CENTRE FOR DISTANCE & ONLINE EDUCATION

UNIVERSITY OF JAMMU

JAMMU

SELF LEARNING MATERIAL

MA ECONOMICS

SEMESTER-II

SESSION : 2025-26 ONWARD

NO. OF CREDITS : 6

Subject : ADVANCED MICRO ECONOMICS **UNIT : I - IV**

COURSE CODE : ECO-201 **LESSON : 1 - 23**

Dr. Neelam Choudhary

Course Coordinator

http:/www.distanceeducationju.in

Printed & Published on behalf of the Centre for Distance and Online Education, University of Jammu by the Director, CDOE, University of Jammu, Jammu.

ECONOMICS

Course Contributors

Dr. Aroon Sharma

Dr. Rupinder Kour

Dr. Neelam Choudhary

Editing & Proof Reading by

Neelam Choudhary

Printed by .... Printers/2025/500 Books

M.A Economics under Non-CBCS (CDOE)

Semester-I1

Syllabus for the examination to be held in May 2026, 2027 & 2028

Course Code : ECO-201	**Title – Advanced Microeconomics**
Credits : 6	**Max. Marks : 100**
Time: 3 Hours	**Internal Assessment : 30**
	Semester Exam : 70

Course Outcomes

- To introduce students to various theories of distribution, Welfare Economics, general equilibrium in closed systems, choice under risk and uncertainty
- To get familiarised with new developments in the theory of markets and information
- To leam to apply Mathematical tools in the special economic situation

UNIT-1 : Alternative Theories of Firm and Information Technology

Critical evaluation of marginal analysis, Baumol's sales maximisation model, price output determination, optimal advertising expenditure, choice of output of various products and input combinations, pricing and changes in overhead costs, Williamson's model of managerial discretion, Maris' model of managerial discretion, full cost pricing rule. Bains' limit pricing theory and its recent developments including Sylos Labini model of the firm behavioral model of the firm information technology systems competition, problem of complements, lock in, network externalities and its effects and implications, rights management and sharing of intellectual property.

UNIT-II : Game Theory and General Equilibrium

Zero Sum Game. Non Zero Sum Game The payoff matrix of a game, definition of Nash equilibrium, pure and mixed strategies, prisoners' dilemma, repeated games- applications, sequential games, sub-game perfect equilibrium examples- entry deterrence Partial and general equilibrium, the Walrasian system, existence, uniqueness and stability of equilibrium A graphical illustration of the path to general equilibrium, algebraic treatment of a two person-two-good exchange model, graphical treatment of the two factor, two commodity two consumer general equilibrium system, static properties of a general equilibrium state, general equilibrium and allocation of resources.

UNIT-III : Distribution and Welfare Economics, Externalities and Provisions for Pubic Goods

Marginal productivity theory of distribution, product exhaustion theorem (Clark-Wicksteed-Walras), Pareto optimal conditions, social welfare function, compensation principle, Scitovsky's paradox. Scitovsky's double criterion, theory of second best, Arrow's impossibility theorem, Externalities-definitions externalities and market failure, Coase theorem, production Externalities- pollution vouchers, property rights and the tragedy of the commons, Public goods-definitions and characteristics, when to provide a public good, optimal provision for a public good, private provision for public good and the free-rider problem, the Clark-groves tax and its problems.

UNIT-IV : Choice under Risk and Uncertainty and Economics of Information

Expected utility theory, Von Neumann-Morgenstern method of constructing utility index under risky situations, the St. Petersberg Paradox and Bernoulli's hypothesis, Alias's paradox and explanations, Attitude towards risk- measures of absolute and relative risk aversion, Certainty equivalent, risk premium, risk averter v/s risk lover. Risk lover and gambling. Risk aversion and insurance. Friedman- Savage hypothesis, Markowitz hypothesis, the investors' choice problem, information and consumer's choice, information and insurance-dealing with asymmetric information. symmetric information- the market for lemons and quality choice, adverse selection, moral hazard, signaling, incentive-systems.

Note for Paper Setting

The term end examination shall be of 70 marks. There shall be two types of questions in each unit, two short answer type (each of 250 words) and two medium answer type (each of 600 words). The candidate will have to attempt one short answer type question and one medium answer type question from each unit. Each short answer type question shall carry 4.5 marks and each medium answer type question shall carry 13 marks.

Reading List

- Henderson J and R Quandt: Microeconomics theory, A Mathematical approach, McGraw-Hill.
- Hader, J:Mathematical theory of economic behavior, Addison-Wesley.
- Koutsyannis, A:Modern microeconomics, Macmillan.
- Mas-collel A, M.D. Whinston& J.R.Greene: Microeconomics theory, OUP.
- Salvatore, D:Microeconomics, UP
- A:Microeconomics, UP
- Varian: Intermediate Microeconomics, East West press.

CONTENTS

Unit-I

Unit-II

Unit-III

Unit-IV

M.A. ECONOMICS	LESSON NO. 1
COURSE NO. ECO 201	UNIT-I

CRITICAL EVALUATION OF MARGINAL ANALYSIS

STRUCTURE

1.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

- Define Marginalist controversy

- Examine Hall and Hitch study
- Discuss Gordon’s criticism of Marginalism
- State points in defense of Marginalism

1.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to:

- Define Marginalist controversy
- Examine Hall and Hitch study
- Discuss Gordon’s criticism of Marginalism
- State points in defense of Marginalism

1.2 INTRODUCTION

Dear learner, in the last semester, you came across the basic concepts of Microeconomics. In this semester, we shall learn advanced concepts of Microeconomics. In this lesson, we shall discuss about Marginalism as an important concept that evolved over a period of time and the criticism faced by it. The neo-classical theory of the firm based on marginalist principle enjoyed wide acceptance till 1938. The criticism began with a study based on empirical evidence of how firms actually behave. In this study, Hall and Hitch (1939) tried to highlight important drawbacks of marginalist principle. This was followed by many more articles and studies, including those in late 1940s by economists like Lester, Machlup, Oliver, Gordon and Friedman. Thus started scepticism about the traditional theory.

1.3 CRITICAL EVALUATION OF MARGINAL ANALYSIS

1.3.1 Basic assumptions of the neo-classical theory

While criticising any theory, we also criticise the assumptions on which it is based. Let us first discuss the main assumptions on which it is based. We shall at the same time also discuss the point of criticism against each assumption.

1. **The entrepreneur is also the owner of the firm:** The first assumption on which the traditional theory is based is the existence of a single

owner- entrepreneur. In other words, there is no separation between owner and entrepreneur. It is further assumed that there is no constraint on the decision making power of the entrepreneur. In other words, the entrepreneur has unlimited time, information and ability that enables him to act with ‘global rationality’ and fulfil the only goal of profit maximisation.The traditional theory is being criticised on this ground, as the single owner- entrepreneur firm is quite unrealistic in today’s highly complex world of corporates. Since there is separation (divorce) of ownership and management, managers are free to pursue goals other than profit maximisation. Moreover, there is neither unlimited (and smooth flow of) information nor time and ability on the part of entrepreneur to be able to act with ‘global rationality’.

2. **Profit maximisation as the single goal of the firm**: The traditional theory believes that irrespective of the market structure, the firms’ goal is only profit maximisation, which they achieve by applying marginalist principle (MC=MR). This assumption has been rejected by the later economists, as due to uncertainty regarding demand and cost conditions, the firms can’t apply marginalist principle. Moreover, even if it is possible to achieve, firms still don’t consider profit maximisation as the only goal. There is a multitude or multiplicity of goals and profit maximisation is just one of them. The alternative goals suggested by the critics are summarised in the following paragraphs.

    - **Managerialism: maximisation of the managerial utility function**. Due to separation of ownership and management, goals pursued by managers need not be profit maximisation. Rather they aim at maximising managerial utility function. Factors that usually enter their utility function include salaries, prestige, market share, job security etc. However, the managerial theorists differ with regard to what maximises this utility function (e.g. Baumol’s sales maximisation, Marris’ maximisation of balanced rate of growth etc.).

    - **Behaviourism: satisficing behaviour** Unlike a certain world assumed by the neo-classical theory, the modern corporate world is quite uncertain. Due to the incorporation of uncertainty in the real world where firms actually behave, the firms can't act with 'global rationality'. Rather they behave with 'bounded rationality', which seems more realistic, as there are various constraints faced by the firms. This gives rise to satisficing (rather than maximising) behaviour. Satisficing behaviour means that the firms pursue the goal of satisfactory growth, sales, profits etc.

- **Long run survival and market share goals** K.W. Rothschild is of the opinion that long run survival is given top priority by today's firms. In other words, the firms want to ensure that there is high probability of their long run survival. Some economists are of the view that in today's world, significant changes have taken place in the market structure, due to which an alternative goal is to capture substantial market share and then to retain it. This is true of today's oligopolistic market structure, in which few firms in the industry dominate and account for bulk (70-80%) of the sales. The firms try to maximise profits, given this constraint of desired market share. If the goal is profit maximisation, firms would choose 'maximum maximorum' i.e. the share that yields maximum profits. This objective is compatible with marginalist principle. But it may not imply maximisation of long run profits.
- **Entry prevention and risk avoidance** Bain, Sylos-Labini, Modigliani and some other economists have suggested that the goal of the firms is to prevent entry of new firms and to avoid risk. Limit pricing is the entry prevention price charged by the existing firms. It is the price set by the existing firms, so that there is no incentive on the part of outer firms to enter the industry. This is again an important feature of today's corporate world ,where perfect competition is not a reality. Firms are in conflict with each other to maximise their profits or pursue other goals. Scholars generally attribute entry prevention goal to three probable reasons i.e. long run profit maximisation, long run survival or avoiding risk associated with the uncertain or unpredictable reaction of new entrants.

Now the question is, to what extent can the managers deviate from profit maximisation goal of the firm? In other words, what is the degree of managerial discretion? It is not so easy to answer this question, as there is a lot of controversy around this issue. Few empirical findings suggest that profits are higher in firms where there is no separation of ownership and management, as compared to the firms where they are separate. This means there is no unlimited discretion to managers. In other words, profit maximisation is still valid. Here we can add the arguments of supporters of profit maximisation, who state that profit maximisers grow better and have better survival than non-profit maximisers.They attribute it to accumulation of financial assets. Only profit maximisers are the fittest, as the rest are eliminated. Moreover, if the goal is profit maximisation, the firm can even fulfil other 'subsidiary' goals easily. However, this is countered by those opposed to the goal of profit maximisation. They argue that in today's dynamic world, firms need to update themselves with regard to new techniques as well as to adopt differentiation and diversification, to avoid elimination even in the absence of profit maximisation.

Based on these arguments for and against profit maximisation as a goal, it can be concluded that multiplicity of goals is pursued by the firms and that the managers have limited discretion. Further, there is a minimum profit constraint on the firms that puts a limit on the pursuing of other goals. How high the constraint is depends on conditions like competition. If there are strong barriers to entry (e.g. strong preferences, government laws, absolute cost advantage, absolute capital requirements, strong economies of scale etc.), there are better chances of the firms' survival even if they are not profit maximisers.

3. **Marginalist principle** The traditional theory assumes the role of marginalist principle i.e. the intersection of MC and MR in decision making with regard to profit maximisation. This equalisation of MC with MR has to be done in each period. However, critics argue that even if long run profit maximisation is a goal of the firm, it need not be

ensured through MC=MR condition in the short run. Moreover, profit maximisation as a goal is objectionable. We have already discussed this in detail.

4. **Certainty** In the early stages of the development of theory of firm, it was assumed that there is complete certainty. The firm has complete knowledge not only about past and present conditions, but also about future happenings. The firm has perfect knowledge about the demand and cost conditions as well as the environment. Later on, it was revised by the neo-classicists. To tackle the element of uncertainty, they adopted probabilistic approach, which means that the information is limited to only a probability distribution of all possible outcomes. It was further assumed in the modified version that profit is not a single valued thing to be realised with certainty, but can take any value from the range of values provided according to the probability realised. The probabilities are known to the firms subjectively. However, this probabilistic approach to uncertainty is also being criticised on the following grounds:

    - It requires a lot of knowledge, computational ability and information on the part of entrepreneurs, which they don't have.
    - The expected profitability is not the only determinant of the actual decision taken by the firm. It depends on the attitude of the entrepreneurs towards risk. If they are risk takers, they would go for the highly profitable venture, as high profit is usually accompanied by high risk. But what about an entrepreneur who avoids risk? In this case, avoiding risk is his priority, even if it means sacrificing some profits.
    - The probabilistic approach ignores the time -horizon. It must be noted that time horizon plays an important role in decision making. If the firms avoid the uncertainty arising from long run, they would decide on the basis of expected profits in the short run. On the other hand, firms which want returns in the long run would have a different approach.

- The probabilistic approach considers the expectations of the entrepreneurs as exogenous to the firm (as they are subjectively determined by the entrepreneurs).But critics are of the view that these are influenced by many factors internal to the firm and hence are endogenous to the firm.

5. **Entry considerations** The traditional theory's approach towards entry is also objectionable. It only deals with actual entry that takes place in the long run. Potential entry and its implications are ignored. Moreover, the entry conditions are dependent on the type of market structure. Under Perfect Competition and Monopolistic Competition, there is free entry. In Monopoly, entry is blocked, that keeps the monopolist unthreatened by a potential entrant and hence he continues to earn supernormal profits even in the long run. Likewise, Oligopoly is also a closed model. Hence potential entry, which plays a role in decision making by firms is ignored.

6. **The static nature of traditional theory** The traditional theory assumes that time periods are independent. It means that decisions taken by a firm in one period do not affect the behaviour of the firm in another period. So the marginal analysis is applied both in the short run as well as the long run. But the assumption of time horizon consisting of identical and independent time periods is said to be the most important shortcoming of this model. Critics argue that decisions are temporally interdependent, as past decisions influence present decisions which have a bearing on future decisions. But since the classical theory ignores this and assumes temporal independence of decisions, it is static in nature.

1.3.2 The Hall and Hitch study and the full cost pricing principle

The 'Oxford Economists Research Group' report, also known as Hall and Hitch report was the first attack on Marginalism. The empirical study, which was carried out on a sample of 38 entrepreneurs (who belonged to efficiently managed enterprises) by interviewing them, concluded that firms do not behave as stated by the neo-classical theory. The findings showed that firms don't behave atomistically, Oligopoly is the most dominant market structure and that Oligopolistic interdependence could not be dealt with by the traditional theory.

As far as the hypothesis of profit maximisation is concerned, the firms don't aim at maximisation of short run profits, but long run profits. For this, they set a price that covers average variable cost (direct cost), average fixed cost (overhead cost) and a normal or satisfactory margin of profit (usually 10%). This is called full cost pricing, average cost pricing, cost plus pricing, mark up pricing etc.

Check Your Progress-I

Answer the questions in the space given below:

Q1. What were the basic assumptions of neo-classical theory?

____

____

____

____

Q2. Discuss the main points of criticism against the traditional theory of the firm.

____

____

____

____

Q3. What is Marginalism? What was its role in the traditional theory of firm? How do the later models of firms comment on its use by the firms?

____

____

____

____

$$P = AVC + AFC + \text{Normal profit margin}$$

According to Hall and Hitch, firms abandon Marginalism in favour of full cost principle due to the following reasons:

1. If they fix a price above average cost, new firms will (or may) enter, lured by supernormal profits .

2. The firms don't know about their MR, AR and MC with certainty. How can they think of fixing price at the point at which MC=MR?

3. Firms believe that average cost price is the right price to be charged at all times, as it covers the whole cost as well a fair profit. It is also based on the moral conviction of the firms, as it justifies what price ought to be charged.

4. There is a tendency in the prices of manufactured goods to stay sticky. We are all familiar with the concept of kinked demand curve. It is a demand curve drawn on the basis of a particular reaction pattern of Oligopolistic firms that their rivals don't react if they resort to price hike, while they react if the firms in question cut prices. This gives rise to a kink in the demand curve, with the upper segment highly elastic, while the lower one relatively less elastic. Price is fixed at the kink and tends to be rigid.

Hall and Hitch try to integrate full cost pricing principle with kinked demand curve model. Fig.1 shows that if the firm decides to produce OM output, the firm will try to work out the cost (fixed and variable plus profit margin) likely to be incurred on the production of this output. Once this cost is determined, the firm will fix its price. In Fig. 1, full cost is KM, while price is OP. KM=OP.

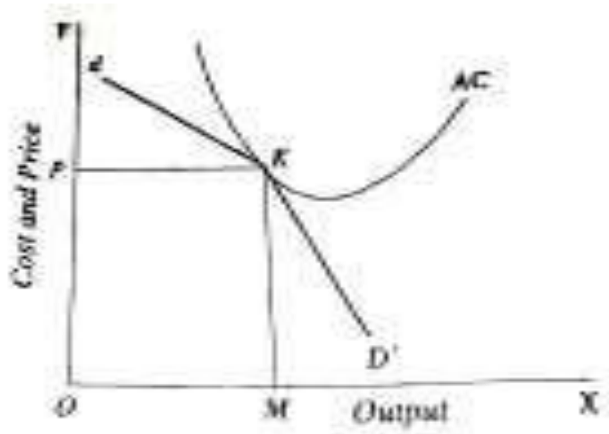


Fig.1 : Average-Cost Pricing and Kinked Demand Curve

Hall and Hitch explain the rigidity of prices in the following way:

1. Usually, increase or decrease in demand may shift the kink , but price remains unchanged. However, if there is a large decrease in demand and demand remains low for some time, prices may fall due to the irrational behaviour of one or more entrepreneurs. It may force others to cut prices.

2. The 'full cost' may undergo revaluation if the average costs of all firms shift by similar amount.

Generally speaking, say Hall and Hitch, Oligopolistic firms do not keep their prices high. As high prices are usually associated with low levels of output, in Oligopoly, firms which produce large amount of output (by keeping plant running as full as possible) don't keep prices high. They favour price stability and are also constrained by the kink, that prevents price rise.

1.3.3 Gordon's critique of Marginalism

Gordon joined the attack on Marginalism in 1948. His criticism may be summarised as:

1. **Real world complexity and changes in economic environment of business.** The real world of firms is highly complex. The demand and cost are determined by many variables. Moreover, the firm also operates though policy variables like price, output, advertising, diversification etc. Since the determinants are subject to change, the entrepreneurs can't think of doing marginal adjustments (by equalising marginal magnitudes i.e. MC and MR) for all the changes. Likewise, the economic environment is dynamic in nature which renders any past experience irrelevant for today. Under these circumstances, the only rational approach is to set a price at Average Cost.

2. **Uncertainty.** In a world of uncertainty, it is not possible to estimate future demand and cost. Uncertainty demands adequate information about expectations of businessmen as well as setting some goals other than profit maximisation. But according to Gordon, the marginal analysis tools can't explain either expectations of profits or alternative goals.

3. **Average cost pricing.** There is sufficient empirical evidence to prove that firms actually adopt average cost pricing rule and not the marginalist principle. The main argument is that most of the firms are multiproduct firms. It is not possible for a firm to know about the Marginal Cost of different products. This renders marginal principle inapplicable. Further, while dealing with actual situations or local problems, the firm managers don't first apply marginalist principle and then resolve the disputes. Under these circumstances, average cost pricing rather than profit maximisation (based on marginalism) seems more realistic, as the aim is to satisfy the demand.

4. **Multitude of goals.** Now a days, the firms have many goals, including peaceful labour relations , maintenance of goodwill etc. These goals may or may not be compatible with profit maximisation. If there is conflictual relation between these goals and profit maximisation, firms may forego profit maximisation as the main goal, in particular if they have adopted full cost pricing rule.

5. **Subjectively determined demand and cost.** For the application of marginalist principle, demand and cost curves must be objectively known. But as Machlup accepts, if businessmen conceive these curves subjectively, Marginalism reduces to a tautology, as any observed behaviour of the firm can mean move towards profit maximisation, as the curves are subjectively conceived and hence can be equalised.

1.3.4 In defense of Marginalism

Marginalism still has many supporters, which makes it difficult to declare Marginalism as invalid. The main arguments in defense of Marginalism are:

- Empirical evidence.
- Validity on the basis of predictions.
- Profit maximising firms are the fittest.
- Assumptions are realistic.
- Limitations of evidence against marginalism.
- Profit maximisation is the goal.

Let us now discuss these in detail:

Empirical evidence. In 1950s, empirical studies were conducted by Earley on a sample of 110 well managed firms in USA. The conclusion was contrary to what Hall and Hitch had reported. Earley reported that in fact the modern accounting methods enabled the firms to get knowledge of marginal costs and revenues, which made it possible to apply marginalism. However, as the views are conflicting, it can't be concluded that all the firms follow this principle. According to Barback, it may not be used by all the firms, as all of them may not possess the information. Few other firms don't apply it in the short run, as they think it may harm long run profits.

Validity on the basis of predictions. Friedman has defended the theory on 'methodological issues'. According to him, a theory should be judged on the basis of its predictions and not realism of assumptions. Since the theory has produced reasonably good predictions, it should be considered a satisfactory theory.

Check Your Progress-II

Answer the questions in the space given below:

Q1. Write a note on Gordon's attack on marginalism

Q2. How do Hall and Hitch justify the breakdown of marginalism?

Q3. What is full cost pricing rule?

Profit maximising firms are the fittest. Another line of argument is put forward by Alchian and penrose. According to them, The Darwinian principle of the survival of the fittest can be applied to firms. Firms which are profit maximisers are able to build up assets and sustain themselves, whereas non-profit maximisers feel resource constrained and hence may face stagnation. In the process of 'economic natural selection', the profit maximisers, being the fittest are able to eliminate the weaker (non-profit maximising) firms. However, A. Koutsoyiannis doesn't agree with the elimination argument.

She says that to avoid elimination, firms may go for diversification or adopt other policies to survive. Moreover, survival is not a permanent feature of firm. Firms which are the fittest today may not retain that level of fitness tomorrow. Further, there would be no 'fittest' if no firm is profit maximiser, but pursue other goals.

Assumptions are realistic. According to Machlup, the assumptions of the traditional theory are quite realistic. He agrees that marginal costs and revenues might not be objectively known to the firms, but he considers it unnecessary for applying marginalism. He says that what matters is the belief of the firms that what their marginal cost and revenue curves are even if that assessment is subjective.

Limitations of evidence against marginalism. Machlup further argues that he is not satisfied with the empirical evidence given by Hall and Hitch against marginalism. He also rejects Lester's criticism. The three points in defense of marginalism are:

- The concepts of Marginal Cost, Marginal Revenue and elasticity are used by economists, but not by the firms. This difference in terminology might have led to a communication gap, which made researchers conclude that firms set prices equal to Average Cost. Moreover, P=AC is not inconsistent with Marginalism, as in the long run, P=AC. However, this was countered by Hall and Hitch, who argued that while the questionnaires were administered, every concept was made clear to the businessmen, which left no scope for misunderstanding of the concepts.
- Even if the firms don't fully understand the concept of elasticity, they do incorporate it implicitly in pricing methods. However, this doesn't make us conclude that marginalist principle is used, as it is applied even in Average Cost pricing method.
- Firms may not reveal that profit maximisation is their goal, even if it is. A fair policy enjoys wider acceptance as compared to an unfair one. Generally, prices based on Average Cost are considered fair. So firms don't hesitate in saying that profit maximisation is not their goal.

5. **Profit maximisation is the goal** Machlup further argues that profit maximisation is the goal of the firm. For attaining profit maximisation in the long run, the firm uses long run demand and cost curves after incorporating expectations of future conditions and changes in business environment (doctored demand and cost curves).However, Machlup fails to understand the problems associated with this treatment of expectations.

   On the basis of all the above arguments, it can be concluded that the Marginalist controversy is still not resolved. We can conclude with the following points:

   - Given the stickiness of prices due to changing business conditions, Marginalism is not applied at least in the short run.

   - Marginalism faces dilemma of either being unrealistic or being in danger if time, uncertainty or additional goals are incorporated into the single set of 'doctored' long run demand and cost schedules. These doctored curves may reduce Marginalism to a tautology. We have already discussed this point.

   - It not clear whether the pricing policies used by firms are in conflict with marginalism. It is also not clear whether Average Cost pricing is an alternative to marginalist theory.

1.4 LET US SUM UP

Dear learner, in this lesson, we discussed on what grounds the Marginalist principle has been criticised. We also discussed the points in defense of Marginalism.

1.5 KEYWORDS

Marginalism: A theory that states that individuals take decisions on the basis of Margin or additional unit e.g. additional or marginal utility .

Full cost pricing: Under full cost pricing, the price is set in such a way that a mark up is added to total cost, that allows for profit margin.

Neo-classical theory: Neo-classical approach to Economics asserts that an price of a good is driven more by how an individual values it and not just by cost of production.

1.6 EXAMINATION ORIENTED QUESTIONS

1. What is Marginalist controversy? Discuss in detail the different lines of attack against it.
2. Do you agree with Machlup's defense of Marginalism? Justify your answer.
3. In what way has Gordon argued against neo-classical theory of the firm?
4. What is full cost pricing rule? How does it make a departure from traditional theory?
5. Do you agree that Marginalism has broken down? Give arguments to support your answer.
6. Discuss the basic assumptions of neo-classical theory and grounds of criticism put forward by critics.
7. Discuss in detail the empirical studies for and against use of marginal tools. Can you conclude anything regarding the validity of profit maximisation as a goal? What other goals are pursued by the firms apart from profit maximisation?

1.7 HINTS TO CHECK YOUR PROGRESS

CYP-I

Q1. See sub-section 1.3.1

Q2. See sub-section 1.3.1

Q3. See sub-sections 1.3.1 and 1.3.2

CYP-II

Q1. See sub-section 1.3.3.

Q2. See sub-section 1.3.2.

Q3. See sub-section 1.3.2.

1.8 SUGGESTED READINGS

1. Ahuja, H.L., Advanced Economic Theory, 9th edition (1997).
2. Koutsoyiannis, A., Modern Microeconomics, 2nd edition (1979).
3. Misra, S.K. and V.K. Puri, Advanced Microeconomic theory, first edition (2001).

BAUMOL'S SALES MAXIMISATION MODEL

STRUCTURE

2.0 Objectives

2.1 Learning outcomes

2.2 Introduction

2.3 Baumol's sales maximisation model and its rationale

- 2.3.1 Introduction
- 2.3.2 Static models
- 2.3.3 Price output determination.
- 2.3.4 Optimal advertising expenditure
- 2.3.5 Choice of output of various products and input combinations
- 2.3.6 Pricing and changes in overhead costs
- 2.3.7 Emphasis on non-price competition
- 2.3.8 Implications and critical appraisal of Baumol's theory

2.4 Let us sum up

2.5 Keywords

2.6 Examination Oriented Questions

2.7 Hints to Check Your Progress

2.8 Suggested Readings

2.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

Define Baumol's sales maximisation model

Discuss Optimal Advertising expenditure

Explain Non-price competition in Oligopoly etc.

2.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to:

Define Baumol's sales maximisation model

Discuss Optimal Advertising expenditure

Explain Non-price competition in Oligopoly etc.

2.2 INTRODUCTION

Dear learner, Baumol's theory is the first in the series of managerial theories of the firm (or managerialism), which we shall be discussing next. Marris' theory and Williamson's theory are the other two theories. The main focus of the present lesson shall be on Baumol's theory.

2.3 BAUMOL'S SALES MAXIMISATION MODEL OF OLIGOPOLY AND ITS RATIONALE

2.3.1 Introduction

Baumol's sales maximisation model suggests sales maximisation as an alternative goal pursued by the firms, where there is separation of ownership and management. Some scholars argue that Baumol's model was more applicable to the firms that existed in the western world at that time. Baumol, an American economist, himself remained a consultant to large firms and observed that managers are more interested in maximising sales than profits. Given the separation of owners from managers, there is a deviation from profit maximisation as the ultimate goal and given the discretion on the part of managers, their self interest gets served in maximising sales and not profits. So sales maximisation seems more plausible.

Baumol further clarifies that managers are not concerned with physical volume of sales, but the revenue from sales. That is why, it is also known as revenue maximisation model. Moreover, Baumol does acknowledge the role of profits, even if it is incompatible with sales maximisation as a goal. Since profits of today become the basis for survival and growth of firms tomorrow, Baumol argues that management of oligopolistic firms "seek to maximise sales subject to minimum profit constraint".

Sales maximisation as a valid and more realistic goal is based on rationality. Following are the main reasons for which the goal of sales maximisation (and not profit maximisation) is pursued by the firms :

- Growing and high sales are an indicator or index used by financial institutions. So firms with better rating are more likely to be financed by these institutions.
- Profits go to the pockets of shareholders (as dividends), whereas growth of sales overtime gives more satisfaction to the managers.
- Salaries and other (Slack) earnings of the managers are associated with sales and not profits.
- Steady performance with satisfactory amount of profits is generally preferred to spectacular profits in one or two periods, as it is difficult to sustain very high profits.
- Profit figures become available annually, while sales are reported more frequently. So if personnel problems are to be handled, sales figures seem more reliable. If a firm is showing better sales, higher payments are offered to employees, rather than waiting for the profit figures.
- Better sales of a firm strengthens its bargaining or competitive position vis-a vis its rivals. It also shows increasing market share of a firm.

Here, Baumol also makes some observations regarding interdependence of oligopolistic firms. The model assumes that usually, in routine policy making, firms don't consider that decisions taken by them can influence their rivals or invite retaliation. However, some radical decisions like launching of a new advertising campaign or new line of products are exceptions to it.

