the bottom of depression, and
(v) Recovery and expansion.

The five phases of the business cycle are presented in Fig. 12.1. The steady growth line shows the growth of the economy when there are no business cycles. The various phases of business cycles are shown by the line of cycle which moves up and down the steady growth line. The line of cycle moving above the steady growth line marks the beginning of the period of ‘expansion’ or prosperity in the economy. The expansion phase is characterized by increase in output, employment, investment, aggregate demand, sales, profits, bank credits, wholesale and retail prices, per capita output and a rise in standard of living. However, growth rate eventually slows down and reaches its peak. The phase of recession begins when the downward slide in the growth rate becomes rapid and steady. Output, employment, prices, etc., register a rapid decline, though the realized growth rate may still remain above the steady growth line. So long as growth rate exceeds or equals the expected steady growth rate, the economy enjoys the period of prosperity—high or low. But, when the growth rate falls below the steady growth rate, it marks the beginning of depression in the economy. When depression continues and hits the bottom, it marks a period of trough. After the period of trough, the economy begins to recuperate and treads on the path of recovery. The process is continuous.
Let us now describe in some detail the important features of the various phases of business cycle, and also the causes of turning points.

**Prosperity: Expansion and Peak:** The prosperity phase is characterized by a rise in the national output, rise in consumer and capital expenditure, rise in the prices of raw materials and finished goods, and rise in the level of employment. In the later stages of prosperity, however, inputs start falling short of their demand. Additional workers are hard to find. Hence additional workers can be obtained by bidding a wage rate higher than the prevailing rates. Labour market becomes a seller’s market. A similar situation also appears in other input markets. Consequently, input prices increase rapidly leading to increase in output and employment. Cost of living increases at a rate relatively higher than the increase in household incomes. Hence consumers, particularly wage earners and households of fixed income class, review their consumption expenditure. Consumers’ resistance gets momentum. Actual demand stagnates or even decreases. The first and most serious impact of decline falls on the demand for new houses, flats and apartments. Incidentally, this is what happened in the US in 2007–08 causing economic depression. Following this, demand for cement, iron and steel, and construction labour tends to halt. This trend subsequently appears in other durable goods industries like automobiles, refrigerators, furniture, etc. This marks reaching the peak.

**Turning-Point and Recession:** As already mentioned, once the economy reaches the peak, increase in demand is halted. The demand even starts decreasing in some sectors, for the reason stated above. Producers, on the other hand, unaware of this fact continue to maintain their existing levels of production and investment. As a result, a discrepancy arises between output supply and demand: supply exceeds demand. The widening of discrepancy between supply and demand is so slow that it goes unnoticed for some time. But, the persistence of this problem makes the producers believe that they have indulged in ‘over-investment’ and over-production. Consequently, future investment plans are given up; orders placed for new equipments, raw materials and other inputs are cancelled. Replacement of worn-out capital is postponed. Demand for labour tends to decrease; and...
temporary and casual workers are laid off in a bid to bring demand and supply in balance. The cancellation of orders for inputs by the producers of consumer goods creates a chain-reaction in the input market. Producers of capital goods and raw materials cancel their orders for their input. This is the turning point and the beginning of recession.

Since demand for inputs has decreased, input prices, e.g., wages, interest, etc., show a gradual decline leading to a simultaneous decrease in the incomes of wage and interest earners. This ultimately causes demand recession. On the other hand, producers lower their price in order to get rid of their inventories and also to meet their financial obligations. Consumers, in turn, expect a further decrease in price, and hence, postpone their purchases. As a result, the discrepancy between demand and supply continues to grow. When this process gathers speed, it takes the form of irreversible recession. Investments start declining. The decline in investment leads to decline in employment, income and consumption. When investments continue to decline, production and employment continue to decline causing in further decline in demand for both consumer and capital goods. Borrowings for investment decreases; bank credit shrinks; stock prices decrease; unemployment increases even though there is a fall in wage rates. At this stage, the process of recession is complete. When growth rate falls below the steady growth line, the economy enters the phase of depression.

Depression and Trough. During the phase of depression, economic activities slide down their normal level. The growth rate becomes negative. The level of national income and expenditure declines rapidly. Prices of consumer and capital goods decline steadily. Workers lose their jobs. Debtors find it difficult to pay off their debts. Demand for bank credit reaches a low ebb and banks experience mounting of their cash balances. Investment in stock becomes less profitable and less attractive. At the depth of depression, all economic activities touch the bottom and the phase of trough is reached. Even the expenditure on maintenance is deferred in view of excess production capacity. Weaker firms are eliminated from the industries. At this point, the process of depression is complete.

How is the Process Reversed? The basic factor that reverses the recessionary trend is the limit to which an economy can shrink. When the economy hits the bottom and stays there for some time, it marks the end of pessimism and beginning of optimism. This reverses the process. The process of reversal generally begins in the labour market. The widespread unemployment forces workers to work at wages less than the prevailing rates. The producers anticipating better future try to maintain their capital stock and offer jobs to some workers here and there. They do so also because they begin to take an optimistic view of the situation due to the halt in decrease in price in the trough phase. Consumers on their part expecting no further decline in price begin to resume their postponed consumption and hence demand picks up, though gradually. Bankers having accumulated excess liquidity (idle cash reserve) try to salvage their financial position by lowering the lending rate and by investing their funds in securities and bonds, even if rate of return is very low. Consequently, investment picks up and employment gradually increases.
Following this recovery in production and wage income, demand for both consumer and capital goods starts picking up. Since banks have accumulated excess cash reserves, bank credit becomes easily available and at a lower rate. For all these reasons economic activities get accelerated. Due to increase in income and consumption, the process of multiplier gives further impetus to the economic activities, and the phase of recovery gets underway, depending on the speed of recovery.

**The Phase of Recovery.** As the recovery gathers momentum, some firms plan additional investment; some undertake renovation programmes; and some undertake both. These activities generate construction activities in both consumer and capital goods sectors. Individuals who had postponed their plans to construct houses undertake this task now, lest cost of construction should mount. As a result, more and more employment is generated in the construction sector. As employment increases despite wage rates moving upward, the total wage incomes increase at a rate higher than the employment rate. Wage income rises and so does the consumption expenditure. Businessmen realizing a quick return with high profitability, speed up the production machinery.

Over a period, as factors of production become more fully employed, wages and other input prices move upward rapidly, though not uniformly. Investors, therefore, select the best of the alternative investment opportunities. As prices, wages and other factor prices increase, a number of related developments begin to take place. Businessmen start increasing their inventories, consumers start buying more and more of durable goods and variety items. With this process catching up, the economy enters the phase of expansion and prosperity. The cycle is thus complete.

### 12.5 INFLATION

Inflation means generally an unduly high and persistent rise in the general level of prices over a period of time. However, there is no universally acceptable definition of inflation. The definition of inflation has been changing over time depending on the perception of the economists. For example, according to Pigou, a neo-classical economist, “Inflation exists when money income is expanding more than in proportion to increase in earning activity.” To Coulborn, inflation is a situation of “too much money chasing too few goods”. Modern economists have tried to define inflation more meaningfully. According to Ackley, “Inflation is a persistent and appreciable rise in the general level or average of prices.” Harry G. Johnson defines inflation as “a sustained rise in prices.” According to Samuelson, “Inflation denotes a rise in the general level of prices.” Bronfenbrenner and Holzman have suggested a number of alternative definitions of inflation which are mostly modified versions of earlier definitions. Their alternative definitions make things more fuzzy rather than adding clarity to inflation.
However, economists seem to agree that inflation means a ‘persistent’ and ‘appreciable’ increase in the general level of prices. The terms like ‘persistent’, ‘appreciable’, ‘sustained’, ‘considerable’, ‘continuing’ and ‘prolonged’ are not precisely defined. In practice, however, the term ‘persistent’ implies that the price rise exhibits a secular trend or continues to rise over a period of one to two years, and does not respond to anti-inflationary policies. The term ‘appreciable’ is more ambiguous because it does not specify as to what rate of increase in the price level is to be considered as ‘appreciable’ or ‘considerable’. It varies from country to country and from time to time. For example, in India, a 7–8% annual inflation in 1970s and 1980s was not so appreciable but a 6% inflation April–May 2007 and 7–8% in 2013 had become highly problematic—a matter of great concern for both the government and the RBI.

Annual consumer inflation in India declined to 3.69 percent in August of 2018 from 4.17 percent in July and below market expectations of 3.86 percent. It is the lowest inflation rate since October of 2017, mainly due to a sharp slowdown in food cost.

Methods of Measuring Inflation

There are two common methods of measuring inflation: (i) by computing change in Price Index Numbers (PIN), and (ii) by comparing the change in GNP Deflator. The two methods are explained here briefly.

Wholesale PINs Method

The formula used for measuring the rate of inflation through the changes in the PINs is given below.

\[
\text{Rate of inflation} = \frac{\text{PIN}_t - \text{PIN}_{t-1}}{\text{PIN}_{t-1}} \times 100
\]

where \(\text{PIN}_t\) is the wholesale price index number for the year selected for measuring inflation and \(\text{PIN}_{t-1}\) is the \(\text{PIN}\) in the preceding year.

GNP Deflator Method

The GNP deflator method uses the ratio of nominal GNP in a year to the real GNP of that year. The GNP deflator is defined as follows:

\[
\text{GNP deflator} = \frac{\text{Nominal GNP}}{\text{Real GNP}}
\]

where Nominal GNP is GNP at current prices and Real GNP is GNP at constant prices.

12.5.1 Factors Causing Inflation

The modern theories of inflation follow the theory of price determination. That is, the general price is determined by aggregate demand for and aggregate supply.
of goods and services and the variation in the aggregate price level is caused by the level of shift in the aggregate demand and aggregate supply curves. The modern theory of inflation is, in fact, a synthesis of classical and Keynesian theories of inflation. The modern analysis of inflation shows that inflation is caused by one or both of demand and supply-side factors. The factors which operate on the demand side are called demand-pull factors, and those operating on the supply-side are called cost-push factors. Accordingly, there are two kinds of inflation:

(i) Demand-pull inflation and (ii) Cost push inflation. The two kinds of inflation are discussed here briefly. However, a section of economists argues that demand-pull and cost-push factors interact to cause inflation. The theory of inflation based on interaction of demand-pull and cost-push factors has been discussed subsequently.

Demand-Pull Inflation

The demand-pull inflation occurs when the aggregate demand increases at a much higher rate than the aggregate supply. In other words, demand-pull inflation occurs when, given the aggregate supply, aggregate demand increases substantially. Increase in aggregate demand may be caused by (a) monetary factors, i.e., increase in money supply and/or (b) real factors, i.e., increase in demand for real output. Let us now see how monetary and real factors cause inflation.

(a) Increase in Money Supply and Demand-Pull Inflation: One important reason for demand-pull inflation is increase in money supply in excess of increase in potential output. Whether increase in money supply in excess of increase in output is the only cause of inflation is a controversial issue. But the fact is that monetary expansion in excess of increase in the level of output is one of the most important factors causing demand-pull inflation.

(b) Real-Factor Demand-Pull Inflation: Real-Factor demand-pull inflation can be caused by any or many of the following real factors.

(i) Increase in the government expenditure without change in tax revenue;
(ii) Cut in tax rates without change in the government expenditure;
(iii) Increase in investment;
(iv) Increase in consumer demand;
(v) Increase in exports given imports; and
(vi) Decrease in imports given the exports.

The first four factors straightaway increase the level of disposable income. Increase in aggregate income increases aggregate demand causing demand-pull inflation. For example, suppose that the government increases its spending financed through borrowings abroad. The rise in the government spending generates additional demand and, therefore, aggregate demand increases. Since there is, by assumption, full employment, additional resources can be acquired only by bidding a higher price. This pushes the prices up without increase in the output. Therefore,
the transaction demand for money increases. In order to meet the additional transaction demand for money, people sell their financial assets—bonds and securities. Consequently, bond and security prices go down and the rate of interest goes up. In the product market, prices increase to such an extent that the additional government spending is absorbed by the price rise. This is how other real factors also cause inflation.

**Cost-Push Inflation**

There are instances of inflation which could not be fully explained by the demand-side factors. The 1958-recession in the Western countries is a famous instance. During the period of recession, the aggregate demand had declined. Yet there was no decrease in the general price level. Instead, it tended to rise. In recent times, it is a common experience that prices generally do not decrease during the period of recession. Furthermore, even when there is stagnation in the economy and there is no inflationary pressure, the general price level generally continues to increase. It implies a situation of stagflation. The investigation into this kind of phenomenon, particularly for the 1958-puzzle, had led to the emergence of *supply-side theories of inflation*, popularly known as *cost-push theory* and *supply-shock theory* of inflation.

Cost-push inflation is generally caused by monopolistic groups of the society, like labour unions and firms in monopolistic and oligopolistic market setting. Strong labour unions often succeed in forcing money wages to go up causing prices to go up. This kind of rise in price level is called *wage-push inflation*. Also, firms enjoying monopoly power have also been found using their monopoly power to raise prices which in turn leads to rise in the general price level. The monopolistic and oligopoly firms push their profit margin up causing a rise in the general price level. This kind of inflation is called *profit-push inflation*. Another kind of cost-push inflation is said to be caused by supply shocks, i.e., decrease in the aggregate supply. This is called *supply shock inflation*. Minimum-wage legislation and administered prices are other supply-side factors which not only keep price level up but also create conditions for increase in the prices. In this section, we will discuss briefly these kinds of cost-push inflation.