2.3.2 Static Models

To discuss the theory of sales maximisation, Baumol has developed both static and dynamic models. But the models of our concern are static models. Following are the assumptions on which they are based:

- It is a single period decision making.
- During this period, the firm aims to maximise sale revenue subject to profit constraint. But the firm is not concerned about the implications of decision taken during this period on subsequent periods.
- The exogenously determined minimum profit constraint must be realised by the firm, lest the managers should lose their jobs. In other words, if profits attained do not reach even the minimum level, share prices will fall, shareholders will tend to sell shares and take over raiders will be attracted by lower prices of shares. This means an exit for the managers.
- The conventional cost and revenue functions are assumed. The cost curves are U-shaped, whereas the demand curve is negatively sloped. The four models discussed here are:
  - Single product model, without advertising.
  - Single product model, with advertising.
  - Multiproduct model, without advertising.
  - Multiproduct model, with advertising.

Let us now discuss different aspects of Baumol's model in detail.

2.3.3 Price output determination (Single product model without advertising)

The graphical explanation of price output determination is shown below:

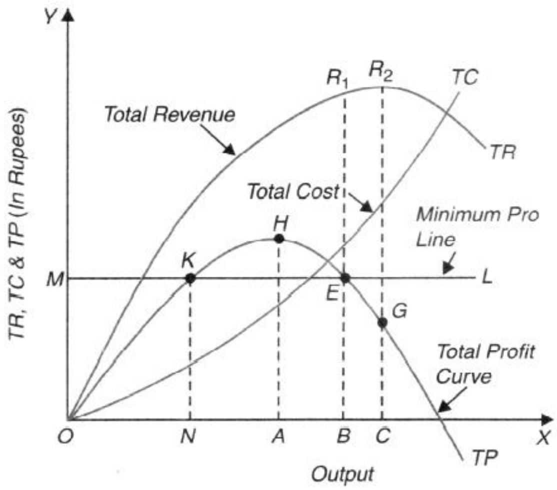


Fig. 2.1 : Prof. Baumol's Sales Maximisation Hypothesis

Source : Modern Microeconomics : Theory and Application by H.L. Ahuja (Revised Edition 2013)

As shown in Fig 2.1, there are three curves in question, TR, TC and TP. As you have learned in the earlier semester, TR refers to Total Revenue and TC refers to Total Cost. TP is the total profit curve. We are representing TR, TC and TP (in rupees) on the Y-axis and output on the X-axis. TR and TP are inverse U –shaped, which means theyobserve a rising phase , peak and a decline. As far as TC is concerned, unlike in the short run, Long run TC curve passes through origin. It means the TC drawn here refers to long run. Total profit refers to the positive vertical distance between TR and TC. Here OM is the minimum total profits to be obtained by the firm. So ML is the minimum profit line.

As shown in the figure, there are different options for the oligopolistic firm with regard to output to be produced. If the aim is profit maximisation, it should produce OA output, as it corresponds to the peak of profit curve. The maximum attainable profits are AH. But this model advocates sales maximisation as the main goal. If total revenue maximisation is the goal, the firm produces OC corresponding to which revenue is CR_2. However, the firm can't produce this output OC, as profits earned are not onlyless than maximum attainable profitAH, but also below minimum profits to be attained. It should be noted that the firm's TP curve cuts minimum profit line at point E. If OC is produced, CG profits are earned, which does not fulfil the minimum profit constraint criterion. So the firm behaves realistically by producing OB level of output, which may not correspond to maximum revenue (as $BR_1 < CR_2$), but at which minimum profits are attained. Another point where the TP curve cuts MLline is at K. But this is also not the desired level of output, as here total revenue is much less. To conclude, the firm produces between OA and OC (i.e. OB). It produces more than a profit maximising output, but less than a revenue maximising output. Profits earned are equal to BE.

Price charged at output OB = Total Revenue/output = $\frac{BR_1}{OB}$

But this level of output corresponds to the given profit constraint. If there is a shift in the minimum profit line, changes need to be made in the output to be produced.

2.3.4 Optimal advertising expenditure (Single product model with advertising)

The observed behaviour of an oligopolistic firmshows that these firms compete more with regard to variables other than price than price .You are familiar with the concept of non-price competition. Here advertising is included as a policy variable to see how it affects the firm's decision. Expenditure incurred on advertisements is expected to raise revenue, as it shifts demand curve to the right.

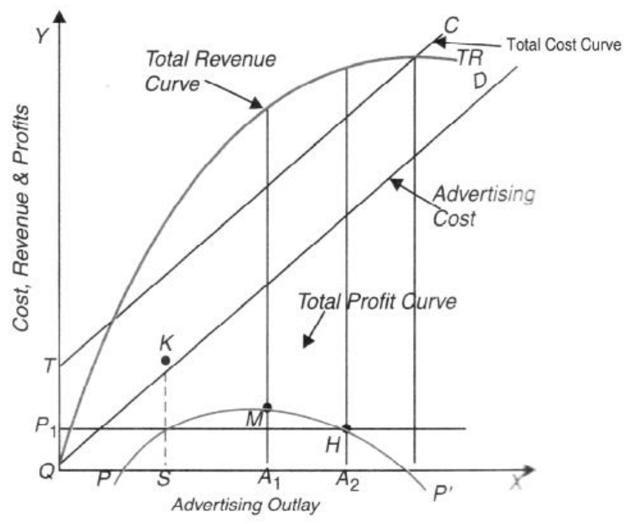


Fig. 2.2 : Optimal Advertising Outlay with Sales Maximization and Profit Maximization

Source : Modern Microeconomics : Theory and Application by H.L. Ahuja (Revised Edition 2013)

As discussed in the earlier model, here also the goal is sales maximisation subject to profit constraint which is exogenously determined. Here Profit, cost and revenue are taken along the Y-axis, whereas advertising outlay is taken on the X-axis. Now the question is to determine the optimal advertising expenditure. Baumol assumes that 'the increase in advertising outlay will always raise the physical volume of sales, though after a point, these will increase at a 'diminishing rate'. This shall also raise the revenue out of sales, given the prices. Here apart from the earlier variables (discussed in fig. 2.1), advertising outlay is added. It is added as curve OD, drawn at an angle of 45° with the X-axis (which is in a way a transformation of advertising outlay taken on X-axis into advertising cost on the Y-axis). When advertisement cost is added to other (fixed and variable) costs, total cost is derived. Thus total profit is calculated (total revenue-total cost).

As shown in the figure, the advertising outlay corresponding to maximum profits is OA_1 while that to maximum revenue is OA_2 (constrained revenue maximisation). Here the firm would choose to spend OA_2. Hence a firm which pursues goal of sales maximisation spends more on advertisements than a profit maximising firm. Here there is no possibility of unconstrained revenue maximisation. In other words, the profit constraint is always operative once advertising outlay enters the model. This is because, unlike reduction in price, advertisement outlay always increases sales. 'So it will always pay the sales maximiser to increase his advertising outlay, until he is stopped by the profit constraint'.

This conclusion also applies to the questions of optimal product adjustment and the optimal amount of special services to be provided by an oligopolist to buyers.

Check Your Progress-I

Answer the questions in the space provided.

Q1. What is Baumol's model of sales maximisation? What are the reasons for sales maximisation being more realistic and valid goal?

____

____

Q2. What are the basic assumptions of static models of Baumol? Name the four models.

Q3. Discuss optimal advertising expenditure. Also explain how is the model of sales maximisation with advertising different from the one in which advertising is not added as an instrument.

2.3.5 Choice of output of various products and input combinations (Multiproduct model without advertising)

In the real world, an oligopolistic firm produces many products. So the model now needs to be modified from a single product model to a multiproduct model. Moreover, it uses many inputs. For the sake of simplicity, Baumol has used two goods model, good X and Y. Here the transformation curve or production

Possibility curve (PPC) has been drawn. It is concave shaped. It shows the various combinations of two goods X and Y that can be produced, with given fixed outlay or R_1, R_2 ....etc. are the iso-revenue curves. As you can recall, an Iso-revenue curve is a curve that shows various combinations of two products, whose sales yield the same revenue. So revenue remains the same along one Iso-revenue curve. Higher revenues are shown on higher curves. Total revenue is maximised at point E. At this combination of two goods X and Y (OM and ON respectively), the highest possible Iso-revenue curve is tangent to the PPC. This is also the combination corresponding to maximum profits, as given the costs, maximum revenue also means maximum profits.

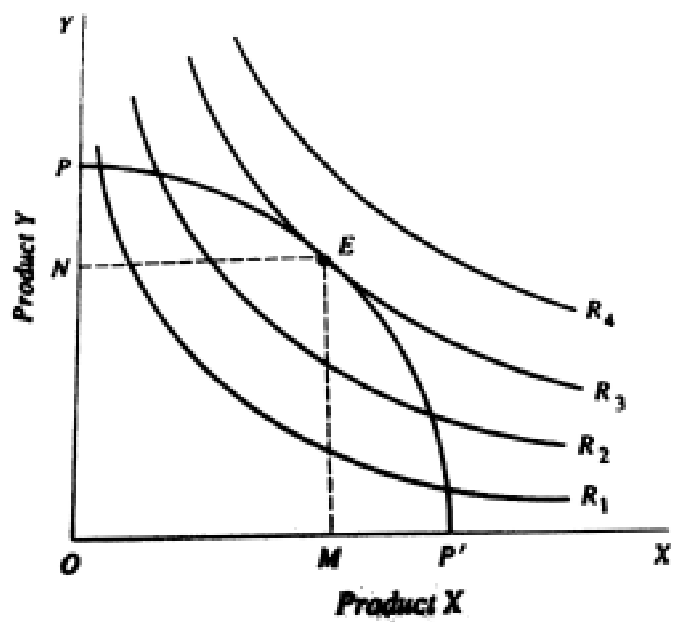


Fig. 2.3 : Choice of Outputs of Two Products under Sales Maximization

So Baumol makes the following generalisation:

"Given the level of expenditure, the sales maximising firm will produce the same quantity of each output, and market it in the same way, as does the profit maximser."

There is another significant conclusion regarding the allocation of outlay if minimum profit constraint is to be followed. In the earlier model, we have discussed the difference between maximum profits and the minimum profit constraint in the sales maximisation model. According to Baumol, the minimum level of profits is viewed by the sales maximising firm "as a fund of sacrificable profits which is to be devoted to increasing revenues as much as possible," by increasing the output of the products. Because extra quantity of a product produced beyond profit maximising point will yield negative marginal profits; so every time the firm increases the output of some product to increase its revenue, it must utilise its fund of sacrificable profits. It is to be allocated among different outputs, markets, inputs etc. so as to maximise sales revenue.

To achieve this, the ratio of MRP to MPY needs to be equal for all the goods. Here MRP means Marginal Revenue product and MPY means Marginal Profit Yield. In other words, the Marginal Revenue Product of a rupee of profit sacrificed from a product becomes equal to the marginal Revenue Product of a rupee of profit from another product.

Or

$$\frac{MRP}{MPY} \text{ of } X = \frac{MRP}{MPY} \text{ of } Y$$

Baumol states: "The relationship indicates that in the sales maximising firm, relatively unprofitable inputs and outputs are to be avoided, whatever the level of outlay and total revenue."

2.3.6 Pricing and changes in overhead costs

The conventional theory asserts that the overhead costs (or fixed costs) do not cause a change in output or price, as fixed costs don't vary with output. But Baumol argues that in actual practice, changes in overhead costs do affect the price and output. He remarks : "This piece of received doctrine is certainly at variance with business practice where an increase in fixed costs is usually the occasion for serious consideration of a price increase."

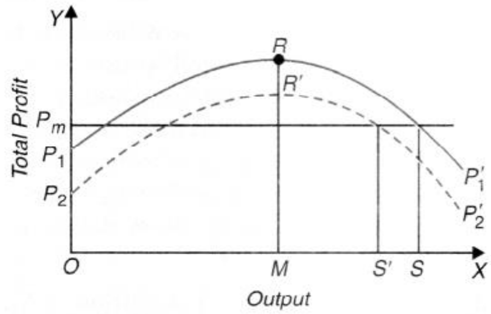


Fig. 2.4 : Increase in overhead costs lowers the level of output and raises price

This may not arise in case of a profit maximising firm, but certainly arises in case of a revenue maximising firm with minimum profit constraint. If the firm chooses to be in equilibrium in the latter case, it can't earn profits beyond a minimum required. In this case, if overhead costs rise, the total costs would rise, due to which profits fall below the minimum acceptable level. So the firm is left with only one choice that is to reduce output and hence raise the price to reach the minimum level. Fig. 2.4 explains it in detail. Here Total Profit curve has been drawn on the Y-axis, while on the X-axis, output is represented. As we can see, OPm is the minimum profit constraint. P_1P_1' is the original profit curve. A sales maximiser will produce output OS, while a profit maximiser will produce at OM. Let there be an increase in overhead costs equal to P_1P_2. This leads to a uniform downward shift in the profit curve. The new profit curve is P_2P_2'. As we can see, if the firm is a constrained sales maximiser, it would now produce at OS', whereas there is no change in the output produced by a profit maximiser. As you can see, the output still remains the same i.e. OM. The reason is that there has been a uniform shift in the profit curve. The height has been affected, but not the location of peak level of output. So it can be concluded that a sales maximiser can reduce output and increase prices as a result of increase in overhead costs, while it may not be true of a profit maximiser.

2.3.7 Emphasis on non-price competition

As we have noted in all the above models, it is not price competition that affects oligopolists as much as non-price competition. Non-price competition mean reliance on tools like advertising expenditure, product modification, introduction of special services for the customers etc. Businessmen, who are sales maximisers, consider non-price competition as more advantageous as compared to price competition. This is because the effect of price cutting on revenue as well as profits is uncertain. If demand is elastic, price cut would increase revenue and hence profits. On the other hand, if demand is inelastic, price reduction would lead to reduction in revenue. According to Baumol, while the effect of non-price competition weapons mentioned above is uncertain on profitability, it is certain on sales, as these affect sales positively. Hence non-price competition is more popular in sales maximisation model, as compared to a profit maximisation model.

2.3.8 Implications and critical appraisal of Baumol's theory

The implications of the theory are :

- Price would be lower and output higher, as compared to a profit maximisation model. This is because here TR is maximised (i.e. MR=0), rather than profits (where MR >0).
- More advertisement expenditure is incurred.
- Prices stay sticky and firms indulge in non-price competition.

Criticism. But the theory is not without criticism. Following are the points of criticism :

- According to the critics, in the long run, Baumol's model and the conventional profit hypothesis give identical results, as the firms earn only normal profits in the long run and minimum profit constraint is likely to co-incide with it.
- This model does not distinguish between firm's equilibrium and industry equilibrium.
- It is not always correct to expect revenue to increase with advertisements. Sometimes, they affect sales negatively as well, in particular in case of excessive advertising.
- It ignores actual competition as well as threat of potential competition.
- Baumol's argument of sales maximisation model being better in terms of welfare improvement also doesn't seem valid, as argued by many scholars.

2.4 LET US SUM UP

Dear learner, in this lesson, we discussed in detail Baumol's sales maximisation model.

2.5 KEYWORDS

Overhead Costs : These costs may not be directly related to production, but are required for the working of a production process e.g. rent, insurance etc.

Non-price competition : In certain market structures like Oligopoly, Monopolistic Competition, the firms don't compete on price, but non-rice factors play a more important role in attracting the customers e.g. different labelling, design, colour, service quality etc.

MRP : Marginal Revenue Product refers to the additional made to total revenue (MR) due to an additional unit of a factor.

MPY : Marginal Profit Yield means the extra profit earned by a firm on account of selling an additional unit of output.

2.6 EXAMINATION ORIENTED QUESTIONS

1. Explain Baumol's sales maximisation model in detail. In what ways is the sales maximisation hypothesis better than conventional model?
2. Discuss Baumol's model with and without advertisement. Does it better explain the behaviour of a business firm than that postulated by the traditional theory?
3. How does increase in overhead costs impact price output determination under Oligopoly? Is there a difference between a profit maximiser and a sale maximiser in this regard?
4. What do you mean by optimal advertising expenditure?
5. Is Baumol's model of sales maximisation constrained or unconstrained?
6. Justify your answer.

7. What reasons, according to Baumol, are responsible for pursuing sales maximisation as a goal?

8. What is the difference between a sale maximiser and a profit maximiser with regard to choice of output of two products?

2.7 HINTS TO CYP

CYP-I

Q1. See sub-section 2.3.1

Q2. See sub-section 2.3.2

Q3. See sub-section 2.3.4

2.8 SUGGESTED READINGS

1. Ahuja, H.L., Advanced Economic Theory, 9th edition (1997).
2. Koutsoyiannis, A., Modern Microeconomics, 2nd edition (1979).
3. Misra, S.K. and V.K. Puri, Advanced Microeconomic theory, first edition (2001).
4. Dwivedi, D.N. Microeconomic theory, 3rd edition (1999).

M.A. ECONOMICS **LESSON NO. 3**

COURSE NO. ECO 201 **UNIT-I**

O.E. WILLIAMSON'S THEORY OF MANAGERIAL DISCRETION

STRUCTURE

3.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

- Define Williamson's theory of managerial discretion
- Discuss Factors affecting utility maximisation

- Explain Basic concepts and relations
- Present Formal model
- State Equilibrium of the firm through graphical presentation

3.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to:

- Define Williamson's theory of managerial discretion
- Discuss Factors affecting utility maximisation
- Explain Basic concepts and relations
- Present Formal model
- State Equilibrium of the firm through graphical presentation

3.2 INTRODUCTION

Dear learner, another managerial theory of the firm is that given by Williamson. It was first developed in Williamson's paper published in 1963.Later, it was extended in the book *The Economics of Discretionary Behaviour,* published in 1964.In this lesson, we shall discuss this theory in detail.

3.3 O.E. WILLIAM'S THEORY OF MANAGERIAL DISCRETION

According to Williamson's managerial theory of the firm, managers are motivated by their self interest, which lies in maximisation of their utility function. While all the managerial theories talk of maximisation of utility function of managers, the difference lies in how they approach the phenomenon. This objective, according to Williamson, is to be fulfilled to the constraint that after tax profits should be sufficient to pay the shareholders acceptable dividends as well as to be able to finance 'economically necessary investments', which is apart from the discretionary investment expenditure made by the managers. However, this is possible only in a corporate form of business organisation, where there is dichotomy between management and owners.

As per Cohen and Cyert : " The separation of ownership and management functions permit the managers of a firm to pursue their own self-interest, subject only to their being able to maintain effective control over the firm. In particular, if profits are at any time, at an acceptable level, if the firm shows a reasonable rate of growth over time and if sufficient dividends are paid to keep the stockholders happy, then the managers are fairly certain of retaining their power."

3.3.1 Factors affecting utility maximisation

Salary, security, power, status, prestige, job satisfaction, professional excellence etc. are the most important factors. Of these, only first variable is measurable, while all others are non-measurable (non-pecuniary). But it is important to make these operational. So they need to be expressed in terms of other 'measurable variables'. 'Expense preference' is the name given to such an adjustment. It is defined as the satisfaction managers derive from certain types of expenditure. For this purpose, Williamson has added staff expenditure, managerial emoluments and discretionary investment in the formal model, as these reflect the aforementioned non-measurable variables.

According to Williamson, the utility function depends on the following factors:

The salaries and other forms of monetary compensation These influence the private consumption expenditure and hence the standard of living of the managers . So these are important, though not the only determinant of utility function of a manager.

Number of staff under the control of managers How much staff is being controlled by the manager is another factor affecting the utility function of managers. It is because the salaries and other monetary compensations are positively related to the staff. It thus increases the status and hence gives more power to the manager. Since salaries depend on the number of staff, in his formal model, Williamson has taken these two variables (salaries+ no. of staff) as one i.e. 'monetary expenditure on the staff'. Staff increases are, according to Dr. A. Koutsoyannis, to certain extent , equivalent to promotion, since they increase the range of activities and control of managers over resources. It is also in a way, a measure of 'professional success', as increasing staff means an expansionary phase is going on in the particular field controlled by the manager.

Management Slack It consists of those 'non-essential management perquisites' likewellfurnished offices,luxurious companycars etc., which is not essential for the smooth functioning of the firm. Scholars call it economic rent or payments with zero productivity, as if withdrawn, it won't make managers quit the enterprise. It is enjoyed by the managers due to the strategic position of managers. Since most of it is invisible, it is not usually a centre of attraction of shareholders or the labour force.

Discretionary Investment Expenditure Another determinant of utility function is the magnitude of discretionary investment expenditure incurred by the managers. It should be noted that it does not include minimum investments required for the normal functioning of the firm. As we have noted in the beginning of this model, utility maximisation is subject to a minimum profit constraint (economically necessaryinvestment+satisfactorydividend policy). The minimum required investment has already been included in the profit constraint. This discretionary expenditure is also an important source of pleasure for the managers and top executives, as they can spend the way they like, promote the projects of their choice etc.

3.3.2 Basic concepts and relations

a) The demand for the firm : A downward sloping demand curve is assumed. The function that defines it is :

$$X = f1\ (P,S,e)$$

Or $$P = f2\ (X,S,e)$$

Where X =output

P = price

S = Staff expenditure

e = condition of the environment or a demand shift parameter reflecting autonomous changes in demand.

The demand is assumed to be negatively related to price and positivelyrelated to staff expenditure and to the shift factor 'e'.

Thus, $\frac{\partial p}{\partial x} < 0$; $\frac{\partial p}{\partial s} > 0$; $\frac{\partial p}{\partial e} > 0$

This is because, an increase in staff expenditure is expected to cause a positive impact on price change. Same holds true about any change in 'e'[1].

b) The production cost Production cost is assumed to be an increasing function of output.

$$C = f3\,(X)$$

Or $$\frac{\partial c}{\partial x} > 0$$

c) Concepts of Profit

Actual profit : Actual profit is revenue minus cost minus staff expenditure

Or $\angle = R\text{-}C\text{-}S$

Reported profit : It is the actual profit minus the tax deductible managerial emoluments. It is actually reported to the tax authorities.

Or $\angle_R = \angle\text{-}M$

Or $\angle_R = R\text{-}C\text{-}S\text{-}M$

Minimum profit : The post tax profits that must be maintained, so that the job security of the top management is not threatened, refers to the minimum profit constraint of the firm. This is denoted by $\angle_0$

$$\angle_0 \leq \angle_R - T \text{ Or}$$

$$\angle_R \geq \angle_0 \mp$$

Where T=Tax.

Further, the tax function has been defined as:

$$T = \bar{T} + t.\angle_R$$

Where t = marginal tax rate

$\bar{T}$ = a lump sum tax.

Discretionary Investment. It is derived as:

$$I_D = \angle_R - \angle_0 - T$$

In other words, whatever is left after subtracting minimum profit and taxes from the reported profit is discretionary investment.

Discretionary profit : After we deduct the minimum profit and tax from the actual profit, we get discretionaryprofit.

$$\angle_D = \angle - \angle_0 - T$$

3.3.3 Formal Model

In the simplified model, that shall be discussed below, Managerial emoluments have been assumed to be zero (M=0). In other words, there is no difference between actual profit and reported profit.

The model may be presented as:

Maximise $U=f(S, I_D)$

Subject to

$$\angle \geq \angle_0 + T$$

Also, there is no differencebetweendiscretionaryinvestment and discretionary profit.

So

$$U=f[S, (\angle - \angle_0 - T)]$$

Further, if it is assumed that there is no lump sum tax,

$$U=f[S, (1-t)\angle - \angle_0]$$

Where

$(1-t)\angle - \angle_0 = \angle_D$ is the discretionary profit.

Check Your Progress-I

Answer the questions in the space provided :

Q1. What is the constraint faced by the managers in Williamson's model of utility maximisation?

Q2. What are the main determinants of utility function, according to Williamson?

Q3. Discuss different concepts of profit in Williamson's model.

3.3.4 Equilibrium of the firm : Graphical presentation

For graphical presentation of the model, we need two tools i.e. indifference map of the managers and the discretionary profit-staff curve. The indifference curve has been drawn between two variables staff expenditure (on the X-axis) and discretionary profit (On the Y-axis). The indifference curve here shows the different combinations of S and $\angle_D$, which give the same satisfaction to the mangers. The indifference curve has been drawn conventionally i.e. is a convex shaped curve, which shows a diminishing MRS S, $\angle_D$. Moreover, the indifference curves do not touch either axis, which means that the choice of managers has been restricted to positive level of both the variables. This map has been shown in Fig. below :

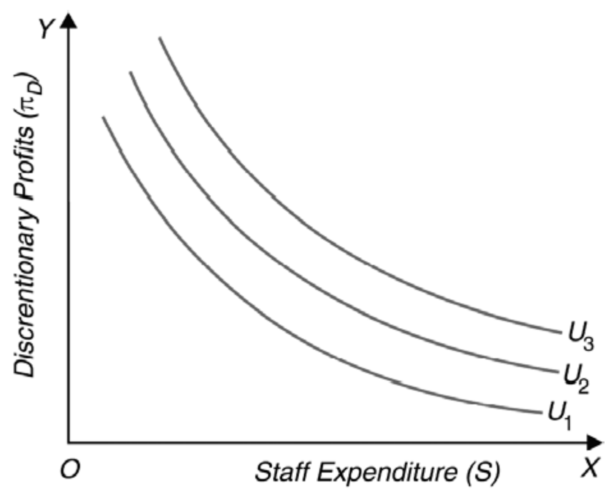


Fig. 3.1 : Indifference curves of managers

As shown in the figure, U1, U2, U3 etc. are the indifference curves of managers.

Further, the relationship between discretionary profit and staff expenditure can be expressed in the form of Fig. 3.2 drawn below. This relationship is determined by the profit function

$\angle$ =f(X) =f (P,S,e) (As 't' and '$\angle_0$' are exogenously determined).

If we assume optimal choice of output (chosen according to marginalistic principle) and given the environment 'e', the relationship between ' and 'S' can be expressed as :

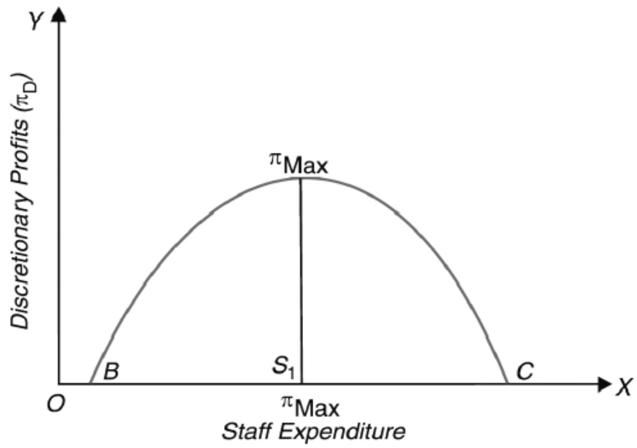


Fig. 3.2 : Relationship between π_D and S

As is shown in the figure above, at initial level of production and up to the maximum level of profit BS, both the discretionary profit as well as staff expenditure increase. But after this, the staff expenditure continues to increase, while discretionary profits fall. Further, it should be noted that if production is increased beyond point 'C', even the minimum profit constraint is not fulfilled. Likewise, points to the left of 'A' are also not feasible.

The equilibrium of the firm in Williamson'smodel is determined at the point of tangency of profit- staff curve with the highest possible managerial indifference curve. It is shown in Fig. 3.3 below.

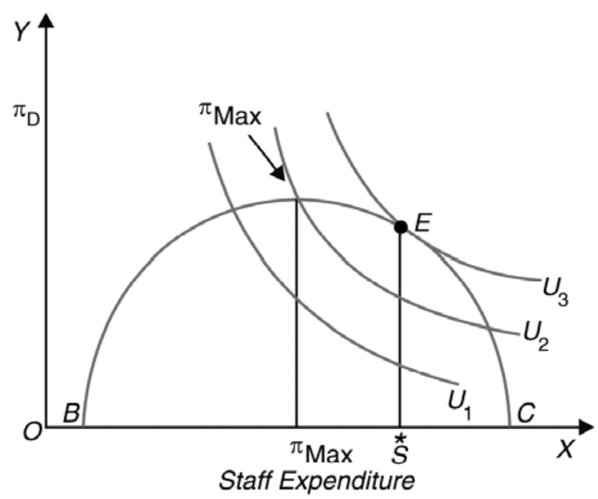


Fig. 3.3 : Equilibrium of the firm

E is the point of equilibrium. As we can observe, here, the total discretionary profit ES_2 is and staff expenditure is OS_2. Had it been a profit maximisation model, profits earned would have been BS_1 and staff expenditure, OS_1 So staff expenditure is more and discretionary profits earned are less in this model as compared to a profit maximisation model.

There are certain observations made by the scholars regarding the predictions of Williamson's model. These are:

- Under perfect competition model, the predictions of Williamson's model match with those of a profit maximisation model. Under Oligopoly and Monopoly (weak competition),the expenditure of advertisements, managerial luxuries etc. shall be greater and hence a greater amount of management slack shall be absorbed as cost, which will lead to an upward trend in prices. On the other hand, if demand weakens, this expenditure shall be reduced, thus reducing the cost and hence price.
- An increase in profit tax rate results in increase in staff expenditure and output. This is because, when profit tax rate increases, managers have to look for alternative ways of maximising utility rather than earning profits. So they spend more on staff expenditure and increase output, so that less profits are reported. On the other hand, if instead of a tax rate, lump sum tax is imposed, staff expenditure will be reduced, as minimum profit constraint shall be raised; and to earn higher profits, there is a need to reduce expenditure. However, a profit maximising firm will neither increase nor reduce expenditure in this case.

To conclude, we can comment that Williamson's predictions are based on empirical evidence of the firms' reactions to changes in taxation, changes in expenditure on staff etc. According to him, these things can be better explained only by his managerial theory of the firm. However, this model fails to consider interdependence among oligopolistic firms or where there is strong rivalry.

3.4 LET US SUM UP

Dear learner, in this lesson, we discussed Williamson's model of utility maximisation in detail.

3.5 KEYWORDS

Managerial discretion : Unlike the conventional models, models like Williamson's model acknowledge the discretion or freedom on the part of managers to fulfil certain personal goals apart from profit maximisation, if the former align with their objectives.