(i) **Wage-Push Inflation:** Wage-push inflation is attributed to the exercise of monopolistic power by the labour unions to get their money wages enhanced more than the competitive labour market conditions would permit. Organized and powerful labour unions exercise their monopoly power and force their employers to increase their money wages above the competitive level without matching increase in labour productivity. Increase in money wages causes an equal increase in the cost of production. The increase in cost of production forces the firms to increase the price. This trend causes the aggregate supply curve to shift backward. A backward shift in the aggregate supply causes an upward movement in the price level.
However, every rise in the money wages is not always inflationary. The rise in money wages due to the following factors is not said to be inflationary.

(i) Increase in wage rate due to increase in productivity,
(ii) Rise in wage rates due to inflation,
(iii) Wage rise where unionized wage bill is very small, and
(iv) Wage rise due to shortage of labour supply.

(ii) Profit-Push Inflation: In contrast to wage-push inflation, profit-push inflation is caused by the use of monopoly power by the monopolistic and oligopolistic firms to enhance their profit margin which results in the rise in price and inflation. Today, monopolistic competition and oligopolies characterize the real market situation all over the world. The monopolies, monopolistically competitive firms and oligopolies tend to increase their price to increase their profit margin, given the market conditions, of course. Therefore, a profit-push type of inflation is certainly a theoretical possibility.

Profit-wage Spiral: It may be added here that wage-push and profit-push inflation go hand in hand, whichever may be the leading cause. Labour unions may be the first to force wage rate to go up but firms raise the price level often more than proportionately. Or else, monopolistic firms may be the first to push the product price up forcing labour unions to demand a higher wage rate. It then has its repercussions on the money wages. Higher prices and profits induce demand for higher wages. The powerful labour unions force their employers to raise their wages. Following the wage hike, firms raise the product prices. When this process gets going, it takes the form of ‘profit-wage spiral’.

(iii) Supply-Shock Inflation: Supply shock is generally caused by unexpected decline in the supply of major consumer goods or key industrial inputs. For example, vegetable and foodgrain prices shot up in India by more than 100 per cent in the last quarter of 2013, making a big election issue in 2014. Prices of some key industrial inputs like coal, steel, cement, oil, basic chemicals, etc., go up because of short supply caused by labour strikes, natural calamities, etc. Also, rise in the price may be caused by supply bottlenecks in the domestic economy or international events (generally wars) causing bottlenecks in the movement of internationally traded goods and causing thereby shortage of supply and rise in imported industrial inputs.

12.5.2 Factors Causing Deflation

Deflation is caused by a decline in demand. When the demand is less then the business enterprises are forced to lower the prices of products which can lead to stiff competition. Deflated is also caused by technology changes, such as more efficient computer chips. Deflation can also be caused by exchange rates. Let us go through some of the factors which cause deflation:
1. Change in Structure of Capital Markets
When a similar kind of product/service is being sold in the market by different enterprises then the price of that product is likely to be lower in the market due to stiff competition. Time and again, the capital structure of the economy will transform and enterprises will have easier access to debt and equity markets, which can be utilized to fund new business enterprises or improve productivity.

2. Increased Productivity
The introduction of innovative measures and new processes assists in increasing efficiency, which ultimately leads to lower prices. It is to be noted that some innovations have an impact only affect the productivity of certain industries; others may have a profound effect on the entire economy.

3. Decrease in Currency Supply
As the currency supply decreases, prices will decrease so that people can afford goods.

4. Austerity Measures
Deflation can be the outcome of decreased governmental, business, or consumer spending, which implies that the government spending cuts can lead to periods of significant deflation.

5. Deflationary Spiral
Once depression comes into the picture, it becomes difficult to control its after effects on the economy of a country. Firstly, when consumers begin reducing their spending amount, the profits of business houses decrease. This implies that businesses have to lessen wages and reduce their own purchases.

12.5.3 Control Measures
Economists agree that inflation beyond a moderate rate is bad and can often prove disastrous and, therefore, it must be kept under control. The various policy measures generally adopted for controlling inflation can be classified as follows.

(i) monetary policy measures,
(ii) fiscal policy measures,
(iii) price and wage control measures, and
(iv) indexation.

These measures of controlling inflation are discussed here briefly.

Monetary Measures
Monetary measures which are generally used to control inflation include (i) bank rate policy, (ii) variable reserve ratio and (iii) open market operation.

(i) Bank Rate Policy: The bank rate as a measure of inflation control works in two ways. One, during the period of inflation, the central bank raises the bank rate. This increases the cost of borrowing which reduces commercial banks’
borrowing from the central bank. This reduces banks' ability to create credit through the process of credit creation. As a result, flow of money from the commercial banks to the public gets reduced. Therefore, inflation is controlled to the extent it is caused by the bank credit. Two, bank rate sets the trend for the general market rate of interest, particularly in the short-term money market. When bank rate is increased with a view to controlling money supply and, thereby, inflation, commercial banks increase their lending rates and other market rates follow suit. In general, the cost of borrowing goes up. This slows down the monetary flows to the society.

This method, however, does not work effectively, if (i) commercial banks have excess liquidity, (ii) they have alternative sources of creating reserves, (iii) they are free not to raise their lending rates following the increase in the bank rate, and (iv) future expectations regarding the market prospects is optimistic.

(ii) Variable Reserve Ratio: Commercial banks are required to maintain a certain proportion of their total demand and time deposits in the form of cash reserves, called cash reserve ratio (CRR). The cash reserve ratio (CRR) is generally determined and imposed by the central bank which it uses as a weapon to control money supply. To control inflation, the central bank raises the CRR. When central bank raises the CRR, it reduces the lending capacity of the commercial banks. As a result, flow of money from the commercial banks to the public decreases. In the process, it halts the rise in prices to the extent it is caused by the banks credits to the public. This measure of controlling inflation has the same limitations as the bank rate policy.

(iii) Open Market Operations: Open market operations refer to sale and purchase of the government securities and bonds by the central bank. To control inflation, central bank sells the government securities to the public through the authorized commercial banks. The sale of government bonds through the banks results in the transfer of a part of bank deposits to central bank account. This reduces credit creation capacity of the commercial banks. Therefore, money supply with public is reduced by a multiple of sale proceeds of the treasury bills. As a result, inflation is controlled to the extent money supply with public decreases.

Fiscal Measures

Fiscal measures to control inflation include taxation, government expenditure and public borrowings. Keynesian economists, also called 'fiscalists', argue that demand-pull inflation is caused by excess of aggregate demand over aggregate supply. The excess demand is the result of increase in expenditure by the households, firms and the government, particularly by excessive spending by government. Excess demand, be it due to household or government expenditure, can be effectively controlled by fiscal measures. Therefore, fiscal policy measures are a more powerful and effective weapon to control demand-pull inflation.

In case government expenditure is the cause of excess demand, it can be controlled straightaway by cutting down public expenditure. A cut in public expenditure reduces government demand for goods and services and also the
private income and consumption expenditure arising out of government expenditure multiplier. Therefore, the excess demand decreases more than a given cut in public expenditure. And, in case excess demand is caused by rise in private expenditure, increasing income tax rate is a more appropriate measure to control inflation.

Taxation of incomes reduces the disposable income of the people and thereby consumer demand.

In case of a very high rate of persistent inflation, the government may adopt both the measures simultaneously. It should cut down its own spending, on the one hand, and increase the rate of income taxation, on the other. This kind of policy is known as a policy of surplus budgeting, i.e., the government should spend less than its tax revenue.

Price and Wage Control

In case monetary and fiscal measures prove ineffective in controlling inflation, direct control measures are adopted to control inflation. Direct measures consist mainly of price and wage controls. The price and wage controls go together because price-push and cost-push inflation go hand in hand whatever may be the cause of initial inflation.

Under price control method, a maximum retail price of goods and services is fixed. Price control may be general, applicable to all goods and services or it may be partial, confined to only scarce and essential goods and services. The primary objective of price control is to prevent the price rise of scarce goods and to ration the use of the commodity. Whether price control works effectively and efficiently is also a controversial matter. It is a general experience that price controls lead to black-marketing of goods and unfair distribution of scarce goods and services, especially where administrative machinery is corrupt and inefficient.

Wage control is used where inflation is of cost-push or of wage push nature. Under this method, rise in wage rate is prevented directly by imposing a ceiling on the wage incomes in both private and public sectors. Often ‘wage-freeze’ is applied to control inflation. In case trade unions are powerful and oppose ‘wage-freeze’ effectively, government uses a weaker method called ‘jawboning’. A more sensible and effective method of containing wage-push inflation is known as ‘wage guideposts’ — a plan of action against inflation prepared by common consent and mutual agreement between the representatives of the government, trade unions and the businessmen, for a disciplined and controlled upward movement in the wages and prices. Under this scheme, wage and price rise are monitored by a board of the representatives. However, this method does not work for long, especially when prices continue to rise. Friedman comments, “Guideposts and pleas for voluntary compliance are a halfway [measure] whose only merit is that they can more readily be abandoned than legally imposed controls. They are not an alternative to other effective measures to stem inflation, but at most a smoke-screen to conceal lack of action.”
Indexation

It should be clear from the foregoing discussion that inflation is an intractable problem. Controlling inflation involves the risk of aggravating unemployment problem. However, as discussed above, a high rate of inflation affects different sections of the society in different ways. Economists argue that if controlling inflation is not advisable, its adverse effects on different sections of society can be minimized by a method called indexation. They suggest indexation of prices, wages and contractual obligations with a view to compensating those who lose their real incomes due to inflation. According to Samuelson and Nordhaus, “Indexing is a mechanism by which wages, prices and contracts are partially or wholly compensated for changes in the general price level.” Thus, indexation is not a method of controlling inflation. It is a method of adjusting monetary incomes so as to minimize the undue gains and losses in real incomes of the different sections of the society due to inflation. Its main objective is to manage social discontent and to make inflation easier to live with.

Indexation of wages is most important and a common practice in many countries where wage contracts are long-term contracts and inflation continues to persist. In such cases, compensating workers for the loss of their real income due to inflation becomes unavoidable. Two systems are adopted in general: one is to tie wages to cost-of-living index (CPI) and second is to make a periodic scheduling of wage rise after CPI goes up by a certain percentage point. For instance, in the USA, wages are linked to cost-of-living index and cost-of-living adjustment (COLA) with inflation is made automatically. Here, wages increase automatically following the increase in the cost-of-living index. In India, wage compensation takes the form of ‘dearness allowance’, i.e., compensation for loss of purchasing power of the nominal wages due to inflation. Dearness allowance to the public sector employees in India is linked to consumer price index (CPI) and dearness allowance is granted after every eight percentage point increase in the CPI.

Although economists strongly recommend indexing of wages, debts, taxes, and all other long-term contractual payments, the governments doubt the feasibility and practicability of indexation method for three reasons. First, adjustment in indexation is impracticable in case of frequent supply shocks of great amplitude. Second, economy being a complex system with interlinked and interrelated prices, a reasonable indexing of all prices to the satisfaction of all concerned is an extremely difficult task. Third, the government finds indexation politically undesirable because it does not control inflation, it rather creates a base for its perpetuation.

12.6 BALANCE OF PAYMENT TREND AND ITS IMPLICATIONS ON MANAGERIAL DECISIONS

The BOP of a country is a systematic accounting of all economic transactions between the residents of the country and the residents of the rest of the world over
a period of time, usually a year. In simple words, the BOP of a country is an account of all international receipts and payments for the accounting year. The accounting year may be a calendar year or a financial year. In India, the BOP is prepared for every financial year, although it is also prepared and reported on a quarterly basis.

Although the definition of BOP is quite clear, some terms used in the definition need some clarification.

The term 'systematic accounting' does not refer to any particular system. But the system generally adopted is double entry book-keeping system. Under this system, both receipts and payments of an economic transaction—sale and purchase—are recorded in one entry. We know that each economic transaction involves two flows—inflow and outflow. What comes in, i.e., inflow, is treated as a debit item and what goes out, i.e., the outflow, is treated as a credit item. Under double entry book-keeping system, the values of debit and credit items are recorded simultaneously. For example, in case of exports, goods and services go out of the country and payments received come to the country. What goes out (export) is recorded as a credit item and what comes in (payments received), as a debit item. Both the value of exports and the payments received are recorded in the same accounting entry of the transaction. Similarly, in the case of imports, the value of imports of goods and services is recorded as a debit item and payments for imports are recorded as a credit item. Both the value of imports and the payments made are recorded in the same accounting entry. This is the double entry book-keeping system.

'Economic transactions' include all such transactions that involve the transfer of money (foreign exchange) along with the right of ownership of the transacted goods, money, and assets. Specifically, exports and imports of goods and services, foreign investment, and borrowing from and lending to a foreign country, are all economic transactions. It is important to note here that while in some transactions, there is a physical transfer of goods, services, assets, and money along with the transfer of ownership, in others, no physical transfer is required. For example, if a US resident imports garments from India, they have to be transferred physically along their ownership. But if a US multinational corporation operating in India reinvests its profits in India, there is no physical transfer, i.e., transfer of money.

The term 'resident' refers to the nationality of the persons or organizations carrying out the transaction, not their physical presence in their own country. For example, foreign tourists, diplomats, military personnel, migrant labour and branches of companies in a country are treated as residents of their own nation, not as residents of the country in which they are physically present.

Purpose of Balance of Payments
The purpose of preparing balance of payments (BOP) is similar to that of the balance sheet of business firms. We know that all business firms prepare their
annual (or periodic) balance sheet of their transactions with the rest of society with the purpose of taking stock of their profit and loss and their assets and liabilities. Likewise, all nations having economic transactions with the rest of the world prepare their periodic BOP accounts just to take stock of their receipts and payments, and their assets and liabilities resulting from their international economic transactions. However, the purpose of BOP goes far beyond this limited purpose as it reveals the overall implications of international economic transactions for the economy of a country. The BOP accounts serve the following purposes.

First, the international economic transactions of a country—its exports, imports, and foreign investments—and inflows and outflows of incomes (e.g., remittances) have a serious impact on the levels of its consumption and the stock of capital. The change in the levels of consumption and of capital accumulation affects the national income of the country and its economic growth over time.

Therefore, in order to find out the impact of the monetary inflows and outflows on the economy, a country has to prepare its balance of payments.

Second, in case a country has a large and persistent trade deficit, and capital inflows are not sufficient to make up the trade deficit, it faces a serious foreign payment problem. A country having a large and persistent trade deficit has to borrow from abroad to settle its payments. When the country makes a large borrowing from abroad over a long period of time, it tends to fall into international indebtedness. This can have serious adverse repercussions for the of the economy of the country. This is what had happened in India in 1990–91.

Third, the balance of payments provides useful data for analysing a country’s weaknesses and strengths with respect to foreign trade and its gains and losses from international transactions. The balance of payments contains data on exports and imports, capital inflows and outflows, borrowings and lendings, and foreign exchange reserves for successive years. By analysing the data one can find out whether the economic condition of a country in relation to the world economy is improving or deteriorating. If it is deteriorating, the government will have to formulate appropriate corrective policy measures and implement them on time. ‘One of the basic problems of international economic policy is to find effective means of restoring external balance to the country whose balance of payments is seriously in surplus or in deficit.’ The balance of payments provides useful data and guidelines for the formulation a suitable foreign trade and exchange rate policy.

Specific Uses of BOP for a Business Manager

- A study evaluating the country’s balance of payments can help the business manager in deciding whether he can borrow/raise funds in a particular country or invest his funds there.
- A proper analysis of balance of payments data can disclose trends in factors that underlie the value of the country’s currency in a foreign market.
12.7 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Inflation means generally an unduly high and persistent rise in the general level of prices over a period of time.

2. The main objective of indexation is to manage social discontent and to make inflation easier to live with.

3. Gross Domestic Product (GDP) is defined as the total market value of all final goods and services produced within a country in a given period.

12.8 SUMMARY

- The subject matter of managerial microeconomics consists of theories of demand and supply, theory of production, cost-output relationships, market mechanism, theory of price determination under different kinds of market structure, pricing principles and prices, and capital budgeting and investment decisions.

- Macroeconomics as a positive science has theoretical orientation. As a positive science, it uses macroeconomic models to explain the behaviour of macroeconomic variables—national income, aggregate consumption, savings and investment, price level, and employment.

- The importance of macroeconomics can be viewed from two angles: (i) how it contributes to government’s economic policy formulation for the management of the economy, and (ii) how it guides business managers for taking appropriate business decisions and business planning.

- Both developed and developing countries are constantly confronted with some or the other kind of macroeconomic problems, e.g., recession and depression, stagflation, balance-of-payment deficits, outflow of capital, mounting debt burden or a country falling into debt trap, and so on.

- The overall economic trend in a country is the most important element of its business environment. A growing economy with a considerably high and sustained growth rate in its GDP over a long period of time provides a promising business prospect and builds business confidence.
• Business cycles refer to the periodic booms and slumps in economic activities, generally compared to ‘ebb and flow’.

• Inflation means generally an unduly high and persistent rise in the general level of prices over a period of time. However, there is no universally acceptable definition of inflation.

• The modern theories of inflation follow the theory of price determination. That is, the general price is determined by aggregate demand for and aggregate supply of goods and services and the variation in the aggregate price level is caused by the level of shift in the aggregate demand and aggregate supply curves.

• Deflation is caused by a drop in demand. Fewer shoppers mean businesses have to lower prices, which can turn into a bidding war. It is also caused by technology changes, such as more efficient computer chips. Deflation can also be caused by exchange rates.

• Monetary measures which are generally used to control inflation include (i) bank rate policy, (ii) variable reserve ratio and (iii) open market operation.

• Fiscal measures to control inflation include taxation, government expenditure and public borrowings.

• Indexation of wages is most important and a common practice in many countries where wage contracts are long-term contracts and inflation continues to persist.

• The BOP of a country is a systematic accounting of all economic transactions between the residents of the country and the residents of the rest of the world over a period of time, usually a year.

12.9 KEY WORDS

• Stagflation: It is a condition of slow economic growth and relatively high unemployment, or economic stagnation, accompanied by rising prices, or inflation.

• Capital formation: It is the process of creating or acquiring man-made means of production.

12.10 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short Answer Questions

1. Write a short note on the nature of macroeconomics.

2. What is the importance of macro-economic factors?
3. How is economic growth defined?
4. What are the methods of measuring inflation?

Long Answer Questions
1. Discuss the determinants of economic growth.
2. Describe the phases of business cycles.
3. Explain the factors causing inflation.
4. Evaluate the measures adopted for controlling inflation.
5. 'BOP of a country is an account of all international receipts and payments for the accounting year.' Elucidate the statement.

12.11 FURTHER READINGS


Websites
https://blogs.economictimes.indiatimes.com/et-commentary/heres-how-indias-widening-income-distribution-can-be-redressed/
UNIT 13 NATIONAL INCOME

13.0 INTRODUCTION

National income is the final outcome of the economic activities of the people of the nation. National income is the most important macroeconomic variable. The level of national income of a country determines its economic strength to sustain the economic, political, social and international problems. It determines the level of employment, per capita income, aggregate demand and supply of goods and services, and the level of production. What is more important from the viewpoint of business management is the fact that change in national income determines the business environment of the country. Business environment matters a great deal in making business decisions. We begin our discussion by defining the term 'national income' and some basic concepts related to national income measurement.

13.1 OBJECTIVES

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

- Discuss the concept of national income
- Describe the methods of measurement of national income
- Explain the theory of national income determination
- Discuss the sectoral and population distributions
• Examine the concept, uses and limitation of per capita income
• Recall the recent developments in the Indian Economy

13.2 INTRODUCTION TO NATIONAL INCOME

National income can be defined as the sum of money value of all final goods and services produced in a country over a period of one year. Some terms used in the definition of national income need to be clarified. The term ‘money value’ means the value estimated at the current price of the goods and services. As regards the term ‘final goods and services’, productive activities create many goods and services. While some goods and services are final, some are used again in the process of production, and some take the form of capital. In all cases, while measuring national income only final goods and services are taken into account.

National income can be defined also as the sum of all factors incomes. Factors of production, viz., land, labour, capital and entrepreneurship, are used to produce goods and services. Use of factors of production generates factor incomes in the form of rent, wages, interest and profit. These factor incomes constitute the national income of a country. The sum of all these factor-incomes gives the estimate of national income.

National Income: Some Basic Concepts

Measuring national income is an extremely complex and complicated task as it involves both conceptual and practical problems. Therefore, before we discuss the methods of estimating national income, it is useful to understand the concepts used in estimating national income. In this section, we take the view of some concepts used in estimation of national income.

1. Gross Domestic Product (GDP). Gross domestic product (GDP) is the measure of the total market value of all final goods and services produced in the domestic economy during a period of one year plus income earned by the foreigners in the country minus income earned by countrymen from abroad.

2. Gross National Product (GNP). GNP is the measure of the total market value of all final goods and services produced in the domestic economy during a period of one year plus incomes earned abroad by the citizens minus income earned by the foreigners in the country.

3. Economic and Non-economic Products. In estimating national income only economic products are included. Goods and services that are produced to be sold at market price and goods and services that are produced by the government and public organizations are treated as economic products. Thus, economic products include both marketable and non-marketable goods. And, non-economic products include services rendered to self, to
family, to relations and to neighbours. Non-economic products are not included in national income.

4. **Intermediate and Final Products.** Products (goods and services) that are used in the process of further production are considered to be *intermediate product* and products that are consumed by the final consumers are considered to be *final products*. The same product may be an intermediate or a final product. For examples, when wheat produced by farmers are consumed by themselves, it is treated as *final good*, but when it is sold to bread companies, it is treated as *intermediate product*. Likewise, when services provided by the government, e.g., transport, telephonic, postal, railway services, etc. are used by the consumers, they are treated as *final goods* and when used in the process of production, then these services are treated as *intermediate goods*. In estimating national income, only final products are taken into account.

5. **Transfer Payments.** Payments made by the people to other people, organizations or to the government without any equal transfer in return are treated as *transfer payment*. For example, gifts paid to relatives and friends, donations given to social organization, and taxes paid to government authorities, etc., are transfer payments. Such payments are not taken into account in national income estimates.

**National Income as Money Flow**

We have defined national income from the angle of *product flows*. The same can be defined in terms of *money flows*. While economic activities generate flow of goods and services, on the one hand, they generate money flows, on the other, in the form of factor payments—wages, interest, rent, profits, and earnings of self-employed. Thus, national income may also be estimated by adding the factor earnings and adjusting the sum for indirect taxes and subsidies. The national income thus obtained is known as *national income at factor cost*. It is related to money income flows.

The concept of national income refers to the income of the society as a whole. It differs fundamentally from the concept of *private income*. Conceptually, national income refers to the money value of the entire volume of final goods and services resulting from all economic activities of the country. This is not true of private income. Also from the calculation point of view, there are certain receipts of money or of services and goods that are not ordinarily included in private incomes but are included in the national incomes, and *vice versa*. National income includes, for example, employer’s contribution to the social security and welfare funds for the benefit of employees, profits of public enterprises and services of owner occupied houses. But it excludes the interest on warloans, social security benefits and pensions. These items are, however, included in the private incomes. The national income is, therefore, not merely an aggregation of the private incomes.
13.2.1 Methods of Measurement

National income of a country is created by its people participating in different kinds of economic activities and producing goods and services. For measuring national income, an economy is viewed from three different angles.

1. The national economy is considered as an aggregate of productive units of different sectors such as agriculture, mining, manufacturing, trade and commerce, services, etc.

2. The whole national economy is viewed as a combination of individuals and households owning different kinds of factors of production which they use themselves or sell factor-services to make their livelihood.

3. The national economy may also be viewed as a collection of consuming, saving and investing units (individuals, households, firms and government).

Following these notions of a national economy, national income may be measured by three different corresponding methods:

1. **Net product method**—when the entire national economy is considered as an aggregate of producing units;

2. **Factor-income method**—when national economy is considered as a combination of factor-owners and users;

3. **Expenditure method**—when national economy is viewed as a collection of spending units.

The procedures which are followed in measuring the national income in a closed economy—an economy which has no economic transactions with the rest of the world—are briefly described here. The measurement of national income in an open economy and adjustment with regard to income from abroad will be discussed subsequently.

### Net Output or Value Added Method

The net output method is also called the *value added method*. In its standard form, this method consists of three stages: "(i) estimating the gross value of domestic output in the various branches of production; (ii) determining the cost of material and services used and also the depreciation of physical assets; and (iii) deducting these costs and depreciation from gross value to obtain the net value of domestic output....". The net value of domestic product thus obtained is often called the *value added or income product* which is equal to the sum of wages, salaries, supplementary labour incomes, interest, profits, and net rent paid or accrued. Let us now describe the stages (i) and (ii) in some detail.

(a) **Measuring Gross Value.** For measuring the gross value of domestic product, output is classified under various categories on the basis of the nature of activities from which they originate. The output classification varies from country to country depending on (i) the nature of domestic activities, (ii) their significance in aggregate economic activities, and (iii) availability of requisite data. For example, in the US,
about seventy-one divisions and subdivisions are used to classify the national output; in Canada and the Netherlands, classification ranges from a dozen to a score; and in Russia, only half a dozen divisions are used. According to the CSO publication, fifteen sub-categories are currently used in India.

After the output is classified under the various categories, the value of gross output is computed in two alternative ways: (i) by multiplying the output of each category of sector by its respective market price and adding them together or (ii) by collective data about the gross sales and changes in inventories from the account of the manufacturing enterprises and computing the value of GDP on the basis thereof. If there are gaps in data, some estimates are made thereof and gaps are filled.

(b) Estimating Cost of Production. The next step in calculating the net national product is to estimate the cost of production including depreciation. Estimating cost of production is, however, a relatively more complicated and difficult task because of non-availability of adequate and requisite data. Much more difficult is the task of estimating depreciation since it involves both conceptual and statistical problems. For this reason, many countries adopt factor-income method for estimating their national income.

However, countries adopting net-product method find some ways and means to calculate the deductible cost. The costs are estimated either in absolute terms (where input data are adequately available) or as an overall ratio of input to the total output. The general practice in estimating depreciation is to follow the usual business practice of depreciation accounting. Traditionally, depreciation is estimated at some percentage of capital, permissible under the tax-laws. In some estimates of national income, the estimators deviate from the traditional practice and estimate depreciation as some ratio of the current output of final goods.

Following a suitable method, deductible costs including depreciation are estimated for each sector. The cost estimates are then deducted from the sectoral gross output to obtain the net sectoral products. The net sectoral products are then added together. The total thus obtained is taken to be the measure of net national product or national income by net product method.