Discretionary profit : It refers to the income left over and above the essential costs and which can be used by the owner of as firm at his discretion e.g. for saving, luxury etc.

Shift factors : Note : 1. As you can recall from the earlier class, shift factors are factors other than price, that affect quantity demanded of a commodity. These cause an upward or downward shift in the demand curve. An upward shift increases demand. If there is no change in supply, price rise takes place. Opposite is true in case of a downward shift

3.6 EXAMINATION ORIENTED QUESTIONS

1. Discuss Williamson's model of utility maximisation in detail.
2. Discuss the different concepts of profit and their relations in Williamson's model.
3. How does a firm attain equilibrium in Williamson's model?
4. What analytical tools have been used by Williamson to explain equilibrium of the firm?
5. Do you think that Williamson's model satisfactorily explains equilibrium of the firm under oligopolistic structure? Does it take into consideration interdependence among firms?

6. What is the impact of profit tax rate and lump sum tax on staff expenditure and output in Williamson's model? How does a profit maximisation model differ from it?
7. Give a graphical representation of Williamson'smodel.
8. Whatare the factors affecting utility maximisation according to Williamson?
9. What is management slack?
10. What is discretionary investment? How does it differ from necessary investment?
11. What is a satisfactory dividend policy? Why is it required?

3.7 HINTS TO CYP

CYP-I

Q1. See sub-section 3.3.2

Q2. See sub-section 3.3.1

Q3. See sub-section 3.3.2

3.8 SUGGESTED READINGS

1. Ahuja, H.L., Advanced Economic Theory, 9th edition (1997).
2. Koutsoyiannis, A., Modern Microeconomics, 2nd edition (1979).
3. Misra, S.K. and V.K. Puri, Advanced Microeconomic theory, first edition (2001).

M.A. ECONOMICS **LESSON NO. 4**

COURSE NO. ECO 201 **UNIT-I**

MARRIS' THEORY OF THE MANAGERIAL ENTERPRISE

STRUCTURE

4.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

- Discuss Marris' model of managerial enterprise
- Define Utility functions
- Explain the Constraints
- Present a critique of the model

4.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to:

- Discuss Marris' model of managerial enterprise
- Define Utility functions
- Explain the Constraints
- Present a critique of the model

4.2 INTRODUCTION

Dear learner, earlier in this unit, we have learned two theories of managerial discretion. In this lesson, we shall be discussing Marris' model. Marris too is of the opinion that profit maximisation is not the only goal of the firm. Let us now discuss the model in detail.

4.3 MARRIS' THEORY OF THE MANAGERIAL ENTERPRISE

4.3.1 Goal and utility function

According to Marris, the goal of the firm is the maximisation of the balanced rate of growth of the firm i.e. maximisation of the rate of growth of demand for the firm's products and of the growth of its capital supply:

Maximise $g = g_D = g_C$

Where g = balanced growth rate

g_D = growth of demand for the products of the firm

g_C = growth of the supply of capital

According to Marris, the utility function of shareholders is different from that of managers. Variables like profits, size of output, size of capital, market share and public image enter the utility function of the owners, while managers' utility function includes salaries, job security, power and status. However, there is not much difference between the utility functions of managers and shareholders, as in both the functions, most of the variables have strong relationship with a third variable i.e. size of the firm. Moreover, in each function, it is reasonable to assume that maximising the long run growth of any indicator is equivalent to maximising the long run growth rate of others.

Let us now discuss the utility functions of managers and owners separately:

It is assumed that for maximising the growth of the firm, owners want to maximise the growth of supply of capital.

$$Uo = f(g_C)$$

Where Uo = utility of owners

g_C = rate of growth of capital

On the other hand, managers believe that growth of demand for the products of the firm is an appropriate indicator of the growth of the firm. The utility function of managers is:

$$UM = f(g_D, s)$$

Where UM = utility of managers

g_D = rate of growth of demand for the products of the firm

s = measure of job security

Here certain things need to be noted. According to Marris, decision making capacity of the managerial team sets a constraint to rate of growth of demand for the products of the firm. He further considers job security as a weighted average of the liquidity ratio, the leverage/debt ratio and the profit retention ratio. These together determine the financial policy of the firm. Further, he assumes that job security has a saturation point. Till that saturation is reached, the Marginal Utility from increase in job security is infinite, whereas once saturation is reached and above that, MU is zero. Job security is an exogenously determined constraint. If we consider this constraint, the above function becomes:

$Um = f(g_D)$

Where $\bar{s}$ = security constraint

4.3.2 Constraints

In the initial model, Marris assumes two constraints:

- The managerial team constraint
- The job security constraint

The managerial team constraint. Here Marris uses the approach made by Penrose. Penrose's thesis is related to a limit on the rate of efficient managerial expansion. At any given time, there is a ceiling on the expansion of the top management. According to Penrose, hiring new managers can expand the managerial capacity, but that too has a limit. Decision making and planning need teamwork, for which cooperation and coordination among different managers is required. However, in this regard, experience makes a lot of difference. In other words, a new manager takes time to get adjusted to a new environment. He is not fully ready to contribute towards smooth functioning of the organisation immediately after joining.

Likewise, in the Research and Development (R&D) department, new ideas and products affect the growth of demand for the products of the firm. But here also, the new scientists, engineers, designers etc. need time to adjust. Here again, simply by hiring more of these experts won't lead to expansion, unless they have stayed with the organisation for some time.

Hence these above two variables set a limit to the rate of growth of demand (g_D) and rate of growth of capital supply (g_C).

The job security constraint. Another constraint is the job security constraint. According to Marris, there is a definite disutility that managers attach to the risk of being dismissed. They usually prefer service contracts, generous pension schemes and dislike policies which put their job at risk. For this, they adopt a prudent financial policy, as they believe that any weakness of the firm on this front may lead to their dismissal. Here the managers take policies to avoid financial failure or bankruptcy and to avoid the firm's being taken over by take over raiders. The former puts the managers' job at risk, as the shareholders may think that a change in top management may be better for the firm. The latter again puts them at risk of being dismissed, as take over raiders may also think of changes at the top level.

There are different ways in which risk of dismissal can be avoided by the managers e.g., they avoid risky ventures. Rather they want steady performance of the firm, due to which they become risk avoiders. Another way is to choose an optimal policy with regard to three crucial ratios. These are:

$$\text{Liquidity ratio} = a1 = \frac{\textit{Liquid assets}}{\textit{Total assets}} = \frac{L}{A}$$

$$\text{Leverage or debt ratio} = a2 = \frac{\textit{Value of Debts}}{\textit{Total assets}} = \frac{D}{A}$$

$$\text{Retention ratio} = a3 = \frac{\text{Retained profits}}{\textit{Total profits}} = \frac{\square R}{\square}$$

Let us now explain these three ratios and their importance for managers. The debt ratio has been defined above. The managers wish to avoid insolvency and bankruptcy. So they don't resort to excessive borrowings. Likewise, liquidity ratio is also very important. It should be neither too high nor too low, as too high a ratio may endanger the job of managers by making it attractive to take over raiders and too low may render the firm insolvent. That is why, Marris advocates an optimal liquidity ratio. He has assumed in his model that the firm operates in that region where there is positive relation between liquidity and security. So an increase in liquidity increases security. The third ratio is retention ratio, which means the ratio of retained profits to total profits. Though, according to Marris, this is very important, as it becomes the basis for growth of capital, there is a limit to which profits can be retained. Distributed profits should not be too low to lead to a fall in price of shares, which makes possible its takeover by take over raiders. Low profits can make the shareholders take a decision to replace the managers.

Marris further states that the three financial ratios are subjectively combined by the managers into a single parameter $\bar{a}$, which is referred to as 'financial security constraint'. According to Marris, this constraint is exogenously determined by the top management, depending on their risk attitude. It is a weighted average of the three ratios, weight depending on the subjective decisions of managers.

As far as the relationship of $\bar{a}$ with the three ratios is concerned, Marris says that it is negatively related to a1, while positively to a2 and a3. In other words, to raise the value of $\bar{a}$, either the liquidity ratio is reduced or debt ratio should increase (which means more of loans or borrowings), or more profits are retained. On the other hand, more liquidity, reduction in borrowings (external finance) or increase in distributed (and hence less of retained) profits or a combination of these lead to reduction in $\bar{a}$. Ahigh value of $\bar{a}$ means that managers are risk takers and a low value means that they avoid risk. Further, there is a negative relation between job security and $\bar{a}$. More the value of $\bar{a}$, more is the risk taking attitude and hence more vulnerable is the situation of the firm. This endangers the position of the managers.

Check Your Progress-I

Answer the questions in the space provided :

Q1. What is the goal of the firm according to Marris?

____

____

____

____

Q2. What is the difference between Uo, UM and Um?

____

Q3. What are the two constraints faced by the firm?

Q4. What is $\bar{a}$? Is it determined exogenously or endogenously?

Q5. What is the relationship between :

- $\bar{a}$ and the three ratios
- $\bar{a}$ and the job security of managers.

4.3.3 The model : Equilibrium of the firm

We have discussed in the beginning the utility functions of managers and shareholders. As you can recall, they are:

$U_o = f\,(g_C)$

Where Uo=utility of owners

g_C = rate of growth of capital

UM=f (g_Ds)

Where UM=utility of managers

g_C =rate of growth of demand for the products of the firm

s = measure of job security

The firm is in equilibrium , when maximum balanced growth rate is attained

i.e. $g_D = g_C = g^*$ maximum.

As given in the above equation, we need to know the factors on the demand and supply side. In other words, what are the factors that determine g_C and g_D Marris expresses these in terms of two variables, the diversification rate, 'd' and the average profit margin, 'm'.

For determining the point of equilibrium, we need to discuss the following:

- The instrumental variables
- The rate of growth of demand
- The rate of growth of capital supply.

Let us now discuss these in detail :

4.3.3.1 The instrumental variables

The firm first determines its financial policy (i.e. value of financial constraint) and then determines the rate of diversification and the average profit margin, so as to maximise balanced growth rate. All the variables discussed here are subjectively determined by the firm.

The policy variables are :

- $\bar{a}$ i.e. the freedom of choice of the financial policy of the firm. The three ratios (discussed in section 4.3.2) affect the rate of growth of firm.

- Diversification rate ‘d’. The firm can diversify by changing the style of existing products or add or change products in its existing portfolio.

- Marris doesn’t consider price as a policy variable. It is an oligopolistic market structure. He does acknowledge the existence of a price war or non-price competition like advertising, product variation etc. or sometimes a tacit collusion, that ultimately results into a price structure. But he considers price as well as production costs as given (parameter) rather than a policy variable, in which firm is free to decide.

- The firm can exercise discretion with regard to determining ‘A’ and ‘R&D’ i.e. expenditure on advertising and Research and Development respectively. Higher the A or R&D or both, lower shall be the profit margin and vice versa.

In Marris’ model, Average Cost pricing rule is implicit:

$$\bar{P} = \bar{C} + A + (R\&D) + m$$

Where $\bar{P}$ = Price, (Given).

$\bar{C}$ = Production costs (Given)

A = Advertising and other sale promotion expenditures.

m = Average profit margins.

$\therefore \quad m = \bar{P} - \bar{C} - A - (R\&D)$

It should be noted that ‘m’ is a residual.

4.3.3.2 The rate of growth of demand

This model assumes that the only way in which the firm grows is by diversification. Hence it rules out the ‘merger’ or ‘take over’ options.

Now $g_D = f1(d, k)$

Where d = diversification rate

K= percentage of successful new products.

There are two ways in which diversification can take place. These are 'differentiated diversification' and 'imitative diversification'. The former refers to the introduction of a new product (which creates new demand), with no close substitute. As it doesn't provoke other firms to retaliate, firms prefer this the most. Imitative diversification, on the other hand, means when a product introduced is a close substitute of already existing products. As this may lead to retaliation by rivals, given the uncertain reaction pattern of oligopolistic firms, it is less popular than the other one. Higher the value of 'd', higher shall be the value of g_D.

Now, 'k', the proportion of successful new products depends on 'd', price of products (p), expenditure incurred on advertisements (A), R&D expenditure and intrinsic value of the products.

K=f2 (d, P, A, R&D, intrinsic value)

Here Marris has borrowed the concept of 'intrinsic value' from Galbraith's and Penrose's thesis, which states that the firm can sell almost anything to the consumers, even against the resistance of consumers, provided an appropriate sale campaign is followed . It is implicitly attached with price.

As given above, 'k' depends on 'A' and 'R&D'. K is a positive function of these two, as more expenditure on advertisements and/or on R&D would lead to higher proportion of successful new products. Marris uses 'm' as a proxy for these two variables. As 'm' is a negative function of these two (see section 4.3.3.1), 'k' is negatively correlated with 'm'.

Further, 'k' depends on 'd', the rate at which new products are introduced in each period. Though g_D is positively dependent on 'd', there is a limit to which 'd' can increase. g_D increases at a decreasing rate, as 'd' increases. This is because if new products are introduced without taking into consideration the adequacy of the firm regarding personnel required to handle the development and marketing issues or time required to research the product, it may result into more failures. In that case, 'k' i.e. the proportion of successful products would be less. So all these factors need to be taken into account.

gD = f1 (d, m)

$$\frac{\partial gD}{\partial d} > 0 \text{ (but declining)}$$

$$\frac{\partial gD}{\partial m} < 0$$

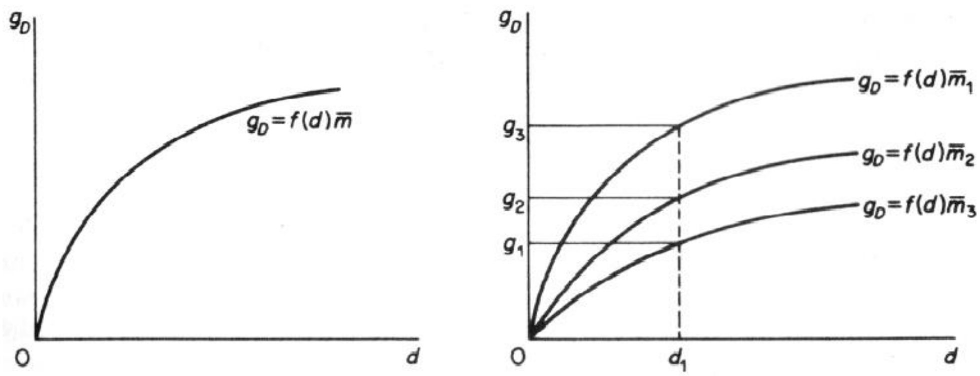


Fig. 4.1 :

Fig. 4.1 : $g_D = f(d)$ *given m* Fig. 4.2 : $g_D = f(d,m)$

Fig. 4.1 and Fig. 4.2

Figures drawn above explain g_D curve. Fig. 4.1 shows that along one g_D curve, the variable 'm' remains fixed or constant. As we can see, g_D is drawn as an increasing function of 'd' (though at a diminishing rate). Fig. 4.2 explains the shifting g_D curve. As we can see, g3>g2>g1. But $\bar{m}3 > \bar{m}2 > \bar{m}1$. This shows that as 'm' will increase, g_D will see a downward shift.

Check Your Progress-II

Answer the questions in the space provided.

Q1. What are the policyvariables in Marris' model? Whyis price not considered a policyvariable?

____

____

Q2. How does g_D depend on 'd' and 'k'?

Q3. What are the two types of diversification? Which is preferable?

Q4. Does every new product introduced by the firm result into success? Justify your answer.

4.3.3.3 The rate of growth of capital supply : g_C

The shareholders aim at maximisation of rate of growth of corporate capital, which the firms consider as a measure of the size of the firm. Corporate capital is defined as "the sum of fixed assets, inventories, short term assets and cash reserves."

Marris argues that the source of finance can be internal as well as external sources. External sources can be issue of new bonds, bank loans etc., while profit is the main internal source. As we have mentioned earlier, there is an upper limit to debt. So profits are used the most to finance the firm's growth requirements. We have also discussed earlier that there is an upper limit to retention ratio (see section 4.3.2). So though retained profits form the basis for growth of firms, the retention ratio should not be so high that it results into an unsatisfactory dividend policy, as that may endanger their (managers') own position.

According to Marris,

$$g_C = \bar{a}(\square)$$

where $\bar{a}$ = the financial security coefficient

$\square$ = level of total profits $\frac{\partial y}{\partial x}$

Initially, in this model, security coefficient is held constant. During this stage, the g_C and profits are not in conflict. Rather higher profits imply higher rate of growth of the firm.

Further, g_C is expressed in the form of policy variables d and m.

$$\square = f3\left(m, \frac{K}{X}\right)$$

Here $\frac{K}{X}$ is the capital output ratio of the firm. It reflects the efficiency of performance of the firm. So level of total profits depends on capital output ratio as well as m.

As we can see, total profits are a positive function of m.

$$\therefore \quad \frac{\partial \square}{\partial m} > 0$$

But the relationship between $\square$ and $\frac{K}{X}$ is not so simple. The latter is a function of diversification rate 'd'. If we assume that there is an upper limit on human and capital resources ,

$$\frac{K}{X} = f4\,(d)$$

So, given the capital constraint, Output X varies positively upto a certain level of ‘d’. It reaches a maximum, before declining. As long as the R&D team’s skills as well as those of management team are being utilised, it pays the firm to introduce new products. But after reaching optimum point, the growth suffers from decline, as the managerial team as well as R&D team have their own limitations and get fatigued, so that any new product’s introduction shall not contribute towards the growth of the firm. So the efficiency declines.

We can also write as

$$\square = f3\ (m, d)$$

So the relationship between level of total profits and rate of diversification is initially positive, reaches a maximum and finally declines as ‘d’ increases.

We can also write

$$g_C = \bar{a}\,[f3\ (m, d)]$$

which means that the rate of growth of capital is determined by three factors, financial policies of the managers, the average rate of profits and the diversification rate.

Initially, Marris assumes that $\bar{a}$ is a constant parameter, while g_C is a positive function of ‘m’.

$$\frac{\partial g_C}{\partial m} > 0$$

As far as relation between g_C and ‘d’ is concerned, it is not monotonic. g_C is positively related to d till the optimal utilisation of R&D team and managerial skills. After this, indiscriminate hastening of the process of diversification leads to hurried and inefficient decision making that ultimately results into low profits and hence a decline in g_C.

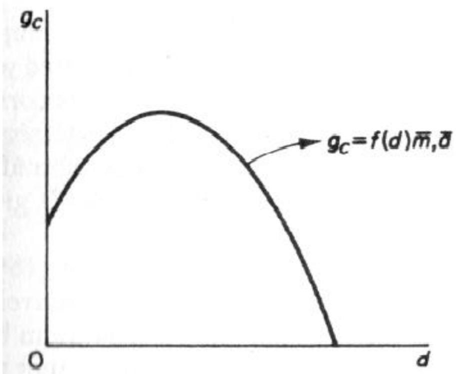


Figure 4.3 $g_C = f(d)$, given $\bar{m}$ and $\bar{a}$

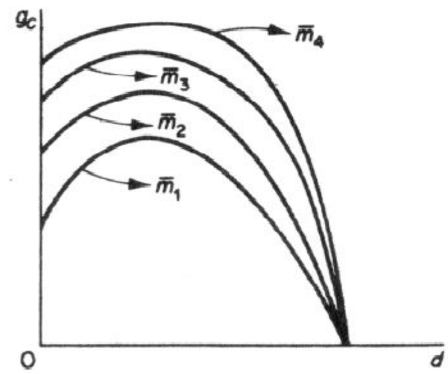


Figure 4.4 $g_C = f(d, m)$, given $\bar{a}$

Fig. 4.3 and 4.4

The figures drawn above show the relation between g_C and d (given m and $\bar{a}$) and g_C and d and m (given $\bar{a}$) respectively. The fig. 4.3 shows g_C on the Y-axis and 'd' on the X-axis. As we can see, g_C first rises, then reaches maximum and finally falls. Fig 4.4 shows the shifting of g_C curve, m being the shifting parameter. As we can see, m3>m2>m1. Higher the 'm', the further gc will be from the origin.

It should be noted that here we have not considered the implications of change in $\bar{a}$. Any increase in it shall shift g_C upward. Likewise, a decrease in it shall shift g_C downward.

Equilibrium of the firm : We have discussed above the different variables , the knowledge of which shall now enable us to determine the point of equilibrium of the firm.

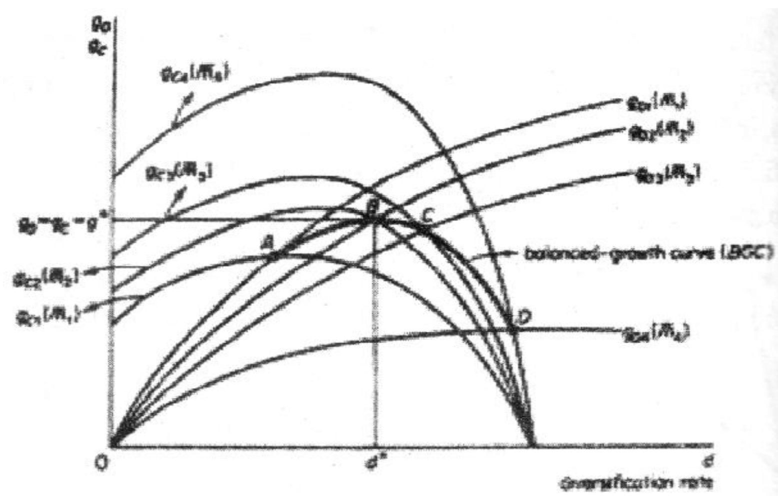


Figure 4.5

Here it is assumed by Marris that one of the policy variables (m or d) is endogenously determined by the managers. The equilibrium has been shown graphically in Fig. 4.5.

Here, as we can see, the fig 4.2 and fig. 4.4 have been superimposed. Each g_D as well as g_c corresponds to a particular value of 'm'. If we join the points of intersection of g_D and g_c corresponding to the same m, we get curve ABCD. This is the BGC or Balanced Growth Curve. According to Marris, the highest point on this curve is the point of equilibrium. If the firm has $\bar{a}$, this point shall be consistent with that policy. g* corresponds to point 'B' on the curve, which shows a unique pair of values of 'm' and 'd'.

Marris argues that as long as $\bar{a}$ is constant, growth and profit are not competing goals.

4.4 A CRITIQUE OF MARRIS' MODEL

Though Marris' model has made significant contributions to better understanding of the variables that enter the decision making process of firms, the model also suffers from certain drawbacks. Let us first understand the major contributions (implications):

- The financial policy variable $\bar{a}$ has been incorporated in the decision making process of the firm. Though $\bar{a}$ is exogenously determined, it does reflect the risk attitude of the managers. So in a way, it is endogenous as well (to some degree). The risk attitude of the managers is influenced by past performance of the firm. If the past performance of the firm is satisfactory, managers tend to be more adventurous and risk takers, as success rate of newly introduced products allows them to be risk lovers. On the contrary, they tend to avoid risk, if the past record of the firm in this regard is not good.
- The model has tried to reconcile the goals of two (hitherto) mutually conflicting entities i.e. managers and shareholders. Here the concept of BGC holds relevance. However, the maximisation of utility functions of both is based on the assumption that variables in both these functions are correlated with size and rate of growth. However, it may be true in periods of steady growth, while not in case of recession.
- As long as $\bar{a}$ is constant, Marris' model states that profits and growth are not competing goals. However, later, when this assumption is relaxed, they are projected as competing goals. This is what happens in real world. So like other models, this model too takes into consideration multiple goals being pursued by a real world firm. The owners sacrifice some profits

to have maximisation of growth of the firm, as they understand that all the goals can't be achieved simultaneously.

Some of the shortcomings of the model are :

- Production costs and prices are given. In this regard, A. Koutsoyiannis writes: "Oligopolistic interdependence is not satisfactorily dealt within Marris' model".
- The model states that by introducing new products, the firm can grow continuously. But according to the critics, the firms can't just sell anything to the consumers. Firms can't have unlimited influence over the consumers.
- The model has been confined to the consumer goods' firms only.
- Critics have also criticised the model on another ground i.e. lumping together of 'A'and 'R & D'. According to them, it is not possible to combine these two, as these are not equally effective in any given period.
- Marris fails to consider the interdependence of firms with regard to launch of new products. He talks of 'differentiated diversification' which does not lead to retaliation. But at the same time, he agrees that the product shall be imitated by other firms in the long run. So there develops an interdependence with regard to different things, including prices, which is ignored by Marris.
- All the firms don't have R&D departments. Sometimes, they imitate what others have invented, at other times they pay royalty to the firms who hold patents. This has been ignored by Marris.
- Welfare implications of this model are also not clear. If all the firms engage in introduction of new products and too much expenditure is incurred on advertisements, it may lead to misallocation of resources. On the other hand, consumers get access to new products i.e. variety, which increases their welfare.

4.5 LET US SUM UP

In this lesson, we discussed Marris' model in detail.

4.6 KEYWORDS

Instrumental/policy variables: At the firm's level, policy variables mean the variables which play a crucial role in firm's decision making . Firms can decide which of these to be considered or changed first to fulfill its objectives**.**

Rate of diversification: It means the rate or speed at which new products get added to or replace the old ones in a product mix.

4.7 EXAMINATION ORIENTED QUESTIONS

1. Discuss in detail Marris' model of managerial enterprise. Also give a critique of it.
2. What are the major constraints in Marris' model?
3. How does a firm attain equilibrium in Marris' model? What is the significance of BGC?
4. How does a firm reconcile the goals of managers and shareholders?
5. What are the implications of Marris' model? What are its major shortcomings?

4.8 HINTS TO CYP

CYP-I

Q.1. See Section 4.3.1

Q.2. See Section 4.3.1

Q.3. See Section 4.3.2

Q.4. See Section 4.3.1

Q.5. See Section 4.3.2

CYP-II

Q.1. See Section 4.3.3.1

Q.2. See Section 4.3.3.2

Q.3. See Section 4.3.3.2

Q.4. See Section 4.3.3.2

4.9 SUGGESTED READINGS

1. Koutsoyiannis, A., Modern Microeconomics, 2nd edition (1979).
2. Misra, S.K. and V.K. Puri, Advanced Microeconomic theory, first edition (2001).

M.A. ECONOMICS LESSON NO. 5

COURSE NO. ECO 201 UNIT-I

AVERAGE COST OR FULL COST PRICING THEORY

STRUCTURE

5.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

- Discuss Average cost pricing: Andrews' version
- Define Cost plus pricing analysis and fix-price marke.

5.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to :

- Discuss Average cost pricing: Andrews' version
- Define Cost plus pricing analysis and fix-price market

5.2 INTRODUCTION

Dear learner, in this lesson, we shall be discussing Andrews' version of Average cost pricing and the flex vs. Fix price market.

5.3 AVERAGE COST PRICING : HALL AND HITCH'S VERSION

This has been discussed in lesson 1 (see section 1.3.2)

5.4 AVERAGE COST PRICING: ANDREWS' VERSION

This version of full cost pricing is similar to that of Hall and Hitch. Andrews uses the concept of 'costing margin' which he defines as: "an addition to constant Average direct cost and will normally tend to cover the costs of the indirect factors of production and provide a normal level of net profit." Though costing margin is similar to the concept of mark up used by Hall and Hitch, there are certain differences. These are:

- Mark up is inflexible or constant, while the concept of costing margin has been granted some flexibility by Andrews, in reaction to competitive and market forces. This flexibility brings the model closer to pricing based on the principle of profit maximisation.
- Hall and Hitch have used kinked demand curve in their model, whereas it has not been used by Andrews.
- Hall and Hitch suggest that cost curve is U shaped, as Average Cost varies with output. However, Andrews argues that for a large relevant level of output, cost remains constant.

Let us now discuss the main points of Andrews' theory:

- Price fixed by the businessmen is equal to the estimated average direct cost of production plus a costing margin.
- Average direct cost (variable cost) curve is a horizontal line over a long stretch, if the prices of direct (variable) factors remain constant.

- The costing margin will cover the cost per unit incurred on the indirect (fixed) factors of production and on the normal rate of profit. Once fixed, the costing margin will remain constant. However, permanent changes in the fixed factors shall result in variation in costing margin. Andrews also attributes changes in costing margin to competitive and market forces.

Given the prices of variable factors, price shall remain constant irrespective of the level of output.

Once the price is fixed, the given product, which has a well established market, shall be sold at that price to the buyers.

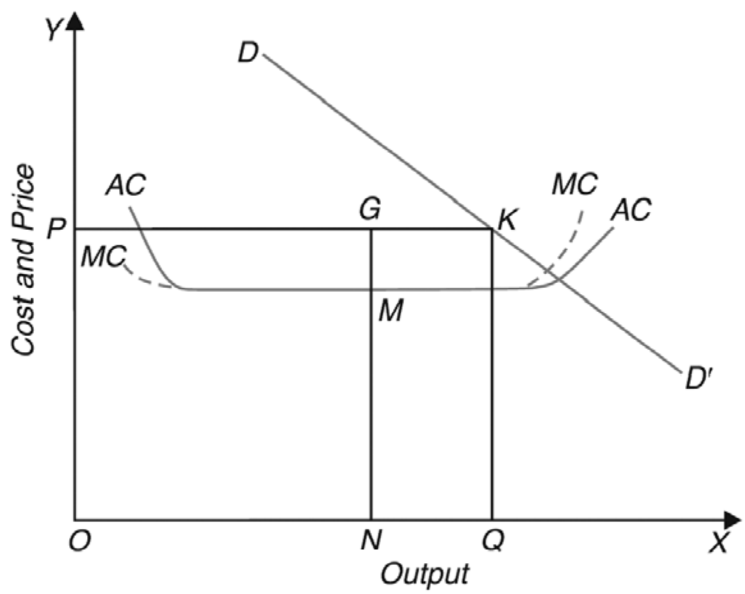


Fig. 5.1 : Price Fixation on the basis of Cost-Plus

The figure above shows Andrews' version of full cost pricing. As we can see in the figure, AC is actually Average direct cost (or variable cost). It is a saucer shaped curve, which means that it has a decreasing and increasing phase, but a relatively significant stretch of constancy. MC curve has the usual shape. It lies below AC, when AC is falling; lies above AC when the latter is rising and coincides with it when it is constant. According to Andrews, for determining the price, the firm needs to calculate indirect costs plus normal profits along with the direct costs. The fixed amount of money thus determined shall remain constant in the short run. The costing Margin, says Andrews, is obtained from this fixed sum of money by dividing it by "some chosen output".