Factor-Income Method

This method is also known as income method and factor-share method. Under this method, the national income is calculated by adding up all the “incomes accruing to the basic factors of production used in producing the national product”. Factors of production are conventionally classified as land, labour, capital and organization. Accordingly, the national income equals the sum of the corresponding factor earnings. Thus,

National income = Rent + Wages + Interest + Profit

However, in a modern economy, it is conceptually very difficult to make a distinction between earnings from land and capital, on the one hand, and between
the earnings from ordinary labour and entrepreneurial functions, on the other. For
the purpose of estimating national income, therefore, factors of production are
broadly grouped as labour and capital. Accordingly, national income is supposed
to originate from two primary factors, viz., labour and capital. In some activities,
however, labour and capital are jointly supplied and it is difficult to separate the
labour and capital contents from the total earnings of the supplier. Such incomes
are termed as mixed incomes. Thus, the total factor-incomes are grouped under
three categories; (i) labour incomes, (ii) capital incomes and (iii) mixed incomes.

(a) Labour Incomes. Labour incomes included in the national income have three
components: (a) wages and salaries paid to the residents of the country including
bonus and commission, and social security payments; (b) supplementary labour
incomes including employer’s contribution to social security and employee’s welfare
funds, and direct pension payments to retired employees; (c) supplementary labour
incomes in kind, e.g., free health and education, food and clothing, and
accommodation, etc. Compensations in kind (food and clothes) to domestic
servants and such other free-of-cost services provided to the employees are
included in labour income. War bonuses, pensions, service grants are not included
in labour income as they are regarded as ‘transfer payments’. Certain other
categories of income, e.g., incomes from incidental jobs, gratuities, tips, etc., are
ignored for lack of data.

(b) Capital Incomes. According to Studenski, capital incomes include the
following kinds of earnings:

(a) dividends excluding inter-corporate dividends;
(b) undistributed before-tax-profits of corporations;
(c) interest on bonds, mortgages, and saving deposits (excluding interests on
war bonds, and on consumer-credit);
(d) interest earned by insurance companies and credited to the insurance policy
reserves;
(e) net interest paid out by commercial banks;
(f) net rents from land, buildings, etc., including imputed net rents on owner-
occupied dwellings;
(g) royalties and
(h) profits of government enterprises.

The data for the first two items are obtained mostly from the firms’ books of
accounts submitted for taxation purposes. But the definition of profit for national
accounting purposes differs from that employed by taxation authorities. Some
adjustments in income tax data therefore, become, necessary. The data adjustments
generally pertain to (i) excessive allowance of depreciation made by the firms;
(ii) elimination of capital gains and losses since these items do not reflect the
changes in current income and (iii) elimination of under or over-valuation of inventories on book-value.

(c) Mixed Income. Mixed incomes include earnings from (a) farming enterprises, (b) sole proprietorship (not included under profit or capital income) and (c) other professions, e.g., legal and medical practices, consultancy services, trading and transporting, etc. This category also includes the incomes of those who earn their living through various sources as wages, rent on own property, interest on own capital, etc.

All the three kinds of incomes, viz., labour incomes, capital incomes and mixed incomes, added together give the measure of national income by factor-income method.

Expenditure Method

The expenditure method, also known as final product method, measures national income at the final expenditure stages. In estimating the total national expenditure, any one of the two following methods are used: (i) all the money expenditures at market price are computed and added up together and (ii) the value of all the products finally disposed off are computed and added up, to arrive at the total national expenditure. The items of expenditure which are taken into account under the first method are (a) private consumption expenditure; (b) direct tax payments; (c) payments to the non-profit-making institutions and charitable organizations like schools, hospitals, orphanages, etc. and (d) private savings. Under the second method, the following items are considered: (a) private consumer goods and services, (b) private investment goods, (c) public goods and services and (d) net investment abroad. The second method is more extensively used because the data required in this method can be collected with greater ease and accuracy.

Treatment of Net Income from Abroad. We have so far discussed methods of measuring national income of a ‘closed economy’. But most economies are open in the sense that they carry out foreign trade in goods and services and financial transactions with the rest of the world. In the process, some nations make net income through foreign trade while some lose their income to foreigners. The net earnings or losses from foreign trade change the national income. In measuring the national income, therefore, the net result of external transactions is adjusted to the total. Net incomes from abroad are added to, and net losses from the foreign transactions are deducted from the total national income arrived at through any of the above three methods.

Briefly speaking, all exports of merchandise and of services like shipping, insurance, banking, tourism, and gifts are added to the national income. And, all the imports of the corresponding items are deducted from the value of national output to arrive at the approximate measure of national income. To this is added the net income from foreign investment. These adjustments for international transactions are based on the international balance of payments of the nations.
13.3 THEORY OF NATIONAL INCOME DETERMINATION

In the preceding section, we have discussed the methods of estimating national income. The method of estimating income does not reveal how the level of national income is determined. In this section, we proceed to discuss the theory of national income determination. The theory of national income determination explains how equilibrium level of national income is determined. This theory was propounded by a British economist, John Maynard Keynes, in 1936 in his book *The General Theory of Employment, Interest and Money*. The Keynesian theory of income determination has been discussed in this chapter. For sake of analytical convenience, the Keynesian theory of income determination is presented in the framework of three models:

1. Two-sector model: households and business sector,
2. Three-sector model: households, business and government sectors, and

The fundamental principle of the Keynesian theory of national income determination is that the national income is determined by aggregate demand and aggregate supply and equilibrium level of national income is determined at the level where aggregate demand equals aggregate supply. The Keynesian theory of income determination has been discussed in this unit under the above three models. We begin our discussion on the theory of income determination with two-sector model.

13.3.1 Two-Sector Model

First, there are only two sectors in the economy: households and firms. Households own the factors of production and they sell their factor services or provide them to the firms to earn their living in the form of factor payments—wages, rent, interest and profits. Also, households are the consumers of all final goods and services. The firms, on the other hand, hire factor services from the households and produce goods and services which they sell to the households.

Secondly, there is no government. Or, if government is there, it does not perform any economic function; it does not tax, it does not spend and it does not consume.

Thirdly, the economy is a closed one: there is no foreign trade. It implies that there is no outflow or inflow of goods and services to and from foreign countries.

Fourthly, there are no corporate savings or undistributed (or retained) corporate profits, i.e., the total corporate profit is distributed as dividends.

Finally, prices of all goods and services, supply of labour and capital, and the state of production technology are given and remain constant.
As noted above, according to Keynes, national income of a country is determined by two factors: (i) aggregate demand (AD) and (ii) aggregate supply (AS) of goods and services. And, the equilibrium level of national income is determined where AD equals AS. Before we illustrate graphically the determination of national income, let us explain the concepts of aggregate demand and aggregate supply.

**Aggregate Supply**

The *aggregate supply* (AS) refers to the total value of goods and services produced and supplied in an economy per unit of time. Aggregate supply includes both consumer goods and producer goods. The goods and services produced per time unit multiplied by their respective (constant) prices give the total value of the national output. This is the aggregate supply in terms of money value.

**Aggregate Supply Schedule.** The Keynesian aggregate supply schedule or aggregate supply curve is drawn on the assumption that total income is always spent – no part of it is retained or withheld. By Keynesian assumption, total expenditure is always equal to the total income. If income increases, expenditure increases by the same amount. This relationship between income and expenditure is shown by a 45° line. This line is also called *aggregate supply schedule*. In the Keynesian theory of income determination, aggregate income equals consumption (C) plus savings (S). Therefore, AS schedule is generally named as C+S schedule. The aggregate supply (AS) curve is also sometimes called ‘aggregate expenditure’ (AE) curve.

**Aggregate Demand**

The aggregate demand is an *ex-post* concept. It implies effective demand which equals actual expenditure. The aggregate effective demand is equal to the aggregate expenditure made by the society per unit of time. Aggregate demand (AD) consists of two components:

(i) aggregate demand for consumer goods (C) and
(ii) aggregate demand for capital goods (I).

Thus, \( AD = C + I \) \hfill (13.1)

**Aggregate Demand Schedule**

The aggregate demand AD schedule is also called C + I schedule. In the Keynesian framework, investment (I) is assumed to remain constant in the short-run. But, consumption (C) is treated to be a constant function of income (Y). Pending detailed discussion on the consumption function till the next section, let us let us note here that the Keynesian consumption function is given as

\[ C = a + bY \] \hfill (13.2)

where \( a \) denotes a constant denoting \( C \) when \( Y = 0 \) and \( b \) denotes a constant ratio of income consumed, i.e., \( b = \Delta C/\Delta Y \).
By substituting Eq. (13.2) in Eq. (13.1), \( AD \) function can be expressed as

\[
AD = a + bY + I
\]  
...(13.3)

Let us now illustrate the construction of the \( C + I \) schedule by assuming:

(i) \( C = 50 + 0.5Y \), and

(ii) \( I = 50 \) billion

By substituting consumption function and \( I = 50 \), \( AD \) function given in Eq. (13.3) can now be written as

\[
AD = 50 + 0.5Y + 50 = 100 + 0.5Y
\]

An aggregate demand schedule based on the above assumptions is given in Table 13.1. It shows the relationship between income and aggregate demand \( (C + I) \). The

<table>
<thead>
<tr>
<th>Income ( Y )</th>
<th>( C = 50 + 0.5Y )</th>
<th>( I = 50 )</th>
<th>( C + I ) Schedule</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>50 + 0 = 50</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>50</td>
<td>50 + 25 = 75</td>
<td>50</td>
<td>125</td>
</tr>
<tr>
<td>100</td>
<td>50 + 50 = 100</td>
<td>50</td>
<td>150</td>
</tr>
<tr>
<td>150</td>
<td>50 + 75 = 125</td>
<td>50</td>
<td>175</td>
</tr>
<tr>
<td>200</td>
<td>50 + 100 = 150</td>
<td>50</td>
<td>200</td>
</tr>
<tr>
<td>250</td>
<td>50 + 125 = 175</td>
<td>50</td>
<td>225</td>
</tr>
<tr>
<td>300</td>
<td>50 + 150 = 200</td>
<td>50</td>
<td>250</td>
</tr>
<tr>
<td>350</td>
<td>50 + 175 = 225</td>
<td>50</td>
<td>275</td>
</tr>
<tr>
<td>400</td>
<td>50 + 200 = 250</td>
<td>50</td>
<td>300</td>
</tr>
</tbody>
</table>

Two-sector Model of National Income Determination: Algebraic Treatment

In preceding sections, we have presented the Keynesian theory of income determination in its simplest form and have explained the derivation of the consumption and saving functions. In this section, we present the two-sector model of income determination in its formal form.

As stated above, equilibrium level of national income is determined at the level of income at which aggregate demand for output \( (C + I) \) is equal to aggregate supply of incomes \( (C + S) \). Thus, equilibrium condition of national income is given as

\[
\text{Aggregate Demand} = \text{Aggregate Supply},
\]

or

\[
C + I = C + S
\]  
...(13.4)

Since \( C \) is common to both the sides, the equilibrium conditions can also be stated as

\[
I = S
\]  
...(13.5)
Given these conditions of equilibrium, there are two alternative ways to show the determination of national income:

(i) by using aggregate demand \((C + I)\) and aggregate supply \((C + S)\) schedules, and

(ii) by using only saving \((S)\) and investment \((I)\) schedules.

The two approaches are known as **income-expenditure approach** and **saving-investment approach**, respectively.

### 13.3.2 Three-Sector Model

We have explained above income determination in a simple, two-sector economy. A real economy, however, also contains two other major sectors, viz., the government sector and the foreign sector. In this section, we explain income determination in a **three-sector model** including household, business and government sectors. The **four-sector model** including three domestic sectors and the foreign sector will be discussed in the next section.

Let us assume an economy consisting of three sectors, viz., households, firms and government. Inclusion of the government into the model brings in two variables:

(i) **government expenditures**, and (ii) taxation. While government expenditures are injections to the income stream, taxation is withdrawal from the income stream. Government expenditure adds to the aggregate demand, and taxation, on the other hand, reduces the aggregate demand. How the level of national income is affected by the inclusion of government expenditure and taxation into the model is discussed below.

To begin with, let us incorporate government sector into our two-sector model, assuming that the government follows a **balanced budget** policy, i.e., government expenditure \((G)\) equals the amount of taxes \((T)\), and both \(G\) and \(T\) are exogenously determined. Let us now redefine the parameters of the aggregate demand and aggregate supply for the three-sector model.

Aggregate demand \((AD)\) in the three-sector model can be expressed as

\[
AD = C + I + G
\]

and aggregate supply \((AS)\) as

\[
AS = C + S + T
\]

The equilibrium of national income is determined where

\[
AS = AD
\]

Thus, at equilibrium,

\[
Y = C + S + T = C + I + G
\]

Since \(S = I\) and \(T = G\), at equilibrium, \(Y = C + I + G\)

...(13.6)

In Eq. (13.6), \(C = a + bY_d\) where \(Y_d\) is disposable income, i.e., \(Y_d = Y - T\) and

\[
T = \text{lump-sum tax}
\]
By substituting consumption function $a + bY$ for $C$ in Eq. (13.6) and $Y - T$ for $Y_s$ in the resulting equation, the equilibrium equation can be written as

$$Y = a + b(Y - T) + I + G \quad \ldots(13.7)$$

Solving for $Y$, we get equilibrium level of national income as

$$Y = a + bY - bT + I + G$$

$$Y(1 - b) = a - bT + I + G$$

$$Y = \frac{1}{1-b}(a - bT + I + G) \quad \ldots(13.8)$$

Alternatively, the equilibrium condition of national income can be derived by saving-investment approach as follows. Consider the equilibrium condition given as

$$C + S + T = C + I + G$$

Since $C$ on each side gets cancelled out, equilibrium equation can be rewritten as

$$S + T = I + G$$

### 13.3.3 Four-Sector Model

We have been concerned so far with the theory of income determination in a 'closed economy', i.e., on economy without foreign trade. In this section, we will discuss income determination in an 'open economy'. An open economy is one which carries out economic transactions with the rest of the world—mainly foreign trade. Inclusion of foreign trade makes the model a complete, four-sector model.

In foreign trade, exports are injections and imports are outflows from the circular flows of incomes. In the national income analysis, however, only trade balances ($X - M$) is taken into account. That is, only $X - M$ is incorporated in the four-sector model of income determination. Injections increase the level of income and withdrawals decrease it. Therefore, if $X > M$, there is net injection and national income increases, and if $X < M$, there is a net withdrawal which causes national income to decrease.

To begin the analysis of income determination with foreign sector, let us first specify the export and import functions.

#### Export Function

Exports of a country depend on such factors as: (i) prices of domestic goods in relation to prices of goods in importing countries; (ii) tariffs and trade policies of importing countries; (iii) export subsidies; (iv) income elasticity for imports in the importing countries, and (v) level of imports by the domestic economy, etc.

The factors other than (i) and (iii) are beyond the control of the economy and hence are not included among the policy variables. For the sake of simplicity, therefore, it is assumed that exports ($X$) are determined by the factors operating
outside the economy. In the income determination model, therefore, \( X \) is treated as an autonomous variable and assumed to be given as.

### Import Function

Imports of a country, like its exports, are determined by (i) import prices in relation to domestic prices; (ii) the level of domestic tariffs; (iii) domestic trade policy; (iv) income-elasticity of imports; (v) the level of incomes and (vi) the level of exports.

For analytical purpose, however, a simplifying assumption is made that imports (\( M \)) depend on the level of domestic income (\( Y \)) and on the marginal propensity to import (MPM). Under this assumption, import function is expressed as

\[
M = \bar{M} + gy
\]  

(13.9)

where \( \bar{M} \) = autonomous imports, \( g = \Delta M/\Delta Y = MPM \) (assumed to be constant).

Having specified the \( X \) and \( M \) functions, we may now specify the four-sector model of income determination as given below.

\[
Y = C + I + G + (X - M) \quad \text{...(13.10)}
\]

where

\[
C = a + bY_d
\]

\[
I = I_d
\]

\[
G = G_d
\]

These variables are constant

\[
x - \bar{X} \]

\[
Y_d = Y - \bar{T}
\]

\[
M = \bar{M} + gY
\]

By incorporating these variables and functions in Eq. (13.10), the equilibrium equation can be written as

\[
Y = \frac{a}{1 - b + g} + \frac{b}{1 - b + g}Y_d + \frac{I}{1 - b + g} + \frac{G}{1 - b + g} + \frac{x}{1 - b + g} + \frac{(X - M)}{1 - b + g} \quad \text{...(13.11)}
\]

The term \( 1/(1 - b + g) \) in Eq. (13.11) is the foreign trade multiple, on the assumption that consumption and imports are both a linear function of domestic income. This can be proved as follows. Let exports (\( X \)) increase by \( \Delta X \), while other variables remain constant. The equilibrium level of national income can then be written as follows.

\[
Y + \Delta Y = \frac{1}{1 - b + g} (a - bT + \bar{T} + \bar{G} + X - \bar{M} + \Delta X) \quad \text{...(13.12)}
\]

or

\[
Y + \Delta Y = \frac{1}{1 - b + g} (a - b\bar{T} + \bar{T} + \bar{G} + X - \bar{M}) + \frac{1}{1 - b + g} \Delta X
\]
By subtracting $Y$ from both sides, we get

$$\Delta Y = \frac{1}{1 - b + g} \Delta X$$

...(13.13)

By rearranging Eq. (13.13), we get foreign trade multiplier ($\Delta Y/\Delta X$) as

$$\frac{\Delta Y}{\Delta X} = \frac{1}{1 - b + g}$$

...(13.14)

Equation (13.14) can be alternatively written as

$$\frac{\Delta Y}{\Delta X} = \frac{1}{1 - (b - g)}$$

...(13.15)

where $b$ is MPC and $g$ is MPM (marginal propensity to import).

Eq. (13.15) reveals that if $b = g$, then foreign trade multiplier will be equal to unity.

---

**Check Your Progress**

1. What are non-economic products?
2. Define expenditure method of determining national income.
3. State the fundamental principle of the Keynesian theory of national income.

---

**13.4 SECTORAL AND POPULATION DISTRIBUTIONS**

There are basically three sectors into which the Indian economy is classified:

- Agriculture and allied
- Industry and Services

The agriculture sector comprises the following:

- Agriculture (agriculture proper & livestock)
- Forestry and logging
- Fishing and related activities

The industry sector comprises the following:

- Manufacturing (both registered and unregistered)
- Electricity
- Gas
- Water supply
- Construction

The services sector comprises the following:

- Trade
- Repair
• Hotels and restaurants
• Transport
• Storage
• Communication and services related to broadcasting
• Financial, real estate and professional services
• Community, social and personal services

The sectoral distribution of income shows the percentage distribution of national income or GDP among different sectors viz., agriculture (primary) industry (secondary) and services (tertiary).

Table 13.2: Annual real GVA and GDP growth (per cent)

<table>
<thead>
<tr>
<th>Services</th>
<th>2014-15</th>
<th>2015-16</th>
<th>2016-17</th>
<th>2017-18 (tea base)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture, forestry &amp; fishing</td>
<td>4.2</td>
<td>6.7</td>
<td>6.9</td>
<td>2.1</td>
</tr>
<tr>
<td>Industry</td>
<td>7.5</td>
<td>8.8</td>
<td>5.8</td>
<td>4.4</td>
</tr>
<tr>
<td>Mining &amp; quarrying</td>
<td>11.1</td>
<td>10.5</td>
<td>11.0</td>
<td>7.4</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>8.1</td>
<td>10.8</td>
<td>11.8</td>
<td>4.3</td>
</tr>
<tr>
<td>Electricity, gas, water supply &amp; other utility services</td>
<td>1.1</td>
<td>3.0</td>
<td>3.3</td>
<td>1.5</td>
</tr>
<tr>
<td>Construction</td>
<td>4.7</td>
<td>5.0</td>
<td>1.7</td>
<td>5.6</td>
</tr>
<tr>
<td>Services</td>
<td>9.7</td>
<td>9.7</td>
<td>7.7</td>
<td>8.3</td>
</tr>
<tr>
<td>Trade, Hotel, Transport, Storage, communication &amp; service related to broadcasting</td>
<td>0.0</td>
<td>10.2</td>
<td>7.4</td>
<td>8.7</td>
</tr>
<tr>
<td>Financial, and computer &amp; related services</td>
<td>11.1</td>
<td>10.0</td>
<td>5.7</td>
<td>7.2</td>
</tr>
<tr>
<td>Public administration, defence &amp; other services</td>
<td>0.1</td>
<td>6.9</td>
<td>10.3</td>
<td>9.1</td>
</tr>
<tr>
<td>GVA at basic prices</td>
<td>7.2</td>
<td>7.9</td>
<td>6.6</td>
<td>6.3</td>
</tr>
<tr>
<td>GDP at market prices</td>
<td>7.5</td>
<td>8.8</td>
<td>7.1</td>
<td>6.3</td>
</tr>
</tbody>
</table>

Source: Based on data from CSO

Population Distribution

This is basically concerned with the distribution of national income among its population.

As per the Census 2011, the workers have been classified by the type of economic activity into broad the type of economic activity into nine broad categories as per National Industrial Classification, 1998. Distribution of main workers by industrial category shows that agriculture sector still employs largest number of workers. The dependence on agriculture is brought out by the fact that of the 313 million main workers in the country, 166 million (56.6%) has been engaged in ‘Agricultural and allied activities’. This is followed by ‘Manufacturing Industries’, which employed about 42 million (13.4%). There are 31.1 million workers in the services sector forming 10% of the total main workers with similar number engaged in ‘Wholesale retail trade and repair work, Hotel and restaurant.

Based on research from Forbes and economist Chancel-Piketty’s paper ‘Indian Income Inequality, 1922-2014: From British Raj to Billionaire Raj’., it...
has been observed that from a share of 6% of national income in 1982-83, the share going to the top 1% reached nearly 22% in 2013-14, of which nearly 9% went to the top 0.1%, nearly 4% to the top 0.01%, and almost 2% to the top 0.001%. Clearly, income concentration had increased rapidly since the early 1980s. The remaining 90% of the population suffered significant declines in income share.

The middle 40% of the national income distribution possessed 46% of total income in 1982-83, declining to 30% in 2013-14. The bottom 50% of the population worsened from nearly 24% in 1982-83 to 15% in 2013-14.

13.5 PER CAPITA INCOME: DEFINITION, AND CALCULATIONS

Per capita income or average income measures the average income earned per person in a given area (city, region, country, etc.) in a specified year. It is calculated by dividing the area’s total income by its total population.

Per capita income is national income/total population.

Per capita income is often used to measure an area’s average income. This is used to compare the wealth of one population with those of others. Per capita income is often used to measure a country’s standard of living. It is usually expressed in terms of a commonly used international currency such as the Euro or United States dollar, and is useful because it is widely known, is easily calculable from readily available gross domestic product (GDP) and population estimates, and produces a useful statistic for comparison of wealth between sovereign territories. This helps to ascertain a country’s development status. It is one of the three measures for calculating the Human Development Index of a country.

13.5.1 GDP and GNP

We have briefly discussed these concepts earlier in the unit.

Gross National Product (GNP)

Of the various measures of national income used in national income analysis, GNP is the most important and widely used measure of national income. It is the most comprehensive measure of the national productive activities in an open economy. The GNP is defined as the value of all final goods and services produced during a specific period, usually one year, plus incomes earned abroad by the nationals minus incomes earned locally by the foreigners. The GNP so defined is identical to the concept of gross national income (GNI). Thus, \( GNP = GNI \).

The difference between the two is only of procedural nature. While GNP is estimated on the basis of product-flows, GNI is estimated on the basis of money income flows, (i.e., wages, profits, rent, interest, etc.).
Gross Domestic Product (GDP)

The Gross Domestic Product (GDP) is defined as the market value of all final goods and services produced in the domestic economy during a period of one year, plus income earned locally by the foreigners minus incomes earned abroad by the nationals. The concept of GDP is similar to that of GNP with a significant procedural difference. In case of GNP, incomes earned by the nationals in foreign countries are added and incomes earned locally by the foreigners are deducted from the market value of domestically produced goods and services. But, in case of GDP, the process is reversed—incomes earned locally by foreigners are added and incomes earned abroad by the nationals are deducted from the total value of domestically produced goods and services.

Net National Product (NNP)

NNP is defined as GNP less depreciation, i.e.,

\[ NNP = GNP - \text{Depreciation} \]

Depreciation is that part of total productive assets which is used to replace the capital worn out in the process of creating GNP. Briefly speaking, in the process of producing goods and services (including capital goods), a part of total stock of capital is used up. ‘Depreciation’ is the term used to denote the worn out or used up capital in the process of production. An estimated value of depreciation is deducted from the GNP to arrive at NNP.

The NNP, as defined above, gives the measure of net output available for consumption and investment by the society (including consumers, producers and the government). NNP is the real measure of the national income. \( NNP = NNI \) (net national income). In other words, NNP is the same as the national income at factor cost. It should be noted that NNP is measured at market prices including direct taxes. Indirect taxes are, however, not a part of actual cost of production. Therefore, to obtain real national income, indirect taxes are deducted from the NNP. Thus, \( NNP - \text{indirect taxes} = \text{National Income} \).

Some Accounting Definitions

(a) Accounting Identities at Market Price

\[ GNP = GNI \text{ (Gross National Income)} \]
\[ GDP = GNP - \text{net income from abroad} \]
\[ NNP = GNP - \text{depreciation} \]
\[ NDP \text{ (Net Domestic Product)} = NNP - \text{net income from abroad} \]

(b) Some Accounting Identities at Factor Cost

\[ GNP \text{ at factor cost} = GNP \text{ at market price} - \text{net indirect taxes} \]
\[ NNP \text{ at factor cost} = NNP \text{ at market price} - \text{net indirect taxes} \]
NOTES

13.5.2 Uses and Limitations

In this section, we will discuss the uses and limitations of per capita income.

Uses of Per Capita Income

Per capita income is a great measure of ascertaining the financial stability of two or more regions through the comparison of the different incomes. In fact, regional income data is mostly calculated through per capita income because it gives an accurate picture of the average level of income of the population of the particular area. The per capita income related comparisons can be used for economic stability, investment and aid related appeals in the international community.

- **Wealth and Stability**: Per capita income is the ratio of the amount of the region’s income divided by the population of the region. This indicates that if this ratio rises then the population of the region on an average is more stable and comfortable than in the past. Similarly, if this ratio decreases then it shows that the income levels of the region has been on the decline. This follows the assumption that the inflation remains the same.