There are different ways in which this level of output can be determined. These are :

- Capacity output can be the basis for it. In other words, the firm can fix a percentage of its total capacity as the level of output.
- Past sales over the preceding period or over a number of years in the past can also determine the level of output.
- Sales expected in the future can be another factor in this regard.

However, for a newly established firm or for a newly introduced product by an existing firm, the second factor is irrelevant. Only the first and third hold applicable.

As shown in the figure above, if the firm decides to choose ON level of output for estimating costing margin MG (fixed amount consisting of indirect cost plus normal profit i.e. over and above direct cost MN), full cost is NM + MG i.e. NG. So the price determined is OP (= NG). The price thus determined shall remain the same whatever the amount demanded and the level of output produced in the given period. Now if the demand curve is shown by DD', output demanded is OQ, which shall also be produced. But this shall not alter the price. As long as any change in demand takes place within the horizontal range of cost curve, it will not result into change in prices. Moreover, as long as there is no change in prices of direct and indirect factors, price shall remain the same.

5.5 COST PLUS PRICING ANALYSIS AND FIX-PRICE MARKET

Let us now understand the implications of the full cost pricing analysis done above:

We have seen that in full cost pricing rule, prices are not determined on the basis of equality of MC and MR, rather full cost is the basis. Secondly, unlike in the traditional market, where firms take prices as given, here the oligopolistic firms administer the price. Thirdly, once price is determined, if a mismatch is found between demand and supply, no adjustments are made in price to take the two variables back to equilibrium. Here we can draw comparison between this analysis and Marshallian analysis. In Marshallian analysis, price is flexible, as any disequilibrium between demand and supply is corrected through upward or downward movement in prices. But here, an oligopolistic market structure (which is known for stickiness of prices) rules out any such change in price. Hicks calls this system as 'fix-price market' as compared to Marshallian 'flex-price market' (or temporary equilibrium method).

Check Your Progress-I

Answer the questions in the space provided:

Q1. What are the major differences between Hall and Hitch's version and Andrews' version of full cost pricing rule?

Q2. What are the different ways in which output can be determined for estimating costing margin?

Q3. According to Andrews' version, how does price remain constant once it is determined?

But now, a question arises : If price doesn't make adjustment, then how is equilibrium restored between demand and supply? This has been discussed below:

Excess Demand: Let us first discuss the case of excess demand. In the figure drawn below, the basic analysis is the same as done earlier. As long as demand is represented by demand curve DD, quantity demanded equals supplied. But as demand curve shifts to D'D', there is excess demand GT (=QQ') at existing price OP.

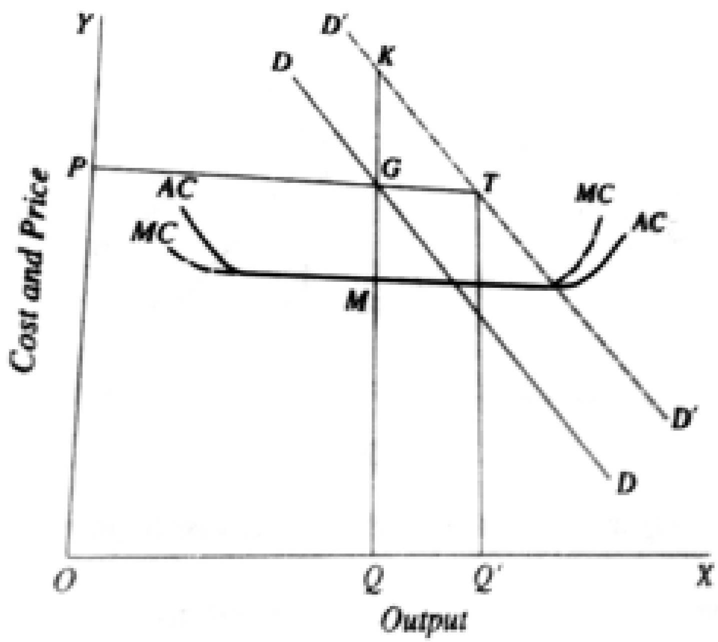


Fig. 5.2 : Average Cost Pricing and Excess Demand

To restore equality between demand and supply, price must rise to QK . But as it is ruled out here, the firm thinks of other options. It may draw upon the existing stock to meet the shortfall, if the change in demand is temporary; it may raise price slightly as a simultaneous step in case demand change is both persistent and large; in case demand increases permanently, the firm increases its productive capacity. If the firm feels that demand change is temporary, new production capacity will become available in near future and/ or goods can be imported, the price is likely to remain below the equilibrium level.

Excess Supply : Let us now explain excess supply as an opposite situation faced by the firm.

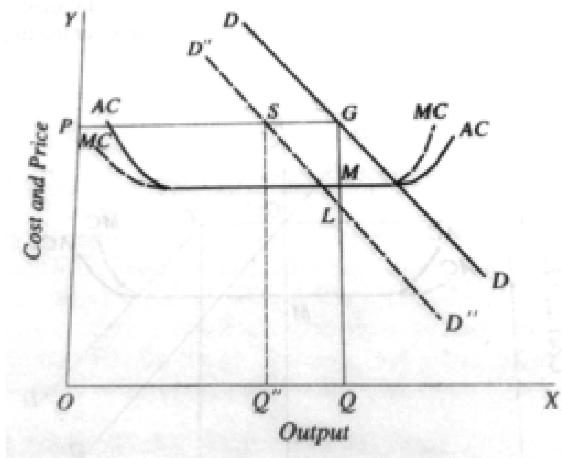


Fig. 5.3 : Cost-Plus Pricing and Excess Supply

The fig. above shows that originally, the quantity demanded and supplied are equal (OQ or PG) at OP price. Let the demand curve shift from DD to D''D.''In case of a flex price market, price would have got reduced to OL. But in a fix price market, the firm won't change price below OP (if there is temporary fall in demand). So at OP, there is excess supply equal to SG (=QQ'').The firm would use this to increase its stock and store it for future use. However, it is not so easy to always resort to this method for handling the issue of excess supply. It is costly to finance this, technological changes maydiscourage the firms from resorting to it, as products may become obsolete etc. So there is a limit to which this can be relied upon.

We have discussed above that full cost analysis also explains stability of prices in an oligopolistic market.

5.6 LET US SUM UP

Dear learner, in this lesson, we discussed in detail average cost or full cost pricing theory.

5.7 KEYWORDS

Kinked demand curve: This demand curve originated in Sweezy's model of Oligopoly in which price rigidity or stickiness was attributed to it. In a kinked demand curve, two different elasticities exist for two parts of the demand curve i.e. above and below the kink. Upper segment is more elastic and lower segment is less elastic.

Average direct cost: Also referred to as variable cost, it means the core costs directly related to output e.g. cost of raw material, wages to staff hired or other costs of production.

Costing margin: It has a meaning similar to Mark up, but is more flexible.

5.8 EXAMINATION ORIENTED QUESTIONS

1. What is full cost pricing or mark up pricing?
2. Discuss how Hall and Hitch integrated Average cost pricing with Kinked demand curve.
3. What are the main features of Andrews' version of full cost pricing?
4. How does Andrews' version differ from that of Hall and Hitch?
5. What is a flex price market? How is equilibrium reached in such a market?
6. What is a fix price market? Discuss excess demand and excess supply situation in such a market.
7. What are the limitations on a firm with regard to use of stock holding?
8. Do you agree that full cost pricing shows stability of prices in an oligopolistic market? Justify your answer.

5.9 HINTS TO CYP

CYP-I

Q1. See sections 5.3 and 5.4

Q2. See section 5.4

Q3. See section 5.4

5.10 SUGGESTED READINGS

- Pindyck R.S., D. Rubinfeld and P.L. Mehta (2012), Microeconomics, Pearson Education, South Asia.
- Salvatore, D. (2010), Microeconomics; Theory and Applications, Oxford University Press, New Delhi

- Cooper, R., & John, A. (2011). Economics-theory through applications, Saylor Foundation. 1
- Shapiro, D., MacDonald, D., & Greenlaw, S. A. (2022). Principles of Microeconomics 3e.2

END NOTES

Note 1: Type of license/Conditions of use: CC BY-NC-SA (Attribution-Non commercial-ShareAlike) .Scan QR code 3.2 to access **Cooper, R., & John, A. (2011). Economics–theory through applications.**

QR code 1

Note 2 . This book is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). "Access for free at openstax.org." **Scan QR code 2.1 to access the e book by Shapiro, D., MacDonald, D., & Greenlaw, S. A. (2022). *Principles of Microeconomics 3e*.**

QR code 2

M.A. ECONOMICS **LESSON NO. 6**

COURSE NO. ECO 201 **UNIT-I**

BAIN'S LIMIT PRICING THEORY AND ITS RECENT DEVELOPMENTS INCLUDING SYLOS LABINI MODEL

STRUCTURE

6.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

- Define the concept of limit pricing

- Explain Bain's model
- Describe Sylos-Labini's model.

6.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to:

- Define the concept of limit pricing
- Explain Bain's model
- Describe Sylos-Labini's model.

6.2 INTRODUCTION

Dear learner, in the earlier semester, we have discussed about the traditional models of Oligopoly. In their models, these traditional theorists like Cournot, Bertrand, Edgeworth and Chamberlin have not considered the entry of new firms. So the no. of firms remains the same during discussion. So these are termed as closed models. In recent literature, economists like Bain, Sylos-Labini, Andrews, Modigliani and Jagdish Bhagwati have discussed how price-output decisions of firms in such a market structure is influenced not only by actual, but also potential entry of firms. These economists argue that firms aim at long run and not short run profits, as profit possibilities get influenced once we take into consideration potential entry (or threat) of new firms. In this lesson, we shall discuss about theory of limit pricing, one such theory, in detail.

6.3 BAIN'S LIMIT PRICING THEORY

This theory was developed by Bain, first in his pioneering work 'A note on pricing in Monopoly and Oligopoly (1949)', followed by his book ' Barriers to new competition'. In this theory, known as limit pricing or entry preventing pricing, Bain argues that firms under collusive Oligopoly don't aim at maximisation of short run profits, but charge a low price, so that entry to the industry doesn't look attractive; and hence entry of new firms is prevented. The price charged is lower than the short run profit maximising price (under Monopoly) and higher than LAC (perfectly competitive market price).

6.3.1 Assumptions

Bain makes the following assumptions in his model:

- In each industry, there is a minimum optimal scale i.e a minimum size of the plant at which economies of scale can be realised.
- Since technology used is same for all the firms, all the firms observe an L shaped LAC. But the established firms have increased their output beyond minimum optimal, whereas the new firms start with sub-optimal and in due course of time, reach minimum optimal level.
- The flat part of LAC determines long run competitive price.
- The market demand curve is known to the existing as well as potential firms. The determinate demand curve for the industry remains unchanged when the existing firms make price adjustments or prices change due to entry of new firms.
- Firms produce very similar products and hence have equal market share.
- The relevant demand curve for each firm shall be a share of the market demand curve.

According to Bain, the level of limit price is set by the following factors:

- The costs of the potential entrants.
- Price elasticity of demand for the product.
- Size of the market (or demand).
- No. of existing firms.
- Level and shape of LAC.

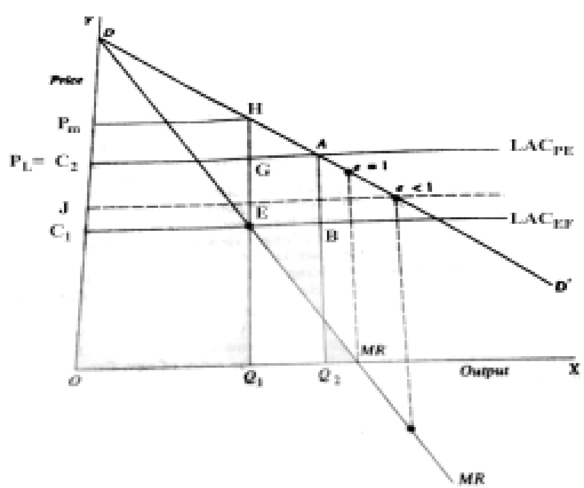


Fig. 6.1 : Limit Pricing Model

The figure above explains Bain's model. Let us assume that DD^1 is the market demand and MR is the Marginal revenue curve. It is a case of collusive Oligopoly.

Let LAC_{EF} be the LAC of existing firms. $LAC_{EF} = LMC_{EF}$, as LAC is constant. If the firms in Collusive Oligopoly wish to maximise profits in the short run, OP_m price is charged (like in Monopoly), where profit per unit is C_1P_m and total profits equal C_1P_mHE. However, as mentioned earlier, price charged is less than OP_m. Let it be OP_L. This corresponds to the LAC_{PE} or the LAC of potential entrants. If this price is charged, firms will still earn profits (C_1C_2AB), though less than in case of Monopoly. Here price charged is limit price, as at this price, no potential entrant can think of entering, as at the existing price, only LAC is covered. In case the new firms enter the industry, price shall fall below OP_L, where not even LAC shall be covered. It is important to mention here that the existing firms sacrifice their short run profits, so as to prevent the entry of new firms.

In a way, Bain has discussed the concept of barriers to entry, by placing the LAC of new firms at a higher level as compared to that of existing firms. In other words, the absolute cost advantage of existing firms over the potential entrants is a barrier to entry. However, if the LAC of potential firms is OJ, price is set at OJ, corresponding to which the elasticity of demand on the demand curve is <1. So in this case, the entry prevention strategy is to charge a price at which MR is <0, so that the new firms are prevented from entering the industry.

However, if Monopoly price is less than limit price, Monopoly price shall be charged by the firms of Collusive Oligopoly.

6.3.2 Barriers to entry and limit price

In his later model, Bain has tried to explain why the Oligopolistic firms set the limit price above perfectly competitive price. The new firm enters the industry when it creates a new productive capacity that is not used before its entry. Further, he gives the concept of 'condition of entry' (or entry gap) which is defined as the margin by which the existing firms can raise price above competitive price, so as to discourage entry of new firms.

$$E = \frac{P_L - P_C}{P_C}$$

Where E = condition of entry, P_L and P_C are limit price and competitive price respectively.

Or $P_L = P_C (1+E)$

Following are the barriers to entry discussed by Bain :

Absolute cost advantage. According to Bain, the established firms enjoy an absolute cost advantage over the potential entrants, due to which the latter find themselves incompetent and hence do not enter. The absolute cost advantage can arise from the following:

- Services of expert managers are available to existing firms, but for availing these, the new firms need to pay more.
- The established firms have better access to sources of raw materials. They either control the supply of these or get at low cost due to bulk purchase or have special arrangements with suppliers of raw materials. A new firm has no such advantage.
- The firms are also able to maintain low cost due to patents or by adopting a superior technique.
- Vertical integration of production process is another way to reduce costs. A new firm, even if desirous of engaging in this process shall not find it financially viable to continue.

Initial capital requirements. Bain also observed that in some industries, initial capital requirements served as an entry barrier. In today's world, establishing a new firm needs a lot of capital, due to sophisticated technology being adopted. But for a new firm with inadequate resources, this becomes a barrier, as it can neither raise funds through capital market, nor are banks so enthusiastic to lend loans to a new firm. Even if the banks agree, they charge a high rate of interest. Under these circumstances, new firms find hard to invest to the extent required and hence are not able to enter.

Economies of scale barrier. Economies of scale act as another important barrier to entry. As we have learned earlier, there are real and pecuniary economies. While in real economies, input per unit of production is reduced, in pecuniary economies , cost of inputs falls. Real economies are like technical, managerial and labour economies. Technical economies mean that the established (big) firms can make use of sophisticated technology. Managerial economies mean that the same managerial input when spread over larger units of output leads to returns. Labour economies mean division of labour and opularizeion, which raises the productivity of labour. Pecuniary economies include low prices of inputs due to bulk buying, low transport costs, low advertising and selling costs etc.

These economies of scale act as barriers to entry, as the established firms are already operating at optimal scale. The potential entrants are aware that these firms have fixed a price above competitive price. They may reduce it to competitive level, that will endanger its (entrant's) survival. Even if the firms reduce output to accommodate these entrants, the latter won't be enjoying the economies of scale required to sustain themselves. Bain observed that in three industries, tractor, automobile and typewriter, these barriers exist.

Product differentiation barrier or preference barrier. Edwin Chamberlin discussed some aspects of product differentiation and its implications for price output decisions. Sraffa too discussed this concept to explain a downward sloping demand curve. However, these economists didn't consider it as an entry barrier. Bain opularize on this. Due to product differentiation, the consumers develop brand loyalty, as so many years of an established firm and so much of investment in maintaining quality accompanied with huge outlay on advertisements are the factors that enable retention of customers to their brand. If in such a scenario, a new firm wants to produce and sell a differentiated good, it will not be easy for it. On the one hand, it will have to incur a lot of expenditure on advertisements. On the other hand, it will have to charge a comparatively lower price than the existing brands, to encourage the customers to buy the product, as it still has to earn the goodwill enjoyed by existing firms. So the firm is in a double disadvantage.

However, the firm may overcome this disadvantage. Whether it succeeds or not depends on the following factors:

- If heavy expenditure is required on advertisements required to opularize the product, barrier is strong and hence difficult to be overcome.
- The barrier is stronger in durable goods, as compared to other goods. Here customers' preferences depend a lot on how popular the brand is as well as to what extent is the goodwill enjoyed by the firm.

- If the markets (both for buying raw materials as well as marketing the products) are extensive and diversified, a strong distribution network is required, which a new firm finds hard to do. To that extent, barrier will be stronger.
- The existing firms' products are preferred also because of the prestige attached or reputation enjoyed. It is missing in case of a new firm.

Bain observed that in five industries, automobile, tractors, typewriters, cigarettes and liquor, there were strong product differentiation barriers.

Check Your Progress-I

Answer the questions in the space provided:

Q1. What is limit pricing? Where is it fixed?

Q2. Why is limit pricing called entry preventing pricing?

Q3. What are the barriers to entry discussed by Bain in his model?

Q4. What are the main assumptions of Bain's model?

____

____

____

____

6.4 SYLOS-LABINI'S MODEL

P. Sylos-Labini further developed the model of limit pricing, in his book 'Oligopoly and technical progress'. In this model, he has specifically dealt with economies of scale barriers to entry. However, the model lacks generalisation like that of Bain, as it is based on certain assumptions restrictive in nature. His model is considered more refined than that of Bain.

6.4.1 Assumptions

- This model is based on the following assumptions:
- The market demand is given with a unitary price elasticity of demand.
- It is a case of homogenous Oligopoly. So the product is homogenous and a unique equilibrium price will be determined.
- The technological rigidity is assumed. In other words, there are technical discontinuities. This is also one of the two features of the market, the other being economies of scale.
- The plants differ technologically i.e. in terms of plant sizes. Further, the cost is a negative function of plant size. There are three possibilities i.e. three different plant sizes are available. Due to technical discontinuities, the LAC curves have to be drawn individually for the three plants, instead of one LAC. So we have three straight line cost curves one above the other; the lowest LAC corresponds to the firm with the largest plant size.
- Price is set by the most efficient firm, which happens to be the largest firm. It is the price leader. Price is set at a level acceptable to all the firms, which are price takers. Price is low enough to prevent entry of new firms.
- All the firms wish to enjoy at least the minimum or normal profits.

- The price leader knows the market demand and cost structure of all plant sizes.
- The new firm, if at all enters shall set up the smallest plant size.
- The established firm expects the potential entrant to join only if the latter thinks that the post-entry price will not fall below his LAC. Moreover, the potential entrant shall enter only if he expects that output produced by existing firms will remain the same in post-entry period, as in the pre-entry period. So he expects the increase in quantity demanded (due to decline in price after its entry) to be attributed to it only. These two important behavioural assumptions made by Sylos are together referred to as 'Sylos postulate'.

6.4.2 Price determination or fixation of limit price

We know that price shall be determined or set by the price leader. At the same time, it should also be noted that since the firms differ in terms of size of plant, there are different prices corresponding to different costs. So each firm has a minimum acceptable price set according to the principle of full cost pricing.

So

$$P_i = \text{ATC}\ (1+r)$$

Where P_i is the minimum acceptable price for the firm with ith plant size.

ATC and r are average cost and minimum or normal profits respectively for the same firm.

We have already discussed the assumptions made by the model. But while

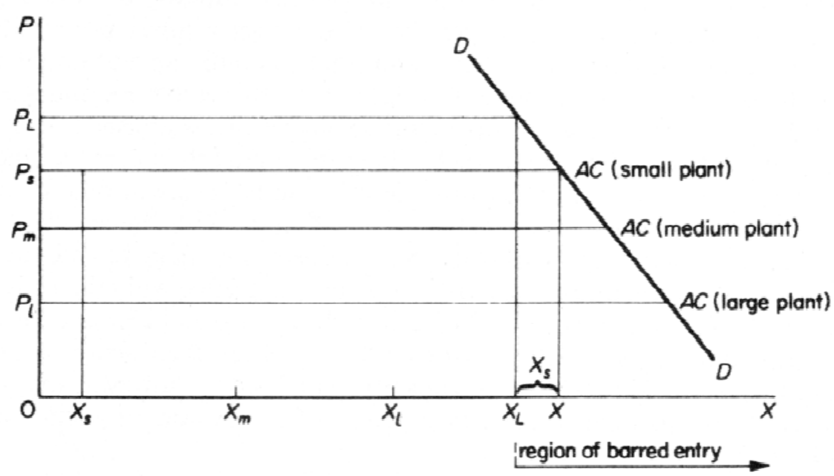


Fig. 6.2 : Determination of Limit Price : Sylo's Model

price has to be set, it is important to remember the two assumptions in particular. These are : prices have to be acceptable to all and secondly, price should deter entry of new firms. With this, let us now discuss the model, after presenting graphically whatever we have discussed so far or shall be part of discussion in the following paragraphs:

As has been shown in the figure, there are three firms operating three plant sizes. The firm with AC_1 as the LAC is the price leader. The firm with AC_3 as the LAC is the least efficient firm. While setting price, the price leader takes note of minimum profits to be earned by all. So the price shall be set above AC_3, so that the costs of all the firms are covered.

But now a question arises: What should be the exact level of price set by the leader? For this, the price setter observes as to what quantity demanded corresponds to the price of the least efficient firm? As we can see, it is OQ (corresponding to P_3). We are given that if the new firm enters, it shall have to establish the size of plant equal to the smallest firm. OQs is the level of output or productive capacity of firm with the smallest plant. So if the firm enters, it shall also produce OQs and its cost curve and price will be AC_3 and P_3 respectively. Now Q_L or the level of output corresponding to which price will be set is Q_L = Q-Qs. The price set is thus P_L. According to Sylos, this prevents the entry, as in the post entry period, output shall increase to Q_L + QS. Since it shall exceed OQ, i.e. the maximum output of the existing firms, price shall fall below AC_3 and hence the new firm shall be at loss.

At price P_L, all the firms are making supernormal profits. But instead of a unique price, there is an upper and a lower limit. The upper limit is P_L and the lower limit is P_3. This is because, at a price higher than P_L, new firms will be attracted; and at a price lower than P_3, the least efficient firm will be at loss. Likewise, there is a limit with regard to output as well. Any output smaller Q_L than shall not prevent entry, as in the post-entry period, output shall increase to <OQ, so that price shall not fall below AC_3. So output must be larger than Q_L, so that entry is prevented.

Check Your Progress-II

Answer the questions in the space provided:

Q1. What are the main assumptions of Sylos-Labini's model ?

____

____

____

____

Q2. What is Sylos' postulate? How does it affect the model?

____

____

____

____

Q3. How is price determined in Sylos' model?

____

____

____

____

6.5 CRITICAL EVALUATION OF LIMIT PRICING MODEL

The model of limit pricing discussed above has made significant contributions to price decisions made by Oligopolistic firms. In fact, Bain was a pioneer in limit pricing. He was the one who for the first time pointed out the implications of potential entry of new firms, that induced the established firms to fix entry preventing price. He also discussed that the firms sacrifice their short run profits to gain maximum profits in the long run.

But at the same time, there are certain grounds on the basis of which it has been opposed by critics.

Let us discuss few points of criticism:

- It is assumed that the potential entrants are small new firms. That is why barriers like absolute cost advantage, capital requirements, scale barriers, diferentiation barrier etc. have been put forward to argue that these newly established firms need to struggle hard to enter. But suppose, the potential entrants are big firms, which are already well established since years. In a way, these are equally competent to handle any challenge; and under these circumstances, these barriers to entry hold no relevance.
- The assumption that firms will keep price constant too looks unrealistic, as the firms in the real world are not so accommodative that they would suffer losses for the sake of a new entrant.
- Another assumption is the price setting by the leader. Here it can be noted that, given the experiences of corporate world, firms do actually retain control over price. Not only this, they also adopt strategies like advertisements etc. This has not been discussed here.

6.6 LET US SUM UP

Dear learner, in this lesson, we discussed Bain's model and Sylos-Labini's model of limit pricing in detail.

6.7 KEYWORDS

Limit pricing : In an Oligopolistic market structure, the firms aim to retain market share and earn profits in the long run than focusing on short run profits. So, the price charged is less then profit maximizing, so that it acts as an entry barrier for outside firms.

Product differentiation : In imperfect competition, firms tend to gain by looking different from others. In other words, they differentiate their product from that of others with an aim to increase brands loyalty and sales.

6.8 EXAMINATION ORIENTED QUESTIONS

1. Discuss in detail Bain's limit pricing model. Also draw a figure to explain the model.
2. Discuss the different entry barriers put forward by Bain in his model.
3. Justify that Bain's limit pricing model and Sylos-Labini's model present a case of collusive Oligopoly.
4. What is the level of output produced by a potential entrant in the post entry period?
5. How is price determined in Sylos' model?
6. How are the small firms' issues taken care of by the most efficient firm in Sylos' model?
7. What are the upper and lower limits in limit pricing in Sylos' model?

6.9 HINTS TO CYP

CYP-I

Q1. See sub-section 6.3.1

Q2. See sub-section 6.3.1 and 6.3.2

Q3. See sub-section 6.3.2

Q4. See sub-section 6.3.1

CYP-II

Q1. See sub-section 6.4.1

Q2. See sub-section 6.4.1

Q3. See sub-section 6.4.2

6.10 SUGGESTED READINGS

- Pindyck R.S., D. Rubinfeld and P.L. Mehta (2012), Microeconomics, Pearson Education, South Asia.

- Salvatore, D. (2010), Microeconomics; Theory and Applications, Oxford University Press, New Delhi
- Cooper, R., & John, A. (2011). Economics-theory through applications, Saylor Foundation. 1
- Shapiro, D., MacDonald, D., & Greenlaw, S. A. (2022). Principles of Microeconomics 3e.2

M.A. ECONOMICS	LESSON NO. 7
COURSE NO. ECO 201	UNIT-II

BEHAVIOURAL THEORY OF THE FIRM

STRUCTURE

7.0 Objectives

7.1 Learning Outcomes

7.2 Introduction

7.3 Behavioural theory of the firm

7.4 Implications of the model for the price behaviour of the firms

7.5 Let us sum up

7.6 Keywords

7.7 Examination oriented questions

7.8 Hints to CYP

7.9 Suggested readings

7.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

- Discuss the behavioural theory of the firm
- State multiple goals of the firm
- Define the determining factors of order of priorities among the goals

7.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to :

- Discuss the behavioural theory of the firm
- State multiple goals of the firm
- Define the determining factors of order of priorities among the goals

7.2 INTRODUCTION

Dear learner, earlier we have discussed different theories of the firm. In this lesson, we shall discuss behavioural theory of the firm.

7.3 BEHAVIOURAL APPROACH

In the behavioural theory of the firm, it is visualized that the firms seek to maximize neither profits, nor sales or utility. So this theory in a way rejects all the earlier models. Here instead of imagining how firms should respond to various situations or how they are expected to respond, their actual behavior is being watched. So it is based on actual decision making by firms. Not only profit maximization, but also certainty gets omitted in this model of the firm. Moreover, the focus here is on the process of decision making that takes place within the firms for predicting the decision of firms regarding output, price, sales etc.

A full scale behavioural theory of the firm has been given by Cyert and March (1963). According to them, in today's highly complex corporate world, it is not realistic to assume that the firm has a single major decision maker i.e. entrepreneur; on the contrary, there are different components of this so called organisational coalition like managers, stockholders, workers, consumers etc. So we can't attribute the process of decision making to only the entrepreneur. Rather all these participate in the decision making of firm. They further argue that it is traditionally assumed (though implicitly) that the goals of the entrepreneur are the goals of the organisation.

Multiple Goals : Cyert and March reject the earlier theory's one goal concept. Rather the firm has multiple goals. They identified five goals which firms in the real corporate world wanted to achieve.

They are :

- Production goal
- Inventory goal
- Sales goal
- Market share goal
- Profit goal

They further argue that each of these goals is related to a particular member of the organisation. The production goal is the concern of those directly related with production i.e. workers. Inventory of finished goods is being taken care of by the salesmen and customers. Sales goal is dealt with by salesmen or others whose survival is dependent on sales. Market share goal is pursued by those who want the comparative strength of this firm to be more than that of others. Finally, profit goal is to be achieved, as profits become the basis for distribution of dividends among the shareholders as well as for increasing the budget on expenses. It also arises on the part of top management, as profits also reflect their performance. It is this profit goal, that is set in terms of aspiration level. Profit goal is also connected with the price output decision of the firm.