- **Aid**: There are a lot of governmental level aid-related agency which provide financial assistance to member countries in need. Since the per capita income showcases the prosperity of the region, it will help these agencies gauge which regions are in need of assistance. For example, if the income of country A is higher than country B, and the price of rice is higher in country A than in country B, then the population of country B are just as prosperous as the population in country A. However, if the per capita incomes are same in both countries A and B and the price of rice is higher in country A, then the increased cost will be an issue. This can be dealt with through financial assistance to the financial worse of country which is country A in this case.

Limitations of Per Capita Income

- The increase in the per capita incomes are due to increase in prices and not physical outputs and this is why it is not really a credible source of measurement.

- Since per capita income is an average, then the national data will cover up the regional variations that might arise in terms of income, employment, and output of the population per head.

- Per capita income, as mentioned before, is an average. This means that even if the average income increases there are chances that the inequalities within the population grows, where the income of one section keeps on increasing while the other section’s becomes stagnant or lower in comparison.
• The per capita income is an economic measure therefore, it might show the increase of incomes but does not reveal the cost at which it is achieved. For example, it might be possible that in order to increase production working hours might have been increased for labour which affects the actual prosperity of the population in qualitative terms.

• Certain indicators of prosperity for example, improvements in life expectancy, health care or education is not reflected through economic measurements like GDP.

13.5.3 Recent Developments in Indian Economy

As per the Economic Survey 2017-18:

• A series of major reforms undertaken over the past year will allow real GDP growth to reach 6.75 percent this fiscal and will rise to 7.0 to 7.5 percent in 2018-19, thereby re-instating India as the world’s fastest growing major economy.

• Due to the launch of transformational Goods and Services Tax (GST) reform on July 1, 2017, resolution of the long-festering Twin Balance Sheet (TBS) problem by sending the major stressed companies for resolution under the new Indian Bankruptcy Code, implementing a major recapitalization package to strengthen the public sector banks, further liberalization of FDI and the export uplift from the global recovery, the economy began to accelerate in the second half of the year and can clock 6.75 percent growth this year.

• India can be rated as among the best performing economies in the world as the average growth during last three years is around 4 percentage points higher than global growth and nearly 3 percentage points higher than that of Emerging Market and Developing Economies.

• The GDP growth has averaged 7.3 per cent for the period from 2014-15 to 2017-18, which is the highest among the major economies of the world. That this growth has been achieved in a milieu of lower inflation, improved current account balance and notable reduction in the fiscal deficit to GDP ratio makes it all the more creditable.

• The agenda for the next year consequently remains full: stabilizing the GST, completing the TBS actions, privatizing Air India, and staving off threats to macro-economic stability. The TBS actions, noteworthy for cracking the long-standing “exit” problem, need complementary reforms to shrink unviable banks and allow greater private sector participation.

• Three areas of policy focus stand out: Employment: finding good jobs for the young and burgeoning workforce, especially for women. Education: creating an educated and healthy labor force. Agriculture: raising farm productivity while strengthening agricultural resilience. Above all, India must continue improving the climate for rapid economic growth on the strength of the only two truly sustainable engines—private investment and exports.
As per Budget 2018-19:

- Budget 2018-19 reflects the Government’s firm commitment to substantially boost investment in Agriculture, Social Sector, Digital Payments, Infrastructure and Employment Generation on the one hand and simultaneously stick to the path of fiscal rectitude by aiming for a reduction of FD by 0.2% of GDP over RE 2017-18.

- Various economic reforms were undertaken in the year which include: implementation of the Goods and Service Tax, announcement of bank recapitalization, push to infrastructure development by giving infrastructure status to affordable housing, higher allocation of funds for highway construction and greater focus on coastal connectivity.

- Other sectoral initiatives undertaken include: measures to revive the construction sector and promotion of exports in textile and apparel industry. Apart from these, the measures that were taken by the Government in the previous years to boost manufacturing, employment generation, improving ease of doing business and transparency via schemes such as Make-in-India, Skill India, direct benefit transfer and measures for financial inclusion were also taken forward in 2017-18.

Check Your Progress

4. What are the components of industry sector?

5. Name the tool which is used as one of the three measures for calculating the HDI of a country.


13.6 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS

1. Non-economic products include services rendered to self, to family, to relations and to neighbours. Non-economic products are not included in national income.

2. The expenditure method of national income calculation in which the national income is calculated at the final expenditure stages.

3. The fundamental principle of the Keynesian theory of national income determination is that the national income is determined by aggregate demand and aggregate supply and equilibrium level of national income is determined at the level where aggregate demand equals aggregate supply.

4. The industry sector comprises the following: manufacturing, electricity, gas, water supply and construction.
5. Per capita income is the tool which is used as one the three measures for calculating the HDI of a country.

6. NDP at factor cost is equal to NDP at market price less net income from abroad.

13.7 SUMMARY

- National income is defined as the money value of all final goods and services produced over a period of one year. Alternatively, national income is also defined as the sum of factors income (wages, rent, interest and profit) generated over a period of one year.

- National income is measured in terms of (i) Gross Domestic Product (GDP), and (ii) Gross National Product (GNP).

- GDP is the measure of the total market value of all final goods and services produced in the domestic economy during a period of one year plus income earned by the foreigners in the country minus income earned by countrymen from abroad.

- GNP is the measure of the total market value of all final goods and services produced in the domestic economy during a period of one year plus income earned abroad by the citizens minus income earned by the foreigners in the country.

- Three methods are used to measure the national income: (i) Net product or Value Added Method, (ii) Factor-income method, and (iii) Expenditure method.

- Under Net Product or value-added method, national income is estimated in three stages: estimating gross value of domestic output, (ii) estimating cost of production, and estimating depreciation of capital. National income is worked out by deducting cost of production and depreciation from the gross value.

- Under factor income method, national income is estimated by adding up all factor incomes. Thus, National Income = Rent + Wages + Interest + Profits. In general, national income is estimated by working out and adding up labour income, capital income and mixed income.

- Under expenditure method, national income is estimated by adding up (i) private consumption expenditure, (ii) direct tax payments, (iii) payments as donation, and private savings.

- Theory of national income determination was formulated by J. M. Keynes. According to Keynes, the equilibrium level of national demand is determined by aggregate demand for and aggregate supply (AS) of goods and services.
Aggregate demand (AD) refers to the total expenditure made by the people of country including households, government, and foreigners. Aggregate supply (AS) refers to the total supply of consumer goods and services and capital goods (machinery, equipments, etc.).

Keynesian theory of income determination is explained under three models: (i) two-sector model including only two sectors—households and firms; (ii) three-sector model including households, firms and government; and (iii) four-sector model including households, firms, government and foreign sector.

Sectoral distribution of income is simply the distribution of income amongst the three sectors of the economy. Population distribution of income reflects the manner in which income is distributed amongst the population of the country.

Per capita income or average income measures the average income earned per person in a given area (city, region, country, etc.) in a specified year. It is calculated by dividing the area’s total income by its total population.

13.8 KEY WORDS

- **National income**: It is defined as the sum of money value of all final goods and services produced in a country over a period of one year.
- **Theory of national income determination**: It refers to the theory which explains how equilibrium level of national income is determined.
- **Aggregate supply**: It refers to the total value of goods and services produced and supplied in an economy per unit of time.
- **Aggregate demand**: It is an ex-post concept which implies effective demand which equals actual expenditure.
- **Per capita income**: It measures the average income earned per person in a given area in a specified year.

13.9 SELF ASSESSMENT QUESTIONS AND EXERCISES

**Short Answer Questions**

1. Define intermediate and final products.
2. What is the difference between national income and private income?
3. List the varied types of income included in the factor income method.
4. Write a short note on sectoral and population distribution of income.
5. What are the limitations and uses of per capita income?
Long Answer Questions

1. Explain the different methods of measurement of national income.
2. Examine, in detail, the Keynesian theory of income determination through different models.
3. Describe the concept of GDP, GNP and NNP.

13.10 FURTHER READINGS


Websites

https://blogs.economictimes.indiatimes.com/et-commentary/heres-how-indias-widening-income-distribution-can-be-redressed/
UNIT 14 ECONOMIC REGULATIONS OF BUSINESS

14.0 INTRODUCTION

The economic framework within which the firm functions plays a crucial role in its efforts to thrive, grow and beat its competitors. Since business is fundamentally an economic activity, the economic environment, both within the country as well as beyond its national frontiers will have a definite influence on its fortunes. It is inconceivable that there could be any institution, bodies or persons who are insulated from the effects of the economic environs. Business has to deal with a vast number of governmental bodies, rules, regulations, and guidelines relating to its statutory responsibilities, the capital market, sources of finance including stock market options, venture capital, offshore funds, disinvestment options, bank funding and so on. In this unit, we will study the concept of economic regulations of business.

14.1 OBJECTIVES

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

- Discuss the antitrust theory and regulations
- Describe the structure-conduct-performance paradigm
- Explain the concept of concentration and its measurement
- Examine externalities and its regulations in managerial economics
14.2 ANTITRUST THEORY AND REGULATIONS

Antitrust laws are often mentioned as ‘competition laws’. These are laws established by the governments to protect the consumers from predacious practices of business. The laws ensure existence of impartial competition in economy which is based open-market. They have developed as the markets have evolved and are in keeping with the recent trends. They safeguard the markets and the consumers from prospective monopolies which will become a cause of commotions to the productivity and competitive flow.

14.2.1 Antitrust Theory

The classical viewpoint on competition maintained that particular agreements and practices of business may result in unreasoned restriction over people’s freedom in business attempting to earn a livelihood. Restrictions were adjudged as acceptable or unacceptable by courts as fresh incidents came to light with the changes in the business environment. Therefore, the courts established exact classifications of agreement, explicit articles, towards foul play of their code on fair economic practices, and courts did not scheme for a central notion of market power. Initial theorists of the field like Adam Smith had precluded power of monopoly on such grounds. He was of the view that any form of monopoly practiced in the market has similar impact as that of a secret in business or the producers. It does not matter how the producers monopolise the market, they could under supply or sell their products more than the actual price and increase their profits.

In his book *The Wealth of Nations*, Adam Smith observed problems related to cartel, but he did not preach any legal methods to curtail them, he observed: ‘People of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices. It is impossible indeed to prevent such meetings, by any law which either could be executed, or would be consistent with liberty and justice. But though the law cannot hinder people of the same trade from sometimes assembling together, it ought to do nothing to facilitate such assemblies; much less to render them necessary.’ At the same time, Smith overruled the mere presence of any kind of corporations whether big, dominating or foul.

Large corporations had become an integral part of market economy towards the middle of the nineteenth century. During this time John Stuart Mill submitted his approach in his book treatise *On Liberty* published in 1859. In his approach, Mill regarded business as a social activity and thought all traders have to consider the interest of others and entire society. Producers supply good as well as bad quality goods and this leaves the buyers to find other alternatives. He preached the doctrine of Free Trade; this was not on the same lines as individual liberty which had been emphasized earlier. All types of restrictions are restraints and these are not good for the trade.
Economic Regulations of Business

NOTES

The Neo-classical viewpoint was a shift from the views of Mill and it focused on the model of competition which was based on free markets and services provided in markets which were free from competition provided maximum social good. According to this model, firms were allowed to enter the market without restrictions and freely compete with firms functioning in the market. The economists felt that in the absence of competition, markets function with greater efficiency. This model is completely contrasting with monopolies and cartels. Conventional economists completely agreed that there was no such thing as perfection in competition hence the objective is always to attain a state of competition which is workable and provided maximum efficiency. After both these contradictory theories, the economists and lawyers connected with the Chicago University advocated an approach which worked towards establishing laws for competition.

Robert Bork was not much in favour of decisions taken by court on antitrust laws and he voiced his opinions in his book *The Antitrust Paradox*. There are views of Harvard Professor Philip Areeda who was keen for the court to adopt more forceful laws and policies over antitrust.

14.2.2 Activities of Business which Require the use of Antitrust Laws

There are numerous debatable activities of business on which the Antitrust laws are used, these include:

- **Allocation of Market**: Suppose company B functions in eastern part of the city and company C functions in western parts of the city so they would continue to function in their specified areas. This will give kind of monopoly to both in their respective areas as cost of setting up a new company is very high hence no new company will provide competition.

- **Bid Rigging**: In this scenario three corporations join hands and form a cartel. In this they equally permit all members to have an equal share and do not try to encroach over others share, thus they do not compete against each other. There is no competition in the market.

Bid rigging further have three sub-types:

(i) **Bid Suppression**: Participants abstain to bid against each other by withdrawing from the bid so that the decided bidder can win

(ii) **Complementary Bidding**: This type is also called courtesy bidding, this takes place in situations where the applicants purposely submit much higher bids and allow the decided bidder to quote less and accommodate all the demands of the buyer. These are most prone to rigging and are developed to provide a false picture for the buyer about the market environment.

(iii) **Bid Rotation**: As the word suggests in this type the participants win the biddings in rotation and follow a cycle so each gets to win.

- **Price Fixing**: Company A and B are the sole producers of a product. The products are completely identical and consumers have no preference other
than the price of the product. Both the companies sell their products at the same price so that they can avoid the conflict because of price. Selling the product at the same price helps in maintaining the margin and results in the customers paying more for the product which they could have got at a lower price.

- **Monopolies:** The formulation of antitrust laws occurred in order to avoid monopolies. Among the most prominent legal cases of antitrust was of Microsoft Corporation, it was charged with undertaking actions which encouraged its monopoly in market and were against the competitive norms of the market. The corporation was compelling its users to install their web browsers for the operating system of Windows.