According to Cyert and March, Following are the features of these goals:

- These five goals guide the firms regarding price, output and sales strategy.
- These are not based on hypotheses, but are the real life examples of firms, based on their (Cyert and March's) studies.
- All the goals must be satisfied.
- There is an implicit or latent order of priorities among these goals, which is reflected in the way search activity takes place.
- It is also reflected in the way aspiration level is changed.
- Variation between firms regarding these implicit priorities is due to the relative Bargaining position of components of the organisation.

Conflict among goals and coalition members. Since the members are desirous of pursuing mutually conflictual goals, the relation is not of perfect harmony. That is why, such an organisation is called a coalition of conflicting interests. However, the conflicts get resolved through persuation and through reconciliation of each other's interests. Also, some side payments are made to different members depending on their demands. If there are sufficient resources at the disposal of the firms, they can make payments to all these members. The theory lays stress on the short run relation between payments and demand.

Organisational slack. A great problem arises when there is inability on the part of firms to meet the demand of all the members, because they are inconsistent, mutually conflicting and resources are not enough. Here the concept of organisational slack holds importance. It is the difference between the payments actually made to these members and that necessary to retain them in the organization. .According to them, these slack payments, which mean payments made to members of the coalition in excess of what is necessary is due to imperfections and frictions in the mutual adjustment of demand and payments. Slack, according to them, is not zero.

Unlike earlier or conventional theory, which puts organisational slack to zero, at least in equilibrium, these scholars argue that there are many instances which show that organisational slack is not zero. For example, stockholders are given adequate returns to retain them, wages are more than required, prices are set lower than required to retain customers and top Management is given much more than minimum required e.g. luxury cars, offices etc., public services are provided in excess.

But they further argue that this organisational slack is a source of stability for the organisation. In case of a favourable external environment, the firm can earn more and hence give more slack payments to the members. In case of unfavourable environment, however, the firm needs to use it as a cushion. In other words, had this organisational slack not existed, the firm would not have been able to provide even the minimum essential payments to the members in case of unfavourable times. That would have endangered the survival of the organisation. So they consider it like a pool of emergency resources.

7.4 IMPLICATIONS OFTHE MODEL FOR THE PRICE BEHAVIOUR OF THE FIRMS

In this model, for each duopolistic firm, there are three sets of goals (profits, sales, production) and three basic decisions with regard to price, output and sales.

In the original model, price was highly sensitive to three things i.e. factors which bring about increase or decrease in magnitude of organisational slack, possible reduction of cost on advertisement and other sale promotion measures and adjustment in profit goals after actual profit achievement . Say for example, if it is found that profits are low, steps would be taken to reduce costs, which will further lead to required changes in rganizational slack and sales promotion. Lower costs would reduce the price of the product. So in this model, there is a close relationship between price, costs and profits and price has been found highly sensitive to factors influencing costs. However, the conventional theory didn't take into consideration any such adjustment in cost conditions, as least cost factor combination is achieved independently of external environment. As argued by Silberston, the important difference between the behavioural and the traditional model is that the former lays emphasis on process of decision making in firms , with continuously changing goals , as the firm learns from its experience and adapts to it; while the latter doesn't, as it is assumed that least possible cost combination has already been achieved.

Check Your Progress-I

Answer the questions in the space provided.

Q1. What are the main goals pursued by a firm according to behavioural theory?

____

____

____

____

Q2. What is the order of priorities in this model? What factors determine this order? Is it implicit or explicit?

Q3. How is the conflict resolved?

Q4. What is organisational slack? How does this model suggest it as an emergency pool of resources?

7.5 LET US SUM UP

Dear learner, in this lesson, we discussed the behavioural theory of the firm discussed by Cyert and March.

7.6 KEYWORDS

Organisational slack: The resources at the disposal of the firm which are not necessary presently but can be used in case of emergency. It is an important component of behavioural theory of the firm.

Multiple goals: Unlike traditional theories of the firm, the behavioural theory asserts that there is not one objective, but multiple objectives of the firm, but the priority or order of fulfillment is the firm's decision.

7.7 EXAMINATION ORIENTED QUESTIONS

1. Discuss in detail the behavioural theory of the firm, with special reference to Cyert and March.
2. What is satisficing? Why does a firm satisfice and not maximise?
3. What is search activity? Why is it required?
4. What is aspiration level? Is it fixed or flexible?
5. What is an organisational coalition? How are the conflicting interest reconciled?
6. Discuss in detail the concept of multiplicity of goals.
7. What is organisational slack? Do you agree that it really plays a stability and adaptive role?
8. How does the behavioural model differ from the traditional theory ?
9. What are the factors to which price is sensitive, according to behavioural theory?
10. Discuss the adjustments that takes place in the policies of the firm according to Cyert and March. Draw a comparison between this model and the traditional theory in this regard.

7.8 HINTS TO CYP

Q1. See section 7.3

Q2. See section 7.3

Q3. See section 7.3

Q4. See section 7.3

7.9 SUGGESTED READINGS

- Pindyck R.S., D. Rubinfeld and P.L. Mehta (2012), Microeconomics, Pearson Education, South Asia.
- Salvatore, D. (2010), Microeconomics; Theory and Applications, Oxford University Press, New Delhi
- Cooper, R., & John, A. (2011). Economics-theory through applications, Saylor Foundation. 1
- Shapiro, D., MacDonald, D., & Greenlaw, S. A. (2022). Principles of Microeconomics 3e.2

M.A. ECONOMICS	LESSON NO. 8
COURSE NO. ECO 201	**UNIT-II**

INFORMATION TECHNOLOGY

STRUCTURE

8.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

- Define Systems competition
- Discuss Lock-in effect

- Describe Switching costs
- Explain Network Externalities
- Define Rights management

8.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to:

- Define Systems competition
- Discuss Lock-in effect
- Describe Switching costs
- Explain Network Externalities
- Define Rights management

8.2 INTRODUCTION

Dear learner, in this lesson, in the earlier lessons, we discussed in detail about different alternative theories of the firm. Now, we shall discuss few concepts regarding Information Technology like Systems and systems competition, problem of complements, network externalities, rights management etc.

8.3 SYSTEMS AND SYSTEMS COMPETITION

To understand, what is Systems and systems competition, let us rely on the definition by Michael L. Katz and Carl Shapiro:

"Many products have little or no value in isolation, but generate value when combined with others. Examples include: nuts and bolts, which together provide fastening services; home audio or video components and programming, which together provide entertainment services; automobiles, repair parts and service, which together provide transportation services; facsimile machines and their associated communications protocols, which together provide fax services; automatic teller machines and ATM cards, which together provide transaction services; camera bodies and lenses, which together provide photographic services. These are all examples of products that are strongly complementary, although they need not be consumed in fixed proportions. We describe them as forming systems, which refers to collections of two or more components together with an interface that allows the components to work together."[1]

The degree of complementarity has increased in today's highly technology driven economies. Unless there are complements available to certain products, their utility can't be maximized. At times, even the product becomes useless. For example, hardware is useless without software, DVD players are useless without content, and operating systems are useless without applications.

It is often discussed about complementarities being as extreme as left and right shoe. The same holds true about many technology products, say again computer hardware and software. So while competition takes place among these players who produce different components to be used together, they worry as much about their complementors, as they do about competitors. For example, it has been said about Apple that its competitive strategy has been to involve relations with software developers, which means that competitive strategy in Information Technology industries is much different than it is in the traditional industries.

8.4 THE PROBLEM OF COMPLEMENTS

To discuss this issue, let us take the example of a Central Processing unit (CPU) and an Operating System (OS). As we know, CPU is the brain of a computer and OS is the software which allows us to have access to the functions of CPU. For being of utility to the end user or consumer, OS needs to be compatible with the CPU (assuming that every OS doesn't work with every CPU).

The sellers of these complementary products face pricing problem. The model is confined to just two aforementioned components, to simplify things (though it may look unrealistic, as so many components are used together). Since these are prefect complements (like left and right shoe), the demand for one depends on the price of the other product apart from its own.

Let P_1 and P_2 be the prices of CPU and OS respectively. Then demand for CPU can be written as :

$$D\ (P_1 + P_2)$$

Further let c_1 be the Marginal Cost of a CPU and F, the fixed cost, then profit maximization problem of CPU manufacturer can be written as:

$$\underset{p1}{\text{Max}}\ (P_1 - c_1)\ D\ (P_1 + P_2) - F_{1.}$$

Similarly, we can write for the OS maker

$$\underset{P2}{\text{Max}}\ (P_2 - c_2)\ D\ (P_1 + P_2) - F_2$$

Let the demand function has the linear form:

$$D(p) = a - bp$$

Let us assume that Marginal costs are so small that they can be ignored

$$\text{Max}\ P_1\ [a\text{-}b\ (P_1 + P_2) - C_1] - F_1$$

$$\underset{p1}{\text{Max}}\ ap^1 - bP_1^2 - bP_1P_2 - F_1$$

If price increases by ΔP_1, Marginal revenue is

$$(a\text{-}2b\ P_1 - b\ P_2)\ \Delta P_1$$

But for profit maximization, MR must be zero

$$a - 2bP_1 - bP_2 = 0$$

$$P1 = \frac{a - bp_2}{2b}$$

Likewise, for OS,

$$P2 = \frac{a - bp_1}{2b}$$

So the profit maximizing or optimal choice of each firm's price depends on what it expects from the other firm to charge for its component (As you can see P_2 finds place in equation of first firm and P_1 for the second) or Nash equilibrium

Solving the system of two equations in two unknowns, we get

$$P_1 = P_2 = \frac{a}{3b}$$

But here, each firm independently and unilaterally sets the price of its component of the system :

Price of the total system is

$$P_1 + P_2 = \frac{2a}{3b}$$

Now if, instead of setting prices independently, the firms merge, the profit maximizing problem is :

$$\max_{p} \; p\,(a - bp)$$

MR due to change in price is

$$(a\text{-}2bp) \quad \Delta P$$

Setting this equal to zero,

$$P = \frac{a}{2b}$$

Here it must be noted that the profit maximizing price set by the integrated firm is less than set by the two independent firms. Since the consumers are better off and the joint profits are more than in the earlier case, it means coordinating the price decision has made all the entities better off.

In general, a merger of two monopolies, that produce complementary products results in lower prices and higher profits, than if the two firms set prices independently.[3]

Check Your Progress-I

Answer the questions in the space provided.

Q1. What do you mean by Systems and Systems competition?

____

____

____

____

Q2. What is the problem of complements?

____

____

____

____

8.4.1 Relationship among complementors

Here we shall discuss the different possibilities that exist with regard to relationships between complementors, so that pricing problem can be solved. We have already seen that joint setting of prices leads to lowering of prices and improves profits, that leads to a desirable solution. There are different cases ranging from fully integrated to fully independent and many intermediate cases in between. We have already discussed independent price setting in the earlier section.

Merger. One possibility is to merge. Here we can discuss the classical model of Cournot, who analyzed the strategic interactions between producers of complementary products, considering a market with two companies: a monopoly zinc producer and a monopoly copper producer. These two supplied a large number of other companies that combined the metals to produce brass. Cournot asked what would happen to the price of brass if the copper and zinc producers merged.

Let us assume that one unit of copper and one unit of zinc combine to create one unit of brass. Competition will push the price of brass down to its cost, which will simply be the sum of the two prices. Demand for brass can then be written as D(p1 + p2). Given our assumptions about the technology, this is also the demand for copper and zinc.

The copper producer, say, wants to maximize the profit of producing copper:

$$\max_{p_1} \; p_1 D(p_1 + p_2).$$

Here we have assumed that the cost of copper production is zero for simplicity.

The zinc producer has the analogous problem;

$$\max_{p_2} \; p_2 D(p_1 + p_2).$$

If the two complementary monopolists merged, they would solve the joint profit maximization problem

$$\max_{p_1, p_2} \; (p_1 + p_2) D(p_1 + p_2).$$

Cournot showed that the complementary monopolists would set prices that were higher than if they merged. The intuition is simple. If the copper producer cuts its price, brass producers will buy more zinc, thereby increasing the profits of the zinc producer. But the zinc producer's additional profits are irrelevant to the copper producer, making it reluctant to cut its price too much. The result is that the copper producer sets a price that is higher than the price that would maximize joint profits.

If, however, the copper and zinc producers merged, the merged entity would take into account that the price of copper affected the demand for zinc and set a lower price for both copper and zinc than independent producers would. Hence, a merger of complementors is a win all the way around: prices fall, making producers *and* consumers better off.[4]

Collaborate. Here, the firms set up a formula for revenue sharing, then one firm sets the price of the joint system. For example, an aircraft manufacturer and an engine manufacturer will agree on a revenue sharing arrangement, then the aircraft manufacturer will negotiate a price for the entire system with customers.

Negotiate. In this case, a firm may commit to cutting its price if the other firm also cuts its price. When DVD players were introduced, what troubled the electronic firms producing these was whether there will be sufficient content available. Similar thinking affected the content providers, who didn't want to produce content unless there were DVD players. Moreover, pricing too was an issue. Both the content providers and DVD players were introduced at relatively low prices, since the participants recognized that a low price for the entire system was critical to ensure its adoption.

Nurture. Here, one firm works with others to reduce their costs. For example, Adobe has been working with printer manufacturers to ensure that they can effectively use its technology.

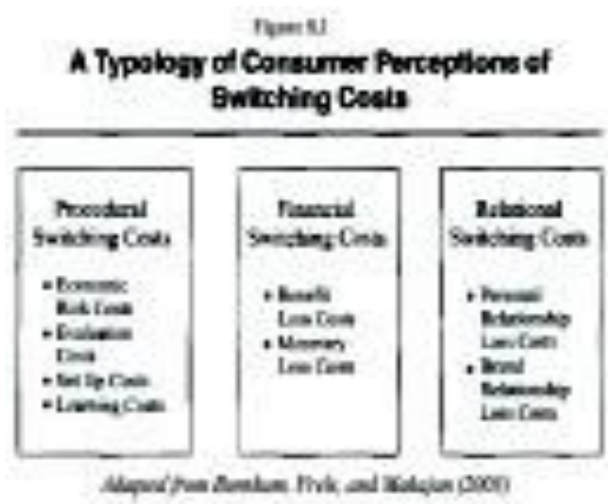
Commoditize. In case of 'commoditize the complement' option, one firm attempts to stimulate competition in the other's market, thereby pushing down prices. This happens in case the lowering of firm's own component's price may or may not increase its revenues, given the elasticity of demand, but lowering the complementor's product does. So it induces the latter to reduce prices. One way to ensure this is to resort to Standardisation**.** Microsoft, for example, has established the Windows Compatibility Lab, to ensure that hardware manufacturers all produce to a common standard. This helps facilitate competition, pushing down the price of hardware.

8.5 LOCK IN AND SWITCHING COSTS

Let us take a very simple example to understand the two concepts. Suppose a person uses a PC, in which different components have been bundled together as a system (say, hardware, software etc.). If now the person wants to switch any one component, he needs to switch others as well (compatibility issue). He not only needs to purchase a new hardware, software but also to learn the new system's working. When the cost of switching is so huge, that the customer is not induced to switch, but continues with the earlier system, he is said to experience 'lock-in'. It is bad for the consumers, but good for the producers of these IT products, as due to inelastic demand of the consumers, they (firms) can charge exorbitant prices and hence extract the maximum consumer surplus.

In a nutshell, lock-ins can be considered as 'inflexibilities' of outcomes.[5]

In many markets, consumers face substantial costs of switching between brands of products that are ex ante undifferentiated.[6] This happens when they move from one product to another in the same category, or when they change their suppliers. Thus, ex-ante homogeneous products become ex-post heterogeneous. While early literature classified switching costs as being either pecuniary (for e.g. termination fees charged for breaking a contract) or psychological (for e.g. the disutility faced by a consumer for buying a brand he/ she is not familiar with), recent work suggests that switching costs are multi-dimensional. Burnham, Frels, and Mahajan build a comprehensive typology of switching costs and identify three major types of switching cost, each with various sub categories: procedural costs (involving mainly expenditures of time and effort and consisting of economic risk, learning and setup costs), financial costs (involving the loss of financially quantifiable resources), and, relational costs (psychological and emotional discomfort due to personal and brand relationship breakages and identity loss)·



However, the types of switching costs, as discussed by Klemperer (1987) include: transaction costs, learning costs, and artificial or contractual costs. “There may be transaction costs in switching between completely identical brands. Two banks may offer identical checking accounts, but there are high transaction costs in closing an account with one bank and opening another with a competitor. Similarly, it is costly to change one’s long distance telephone service, or to return rented equipment to one firm and rent identical equipment from an alternative supplier. The learning required to use one brand may not be transferable to other brands of the same product, even though all brands are functionally identical. A number of computer manufacturers, for example, may make machines that are functionally identical, but if a consumer has learned to use one firm’s product line and has invested in the appropriate software, he has a strong incentive to continue both to buy machines from the same firm and to buy software compatible with them.”

The lock-in effect and the switching costs have various implications for the pricing decisions of the firms producing and selling these products to customers, which then has a bearing on the satisfaction level as well as welfare of the customers.

Different scholars have discussed a tiered pricing mechanism, that can be tried by these firms, whereby the customers, who are old and in which case, brand loyalty exists, can be overcharged once they are locked-in and the new customers, who are expected to yield profits in future and are likely to contribute to future customer base, can be offered low price. But this puts the firms in the so called 'Harvest or invest dilemma'(which means that at the same time, for the identical products, the firm has to bifurcate its customers into two categories, because, it wants to harvest gains from earlier category, while at the same time invest in the new category). This is not only unjustified to penalize its old customers for their loyalty to the brand, but also unsustainable, if the customers can overcome the inertia associated with switching costs.

Many a time, it has been observed that customers used to a particular brand get fatigued and want to try something new, provided switching cost is not too high. Moreover, if with high switching cost, the aim is to retain old customers by making their choice constrained, there can be no certainty about the new customers to be attracted towards this brand only, as their choice base is quite broad.

Switching costs also cause 'Fat cat effect' which means that companies who have a large customer base can resort to high prices, whereas producers with small base need to offer low prices. Many a time, customers are used to an earlier system, with which they feel comfortable or even build relationship with brand. In such a case, if a new company launches its product, it needs to offer very low price to induce these customers to purchase the product. Here switching costs observed by customers act as an impediment to technological progress and innovation.[8]

Check Your Progress-II

Answer the questions in the space provided.

Q1. What is a merger?

Q2. How do complementors collaborate?

Q3. What are lock-in effect and switching costs?

8.6 NETWORK EXTERNALITIES

So far, we know what are externalities. Externalities mean when one person's production or consumption affects other persons. There are positive and negative externalities.

Network externalities are a special kind of externalities in which one person's utility for a good depends on the number of other people who consume this good.[9]

A product is said to exhibit network externalities if its 'Buyer Value' (the Buyer Value of each demander is equal to his Initial Value multiplied by a "Network Externality Factor" that depends on the total number of units sold.) for those who consume it is higher, the greater the number of other consumers who also consume the product. Some of the most striking instances of network externalities occur with products that aid communications. The value to you of a device that sends and receives messages will be larger, the greater the number of people who can receive your messages and send messages to you. For this reason, communications devices like the telegraph, telephone, fax machine, and computers connected to the internet all exhibit strong network externalities. Each of these technologies enjoyed a period of explosive growth as the value of being connected to the network increased at the same time that the network grew. Another source of network externalities is shared infrastructure. If only a few people in your country have automobiles, then it is difficult to find gasoline stations, repair shops, and good roads. As more people acquire automobiles, the shared infrastructure grows and owning an automobile becomes more attractive.[10]

Some scholars have tried to explain it by separating the value received by a consumer due to it into two parts i.e. Autarky value and synchronization value.[11] The former means the value that accrues to a consumer by using a commodity (say Fax machine), when there are no other users of it . On the other hand, synchronization value is derived from being able to interact or communicate with other users of the product. It is synchronization value that matters here. Further, they have also distinguished between direct and indirect network effects. The former means the increase in the number of users of the product, say a printer. Indirect effects are the market mediated effects that result due to the cheap and adequate complementary goods' availability (say toner of a cartridge) due to increase in users of the product (printer).

The aforementioned definitions and examples of Network Externalities make its meaning clear.

8.6.1 Markets with network externalities

This can be understood with the help of a simple demand and supply model. Let there be 1000 people in a market for some good. Let the people be indexed by v, so that v=1,………,1000. Let v measure the 'reservation price' for the good by person v. If the price of the good is p, the number of people who think that the good is worth at least p is 1000-p. Say, for example, if the price of the good is $200, there are at least 800 people who are willing to pay at least $200 for the good. So the number of units sold will be 800. This gives rise to a usual downward sloping curve.

But suppose the good is not a simple good; rather it is a good with network externalities (e.g. fax machine, telephone). In other words, whether a person will purchase it or not depends on the number of individuals who are using it. Let value of the good to v is vn, where n is the number of people who consume it or are connected to the network. The more the people connected to it, more each person is willing to pay to acquire it. The demand function will be different in this case, as we shall see below.

If price is p, there is someone, who is just indifferent between buying the good and not buying it. He is the Marginal Buyer, whose willingness to pay for the good equals its price. Let this person be indexed by $\hat{v}$

∴ $P = \hat{v}\, n$ ……… (i)

Since this marginal person is indifferent, everyone with a higher value of v than must definitely want to buy.

So the number of people who want to buy the good is:

$N = 1000 - \hat{v}$ ……… (ii)

Putting the above two equations together, the condition that determines equilibrium in this market is:

$P = n(1000-n)$ ……….(iii)

The above equation gives us a relationship between the price of the good and the number of users. So it is a kind of demand curve. Here if we draw a demand curve, it shall be drawn between network size (X axis) and willingness to pay

(Y axis). As we can see, the shape is not like that of a standard demand curve. If the no. of people who connect is low, the willingness to pay of the Marginal buyer is low, as there are not manypeople with whom he can communicate. On the other hand, when the no. of people connected is large, the willingness to pay is low, as all those who valued it more than him have already connected. This gives rise to a humped shape of demand curve.

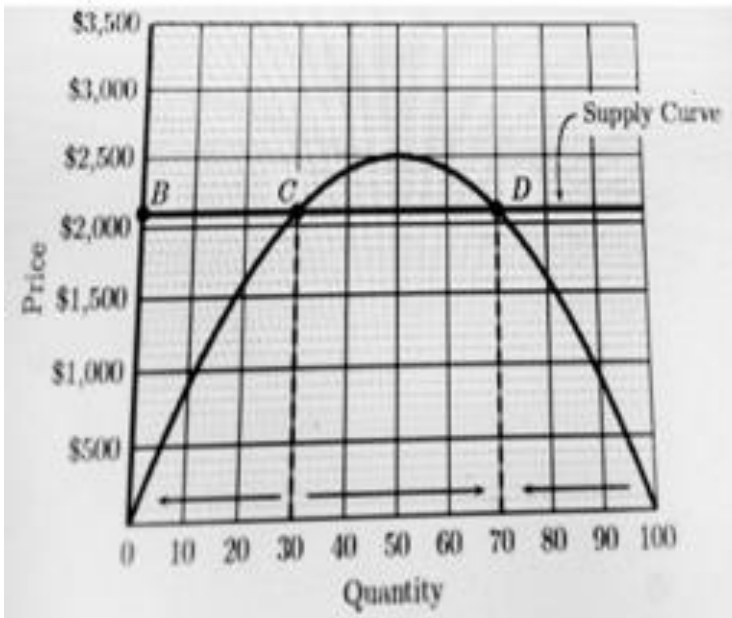


Figure 8.2 : Network Supply and Demand

As mentioned in the figure above, while the demand curve is humped, the supply curve is horizontal, as the cost of adding each additional firm to the network is fixed. Let it be $2,100. To determine the point of equilibrium, we need to see the point of intersection of demand and supply curves. The supply curve crosses the demand curve at three points, B, C, and D, corresponding to outcomes in which the number of firms connected to the network is 0, 30, and 70. Each of these outcomes is an equilibrium quantity when the cost of joining the network is $2,100. If everyone believes that no firms will join the network, then nobody will want to join; if everybody believes that exactly 30 firms will join the network, then exactly 30 firms will want to join; and if everybody believes that exactly 70 firms will join the network, then exactly 70 firms will want to join. Let us try to understand it. The low level equilibrium where nobody is willing to pay isreferred to as "a pessimistic expectations" equilibrium. The middle equilibrium, with a small but positive number, is one where people don't think the network will be big, hence they are not willing to pay much, hence the network is small. The last point of equilibrium, point D shows that Marginal buyer again doesn't want to pay more, as he doesn't value it much, though the no. of people connected is more.

Now a question arises : Out of the three equilibria that are shown here, which is likely to occur? For answering this, we need to add dynamic adjustment mechanism. When the people are willing to pay more than the cost of the good, the size of market expands; and when they are willing to pay less, the size contracts. This is shown by the arrows in the figure above. It means when the demand curve is above the supply curve, the quantity increases and when it is beneath the supply curve, the quantity decreases. This is the adjustment process.

Here we can add one more thing. The low level equilibrium point where no one connects and the high level point where many people connect are both stable equilibria. The middle equilibrium point is unstable and hence can't be considered. But which of the two above points is the final resting point of the system?

For this we need to see how costs change over time in these goods (e.g. Fax, computer etc.)? As has been observed in most of these goods, the cost is high initially, but decreases over time due to technological progress.

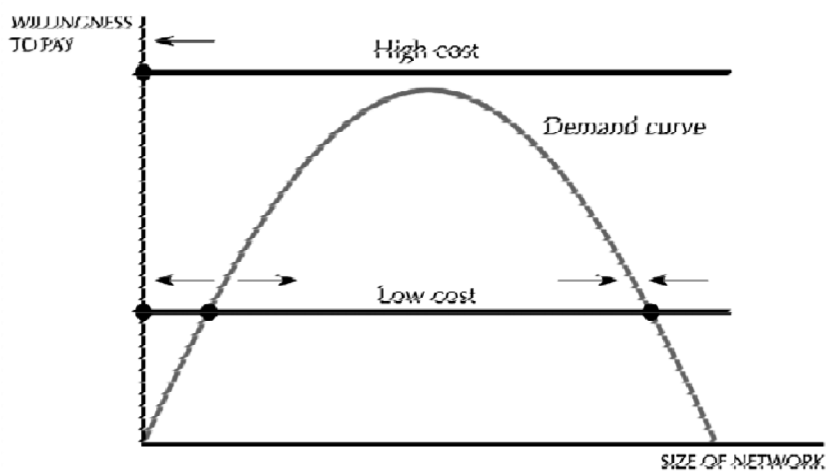


Fig. 8.3

As has been shown in fig. 8.3 above, there is high unit cost initially, where only one stable equilibrium can occur (at origin, where demand=zero). But as cost decreases, there are again two stable equilibria.

Let us add some noise to the system. Let us perturb the number of people connected to the network at point where n=0, through initial discounts or other promotions. As the cost becomes smaller and smaller, one of these perturbations will kick the system up beyond the unstable equilibrium and the dynamic adjustment process will thus make the system move up to high level equilibrium. This can be explained with the help of another figure (8.4) below :

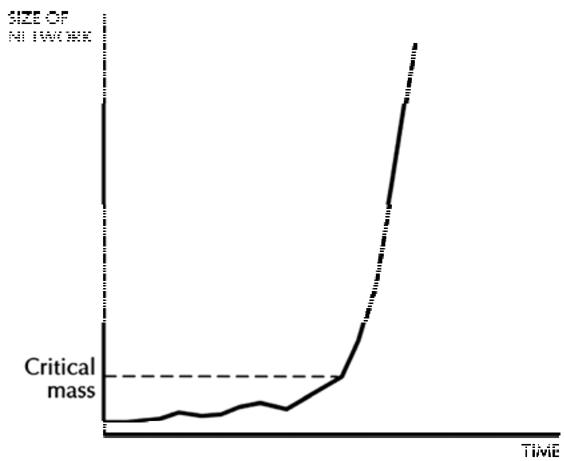
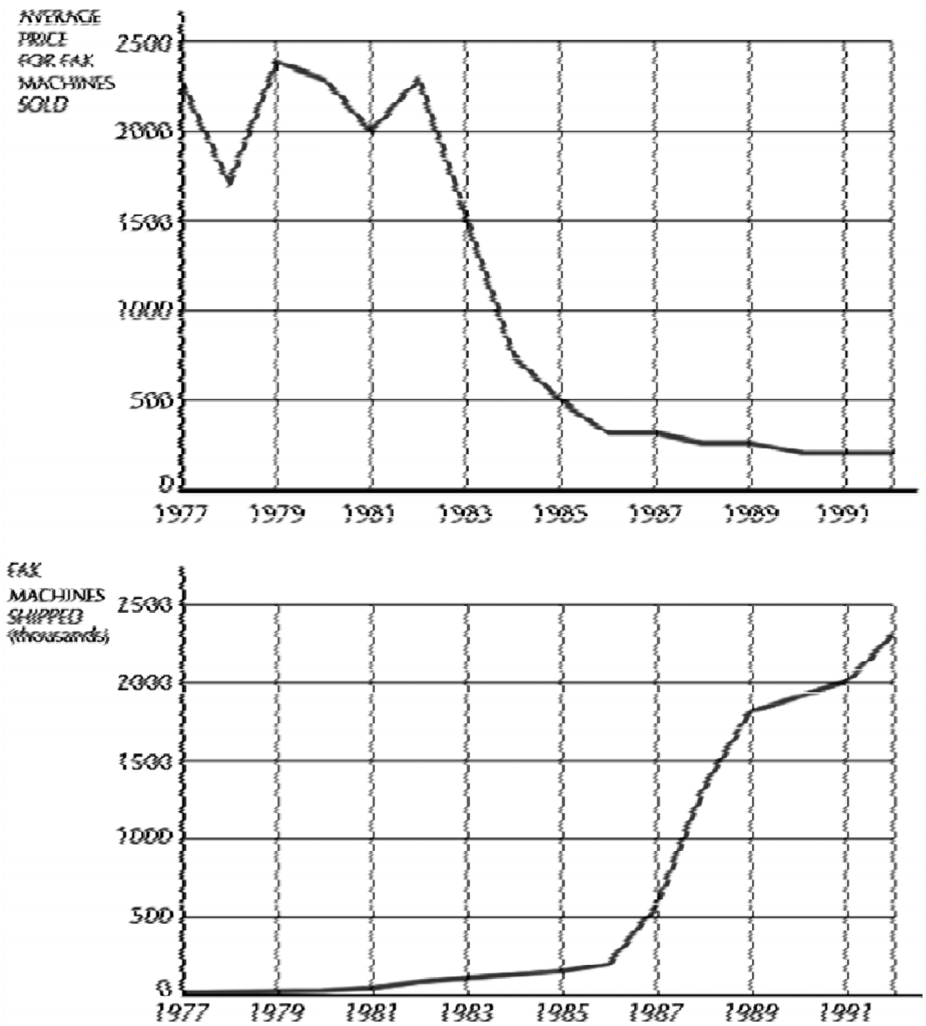


Fig. 8.4

Here we need to mention an important development, critical mass. The system needs a critical mass of customers (minimum no.) to be able to move to a high level of equilibrium. Once this critical mass is reached, the network growth takes off tremendously.