Controllers have to make sure that monopolies are prevented from cultivating an environment which is obviously competitive and market share should be earned with the help of business expertise and new inventions. Earning market share by illegal means is considered wrong.

Following are some of the arrangements which lead to monopolising market share in an illegal manner and are subjective to action by law:

- **Exclusive Supply Agreements:** This happens in case a supplier is stopped from retailing to any other buyer other than the ones mentioned in the contract. These type of agreements kills all the competition from the market and helps in establishing monopoly.

- **Securing the Sale of Two Products:** In cases where a producer has a strong monopoly over a product in the market and desires to acquire monopoly over a second product, the producer in order to acquire monopoly on the second product ties the sale of strong product with the new one and compels the buyers to purchase the new product as well even if it is not needed by them. This compelling is against the antitrust laws.

- **Predacious Pricing:** Though this practice of establishing monopoly is not easy to bring out in the open and, the firm may keep prices low for a long time and maintain the market share so that it can overcome the losses in the long run.

- **Rebuff Deals:** Companies who have monopoly over the market are able to decide the firms with whom they wish to have business dealings. In instances where they use their dominance to curtail competition they are liable for violating the antitrust laws.

- **Mergers and Acquisitions:** The antitrust laws cover the sphere of all the mergers and new businesses which are acquired. There are divided into three types:
  - **Horizontal Mergers:** Here the two companies which are rivals share the same product line and markets to become one single firm.
NOTES

(a) **Vertical Mergers**: These are mergers which are undertaken when two companies at different levels of the production of same product merge their operations.

(b) **Conglomerate Mergers**: These are mergers where between companies operating in different industries for synergy, product extension and market extension benefits.

**Competition Laws in India**

India was among the first developing countries to have a competition law in the form of the Monopolies and Restrictive Trade Practices (MRTP) Act, 1969. The MRTP Act was designed to check concentration of economic power, prohibit restrictive or unfair trade practices and control of monopolies. Then came 1991, which has proved to be a watershed in the history of India’s economic development. The new India required new rules. Hence, the need for a new competition law. Accordingly, the Competition Act was passed in 2002 and amended in 2007. The Competition Commission of India (CCI) was established on March 1, 2009 as an autonomous body comprising of a Chairperson and six members. An appellate body Competition Appellate Tribunal was also set up in May 2009 with final appeal lying to the Supreme Court of India. Subsequently, the MRTP Act was repealed, MRTP Commission established under that act was abolished and its pending cases were transferred to the CCI. Sections 5 & 6 relating to Merger & Acquisition were notified in June, 2011.

The Competition Act, 2002, as amended by the Competition (Amendment) Act, 2007, follows the philosophy of modern competition laws. The Act prohibits anti-competitive agreements, abuse of dominant position by enterprises and regulates combinations (acquisition, acquiring of control and M&A), which causes or likely to cause an appreciable adverse effect on competition within India. The sectors that have been covered are as diverse as infrastructure, finance, entertainment, IT, telecom, civil aviation, energy, insurance, travel, automobile manufacturing, real estate and pharmaceuticals etc.

14.3 THE STRUCTURE – CONDUCT – PERFORMANCE PARADIGM

Structure-Conduct-Performance (SCP) framework suggests a connecting affiliation among the market’s structure (S), the conduct (C) of various firms of that market and their performance (P) in the economy. The paradigm was initially suggested by E. Mason during the 1930s and was further developed in 1950s by J Bain. The neo-classical theory of firm formed the basis of the paradigm and since then it has been the focal point if all the industrial organization studies.

The systematic framework which is used to create a relationship between the structure of the market, conduct of the market and performance of the market
is called Structure, Conduct and Performance paradigm (SCP). As mentioned earlier, in 1959 American Economist Joe S Bain Jr developed the framework and he had included the SCP in his book *Industrial Organization*. The paradigm created by him was of the structure, conduct, and performance in connection to the industry and was different from firms and industries owned by individuals. It suggested and verified blockades to access as a determining factor of industry performance. Market concentration had been examined in connection to power of the market and rate of profit.

Since it was developed, in the field of economics and all fields of business management and controlling the SCP paradigm is regarded as mainstay of the theory of industrial organization. It offers the starting point for analysing various markets and industries. The paradigm helped a lot of researches and academics of economics to base their work on the foundation of SCP, even American academic Michael E. Porter built his works on competition, theories on economics and business strategy on the principles of the paradigm.

On the reasoning of the paradigm it can be understood that performance of industries (that are working towards social good) are dependent on the conduct of the firms functioning in the confines of the specific industry and they are further dependent on the market structure.

### 14.3.1 Elements of the Paradigm

Let’s look at each of the elements of the paradigm in this section.

**Structure:** It involves the fixed aspects which remain considerably same with passing time and result in impacting the behaviour of both suppliers and consumers or either one of them. The system where the market is unable to follow a condition of perfect competition is dependent on extent of factors such as concentration of supply and demand both respectively, product differentiation and barriers in entering the market. Furthermore, the product’s nature and the technology used for manufacturing the product would also play a significant role in determining the structure of the market.

**Conduct:** This refers to the behaviour of the consumers and retailers, their behaviour with each other as well as between themselves have an impact. The behaviour of the firms is based on their strategy which is decided after researching aspects like investment, development, levels of advertisement and collusions.

**Performance:** The third element of the paradigm is performance which is evaluated on the basis of the comparison drawn from the outcomes of all the firms of the industry in terms of the level of efficiency and diverse ratios are employed to evaluate the varied levels of profitability. The varied levels taken under consideration are level of pricing, quality of the product, distribution of resources, levels of efficiency during production and many more.

The markets are made up of retailers and consumers; the ever-changing behaviour of both the constituents has an impact on the markets. The dynamism of
their behaviour makes it difficult to predict and develop a market structure which is consistent. Several problems come up while attempting an explanation of the paradigm and these problems are not just because of changing behaviours in the market but more importantly they are faced as a result of inadequate data and the manifold classifications and addition of markets. Essentially, the real issue in applying the SCP technique for analysing the market and an industry is that there is no demarcated boundary of the particular industry.

As a result of such scenarios of policies of economics, the Harvard school has recognized the power of market as risky, and established a relationship between, the concentration ratio, and the damaging impact on social well-being.

![Fig. 14.1 The SCP Paradigm](image)

14.3.2 Criticisms of the SCP Paradigm

The paradigm provided basis to a number of theories of economics, yet it is not free from criticism:

The paradigm of SCP is drawn from microeconomic theory which is used to observe the pragmatic behaviour of firms. Conversely, economic theory is not able to provide the accurate relationships prevailing among the three elements. For instance, oligopoly theory is considered to be highly uncertain and it fails to generate any precise conclusions. There are many indicators for measuring the structure yet the paradigm emphasis only on concentration. The concern of the SCP is focused on the equilibrium of short run and does not take into account the future evolutions and changes in structure, on both accounts the conduct and performance will also undergo changes. Another criticism directed towards SCP is also about its loose derivation of the theoretical foundations, and as a result it is often referred along with the neoclassical theory.

### Check Your Progress

1. What is complementary bidding?
2. When was the Competition Act passed in India?
3. What is conduct in the SCP paradigm?
14.4 CONCENTRATION: OVERVIEW AND MEASUREMENT

Economic concentration has several denotations in the industrial structure and is often applied to various ‘indexes of concentration’ in order to understand the distinctive features of the structure in industries. According to the economic theory, concentration is a vital factor as it determines the behaviour of the market and also affects its outcome. All other things being equal, when few big firms are responsible for supplying chunk of the market demand it is highly possible for monopolies to be established in comparison to when large firms do not supply the bulk demand. Therefore, in the clarification of business policy, the features of an industry articulated in the concentration index are expected to have an essential role. This connection with extent of monopoly has encouraged several pragmatic studies to be covering the aspect of measurement of concentration.

Industrial concentration refers to retailer concentration. It means that the market is dominated by few large firms and they control manufacturing and selling of products. Two factors affect the extent of industrial concentration, one factor is quantity of firms active in a particular market, and the other factor is percentage of the share of the market demand fulfilled by the firm. In case the number of firms in the market is low then the share of the firms would be large and these big firms would also have larger control on the overall supply. This being the case then the scenario would be referred as high quantity of retailer/industrial concentration. The market power of each firm decides the extent of its quantity industrial concentration. The capability of a firm or retailer to affect the product’s cost in the market is referred to as market power. Market power of a firm is zero in a market situation with perfect competition, thus even the industrial concentration is zero. Industrial concentration is highest in a market where monopolistic practices are followed.

14.4.1 Measuring Industrial Concentration

Keeping a measurement of industrial concentration is essential as it clearly indicates the nature of the market. When there are more number of firms, the market has the character of perfect competition and less number of firms will automatically increase the measurement of the concentration. Another reason for keeping a track on the extent of concentration is for highlighting the percentage of concentration. For instance, if the Concentration Ratio (CR) is 54.2 percent it clearly points out that production of five firms is above fifty percent in a particular industry.

The ways of measuring industrial concentration has two classifications namely absolute and relative.
Absolute Measures of Concentration

The fundamental fact of this classification is that the number of firms is absolute in terms of manufacturing, service or capital in the industry. There are two measurements for this:

(i) The Concentration Ratio (CR): This ratio measures the share of manufacturing, sales and employment of three or five largest firms. In the illustration the analysis of the manufacturing is taken. The share of ten firms is taken into account; assume there are forty firms in total. Total manufacturing of five largest firms is as given below:

Firm 1 manufactures 10,000 Parts
Firm 2 manufactures 12,000 Parts
Firm 3 manufactures 9,000 Parts
Firm 4 manufactures 15,000 Parts
Firm 5 manufactures 14,000 Parts
Total Production of the five biggest firms – 60,000 Parts
Total Production in Industry—1,10,000

Concentration Ratio of the Five firms – (60,000/1,10,000) × 100 = 54.54%
Thus the five firms have an industrial concentration of 54.54%

(ii) The Herfindahl Index: This is another absolute measure for industrial concentration, it is also referred to as Herfindahl–Hirschman Index, HHI, or even as HHI-score. It gives the measure of magnitude of firms in comparison to others in the industry and also indicates the quantity of competition between the firms. It takes its name from its creators Orris C. Herfindahl and Albert O. Hirschman, both are renowned economists and the index is often used cases of in competition law.

In this all firms existing in the industry are taken into account. H-index is explained below:

Assume,

n = Number of firms in the industry.

x = Absolute amount of firm manufacturing.

T = Total manufacturing of the industry.

As stated above the H index takes into account all firms of the industry while calculating. The index value is zero to one. In case the industry has more firms of similar size then the index value is taken as zero. On the other hand, in case industry has one firm then the value is one. The H-index has been explained in the following illustration, assume an industry has four firms totally and each firm manufactures 500, 350, 300, and 200 parts respectively. This makes the value of T as 1350 that means share of each
firm in total manufacturing is \( \frac{x_1}{T} \) that makes the share of each firm as 0.372, 0.262, 0.222 and 0.152 as a result, the value of H index is 0.275. Since this is very close to zero, hence it is clear that concentration levels are low in the industry. The index helps in showing the extent of concentration for all periods.

- **Relative Measures of Concentration**

  The inequality of firms is focused in this measure. This implies that markets which are constituted with firms of uneven proportions have uneven productions as well. The approximate industrial concentration may be measured by using the Lorenz Curve. The Lorenz curve provides the representation of wealth distribution with the help of a graph. In 1905, Max O. Lorenz developed this curve to represent the disparity in the distribution of wealth. In mathematics it is referred as Gini-coefficient.

  Gini Concentration Ratio is the statistical measure centered on Lorenz Curve. The value of Gini coefficient is 0 to 1 in a Gini coefficient of inequality. In situations where all the firms are of same size then the line representing equality and Lorenz Curve will be similar. The value of Gini co-efficient will be zero in such a situation. Increased inequality will result in increasing the value of the Gini coefficient.

**14.4.1 Monopoly Vs Concentration of Economic Power**

A firm’s control over production and distribution of a product is referred to as monopoly. Monopolistic environment may be created in large as well as small scale industries. On the other hand, concentration of economic power refers to centrally controlling essential economic activities such as availability of capital, employment and manufacturing in a major industry or agriculture etc.

The concentration of economic power in India is achieved by promoting monopoly. In fact, it is the means to acquire power and then promote monopolistic activities so that the power can be maintained. The aim of all modern economies is to diffuse concentration of economic power as it hampers social justice.

**14.5 REGULATION OF EXTERNALITIES**

In business, once capital has been sourced, the investors need to determine how much capital will be needed to produce a given output. This is revealed through the capital output ratio. Capital output ratio is the ratio of capital to output over a period of time. In order to calculate this ratio, we need to know the total output produced by a country. But it is known that the production, consumption and investment decisions relating to the output are bound to lead to consequences that are not directly exposed or experienced? These indirect and unseen outcomes may exist and can become problematic when they create large problems. These problems or side effects are referred to by the economists as externalities.
Externalities are one of the reasons that government intervention is particularly required in economic affairs of the state. Also, this is why there is a difference between the private returns of a commodity (to the owners of factors of production) and the costs of the commodity to a society as a whole. Environmental pollution is a good example of an externality. Additionally, externalities are not always negative. For example, a company may be funding a research conducted by its Research and Development Department, but the findings of the research will have positive effects on the society as a whole. Externalities help the government in determining factors such as taxation, grant of subsidies, permits and licenses.