Have a look at fig. 8.2 again. If we want to discuss critical mass with the help of it, we can observe that to get to the high-level equilibrium from the zero equilibrium, it would not be necessary for all 70 firms to agree in advance to join. All that would be needed is to achieve a critical mass of 31 members. That is, the number needed to get just beyond the unstable equilibrium at 30. If Firm 31 joins, then it will be profitable for Firm 32 to join. Firm 33 will be attracted by the other 32 members, and then 34, and so it goes all the way up to 70.

This can be explained through another figure (fig. 8.5) in case of fax machines, which shows how initially, demand was close to zero and with significant price cut in the mid eighties, demand exploded.



8.6.2 Implications of Network Externalities

Network externalities have a lot of implications for the market. Following are few points that make it clear:

1. Critical mass of consumers is a very important issue. Unless that is achieved, it is not possible for the network market to take off. If steps are taken

to stimulate growth early in the life cycle of a product, it is achievable. That is why, many producers now a days offer cheap access to a piece of software or a communications service to create a market, where it didn't exist .How big the market has to be depends more on current conditions of the market like costs, expected benefits to the consumers, nature of the good etc. than the theory.

2. Another implication is the bandwagon effect it creates. Rohlfs (2001) argues that growth in demand generates bandwagon effect, which leads to further increase in demand. Shapiro and Varian attribute it to positive feedback. According to them, "Success feeds on itself."

3. However, it is not always positive in nature. Sometimes, (as argued by Farrel and Saloner, 1986; Srinivasan et al., early in the life cycle of these products, few consumers see utility in its adoption. So many may adopt a 'wait and see' approach, unless there are more adopters. So diffusion may be very slow initially (due to excess inertia) leading to chilling effect, before it catches momentum or the surge in growth.

4. It also depends on the source of externality, local or global. Sometimes before adopting a product, the consumer just takes into consideration his or her immediate social network (local effect), while in some cases, the global effect or a broader perspective is considered. It also influences the decision of firms producing a software. They take into consideration the global adopters, not just local.[12]

Check Your Progress-III

Answer the questions in the space provided.

Q1. What are network externalities?

____

____

____

____

Q2. What are the implications of network externalities?

____

____

____

____

8.7 RIGHTS MANAGEMENT & SHARING OF INTELLECTUAL PROPERTY[13]

There are different ways in which intellectual property (IP) transactions take place e.g. books are sold out-right and also borrowed from libraries. Videos can either be sold or rented. Some software is licensed for particular uses; other software is sold outright. Choosing the terms and conditions under which a piece of intellectual property is offered is a critical business decision. Should you use copy protection? Should you encourage users to share a news item with a friend? Should you sell to individuals or use site license?

Some simple Economics helps to understand the relevant issues. Let's consider a purely digital good, such as an online newspaper, so that we can ignore its marginal cost of production. First let us consider behaviour under some default set of terms and conditions. The owner of the digital good will choose a price and, implicitly, a quantity to sell so as to maximize profit :

$$\max_{y} \; p(y)y \quad \text{..............} \; \text{(i)}$$

This yields some optimal (p^*, y^*).

Now the seller of the good contemplates liberalizing terms and conditions : let's say extending a trial period of free use from 1 week to 1 month. This has two effects on the demand curve. First, it increases the value of the product to each of the potential users, shifting the demand curve up. But it also mayeasilyresult in less of the item being sold, since some users will find the longer trial period enough to meet their needs.

Let us model this by defining the new amount consumed by $Y=by$ where $b>1$, and the new demand curve by $P(Y) = ap(Y)$, where $a>1$. The new profit- maximization problem now becomes.

$$\max_{Y} P(Y)y$$

Note that we multiply price times the amount sold, *y*, not the amount consumed, Y.

Applying the definitions Y=*by* and *P(Y)* = *ap(Y)*, we can write this as

$$\max_{Y} ap(Y)\frac{Y}{b} = \max_{Y} \frac{a}{b} p(Y)Y$$

This maximization problem looks like problem ((i).7) except for the constant *a/b* in front of the max. This will not affect the optiomal choice, so we can conclude that Y* = *y**

This simple analysis allows us to make several conclusions :

The amount of the good consumed, *Y**, is independent of the terms and conditions.

The amount of the good produced is *y**/*b* which is less than *y**.

The profits could go up or down depending on whether *a/b* is greater or less than 1. Profits go up if the increase in value to the consumers who buy the product compensates for the reduced number of buyers.

Here we can give an example of video Rental.

Video stores can choose the terms and conditions under which they rent videos. The longer you can keep the video, the more valuable it is to you, since you have a longer period of time during which you can watch it. But the longer you keep the video, the less profit the store makes from it, since it is unable to rent it to someone else. The optimal choice for the rental period involves trading off these two effects.

In practice, this has tended to lead to a form of product differentiation. New releases are rented for short periods, since the profits from other renters being excluded are very substantial. Older videos are rented for longer periods, since there is less cost to the store from the video being unavailable.

Sharing Intellectual Property

Intellectual property is often shared. Libraries, for example, facilitate the sharing of books. Video stores help people to “share” videos—and charge a price for doing so. Interlibrary loan helps libraries share books among themselves. Even textbooks are shared among students from one term to the next via the resale market.

There is considerable debate in the publishing and library communities about the proper role of sharing. Librarians have established an informal "rule of five" for interlibrary loan: an item may be loaned out up to five times before additional royalty payments should be made to the publisher. Publishers and authors have traditionally been unenthusiastic about the resale market for books.

The advent of digital information has made this situation even more acute. Digital information can be perfectly reproduced, and "sharing" can be taken to new extremes. It is said that a well-known country music singer was once engaged in a vociferous public relations campaign against stores selling used CDs. The problem was that CDs do not deteriorate with replay and it is possible to buy a CD, tape it, and then sell the CD to the used-CD store.

Let us try to construct a model of this sort of sharing phenomenon. We begin with the baseline case in which there is no sharing. In this case a video maker chooses to produce y copies of a video to maximize profit :

$$\max_{y} \; p(y)y - cy - F. \qquad \text{.............} \; \text{(ii)}$$

As usual, $p(y)$ is the inverse demand function, c is the (constant) marginal cost, and F is the fixed cost. Let the profit maximizing output be denoted by y_n, where the n stands for "no sharing".

Now suppose that a video rental market is allowed. In this case the number of videos *viewed* will be distinct from the number of copies produced. If y is the number of videos produced and each video is shared among k viewers, then number of viewings will be $x = ky$. (For simplicity we are assuming that *all* copies of the video are rented in this case.)

We need to specify how the consumers sort themselves into the "clubs" that share the videos. The simplest assumption is that the consumers with high values associate with each other, and the consumers with low-values associate with each other. That is, one club consists of consumers with the *k* highest values, another club consists of the consumers with the next *k* highest values, and so on. (Other assumptions could be used, but this one gives a very simple anaylsis.)

If *y* copies of the video are produced, $x=ky$ copies will be viewed, so the willingness to pay of the marginal individual will be $p(x)=p(ky)$. However, it is clearly the case that there is some inconvenience cost to renting a video rather than owning it yourself. Let us denote this "transactions cost" by t, so that the willingness to pay of the marginal individual becomes $p(x)–t$.

Recall that we have assumed that all copies of the video are shared among *k* users. Therefore the willingness to pay of a *video store* will just be *k* times the willingness to pay of the marginal individual. That is, if *y* copies are produced, the willingness to pay of the video store will be

$$P(y) = k[p(ky)–t.] \quad \text{...........} \quad \text{(iii)}$$

The above equation contains the two key effects that arise from sharing: the willingness to pay goes *down* since more videos are viewed than are produced; but the willingness to pay also goes *up* since the cost of a single video is shared among several individuals.

The profit maximization problem of the producer now becomes.

$$\max_y \ P(y)y - cy - F,$$

which can be written as

$$\max_y \ k[p(ky)-t]y - cy - F,$$

or

$$\max_y \ p(ky)ky - \left(\frac{c}{k} + t\right)ky - F,$$

Recalling that the number of viewings, *x*, is related to the number produced, *y*, via $x=ky$, we can also write the maximization problem as

$$\max \ p(x)x - \left(\frac{c}{k} + t\right)x - F,$$

x

Note that this problem is identical to problem (ii), with the exception that the marginal cost is now *(c/k+t)* rather than c.

The close relationship between the two problems is very useful since it allows us to make the following observation : *profits will be larger when rental is possible than when it is not if and only if.*

$$\left(\frac{c}{k}+t\right) < c$$

Rearranging this condition, we have

$$\left(\frac{k}{k+1}\right) t < c$$

For large *k*, the fraction on the left is about 1. Hence the critical issue is the relationship between the marginal cost of production, *c*, and the transactions cost of renting, *t*.

If the cost of production is large and the cost of renting is small, then the most profitable thing for a producer to do is to produce a few copies, sell them at a high price, and let the consumers rent. On the other hand, if the transactions cost of renting is larger than the cost of production, it is more profitable for a producer to have renting prohibited: since renting is so inconvenient for the consumers, video stores aren't willing to pay much more for the "shared" videos, and so the producer is better off selling.

8.8 LET US SUM UP

Dear learner, in this lesson, we discussed systems competition, lock-in effects and switching costs, network externalities and rights management issues.

8.9 KEYWORDS

Complementors: Complementor means a firm or business organization whose product adds value to another company's /firm's product , as both are consumed by mutual customers e.g. Apple and App store developers.

Switching costs : The different kinds of costs (monetary, psychological etc.)/ inconveneience associated with a customer's shifting to a new brand, product, supplier etc. is called switching cost. The aim is to retain the customers.

Marginal buyer: A marginal buyer is the one who is willing to pay a price just equal to MU of the product.

Critical mass: The minimum threshold required by a company to compete efficiently in the market is called critical mass. Once it is reached, the company becomes self-sustaining.

8.10 EXAMINATION ORIENTED QUESTIONS

1. What is systems competition?
2. What do you mean by problem of complements? Explain.
3. What is lock-in effect? Why is it bad for consumers?
4. What is switching cost? Does it help retain customers?
5. Is it justified to keep customers locked-in? Justify your answer.
6. What are network externalities? What are their implications?
7. What is critical mass? Why is it important? Draw a diagram to show how a network achieves this?
8. What is pessimistic expectations equilibrium? Is it stable? Is it desirable?
9. Explain with the help of a model, equilibrium in case of a market with network externalities.
10. What is a humped demand curve? How is it different from a normal demand curve?
11. Discuss the rights management issues and sharing of intellectual property in case of a digital good.

8.11 HINTS TO CYP

CYP-I

Q1. See section 8.3

Q2. See section 8.4

CYP-II

Q1. See sub-section 8.4.1

Q2. See sub-section 8.4.1

Q3. See section 8.5

CYP-III

Q1. See sub-section 8.6.1

Q2. See sub-section 8.6.2

8.12 SUGGESTED READINGS

Please refer to the sources mentioned in End Notes section.

END NOTES :

1. *Michael L. Katz and Carl Shapiro* ***Systems Competition and Network Effects***

2. *Source: Hal R. Varian* ***Economics of Information Technology*** *(2003)*

3. *This fact, as cited by Varian, was discovered by Cournot.*

4. *This matter/ theory has been discussed, by relying on the source, as mentioned in note 2 above.*

5. *Source: Uwe Cantner, Simone Vannuccini* ***Innovation and Lock-in*** *Jena Economic Research Papers 2016 – 018*

6. *Source: Paul Klemperer* ***Markets With Consumer Switching Costs*** *The Quarterly Journal of Economics, May 1987*

7. *Source: Abhi Bhattacharya* ***Switching Costs and Sustained Competitive Advantage*** *International Journal of Business and Management Invention ISSN (Online): 2319 – 8028, ISSN (Print): 2319 – 801X www.ijbmi.org Volume 2 Issue 9? September. 2013? PP.101-111*

8. *Same as 7 above.*

9. *Hal Varian (2003)*

10. *Source: networks.pdf (author not mentioned)*

11. *S.J. Liebowitz and Stephen .E. Margolis* ***Network Externalities (Effects)***

12. *Source of points 2,3 and 4 Goldenberg et.,al (2010)* ***The chilling effects of network externalities,*** *International journal of research in marketing, Volume 27, issue 1*

13. *Source: Hal Varian* ***Intermediate Microeconomics.***

M.A. ECONOMICS **LESSON NO. 9**

COURSE NO. ECO 201 **UNIT-II**

GAME THEORY : PART-I

STRUCTURE

9.0 OBJECTIVES

The main objectives of this lesson are to enable you to :

- Define theory of games and basic concepts
- State the assumptions of game theory
- Discuss Zero sum game and non-zero sum game.
- Identify the Factors on which adoption of maximin strategy depends
- Evaluate Dominant strategies

9.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to:

- Define theory of games and basic concepts
- State the assumptions of game theory
- Discuss Zero sum game and non-zero sum game.
- Identify the Factors on which adoption of maximin strategy depends
- Evaluate Dominant strategies

9.2 INTRODUCTION

Dear learner, we have already discussed the models of Oligopoly, including collusive and non-collusive Oligopoly. We have also observed that none of the models is able to adequately explain the interdependence among these firms, as there is a lot of uncertainty in an oligopolistic market structure. Since all these models with their analytical tools have failed to explain the decision making process of oligopolists, an alternative is provided by the game theory. In this lesson, we shall discuss few important concepts related to game theory. Some other concepts shall be dealt with in the next lesson.

9.3 THEORY OF GAMES

A new approach to the study of Oligopoly was given by Professors Neumann and Morgenstern in their book 'The Theory of Games and Economic Behaviour', which was published in 1944. This approach discusses solution to problems involving conflicting situations. Oligopoly is not the only problem discussed by it. It is also applicable to other questions of Economics, when uncertainty is present. Not only this, it is applicable beyond Economics to fields like Business Administration, Sociology, Psychology, Political Science and military planning. In general, the theory examines the outcome of a situation involving interactions between parties with conflictual relations. It provides guideline regarding the rational course of action to be taken by an individual, when confronted with an uncertain situation, as the outcome depends not only on his action, but also on others' reaction to it. In the next section, we shall discuss briefly the different concepts, without the knowledge of which, it is not possible to understand what the game theory is.

9.3.1 Basic concepts

(i) Game : A situation in which individuals must make decisions and in which final outcome will depend on what each person decides to do is to be viewed as a game.[1] The basic elements of a game are: players, strategies and payoffs.

(ii) Players: The decision maker is the player. The concept of player changes with the problem under consideration e.g. firm is the player in case of Oligopolistic markets, individuals are players in case of poker games and nations are the players in military operations. The number of players is usually fixed e.g. in our discussion of game theory in Oligopoly, most of the games are two player games, in which both aim to make a choice so as to get the most favourable outcome.

(iii) Strategies: The firm pursues its goals with the help of various instruments or policy variables; the most important of these are: price, quantity of output, style of product, advertising and other sales promotion measures, Research and Development expenditure, change in the number of products etc.

A strategy is a specific course of action available to the player in the game. In other words, it has clearly defined values for the particular variable under consideration. As we mentioned above, there are different policy variables. Strategy here may mean setting a price, keeping price unchanged, raising price, cutting price etc. It may also mean spending a particular amount on advertisements, R&D, etc. With regard to output, it may mean keeping the output constant, increasing or decreasing the level of output etc. Likewise, a strategy with regard to varying the output may include changing the colour of packaging, type of packaging, quality of product etc. The number of strategies open to a player is finite. In case of oligopolistic models, only two strategies are open to a player.

(iv) Payoff : The payoff is the net gain or return to the player at the end of the game. In other words, after the firm in question has taken a strategy and a counter-strategy has been taken by its competitor, whatever gain accrues to the firm in question is the payoff to that firm. Payoffs are usually the monetary returns to firms in the form of profits (if goal is profit maximisation) or sales revenue (if the goal is sales maximisation) or even actual marketshare gained by it (in case the goal is maximisation of market share). Apart from explicit returns, implicit returns like embarrassment, gaining self-esteem etc. are also observed. It is usually assumed that firms are able to rank the payoffs from the most preferred to the least preferred.

(v) Payoff matrix: The payoff matrix of a firm is a table showing the payoffs that accrue to the firm in question, after taking into consideration each possible combination of strategies available to it and its rival. Let us assume that there are two firms, 1 and 2. Let strategies A1, A2......A5 are available with A, among which he has to choose one. Likewise, let B1, B2.........B6 i.e. six strategies are given to B, among which he has to make choice. So payoff matrix of each firm will be 5X6 i.e. 30 payoffs will be included.

		Firm II's strategies					
		B_1	B_2	B_3	B_4	B_5	B_6
Firm I's strategies	A_1	G_{11}	G_{12}	G_{13}	G_{14}	G_{15}	G_{16}
	A_2	G_{21}	G_{22}	G_{23}	G_{24}	G_{25}	G_{26}
	A_3	G_{31}	G_{32}	G_{33}	G_{34}	G_{35}	G_{36}
	A_4	G_{41}	G_{42}	G_{43}	G_{44}	G_{45}	G_{46}
	A_5	G_{51}	G_{52}	G_{53}	G_{54}	G_{55}	G_{56}

Fig. 9.1 : Payoff matrix of Firm 1

Source : Modern Microeconomics (2nd Edition) by A. Koutsoyiannis

Table 9.1 shows the payoff matrix of firm 1, with the given example. As we can see, firm A has five strategies open to it, whereas firm 2 has six possible strategies. The payoff can be expressed in the form 'Gij', where 'G' means gain to firm 1 due to adoption of strategy or move 'i' by firm 1 and counterstrategy (or countermove) 'j' has been adopted by firm 2.Take an example e.g. let us consider box corresponding to payoff G23; it means payoff to firm 1, when it adopts strategy A2 and the counterstrategy taken by rival firm 2 is B3.

Likewise, the payoff matrix of firm 2 can also be drawn in the form of a table. Here one thing needs to be noted. The counter movements made by rivals are probable but not certain. In other words, as discussed in the above example, firm 1 knows that firm 2 can adopt any of the six strategies, but it is not certain about which strategy will be actually chosen by it.

(vi) Non-cooperative versus cooperative games: A game is cooperative if the players can negotiate binding contracts that allow them to plan joint strategies. A game is non-cooperative if the players find that negotiation and enforcement of binding contracts are not possible.[2]

Let us understand with the help of an example. A buyer wants to purchase a furniture item and puts its value at Rs. 25,000. Let the cost of that item be actually Rs. 15000. If negotiations take place between the buyer and the seller, the game will be cooperative, as any price reached between both the parties through bargaining (between Rs. 15001 and Rs. 24999) will make both better off. Likewise, two firms can also decide to undertake a project (to increase productivity) jointly, if either of them is not able to do so individually .But in case of a non-cooperative game , there is no scope for a binding contract. Most of the game theories discussed today in Microeconomics are non-cooperative games. Here the main focus of the firm is on the rival's point of view. The firm tries to understand the rational course of action taken by the rival firm and choose a strategy based on that.

9.3.2 Assumptions of game theory

The game theory is based on the following assumptions :

- Each player (oligopolist), while choosing a strategy, thinks that his rival will choose a strategy that is worst for him. In other words, the firm expects his rival to choose theaction most unfavourable to him. So he plays it safe. So the oligopolist chooses the best of the worst i.e. he chooses the maximum returns among the minimum returns.

- The oligopolist knows the complete set of strategies open to him as well as to his rivals.
- The struggle is of the nature of a 'strictly adversary game' i.e. a game in which the outcome which is favourable from one's point of view is unfavourable from another's point of view.

9.4 ZERO SUM GAME AND NON-ZERO SUM GAME

Two person Zero sum game

Before understanding a zero sum game, let us understand a constant sum game. A constant sum game is one in which the outcome to the two players adds up to a constant amount. For example, if the constant sum of profits is 10, if first players's share is 3, the second player's share will be 7. If first player's share is 6, the second player's share shall be 4. In other words, the total amount 10 has to be divided between the two players. One's gain is thus another's loss. Zero sum game is a special type of constant sum game in which the net gain adds up to zero.

Assumptions of the certainty model. The following assumptions are made in this duopoly model:

- The firms have a given well defined goal. In this example, the goal is market sharemaximisation.
- The strategies open to the firm and its rival are known to the firm.
- The firm knows with certainty how much cost, revenue, profit etc. shall accrue to it as a result of each combination of strategies.
- The total size of the market is not affected by the actions taken by the firms. To make things more clear, one's gain in market share is another's loss. So both the firms try to grab each other's share in the market.
- The firm acts rationally i.e. it expects the worst from its rival.

- Since the goals of the firms are diametrically opposed, there is no scope for collusion.

Let the two firms are firm 1 and 2. Four strategies are open to the first firm, while the second firm has five possible strategies to choose from. Table 9.2 and 9.3 show the payoff matrices of the duopolists. It should be noted that the sum of payoffs in corresponding cells of the two tables adds up to unity or one, as the goal is market share which is to be shared between the duopolists.

Table 9.2 shows the payoffs to firm1. The firm 1 will examine the payoffs row wise, as each row shows the counter strategy adopted by firm 2, in response to its own strategy. As it expects the worst outcome for itself, it looks for the row minima. The minimum returns that it expects as a result of its strategy A1, A2, A3 and A4 are 0.10, 0.30,0.20 and 0.15 respectively. It chooses the best of these i.e. 0.30, which results due to its strategy A2 and countermove B2 by the rival. As the firm chooses the best of the worst, it is called 'maximin strategy'.

The payoffs of firm 2 are shown in table 9.3. Here again, the approach is the same. But here instead of row minima, the column minima is examined, as the payoffs are shown column wise here. As we can see, best of the worst i.e. maximin strategy brings payoff equal to 0.70, that results from strategy A2 and B2 by firm 1 and 2 respectively.

However, in case the payoffs of firm 2 too are to be shown through the same table (see table 9.4), (and usually one table is adequate), minimax strategy i.e. minimum of the maximum is chosen. In other words, the firm 2 examines column maxima and chooses the worst . Again the payoffs result from strategy A2 and B2 being chosen by firm 1 and 2 respectively.

The equilibrium solution is thus strategy A2 by firm 1 and B2 by firm 2. Payoffs to firm1 and 2 are 0.30 and 0.70 respectively. Since both the firms prefer this strategy simultaneously , it is called 'dominant strategy' .Moreover, it doesn't matter who plays first, firm 1 or 2, or whether or not they play simultaneously, as the maximin strategy of one co-incides with the minimax strategy of the other. So the payoff matrix is said to possess an equilibrium point or saddle point, as neither of the two players have an incentive to change the strategy once chosen, as the coincident payoff entry in the table (corresponding to cell A2B2) in table 9.2 is acceptable to both.

Table 9.2

Payoff matrix of Firm I

Firm I's strategies	Firm II's strategies: B_1	B_2	B_3	B_4	B_5
A_1	0·10	0·20	0·15	0·30	0·25
A_2	0·40	0·30	0·50	0·55	0·45
A_3	0·35	0·25	0·20	0·40	0·50
A_4	0·25	0·15	0·35	0·60	0·20

Table 9.3

Payoff matrix of Firm II

Firm I's strategies	Firm II's strategies: B_1	B_2	B_3	B_4	B_5
A_1	0·90	0·80	0·85	0·70	0·75
A_2	$\boxed{0·60}$	$\boxed{0·70}$	$\boxed{0·50}$	0·45	0·55
A_3	0·65	0·75	0·80	0·60	$\boxed{0·50}$
A_4	0·75	0·85	0·65	$\boxed{0·40}$	0·80

Table 9.4

Combined payoff matrix

Firm I's strategies (maximin behaviour)	Firm II's strategies (minimax behaviour): B_1	B_2	B_3	B_4	B_5
A_1	0·10	0·20	0·15	0·30	0·25
A_2	$\boxed{0·40}$	$\boxed{0·30}$	$\boxed{0·50}$	0·55	0·45
A_3	0·35	0·25	0·20	0·40	$\boxed{0·50}$
A_4	0·25	0·15	0·35	$\boxed{0·60}$	0·20

Check Your Progress-I

Answer the questions in the space given below

Q1. What is a payoff?

Q2. What is zero sum game?

Q3. Distinguish between maximin and minimax strategy.

9.4.2 Non-zero Sum Game

Let us now consider the case of a non-zero-sum game or variable-sum game. We shall try to understand whether the maximin strategy that led to a stable equilibrium in zero-sum game holds applicable here as well or a different solution is more desirable.

Let us assume there are two firms, firm 1 and 2. Let price be the instrument variable. There are two strategies open to both the firms; i.e. both can charge either of the two prices, Rs.5 and Rs.3. The firm with lower price expects to gain more profits. But unlike in a zero-sum game, here one's gain is not another's loss. The total market size is affected by the joint action of the rivals. Moreover, each firm has a different cost structure.

Firm I's payoff matrix
(level of profits of I)

		Firm II's strategies	
		$P = 5$	$P = 3$
Firm I's strategies	$P = 5$	$\Pi_A = 90$	$\Pi_A = 50$
	$P = 3$	$\Pi_A = 150$	$\Pi_A = 80$

Firm II's payoff matrix
(levels of profit of II)

		Firm II's strategies	
		$P_B = 5$	$P_B = 3$
Firm I's strategies	$P_A = 5$	$\Pi_B = 110$	$\Pi_B = 120$
	$P_A = 3$	$\Pi_B = 60$	$\Pi_B = 100$

Table 9.5 and 9.6

We have drawn above tables 9.5 and 9.6 to explain the case of a non-zero-sum game. The payoffs of firm 1 and 2 are separately shown in tables 9.5 and 9.6 respectively. Let us consider table 9.5 first. The firm has two strategies to choose from. The payoffs corresponding to these strategies (and given the counter-strategy of rival) are being examined by the firm. If it chooses price 5, the row minima is 50;if it chooses price Rs. 3, row minima is 80. Given the maximin strategy, it will choose price Rs. 3. Likewise, firm 2 also chooses price Rs. 3, as it too resorts to maximin strategy (table 9.6). So a unique equilibrium price Rs. 3 has been chosen by both the firms. It also depicts a case of dominant strategy, but it doesn't mean that this is the best strategy for the firms. In fact, both are worse off, as this strategy doesn't provide optimal solution. Had they chosen an alternative strategy, it would have maximised not only their individual profits but also joint profits. If the firms undergo collusion and charge a higher price of Rs.5, you can see individual profits to improve (90>80,110>100) as well as the joint profits being maximised (as 200>180). This has been depicted in table 9.7 . Hence the maximin strategy does not provide an optimal solution in case of a non-zero-sum game.

Table 9.7

Combined payoff matrix

		Firm II's strategies			
		$P_B = 5$		$P_B = 3$	
Firm I's strategies	$P_A = 5$	$\Pi_A = 90$	$\Pi_B = 110$	$\Pi_A = 50$	$\Pi_B = 120$
		Joint $\Pi = 200$		Joint $\Pi = 170$	
	$P_A = 3$	$\Pi_A = 150$	$\Pi_B = 60$	$\Pi_A = 80$	$\Pi_B = 100$
		Joint $\Pi = 210$		Joint $\Pi = 180$	

Factors on which adoption of maximin strategy depends

The game theory, in particular the conservative maximin strategy has been able to explain and analyse many actions of oligopolistic firms e.g. many of these firms don't adopt a strategy that maximises their profits. On the contrary, they move cautiously, as they expect the worst from their rivals. But whether or not a firm adopts maximin strategy depends on the following factors:

- If duration of rivalry is long, firms know from their experiences the reaction pattern of their rivals. So they are able to predict the reactions of their rivals. If their past experience shows that any move taken by them in accordance with maximin strategy proved disastrous, they avoid it.

- Maximin strategy is adopted more in markets where there are frequent changes in tastes of customers.

- If the firm in question is not having access to information or lacks communication with rivals, the worst is expected and hence the strategy chosen is maximin.

- If imitation of an action (a new product or process) is easy, firms adopt this strategy rather than taking aggressive steps. However, if imitation involves a time lag, firms choose actions that bring them most favourable outcomes rather than expecting the worst from rivals.

9.5 DOMINANT STRATEGIES

In the above section, while discussing zero sum game and non-zero sum game, we have also discussed the concept of dominant strategy, i.e. a strategy that is successful irrespective of the move made by the competing firm or rival. There we used the tool of payoff matrix in tabular form. In this section, we shall explain it through a game tree. We shall take the example of advertising game played by two players, who happen to be the oligopolistic (duopolistic) firms.

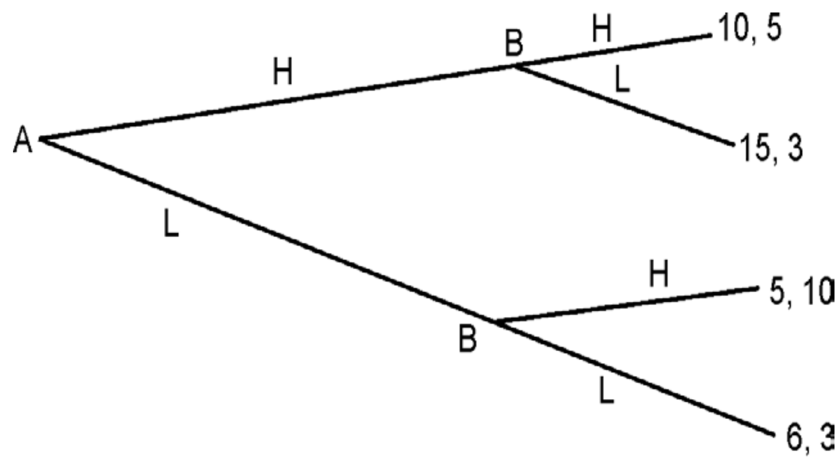


Fig. 9.1 : Dominant Strategy–The Advertising Game

A game tree has been drawn above (Fig 9.1.) Here there are two firms A and B, selling competing products. There are two strategies, whether the advertising budget should be high (H) or low (L) . The action proceeds from left to right. The payoffs are provided in the end e.g. If A chooses 'H' and B also chooses 'H', payoff 10 accrues to A and gain to B is 5. The different possible strategy combinations along with the payoffs to A and B are : A-H, B-H (10,5); A-H,B-L (15,3);A-L, B-H (5, 10); A-L, B-L (6,3).As we can observe, the optimal decision for A is H i.e. spending more on advertisements regardless of what his opponent does. Likewise, firm B's optimal strategy is again H, without worrying about what action A takes. We can thus conclude that high advertising outlay is the dominant strategy here. Such an equilibrium is referred to as equilibrium in dominant strategies.

Table 9.8 : Pay-off Matrix for the Advertising Game

		Firm B's Strategies	
		High Advertising Budget	*Low Advertising Budget*
Firm A's Strategies	*High Advertising Budget*	10, 5	15, 3
	Low Advertising Budget	5, 10	6, 3

Whatever we have discussed through game tree can also be discussed through payoff matrix drawn above (table 9.8). As we can observe from the table, the dominant strategy is High advertising budget by both A and B.

(Note : You can revise the section on payoff matrices or the section dealing with maximin strategy to elaborate this statement and to be able to explain how we reach this outcome).

Check Your Progress-II

Answer the questions in the space given below

Q1. What are the factors on which maximin strategy depends?

Q2. Illustrate the case of a firm's dominant strategy through a game tree.

9.6 LET US SUM UP

Dear learner, in this lesson, we discussed few important concepts related to game theorylike payoff matrices, maximin strategy, dominant strategy, game tree etc.

9.7 KEYWORDS

Zero sum game : Though this concept is widely used in game theory, we also experience it in our day to day life. For example, if there is one vacancy available for a job, only one person can get selected. The net outcome in case of zero sum game is zero, as the winner can win at the cost of loser or gain and loss cancel out each other.

Non-zero sum game : In case of a non-zero sum game, the loss or gain of two players don't match. In other words, one's gain is not necessarily other's loss. There is scope for cooperation.

Maximin strategy: In a game theory, the two players think that the other will do the worst for him. Keeping that in view, he tries to maximise the worst outcome or his strategy is to make the best out of worst.

Minimax strategy: In such a case, the player thinks that the opponent will try to give him maximum loss. This player tries to minimize that loss. For this, he considers all the strategies of his opponent and chooses that strategy for him that suits him.

9.8 EXAMINATION ORIENTED QUESTIONS

1. What is game theory? How does it contribute to oligopolistic market analysis?
2. What is a payoff matrix? Give an example to explain.
3. What is maximin strategy? How does it differ from minimax strategy?
4. Why is maximin strategy the best for firms?
5. What is zero sum game? Explain through payoff matrix.
6. What is a dominant strategy? Explain through both a game tree as well as payoff matrix.
7. Why does maximin strategy not provide optimal solution in case of non- zero sum games?
8. Distinguish between a non-cooperative and a cooperative game.
9. Discuss the concepts of payoff, player and strategy.

9.9 HINTS TO CYP

CYP-I

Q1. See sub-section 9.3.1

Q2. See section 9.4.1

Q3. See sub-section 9.4.1

CYP-II

Q1. See sub-section 9.4.3

Q2. See section 9.5

9.10 SUGGESTED READINGS

1. Ahuja, H.L., Advanced Economic Theory, 9th edition (1997).
2. Koutsoyiannis, A., Modern Microeconomics, 2nd edition (1979).
3. Misra, S.K. and V.K. Puri, Advanced Microeconomic theory, first edition (2001).
4. Pindyck, Robert. S and Daniel l. Rubinfeld, Microeconomics (fifth edition).
5. Pindyck R.S., D. Rubinfeld and P.L. Mehta (2012), Microeconomics, Pearson Education, South Asia.
6. Salvatore, D. (2010), Microeconomics; Theory and Applications, Oxford University Press, New Delhi
7. Cooper, R., & John, A. (2011). Economics–theory through applications, Saylor Foundation. [3]
8. Shapiro, D., MacDonald, D., & Greenlaw, S. A. (2022). *Principles of Microeconomics 3e*.[4]

END NOTES

1. Walter Nicholson, Microeconomic theory: Basic principles and extensions

(Orlando: The Dryden press, 1992).

2. Robert S. Pindyckand Daniel L. Rubinfeld, ‘Microeconomics’ (New Delhi: Prentice Hall of India, 1995)

3. Type of license/Conditions of use: CC BY-NC-SA (Attribution-Non commercial-ShareAlike) .Scan QR code 1 to access **Cooper, R., & John, A. (2011). Economics–theory through applications.**

4. This book is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). “Access for free at openstax.org.” **Scan QR code 2 to access the e book by Shapiro, D., MacDonald, D., & Greenlaw, S. A. (2022).** ***Principles of Microeconomics 3e*****.**

QR code 1

QR Code 2

M.A. ECONOMICS **LESSON NO. 10**

COURSE NO. ECO 201 **UNIT-II**

GAME THEORY : PART-II

STRUCTURE

10.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

- Define the Nash equilibrium
- Describe Pure and mixed strategies
- Explain the prisoner's dilemma
- Relate Prisoners' dilemma and instability of a cartel
- Discuss repeated games
- Define sequential games
- Evaluate entry deterrence

10.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to :

- Define the Nash equilibrium
- Describe Pure and mixed strategies
- Explain the prisoner's dilemma
- Relate Prisoners' dilemma and instability of a cartel
- Discuss repeated games
- Define sequential games
- Evaluate entry deterrence

10.2 INTRODUCTION

Dear learner, in the last lesson, we discussed many concepts of game theory. In this lesson, we shall continue to discuss the remaining topics.

10.3 THE NASH EQUILIBRIUM

It is not always possible to have a dominant strategy for both the players in a game. Here optimal decision made by a firm depends on what strategy has been adopted by his rival. Here the concept of Nash equilibrium holds importance. Nash equilibrium has been referred to a stable situation in which each firm is doing the best it can, given what its competitors are doing. For understanding Nash equilibrium, let us draw another table (table 10.1), which is a modified form of the earlier table (table 9.8 in lesson 9).

Table 10.1 : Pay-off Matrix for Modified Advertising Game

		Firm B's Strategies	
		High Advertising Budget	*Low Advertising Budget*
Firm A's Strategies	*High Advertising Budget*	10, 5	15, 3
	Low Advertising Budget	5, 10	18, 3

The table drawn above differs from the table 9.8 in only one respect i.e. with regard to the payoffs in bottom right hand corner or the cell corresponding to adoption of strategy L by both the firms. Here we can observe that payoffs to firm A as a result of incurring low expenditure on advertisements has increased to 18, whereas the gains to B remain the same i.e.3. This shows that A has been resorting to advertisements as a defensive move, to counteract firm B's advertisements. It is also expensive for it to go for it; that is why low advertisement expenditure has increased its payoffs.

But from this we should not conclude that firm A has chosen low advertisement expenditure as the dominant strategy. In fact, as we shall see, in this case the firm A has no dominant strategy. Whatever strategy is chosen by A shall depend on B. If firm B goes for H, A goes for H and if B goes for L, it is profitable for A also to go for L. On the other hand, B's strategy is dominant. It will choose H no matter how A reacts.

Let us suppose that firm B does not react first; instead it has to be decided simultaneously by both the firms how much (high or low) they want to spend on advertisements. Now A must think from the point of view of B. As the optimal strategy for B is given, A concludes that B shall choose H. So it is advisable for A also to choose H. Both A and B thus choose High advertising expenditure. It is the best strategy for A, given B's decision and the best for B, given A's decision. It is an example of Nash equilibrium. Like dominant strategies, Nash equilibrium also leads to stable equilibrium, as there is no incentive on the part of players to deviate from the Nash strategy chosen.

A beautiful comparison has been drawn between a Dominant strategy and Nash equilibrium in the box below.

Dominant Strategies:	I'm doing the best I can *no matter what you do.* You're doing the best you can *no matter what* I do.
Nash Equilibrium:	I'm doing the best I can *given what you are doing.* You're doing the best you can *given what I am doing.*

Source : **(Pindyck and Rubinfeld, Microeconomics p.466)**

a game may not have a Nash equilibrium at all; while others may have several Nashequilibria.

Let us discuss the case of 'product choice problem' to learn more about this. Let us assume that there are two firms,1 and 2. Let they be the manufacturers of breakfast cereals. Let us assume that two new variants of breakfast cereals can be launched successfully, provided one firm introduces one such product. Also the resources of firms are limited, so that either of them can produce only one of these products.

The payoff matrix has been shown in table 10.2 below:

Table 10.2 : Product Choice Problem

		Firm 2	
		Crispy	Sweet
Firm 1	Crispy	–5, –5	10, 10
	Sweet	10, 10	–5, –5

As we can see from the table, there are two variants: crispy and sweet. Had coordination or collusion been possible, they would have mutually agreed what to produce (and shared the market). But since in the real world, firms usually play a non-cooperative game, this is not possible. But firms can still have access to each other's information, they can get hints dropped by the rivals. Let the firm 1 indicates through a press release that it is introducing sweet cereal. After hearing this news, let firm 2 announce that it shall produce crispy cereal. Since both get to know about each other's proposed action, neither of them likes to deviate from this action. In other words, firm 1 chooses to produce sweet and firm 2, crispy. The payoff matrix shows that the bottom left corner solution i.e. 10,10 is Nash equilibrium in this case. But upper right hand corner solution is also a Nash equilibrium. So there is no unique Nash equilibrium here. The other two cells show that both lose if they produce the same product (sweet, sweet or crispy, crispy).

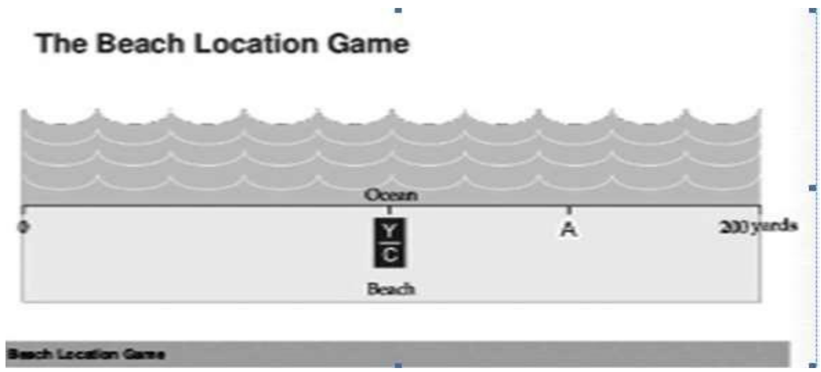


Fig. 10.1

Another example is the 'Beach Location game', where you and your competitor are selling the same soft drink to the people who visit it. The beach is 200 yards long and the sunbathers are spread evenly. Finally, the stable (Nash) equilibrium for both you and your competitor shall be to get located at the same point in the center, as from here neither wants to deviate.

Check Your Progress-I

Answer the questions in the space provided.

Q1. What is a Nash equilibrium? How does it differ from dominant strategy?

____

____

____

____

Q2. What is a product choice problem? What does it depict?

____

____

____

____

Q3. Show Nash equilibrium with the help of beach location game.

____

____

____

____

Q4. Is it possible to have more than one Nash equilibria? Give example.

____

____

____

____

10.4 PURE AND MIXED STRATEGIES

Pure strategies: All the strategies discussed in the various models were pure strategies. In other words, when a firm pursues one course of action like low advertisement expenditure, charging price of Rs. 3, selling crispy product etc., it is a pure strategy. So when a specific choice is made by the players in a game, it is a case of pure strategy.

Mixed strategies: However, there are games in which a pure strategy can't ensure Nash equilibrium. For that, a mixed strategy is required .Here the concept of mixed strategies needs to be introduced. Situations where pure strategies fail to yield a solution, mixed strategies do. Mixed strategies are defined as : "strategies in which players make random choices among two or more possible actions, based on sets of chosen probabilities."

Let us understand this through 'Matching pennies' and 'Battle of the sexes'.

Table 10.3 : Matching Pennies

		Player B	
		Heads	Tails
Player A	*Heads*	1, −1	−1, 1
	Tails	−1, 1	1, −1

In this game (matching pennies), each player chooses heads or tails and the coins are to be revealed at the same time. If the coins match (both are heads or both are tails),A wins and gets a dollar from B. If they don't match, B wins and gets a dollar from A. There is no Nash equilibrium in pure strategies here. If A plays heads, B plays tails. But if B plays tails, A too plays tails. So there is no combination of heads and tails that leaves both the players satisfied; they would like to change the strategies. Now if instead ofplaying justheadsortails, let us assume that theplayers simplyflip thecoins and play heads and tails with probability ½. Here we can get a Nash equilibrium.

Table 10.4 : The Battle of the Sexes

		JOAN	
		Wrestling	Opera
JIM	Wrestling	2, 1	0, 0
	Opera	0, 0	1, 2

'Battle of the sexes' is one of the games in which Nash equilibria take place in both pure and mixed strategies. Let Joan and Jim wish to spend time together . Joan wants to go to Opera, but Jim wants to watch mud wrestling. Now the payoff matrix shows that Joan would most prefer to go to the Opera with Jim, but prefers watching mud wrestling along with Jim to visiting Opera alone. Similar is the case with Jim. It must be noted that here there are two Nash equilibria; Jim and Joan watching mud wrestling together and both of them going to the Opera. Joan prefers the first outcome and Jim, the second. But both are equilibria, as neither Joan nor Jim wants to deviate from his or her decision, given the decision of the other.

This game can also have an equilibrium in mixed strategies. If Jim chooses wrestling with probability 2/3 and opera with probability 1/3 and Joan chooses opera with probability 2/3 and wrestling with probability 1/3, the outcome is random and each will have an expected payoff 2/3. Let us explain this. Let p be the probability of wrestling and (1-p) be the probability of opera. Since Joan is using probability 1/3 for wrestling and 2/3 for opera, the probability that both will choose wrestling is (1/3)p and probability that both will choose opera is (2/3) (1-p). Jim's expected payoff is 2(1/3)p + 1 (2/3) (1-p) = 2/3. Likewise the payoff to Joan is also 2/3.

Now the question is: do these players use these mixed strategies? Here, the answer is no, unless it is a strange couple. The payoffs that result from pure strategy have a minimum value 1. By adopting mixed strategies, it is 2/3. So they prefer a pure to a mixed strategy.

10.5 THE PRISONER'S DILEMMA

We have discussed earlier that in most cases of variable sum games, the conservative or maximin-minimax behaviour of the rivals leads them to sub-optimal solutions, that is worse than what it should be. These are a special case of 'prisoner's dilemma'. Prisoner's dilemma is the most well known example of strategic behaviour. It explains how rivals, while behaving selfishly act against their mutual disadvantage.

The original model has been developed by taking an example of two prisoners, Billa and Ranga, who have jointly committed a bank robbery. The prosecution has not enough evidence against their crime. So the only option is confession. Both have been put in separate jails, so that there is no communication between them. While interrogating them, the police offer them the following (individually):

- If both confess, each will go to jail for 6 years.

- If neither of them confesses, prosecution still have little evidence against them, on the basis of which they shall be imprisoned for 3 years.
- If one confesses and the other does not, then one who confesses shall get imprisoned for 1 year, whereas the one who doesn't confess, gets 12 years' imprisonment.

The payoff matrix is given below in table 10.5

Table 10.5

Table : Payoff Matrix for Prisoners' Dilemma

		Prisoner B			
		Confesses		Does not confess	
Prisoner A: Confesses		A	6 years	A	1 year
		B	6 years	B	12 years
Prisoner A: Does not confess		A	12 years	A	3 years
		B	1 year	B	3 years

As we can see in the table, two strategies are given to both the accused. The outcome or payoff is in the form of length of sentence. The payoffs are: If both Ranga (prisoner A) and Billa (prisoner 2)confess , each will get imprisonment of 6 years. If A confesses (i.e. cooperates with police), while B doesn't, A will get punishment of 1 year, while B of 12 years. Likewise, in response to non-confessing by A, payoffs are 12 years for A and 1 year for B, if the latter confesses. Finally, if B too doesn't confess, they will get punishment of 3 years each.

Now which strategy to adopt? Here the ideal outcome or the best combined solution is the bottom right corner, or both not confessing and getting the minimum sentence. But this is possible only if they were able to communicate, collude or at least trusted each other. None of these seems to exist. Under uncertainty of how the co-accused will react, both of them adopt a strategy independently, but payoff to each depends on how the other has reacted. A thinks that : " if I don't confess, B may confess and get less punishment (instead of 12 years to me)." Similar approach is made by B, who too prefers to confess. Thus by acting selfishly and in self-interest , both cheat each other and end up confessing to crime. Had they trusted each other and been loyal to each other, both of them would have chosen not confessing and hence attained the best solution. But under the given circumstances, the decisions of both is quite rational, as they have chosen best of the worst. However, it should be noted that confessing is dominant strategy for both the prisoners, as it gives a higher payoff irrespective of the strategy of the other. It is a maximin strategy as well. It is also a Nash equilibrium.

10.5.1 Prisoners' dilemma and instability of a cartel

We have discussed above the basic prisoners' dilemma model. This model has wide applications in oligopolistic pricing game. Let us consider the case of a cartel agreed to by two firms. How much output is to be produced and at what price to be sold is mutually decided. Now there are two strategies open to the firms, cheat (i.e. lower the price) and maximise individual profits or cooperate and enjoy joint Monopoly profits. However, if both cheat, cartel would break down and profits would be reduced to competitive level. The payoff matrix has been shown below:

Table 10.6 : Payoff Matrix for Cartel Members

		Firm A	
		Cheat	Cooperate
Firm B	Cheat (lower the price)	A : 5 lakhs B : 5 lakhs	A : 2 lakhs B : 25 lakhs
	Cooperate	A : 25 lakhs B : 2 lakhs	A : 15 lakhs B : 15 lakhs

The table shows that if both cooperate, they get profits of Rs. 15 lakh each. If they don't cooperate i.e. cheat, the profits are competitive due to collapse of the cartel. If either of them cheats, profits to that firm are Rs. 25 lakh or to the other, Rs.2 lakh. A's best strategy is to cheat. Same is true for B. So both of them end up choosing Cheat and hence get profits of Rs. 5 lakh each. If they had cooperated, both would have maximised profits. But under Oligopoly, firms generally face prisoners' dilemma and don't trust each other.

10.6 REPEATED GAMES

In case of prisoners' dilemma, the accused persons don't cooperate with each other, as the game is to be played only once i.e. may be the criminal is given only one chance to confess or not in his life. But in real life, firms play repeated games. Let us now discuss, whether, playing repeated games makes firms find a way out of this dilemma? In other words, can cooperation and coordination be made possible in Oligopoly, generally known for rivalry among the firms?

A repeated game is one in which actions are taken and payoffs received over and over again. With each repetition of prisoners' dilemma, each firm develops a reputation of its own as well as gets to know about the behaviour of its rival firms. This repetition is able to change the likely outcome of the game.

Let us understand this.

Table 10.7 : Pricing Problem

		Firm 2	
		Low Price	High Price
Firm 1	Low Price	10, 10	100, –50
	High Price	–50, 100	50, 50

Imagine you are firm 1. Your competitor is firm 2. Here again, the payoff is 10, 10 (profits) in case both charge a low price, 50,50 in case both charge a high price, loss to one (-50) if only one charges high price and gain to the other (100). But as the two firms don't cooperate in one time game, the outcome here shall be both charging a low price. So far, the discussion is a repetition of what we have discussed in prisoners' dilemma.

Repeated games make a departure with the earlier model. Here instead of one time play, the two firms play again and again. Assume that you and your competitor announce prices simultaneously on the first of every month. Here the decision may change.

In a study, Robert Axelrod asked game theorists to come up with a strategy that can best explain this situation and solve it, given the fact that the game is to be played over and over again. The aim was to look for a strategy that was the most robust i.e. would work best on an average, against all other strategies. The strategy that worked best was surprisingly the simplest i.e. 'Tit for tat' strategy.

The tit for tat strategy has been thus defined by Pindyck and Rubinfeld : "Repeated game strategy in which a player responds in kind to an opponent's previous play, cooperating with cooperative opponents and retaliating against uncooperative ones."

The tit for tat strategy in the context of example taken is : " I start out with a high price, which I maintain , so long as you continue to cooperate (i.e. charge a high price). As soon as you lower your price (i.e. don't cooperate), I follow suit and lower mine. If you later decide to cooperate and raise your price again, I'll immediately raise my price as well."

Now whether this tit for tat strategy works here or motivates my competitor to cooperate again depends on whether this game is played for finite number of times or infinite number of times. In case the game is to be played for infinite period, cooperative behaviour is the rational response. So every month, while announcing prices, my competitor knows that I will go for tit for tat strategy. So he will not try to undercut me, as the short term gain made by him this month shall be more than offset by the cumulative losses in profits that shall accrue to both of us. Here one thing needs to be noted. My competitor need not be sure that I am actually adopting tit for tat strategy. Even if there is a small probability, it is still rational for him not to undercut, as the expected gains from cooperation are much more.

On the other hand, the rational response may change, if the game is to be played a finite number of times, say there are 20 moves to be made. Now my competitor knows that he has to cooperate with me till the 19^{th} move; he decides to undercut me in the last move, as the game will be over and I won't be able to retaliate. But similar thinking goes on in my mind as well. So I may also think that in the last move, I will not cooperate and hence charge a low price. But my rival might have predicted my behaviour. So he may think of undercutting me in the 19^{th} move itself. This holds true for me as well. It means that this behaviour will continue for all the preceding moves. So here the rational outcome for both of us is to charge a low price. In this case, we are back to square one i.e. are still stuck in prisoners' dilemma, which we wanted to overcome.

10.7 SEQUENTIAL GAMES

So far, we have discussed the models in which there was no first mover. In other words, both the players move simultaneously. But in sequential games, players move in turn. The examples may include a decision regarding advertisement by one firm, which is observed by the competitor who then reacts accordingly. Sequential games are easier to understand and analyse than simultaneous games. Logical reasoning plays a great role here.

To understand the game, let us consider the example of payoff matrix in case of product choice:

Table 10.8

Table : Payoff Matrix for Product Choice

		Firm B	
		Sweet	Salty
Firm A	Sweet	–10, –10	40, 20
	Salty	20, 40	10, 10

(Note: The product choice problem has already been discussed. It is expected that you know what it is all about)

The payoff matrix given above has been discussed about two firms A and B, where A has the advantage of being the first mover. A announces launching of sweet biscuits. If B also launches the same, it will not only incur loss of 10 to itself, but also to firm A. So he launches salty, as the payoff here is still 20, though it is half of what has been gained by A. B has to be satisfied with this launch even if it gives him low returns, as he has missed an opportunity of being the first mover. The options have been closed for the follower.

The same can be expressed in the form of 'Extensive form of a game'. Many a time, sequential games are easier to understand with the help of extensive form of a game or a game tree. As we know, the game tree or decision tree shows the payoffs at the end of a branch. To understand the solution, we work backwards from the end. Here the best sequence of moves for firm A is the one in which it earns 40 and firm B earns 20. Hence it launches sweet. This has been shown in fig. 10.2.

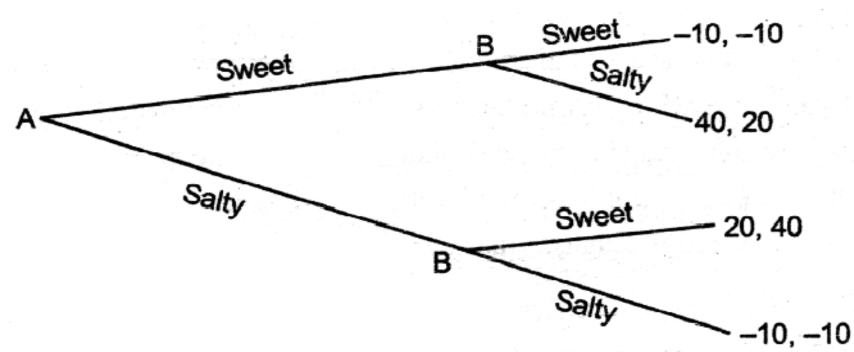


Fig. 10.2 : Product-choice Game

10.7.1 Games of perfect information and backward induction strategies:

Here we will discuss the strategy of backward induction. Suppose there is perfect information. Under perfect information, we can use this technique to solve game and reach equilibrium. Here, we begin by analyzing decision nodes, or by working backwards from the final stage. In fig. 10.3 a, we are given the penultimate nodes, x and y. Both these belong to player 1. At x, player 1 chooses R^1 and at y, he chooses L^{11}. So the game in Fig. 10.3 a can be reduced to 10.3 b, where the decision nodes, x and y have been replaced by the payoffs, that are inevitable, once x and y are reached. Let us now repeat this process on the reduced game. Now both w and z are penultimate decision nodes, belonging to player 2. If w is reached, player 2 chooses r and if z is reached, he chooses *I* . The game gets reduced further and this is shown in fig. 10.3 c, where it is clear that player 1 will choose R. We can thus conclude that player 1 will choose

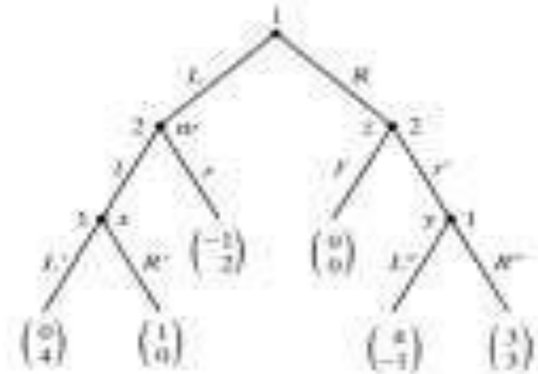
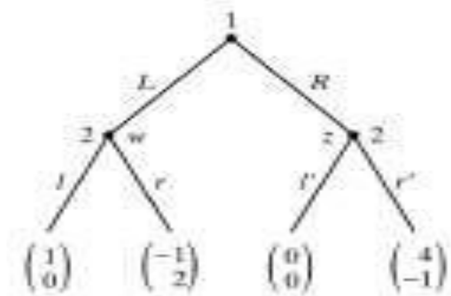


Fig. 10.3(a)

Fig. 10.3(b)



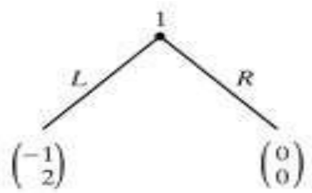


Fig. 10.3(c)

strategy (R,R'', L'') and 2 (r, *l*) .It means 5that player 1 will choose R on first move, R' if decision node x is reached and L'' if decision node y is reached. The outcome or payoff for both is zero.

10.8 GAMES OF IMPERFECT INFORMATION AND SUB-GAME PERFECTEQUILIBRIUM

Though the backward induction technique discussed above is applicable to games with perfect information, It does not, however, immediately extend to other games. Consider, for example, the game in Fig. 10.4 a, in which player 1 has the option of playing a coordination game with player 2. Let us try to apply the backward induction technique to it. As before, the first step is to locate all information sets such that whatever action is chosen at that information set, the game subsequently ends. For the game in Fig. 10.4 a, this isolates player 2's information set, i.e., the point in the game reached after player 1has chosen *IN* and then *L* or *R*. Note that when it is player 2's turn to play, taking either action *l* or *r* will result in the end of the game. Now, according to the backward induction algorithm, the next step is to choose an optimal action for player 2 there. But now we are in trouble because it is not at all clear which action is optimal for player 2. This is because player 2's best action depends on the action taken by player 1. If player 1 chose *L*, then 2's best action is *l*, whereas if player 1 chose *R*, then 2 should instead choose *r*. There is no immediate way out of this difficulty because, by definition of the information set, player 2 does not know which action player 1 has taken. We rely on backward induction, because to determine optimal play early in the game, we first have to understand how play will proceed later in the game. But in the example at hand, the reverse is also true. To determine optimal play later in the game (i.e., at player 2's information set), we must first understand how play proceeds earlier in the game (i.e., did player 1 choose *L* or *R*?).

Thus, in this game (and in games of imperfect information quite generally), we must, at least to some extent, *simultaneously* determine optimal play at points both earlier and later in the game.