An externality exists when the actions or decisions of one person or group impose a cost, or bestow a benefit on some second or third parties. Externalities are sometimes called spillovers or neighbourhood effects. Inefficient decisions result when decision makers fail to consider social costs and benefits.

The presence of externalities is a significant phenomenon in modern life. Examples are everywhere: air, water, land, sight, and sound pollution; traffic congestion; automobile accidents; abandoned housing; nuclear accidents; and passive smoking are only a few of them. Because there are so many externalities that affect the environment, the study of externalities is a major concern of environmental economics.

Externalities are costs and benefits of economic activity that do not enter into the firm’s own calculations of costs or revenues. Thus, in the absence of some sort of non-market correction (usually government action), they lead to inefficient allocation of resources.

External costs and benefits are at the core of the common resource problem. A new producer creates higher costs for all previous entrants, or all producers impose external costs on the general population. In either case, in the absence of regulation, taxes, or property rights for environmental quality, external costs are not borne by the producers who cause them. It is obvious then that the prices of their products do not reflect the social costs of production. Thus, more of these resource-depleting or -polluting goods and services are produced and consumed than would be the case if prices reflected external costs. Hence, societies pollute more than their people would choose if markets reflected all social costs.

In a market with competitive producers, the supply curve represents private marginal costs. But, if this is a polluting industry, the external costs would make the social marginal cost, SMC, higher. If these costs were reflected in the market, the price would jump higher and demand, and therefore, output would be reduced. As less of the offending product would be grown or manufactured, there would be less environmental degradation.

Polluters impose costs on others. If these external costs are reflected in the firms’ costs, the social marginal cost curve, SMC, would prevail; the market price would increase; and output will reduce. But, because the firms do not bear these
costs, their private marginal cost curve, is lower. So, more Q1 of the polluting product is produced and consumed.

Types of Externalities

(I) Detrimental (negative) externalities: Detrimental externalities impose costs on other outside firms, but do not cost the firm anything. Beneficial externalities create benefits outside the firm, but provide no revenue to the firm.

A typical example of a detrimental externality is the pollution. It can be explained as follows:

- A firm’s production process may send smoke into the air. The firm sustains no cost to ‘use’ the air this way. It actually is cheaper for the firm to do so than to make the effort and pay the expense to clean up its exhaust.
- The surrounding area, however, sustains the cost of enduring and cleaning up after the effects of the firm’s smoke. Therefore, there is a cost of this pollution, but it is not paid by the polluting firm; it is imposed on someone else (neighbours).

Social and private costs: The social cost of the polluting firm’s production is the sum of its internal costs (private costs, which the market imposes on the firm) and the external social costs imposed on others by its pollution.

- One way or another, society at large has to pay the full social cost in order for the firm’s production to occur. The market provides no reliable way to force the firm to pay anything, but the private cost portion of total social cost, because there are no private property rights to the air.
- The cost of the detrimental externality is not necessarily imposed by the market upon the firm’s customers, either. They, too, may benefit from the firm’s lower costs.
- If the total social cost could be imposed upon the firm, its cost would be higher and its cost curves would rise. The result would be lower output, at a higher price, reflecting the true social cost of the output.

(II) Beneficial (positive) externalities: Beneficial or positive externalities are positive effects upon others due to private economic activity. For example, a flower farm exists for the purpose of making profit by growing and selling flowers. However, it also is a very attractive thing to see and smell, and so, produces positive externalities for its neighbours and passersby.

- The market does not provide the firm (flower farm) any way to collect the extra value to society of the beneficial externality, even though the firm is responsible for producing it.
- If the firm could collect the value of the beneficial externalities it produces, its perceived demand and marginal revenue would increase, and it would produce more.
Remedies for Externalities

The market provides no remedy for the misallocation of resources caused by externalities. Therefore, non-market means have to be found. The most common one is government action.

- **Government regulation:** In many areas, state and central government apply regulations that force some firms to 'internalize' negative externalities. Significant examples are anti-pollution laws that require firms to pay the costs of remedying negative pollution externalities.

- **Taxation:** Government also can impose a tax upon firms creating negative externalities. The tax collections then are used to reimburse those who actually do pay these external costs. The problems in such an arrangement include accurately measuring the social cost involved; how to see to it that those who really do pay the external costs are the ones who get compensated.

- **Subsidy:** In the case of positive externalities, government can subsidize firms that produce them, thus providing the firms with revenue fully reflecting the social benefit that they produce.

**Taxes on Negative Externality**

Because pollution imposes costs on the general public, the social marginal cost of the pollution ($SMC$) exceeds the private marginal cost ($PMC$) of making the good by the good’s external marginal cost ($EMC$). The private marginal cost is the marginal cost to the producers of the good. The social cost is the cost to everyone, including the public and the producer. In this case, we have the following equation:

\[ SMC = PMC + EMC; \]
\[ SMC > PMC \text{ when } EMC > 0 \]

Assuming there are no positive externalities, the demand curve reflects both the social marginal benefit ($SMB$) and the private marginal benefit ($PMB$) from making the good ($SMB = PMB$).

In the absence of a tax, the free market will produce the good, and sell them for at a price where $PMC = PMB$. But because $SMC = PMC + EMC$, the socially optimal level of production is where $SMC = MB$.

An optimal tax would equal the good’s external marginal cost ($Tax = EMC$). This makes the $SMC$ curve in effect the industry’s $PMC$ curve.

**Subsidy for Positive Externalities**

Because the good creates benefits for more than just its buyers pay, its social marginal benefits ($SMB$) exceed its private marginal benefits ($PMB$) by its external marginal benefit ($EMB$). So, we have the following equation:

\[ SMB = PMB + EMB; \]
\[ SMB > PMB; \text{ and }\]
\[ EMB > 0 \]
We assume there are no negative externalities in this figure. So, the supply curve reflects both the private and social marginal cost of making the good (PMC = SMC).

In the absence of a tax, the free market will produce the good, and sell them at a price where PMC = PMB. But because SMB > PMB, the socially optimal level of production is where SMB = MC.

The socially optimal subsidy is equal to the good’s external marginal benefit (subsidy = EMB).

### Pollution as an Externality

In this section, we will restrict our study to learn about pollution as an externality.

The government regulation of pollution illustrates some of the principles we have been talking about. Pollution is a negative externality. Because of transaction costs, it is too difficult for people to, by themselves, bargain with firms for the optimal level of pollution.

Two solutions have emerged:

1. **Taxing polluters:** One solution has been proposed to tax polluters. Emission charges are common in Europe. Water polluters have to pay a waste disposal tax there is also an emission charge on gasoline. The social cost includes the private cost plus the cost of pollution emitted by cars. An emission tax equal to the difference between the social and private marginal cost will result in the optimal level of driving.

2. **Limiting the amount of pollution:** The second solution is limiting the amount of pollution. In place of a tax, the government can regulate and limit the amount of pollution allowed.

However, if the same limit is imposed on all individuals, we are not taking into consideration that the needs of driving vary from individual to individual. Regarding the preceding example of driving, some people have a greater demand for driving than others. Similarly, some people drive cleaner cars than others.

### The Optimal Level of Pollution

What level of pollution should a regulatory agency allow? Pollution is costly. It is a subject of controversy as to how costly it is. But let us pretend that we are all-knowing and have precise figures on pollution’s cost. Similarly, it is costly to get rid of pollution. Let us also suppose that we know these costs. How then do we decide what is the correct level of pollution? There are two costs society faces:

**Total cost = Total pollution cost + Total pollution abatement cost**

The socially optimal level of pollution minimizes the sum of these two costs. We can use marginal analysis to assist us. Let Q be the quantity of pollution allowed. The marginal cost of increasing Q by one unit is the bad that occurs. Total pollution costs go up. New Delhi gets smoggier. More fish die from acid rain. More people...
have a higher risk of getting sick. The marginal benefit of increasing $Q$ by one unit is the reduction in pollution abatement costs. $Q$ should be increased as long as $MB > MC$. At the optimal $Q$, $MB = MC$. Because the cheapest ways to reduce pollution will be used first, the subsequent units of pollution will cost more to clean up. Similarly, the marginal cost of pollution will increase. This means the higher the level of pollution; the costlier a given addition is in terms of the damage that it does.

**Check Your Progress**

4. Name the factors which affect the extent of industrial concentration.
5. What is HHI?
6. Mention some of the other names for externalities.
7. What represents private marginal costs in a market with competitive producers?

**14.6 ANSWERS TO CHECK YOUR PROGRESS QUESTIONS**

1. Complementary bidding is also called courtesy bidding, this takes place in situations where the applicants purposely submit much higher bids and allow the decided bidder to quote less and accommodate all the demands of the buyer. These are most prone to rigging and are developed to provide a false picture for the buyer about the market environment.

2. The Competition Act was passed in 2002 and amended in 2007.

3. Conduct in the SCP paradigm refers to the behaviour of the consumers and retailers, their behaviour with each other as well as between themselves have an impact. The behaviour of the firms is based on their strategy which is decided after researching aspects like investment, development, levels of advertisement and collusions.

4. Two factors affect the extent of industrial concentration, one factor is quantity of firms active in a particular market, and the other factor is percentage of the share of the market demand fulfilled by the firm.

5. Herfindahl–Hirschman Index, HHI, or even as HHI-score is an absolute measure of concentration which gives the measure of magnitude of firms in comparison to others in the industry and also indicates the quantity of competition between the firms.

6. Externalities are sometimes called spillovers or neighbourhood effects.

7. In a market with competitive producers, the supply curve represents private marginal costs.
14.7 SUMMARY

- The economic framework within which the firm functions plays a crucial role in its efforts to thrive, grow and beat its competitors. Since business is fundamentally an economic activity, the economic environment, both within the country as well as beyond its national frontiers will have a definite influence on its fortunes.

- Business has to deal with a vast number of governmental bodies, rules, regulations, and guidelines relating to its statutory responsibilities, the capital market, sources of finance including stock market options, venture capital, offshore funds, disinvestment options, bank funding and so on.

- Antitrust laws are often mentioned as ‘competition laws’. These are laws established by the governments to protect the consumers from predacious practices of business. The laws ensure existence of impartial competition in economy which is based open-market. They have developed as the markets have evolved and are in keeping with the recent trends.

- There are numerous debatable activities of business on which the Antitrust laws are used: allocation of market, bid rigging, price fixing, monopolies, etc.

- India was among the first developing countries to have a competition law in the form of the Monopolies and Restrictive Trade Practices (MRTP) Act, 1969. The MRTP Act was designed to check concentration of economic power, prohibit restrictive or unfair trade practices and control of monopolies.

- The year 1991 proved to be a watershed in the history of India’s economic development. The new India required new rules. Hence, the need for a new competition law. Accordingly, the Competition Act was passed in 2002 and amended in 2007.

- The Competition Act, 2002, as amended by the Competition (Amendment) Act, 2007, follows the philosophy of modern competition laws. The Act prohibits anti-competitive agreements, abuse of dominant position by enterprises and regulates combinations (acquisition, acquiring of control and M&A), which causes or likely to cause an appreciable adverse effect on competition within India.

- Structure-Conduct-Performance (SCP) framework suggests a connecting affiliation among the market’s structure (S), the conduct (C) of various firms of that market and their performance (P) in the economy. The paradigm was initially suggested by E. Mason during the 1930s and was further developed in 1950s by J Bain. The neo-classical theory of firm formed the basis of the paradigm and since then it has been the focal point if all the industrial organization studies.
Economic concentration has several denotations in the industrial structure and is often applied to various ‘indexes of concentration’ in order to understand the distinctive features of the structure in industries. According to the economic theory, concentration is a vital factor as it determines the behaviour of the market and also affects its outcome.

A firm’s control over production and distribution of a product is referred to as monopoly. Monopolistic environment may be created in large as well as small scale industries. On the other hand, concentration of economic power refers to centrally controlling essential economic activities such as availability of capital, employment and manufacturing in a major industry or agriculture etc.

Externalities are costs and benefits of economic activity that do not enter into the firm’s own calculations of costs or revenues. Thus, in the absence of some sort of non-market correction (usually government action), they lead to inefficient allocation of resources.

Remedies for Externalities include: government regulations, taxation and subsidies.

14.8 KEY WORDS

- **Antitrust laws**: Also known as ‘competition laws’, these are laws established by the governments to protect the consumers from predacious practices of business.
- **Structure-Conduct-Performance (SCP) framework**: It suggests a connecting affiliation among the market’s structure (S), the conduct (C) of various firms of that market and their performance (P) in the economy.
- **Externality**: An economic effect that results from an economic choice but is not reflected in market price.

14.9 SELF ASSESSMENT QUESTIONS AND EXERCISES

Short Answer Questions

1. Give a brief introduction to the development of Antitrust theory.
2. Write a short note on competition laws in India.
3. What are the criticisms against the SCP paradigm?
4. What is the difference between monopoly and concentration of economic power?
5. Mention the solutions to pollution as an externality.
Long Answer Questions

1. What are the activities of business which require the use of anti-trust laws?
2. Explain the elements of the SCP paradigm.
3. Describe the absolute and relative measures of industrial concentration.
4. Discuss the types of and remedies for externalities.

14.10 FURTHER READINGS


Websites

https://blogs.economictimes.indiatimes.com/et-commentary/heres-how-indias-widening-income-distribution-can-be-redressed/