Let us continue with our analysis of the game of Fig. 10.4a. Although we would like to first understand how play will proceed at the 'last' information set, let us postpone it and do the next best thing. Consider moving one step backward in the tree to player 1's second decision node. Can we determine how play will proceed from that point of the game onwards? If so, then we can replace that 'portion' of the game, or subgame, with the resulting payoff vector, just as we did in the backward induction algorithm. But how are we to determine how play will proceed in the subgame beginning at player 1's second information set?

The idea, first developed in Selten (1965, 1975), is to consider the subgame as a game in its own right. (See Fig.10.4 b.) Consider now applying the Nash equilibrium solution concept to the game of Fig. 10.4 b. There are two pure strategy Nash equilibria of this game: *(L, l)*, and *(R, r)*. Let us suppose that when this subgame is reached in the course of playing the original game, one of these Nash equilibria will be played. For concreteness, suppose it is *(L, l)*. Consequently, the resulting payoff vector will be *(*1, 3*)* if the subgame is reached. We now can proceed analogously to the backward induction algorithm by replacing the entire subgame by the resulting payoff vector *(*1, 3*)*. (See Fig. 10.4 c) Once done, it is clear that player 1 will choose *OUT* at his first decision node, because given the behaviour in the subgame, player 1 is better off choosing *OUT*, yielding a payoff of 2, than choosing *IN* and ultimately yielding a payoff of 1.

Altogether, the strategies previously derived are as follows. For player 1: *OUT* at his first decision node and *L* at his second; for player 2: *l* at his information set. Such types of strategies are called subgame perfect equilibrium strategies.

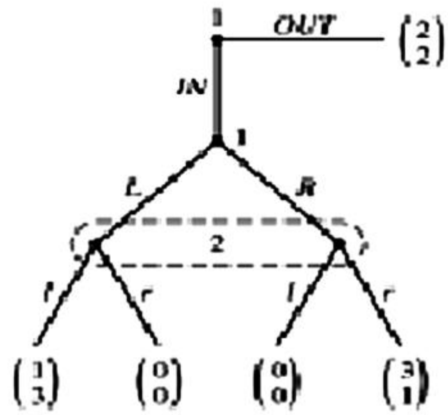


Fig. 10.4 (a)

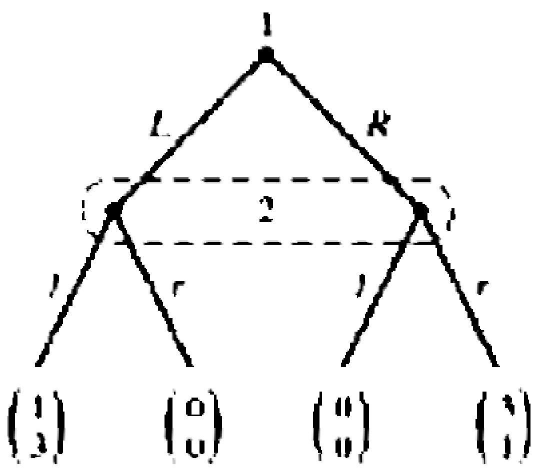


Fig. 10.4 (b)

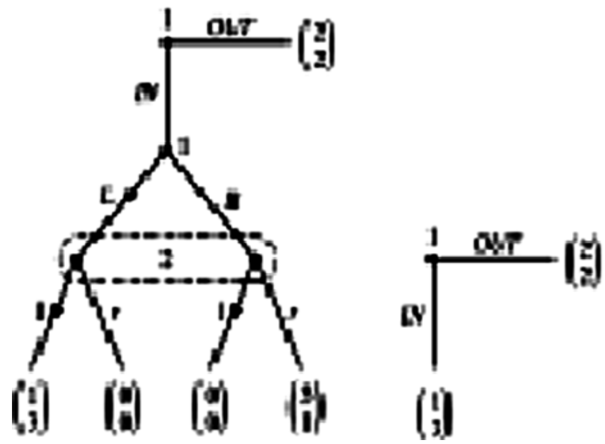


Fig. 10.4 (c, d)

10.8.1 Entry Deterrence

While discussing game theory, we come across the concept of entry deterrence. Let us assume that there is a Monopolist firm facing threat from a potential entrant. It can deter or prevent the entry of the prospective entrant if it can prove to the entrant that its entry will be unprofitable for it (entrant). But this threat is an empty threat as long as the competitor is not convinced. One the threat becomes credible, entry can be deterred. Let us explain this though the payoff matrix below:

Table 10.9

Table : The New Entry Game–Entry Possibilities

		Potential Entrant	
		Enter	Stay out
Incumbent	No warfare	9, 5	19, 1
	Warfare	8, –5	15, 0

Table 10.10

Table : The New Entry Game–Entry Deterrence

		Potential Entrant	
		Enter	Stay out
Incumbent	No warfare	6, 5	15, 1
	Warfare	8, –5	15, 0

As shown in the table 10.9, there are two strategies open to the potential entrant, to enter or not to enter. For the incumbent firm, again two choices are there: whether to indulge in warfare (i.e. reduce price) or not. Now, the payoff matrix shows that if firm B stays out, the payoff is 1 to it and 19 to the firm A. If it decides to enter, payoffs depend on how firm A responds. If it reduces price, there is net loss of 5 to firm B and its own gain is 8. If A doesn't reduce price, payoff to firm B is 5 and that to A is 9. But if firm B enters, the rational strategy for A is not to fight. The equilibrium outcome is thus the upper left corner. Here entry possibilities exist, as the rational strategy for the incumbent is to accommodate the other. Here threats to fight are empty and not taken seriously by the entrant.

But suppose, the incumbent firm commits itself to fight and makes threat credible so as to prevent the entry of the new firm. This is shown in table 10.10. Let firm A decide to increase its production capacity. But it will not use this capacity until the entry of new firm, as it is already producing the profit maximising Monopoly output. Now the payoffs depicted in the table have changed in favour of A due to increase in production capacity by A. If the potential entrant ignores the threat and decides to enter, A's payoff is 6 if it doesn't fight and 8 if it fights. So the rational strategy for A is to fight. Seeing this, firm B decides whether to enter or not to enter, given the strategy of A. Here it is a sensible policy not to enter, as the payoffs accruing to it due to entering and staying out are -5, 0 respectively.

Check Your Progress-II

Answer the questions in the space provided:

Q1. What is the difference between a pure and a mixed strategy?

____

____

____

____

Q2. What is a sequential game? How is different from a simultaneous game?

____

____

____

____

Q3. What is a prisoners' dilemma?

____

____

____

____

Q4. What is the difference between an empty threat and a credible threat?

____

____

____

____

10.9 LET US SUM UP

Dear learner, in this lesson, we discussed topics like Nash equilibrium, product choice problem, pure and mixed strategies, sequential games etc.

10.10 KEYWORDS

Cartel: A form of collusive oligopoly in which the decisions regarding prices, output, market sharing etc. are taken jointly. OPEC is the biggest cartel.

Sequential vs. simultaneous game: In a sequential or consecutive game, one player chooses a strategy first, followed by the second player. On the other hand, in a simultaneous game, neither of the players knows about the strategy of the rival, as they need to play at the same time.

Pure vs. mixed strategies: In a pure strategy, the player is clear about his strategy e.g. if it is a decision regarding whether advertisements are to be given to promote the product, it will be 0, 1. However, in certain cases, there is no pure strategy. In such cases, a probability is assigned to all the possible outcomes e.g. rock, paper and scissors are the three likely outcomes in this game. So the probability of each is 1/3, assuming that there is equal chance of their occurrence.

10.11 EXAMINATION ORIENTED QUESTIONS

1. How does a Nash equilibrium differ from an equilibrium in dominant strategy?
2. Why is Nash equilibrium a better option than the maximin strategy? Give examples to explain.
3. What do you mean by prisoners' dilemma? Explain the basic model.
4. Show with the help of payoff matrix how prisoners' dilemma leads to breakdown of a cartel?
5. Why do the criminals in prisoners' dilemma model confess, whereas that is a sub-optimal solution for both?
6. Is it possible for oligopolistic firms to show cooperative behaviour? Show with the help of repeated games.
7. Under what circumstances, is a tit for tat strategy successful?
8. What is a first mover advantage? Show this with the help of product choice problem.

9. What is entry deterrence?

10. When do entry possibilities exist in case of entry deterrence model?

11. How does an empty threat fail to influence the decision of potential entrant?

12. How does credible threat deter the entry of prospective entrant? Show with the help of payoff matrix.

13. What is a sequential game? Draw in extensive form.

14. What is a decision tree? How do we move along it for getting the solution?

15. What is battle of the sexes? Explain how both pure and mixed strategies lead to solution there. Also show why they opt pure rather than mixed strategies.

10.12 HINTS TO CYP

CYP-I

Q1. See section 10.3

Q2. See section 10.3

Q3. See section 10.3

Q4. See section 10.3

CYP-II

Q1. See section 10.4

Q2. See section 10.7

Q3. See section 10.5

Q4. See section 10.8.1

10.13 SUGGESTED READINGS

1. Ahuja, H.L., Advanced Economic Theory, 9th edition (1997).
2. Koutsoyiannis, A., Modern Microeconomics, 2nd edition (1979).
3. Misra, S.K. and V.K. Puri, Advanced Microeconomic theory, first edition (2001).
4. Pindyck, Robert. S and Daniel l. Rubinfeld, Microeconomics (fifth edition).
5. Boyes, William and Michael Melvin, Textbook of Economics (sixth edition, Indian adaptation)
6. Sloman, John, Economics (Third edition).
7. Pindyck R.S., D. Rubinfeld and P.L. Mehta (2012), Microeconomics, Pearson Education, South Asia.
8. Salvatore, D. (2010), Microeconomics; Theory and Applications, Oxford University Press, New Delhi
9. Cooper, R., & John, A. (2011). Economics–theory through applications, Saylor Foundation. [2]
10. Shapiro, D., MacDonald, D., & Greenlaw, S. A. (2022). *Principles of Microeconomics 3e*.[3]

END NOTE

1. Sections 10.7.1 and 10.8 rely heavily on **Advanced Microeconomic theory by GEOFFREY A. J EHLE**
2. Type of license/Conditions of use: CC BY-NC-SA (Attribution-Non commercial-ShareAlike) .Scan QR code 1 to access **Cooper, R., & John, A. (2011). Economics–theory through applications.**

3. This book is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). “Access for free at openstax.org.” **Scan QR code 2 to access the e book by Shapiro, D., MacDonald, D., & Greenlaw, S. A. (2022). *Principles of Microeconomics 3e.***

QR code 1

QR code 2

M.A. ECONOMICS | **LESSON NO. 11**

COURSE NO. ECO 201 | **UNIT-II**

PARTIAL AND GENERAL EQUILIBRIUM, THE WALRASIAN SYSTEM AND EXISTENCE UNIQUENESS AND STABILITY OF EQUILIBRIUM.

STRUCTURE

11.0 OBJECTIVES

The main objectives of this lesson are to enable you to :

- Define Partial and General equilibrium
- Discuss Walrasiansystem
- Explain Uniqueness and stability of equilibrium

11.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to:

- Define Partial and General equilibrium
- Discuss Walrasiansystem
- Explain Uniqueness and stability of equilibrium

11.0 INTRODUCTION

Dear learner, in the methodology of Economics, the concept of equilibrium occupies an important Place. The equilibrium means a state of balance. The equilibrium is said to exist when two opposing forces working on a particular object are balanced so that there is no tendency to move either in one direction or the other According to G Ackley, 'A system can be said to be in equilibrium when all of its significant variables show no change, and when there are no pressures or forces for change which will produce subsequent change in the values of the significant variables.' The equilibrium has two types (1) Partial equilibrium (2) General equilibrium. In this lesson, we shall be discussing the concept of equilibrium.

11.3 PARTIAL EQUILIBRIUM

Partial equilibrium is an equilibrium related to a part of an economic system.* In partial equilibrium analysis, the price of a commodity is determined, assuming the prices of the other commodities, tastes and preferences of consumers as constant. In other words, the assumption of ceteris Paribus is assumed.

The equilibrium of consumer, a particular firm, a factor of production is determined with the help of partial equilibrium approach. Marshall writes, "The forces to be dealt with are, however, so numerous that it is best to analyze a few at a time and to work out a number of partial solutions as auxiliaries to our main study. Thus we begin by isolating the primary relation of supply, demand and price in regard to a particular commodity. We reduce to inaction all other forces by the phrase 'other things being equal'*

** Ackley. G Micro Economic Theory 1969, pp 14-15*

** lipsey R.G. introduction to positive economic, 3rd edition, P.404*

In the Marshallian analysis of price determination under Perfect Competition, the demand function of a commodity is drawn on the assumption that prices of other commodities remain constant. In supply function, it is assumed that the prices of different factors of production and the production are constant. So, the price of a single commodity is determined by the intersection of demand and supply curves, prices of other goods and resources remaining constant. If there is some change in the data, new demand and supply functions will be reconstructed and corresponding to such a change, new price of the commodity will be determined.

In the partial equilibrium analysis, It is assumed that the changes in single sector do not induce changes in other sectors. This assumption is always violated to some extent for anything that happens in one sector must cause changes in other sectors. What matters is that the changes induced throughout the rest of the economy are sufficiently small and diffuse so that the effect they in turn have on the sector A (Sector in question) can be safely ignored.

11.4 GENERAL EQUILIBRIUM

The general equilibrium deals with not only a part but the entire organism of a given economic system. The partial equilibrium analysis assumed that the different economic units are independent of one another in such a way that a change in one unit has no effect on other economic units. The general equilibrium analysis, on the other hand, considers all the economic units as interdependent and capable of interacting with one another So, the general equilibrium analysis is a multi-market equilibrium. It explains the simultaneous determination of equilibrium prices and quantities of all goods and services after the initial disturbance gets fully adjusted in the whole Economy.

The general equilibrium deals with the over all effect of a disturbance on all the relevant variables which can be directly or indirectly affected by the disturbance. Leftwitch has explained the general equilibrium from the viewpoint of widespread effect of an initial disturbance in the system. According to Leftwitch, "First comes big splash from the disturbance. Particular equilibrium analysis handles the splash. But waves and then ripples are set up from it, affecting the area of splash. The ripples run further and further becoming smaller and smaller, until eventually they dwindle away. The fools of general equilibrium are required for the entire series of adjustments"*

** Lisey G.G introduction to positive economics, 3rd, p 404*

When we look at the economic system as a whole, there is a great deal of inter-relationship among various markets for goods and inputs and there are a large number of decision making agents like consumer, producer etc. All these agents are interested to maximize their goals, So, a comprehensive analysis of the economy which takes into account all interrelationships could be made only with the help of general equilibrium analysis.

11.4.1 Walrasian System

Walras[1] developed a Mathematical general equilibrium model in his work 'Elements of pure Economics' in 1874. He argued that all prices and quantities in all markets are determined simultaneously through their interaction with one another Walras used a system of simultaneous equations to describe the interaction of individual sellers and buyers in all markets , and he maintained that all the relevant magnitudes can be determined simultaneously by the solution of this system.

Walras explained the behaviour of individual decision makers by a set of equations. For example, an individual consumer has double role to play, he buys the goods and sells services of factors to the business firms. So, each consumer has two sets of equation, one showing his demands of the different goods, and the other describing his supplies of factors. The behaviour of the business firm is presented by two equations, one for the quantities of goods that it produces and the other for the demand of factors inputs. The important characteristic of these equations is their interdependence. The solution of simultaneous equations defines the 'unknown' of Walras' model. The prices and quantities of all commodities and all factors are the unknowns of this model.

In Walrasian general equilibrium model, there are as many markets as there are goods and factor inputs. In each market, there are three types of functions namely demand function, supply function and a clearing the market equation, which stipulates that the quantities demanded be equal to quantities supplied. In goods market, the number of demand equations is equal to the number of consumers and the number of supply equations is equal to the number of firms producing the commodity. In the factor market, the number of demand equations is equal to the number of firms

1. Walras l, elements & economic political pure (p 41874)

multiplied by number of commodities they produce. The number of supply functions is equal to the number of consumers who own the factor inputs.

According to Walras, for the existence of equilibrium, the number of independent equations must be equal to the number of unknowns in the system. This is a necessary condition for equilibrium. So, firstly the economy is described by means of a system of equations defining how may equations are required to find solution for the system.

For example, It is assumed that an economy has two consumers (A and B), two factors of production (K and L) and two commodities (Y and Z) are produced by two firms. It is further assumed that each firm produces one commodity, each consumer buys some units of both the commodities and both consumers own some units of both factors. In simple model, the following unknowns are there:

Quantities demanded of Y and Z by consumers	2×2 = 4
Quantities supplied of K and L by consumers	2×2 = 4
Quantities demanded of K and L by the firms	2×2 = 4
Quantities of Y and Z supplied by the firms	2
Prices of commodities Y and Z	2
Prices of factors K and L	2
Total number of ‘unknowns	**18**

To find these unknowns, we have 18 equations. These are :–

Demand functions of consumers	2×2 = 4
Supply functions of factors	2×2 = 4
Demand functions for factors	2×2 = 4
Supply functions of commodities	2
Clearing the market of commodities	2
Clearing the market of factors	2
Total number of equations	**18**

2. Koutsoyinanes, modern microeconomic, II edition, P486

As the number of equations is equal to number of unknowns, one must expect general equilibrium to take place. But this does not mean that general equilibrium exists. This is neither a sufficient nor a necessary condition for the existence of a solution.

In the Walrsian system, one equation is not independent of others (redundant equation). So, The number of independent equations is larger than the unknowns, which deprives the system of a solution. In this model, the absolute level of prices is not determined. General equilibrium theorists have thus adopted the device of choosing arbitrarily the price of one commodity as a numeraire (or unit of account) With this device prices are determined only as ratios, where each price is given relative to the price of numeraire. This indeterminacy can be eliminated by the introduction explicitly in the model of a money market in which money is not only the numeraire but also the medium of exchange and store of wealth.

Even if the number of equations is equal to unknowns, there is still no guarantee that a general equilibrium solution exists. Walras was not able to prove existence of general equilibrium. Arrow and Debru (1954) provided a proof of existence of equilibrium in perfectly competitive markets in which there are no indivisibilities and no increasing returns to scale. Further, in 1971, Arrow and Hahn explained the existence of general equilibrium with limited increasing return and monopolistic competition, without indivisibilities. But these proofs are limited to specific markets and on limited assumptions.

11.5 EXISTENCE, UNIQUENESS AND STABILITY OF AN EQUILIBRIUM

In case of general equilibrium, three problems arise. These are problems of existence, uniqueness and stability of equilibrium in the economic system.

- **Existence**

The first problem is whether a general equilibrium exists or not. The proofs of existence of general equilibrium have been provided by the economists in a perfectly competitive system where discontinuities and increasing returns to scale are absent. According to Koutsoyiannis, 'Perfect competition guarantees the existence of general equilibrium'. Equilibrium exists, when at a certain positive Price, Quantity demanded Q_D = Quantity supplied.

- **Uniqueness Problem**

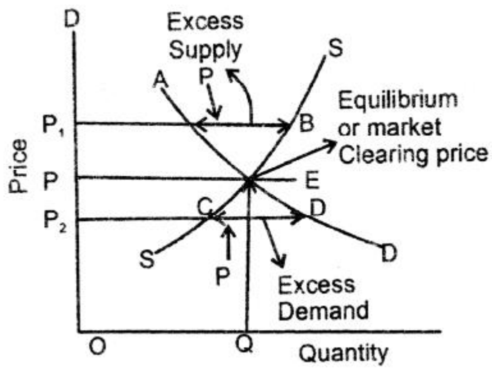
The problem of uniqueness is concerned with the issue that given the existence of generated equilibrium, is it unique ? The uniqueness of equilibrium is related to the slope of the excess demand function, the curve which shows the difference between quantity demanded and quantity supplied at one price.

- **Stability**

The stability problem is concerned with the question that if equilibrium exists is it stable or not? The problems of existence, uniqueness and stability can be discussed through the partial equilibrium analysis of a demand supply model. It is assumed that a commodity is sold in perfectly competitive market. We have consumer as the utility maximiser and firm as the profit maximiser. An equilibrium exists when the quantity supplied is equal to quantity demanded at a certain positive price. At equilibrium price, there is neither excess demand nor excess supply. Thus an equilibrium price can be defined as the price at which the excess demand is zero and market is cleared.

The stability of equilibrium depends upon the slopes of demand and supply functions. The equilibrium is stable if the demand function cuts the supply function from above. Here Prices rise in case of excess demand while fall in case of excess supply (Fig. 11.1). The equilibrium is unstable if the demand function intersects the supply function from below. Here excess demand leads to a price fall, while excess supply (negative excess demand) drives it up (Fig. 11.2).

The uniqueness of equilibrium is related to the slope of excess demand function. The case of unique and stable equilibrium can be explained through Fig. 11.1. The quantity demanded and supplied is measured on X-axis and price is measured on Y-axis. DD and SS are the demand and supply curves respectively. At point E, the demand curve cuts supply curve from above. So the unique and stable equilibrium exist. OQ is the equilibrium quantity and OP is the equilibrium price. At OP Price, there is AB excess supply. It puts the price to fall down to OP level. On the other hand at OP_2 price, there is CD excess demand. It puts the price back to OP level. Thus, system again gets back to the unique and stable equilibrium.



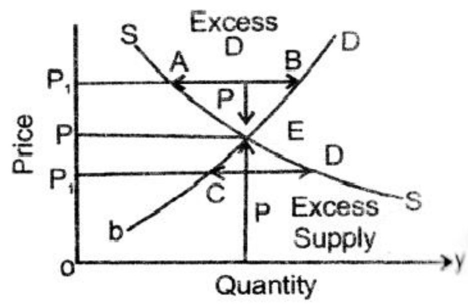


Fig 11.1 (Unique, Stable Equilibrium)

Fig. 11.2 (Unique, Unstable)

The stable equilibrium has been discussed in Fig. 11.3 where the excess demand is measured on the X-axis and price on the Y-axis. EE is the excess demand curve which slopes downward from left to right. The excess demand is zero at OP price level. Stable equilibrium exists at OP price.

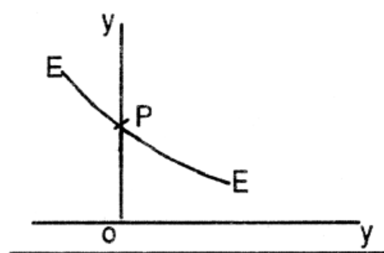


Fig 11.3 Stable equilibrium (slope of E(P) < 0)

The EE intersects Price axis when there is an equilibrium *i.e.*, when EE or E(P) is zero. For equilibrium to be stable, EE must have a negative slope at point of intersection.

Unstable equilibrium has been shown in fig. 11.4, where EE [or E(P)] intersects price axis when E(P) > 0. It is unique, but unstable.

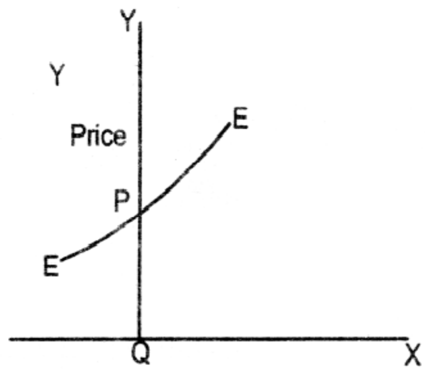


Fig. 11.4 : Unstable equilibrium (slope of E(P) > 0)

In Fig 11.5, we explain the case of multiple equilibria. The demand curve DD cuts the supply curve at two points. At point E, the equilibrium price is OP_1 and equilibrium quantity is OQ_1. At point E', there is stable equilibrium.

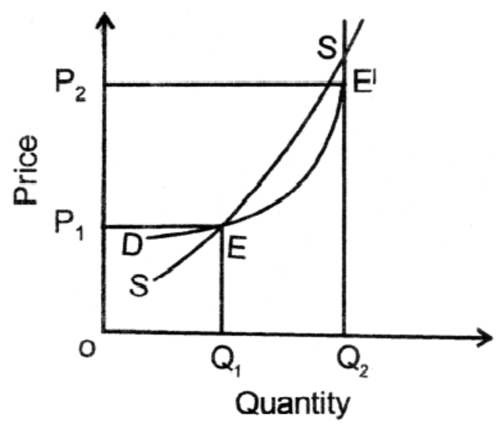


Fig 11.5

In Fig. 11.6, E is the excess demand function that cuts the vertical axis at P_1 and P_2 points. There is stable equilibrium at P_2 and unstable at P_1 price. As we know, the point of equilibrium exists as many times as the EE intersects the price axis.

Here EE cuts price line twice. Hence it is also a case of multiple equilibria.

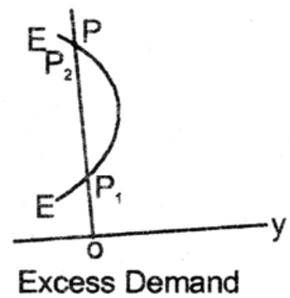


Fig 11.6 : Multiple equilibria

The situation in which no equilibrium exists is explained through Fig. 11.7 and 11.8. In Fig. 11.7, the equilibrium does not exist at a positive price. In Fig. 11.8, EE is the excess demand.

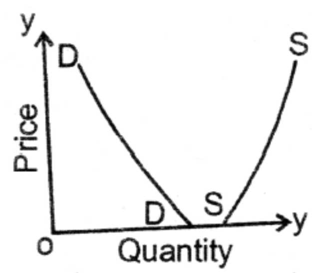


Fig. 11.7

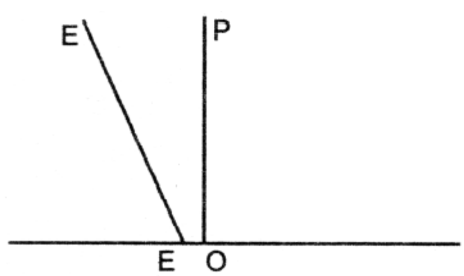


Fig. 11.8

Thus, If excess demand function does not cut the vertical axis at any one price, equilibrium does not exist.

3. Koutsoyianis Modern Micro Economics, I[I] edition, P. 490.

Check Your Progress-I

Answer the questions in the space provided

Q1. Distinguish between partial and general equilibrium.

Q2. What are the unknowns & equations in a Walrasian system?

Q3. What makes an equilibrium unique ? Show through excess demand function EE.

Q4. How do multiple equilibria exist?

Q5. Distinguish between stable and unstable equilibrium.

11.6 LET US SUM UP

Dear learner, in this lesson, we discussed the following:

- In Partial equilibrium analysis, the price of a commodity is determined, assuming the prices of the other commodities, tastes and preferences of consumers are constant.
- The general equilibrium analyses the multi-market equilibrium. It explains the simultaneous determination of equilibrium prices and quantities of all goods and services after the initial distribution gets fully adjusted in the whole economy.
- In Walrasian general equilibrium model, there are as many markets as there are goods and factor inputs. In each market, there are three types of functions namely demand function, supply function and a clearing equation.
- Three problems arise in general equilibrium. i.e existence, uniqueness and stability of equilibrium in the economic system.
- Equilibrium exists when demand and supply intersect at a positive price.

11.7 KEYWORDS

11.8 EXAMINATION ORIENTED QUESTIONS

1. Describe Walrasian general equilibrium Model.
2. Differentiate between Partial and General equilibrium.

3. Do you think general equilibrium analysis is an extension of partial equilibrium analysis? Discuss.

4. Discuss the existence, uniqueness and stability of equilibrium.

5. How do multiple equilibria exist?

6. Explain diagrammatically the non-existence of equilibrium.

7. Is the equality of no. of unknowns and equations in Walrasian system enough for equilibrium? Justify your answer.

11.9 HINTS TO CYP

Q1. See sections 11.3 and 11.4

Q2. See sub-section 11.4.1

Q3. See section 11.5

Q4. See section 11.5

Q5. See section 11.5

11.10 SUGGESTED READINGS

1. Blaug, M,(1976) Economic Theory in Retrospect, Irwin Hicks, H.R (1968), Value and Capital

2. Hansen, B.(1970) A survey of General Equilibrium, Aldine Press, New York. Simpson, D.(1975), General Equilibrium Analysis, Black well.

3. Stigler, G.J(1952), The theory of Price.

4. Walras Leon 1954, Elements of Pure Economics trans. jaffe. Allen & unwin.

M.A. ECONOMICS **LESSON NO. 12**

COURSE NO. ECO 201 **UNIT-II**

A GRAPHICAL ILLUSTRATION OF THE PATH TO GENERAL EQUILIBRIUM

STRUCTURE

12.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

- Graphically present the path to general equilibrium

- Discuss the impact of consumers' tastes on output
- State the impact of entry and exit on labour market

12.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to:

- Graphically present the path to general equilibrium
- Discuss the impact of consumers' tastes on output
- State the impact of entry and exit on labour market

12.2 INTRODUCTION

Dear learner, in the previous lesson, the concepts of partial and general equilibrium were introduced to you. In this lesson we shall explain how a simple economy with perfectly competitive production and factors markets will have a tendency towards a general equilibrium.

12.3 A GRAPHICAL ILLUSTRATION OF THE PATH TO GENERAL EQUILIBRIUM

12.3.1 Assumptions

Before discussing the analysis, we shall take the following assumptions:

1. Two substitute commodities (X and Y) are produced in two competitive industries
2. There are two factors, capital and labour. The quantities of these factors are fixed and markets are competitive.
3. Diminishing Marginal Rate of factor Substitution and decreasing returns to scale prevail in the production process.
4. The industry producing X commodity is less capital intensive than the industry producing Y commodity. In other words, K/L ratio in the former is less as compared to the latter.

5. The consumers want to maximize utility and producers' goal is to maximize profits.

6. Perfect competition is assumed in product and factor markets. HenceL the usual assumptions are made.

7. Initially the system is in equilibrium *i.e.* in all markets, demand is equal to supply at a positive price.

12.3.2 Taste of consumers and its impact on output

Assume that due to a change in external factor like tastes and preference of consumers, the demand curve for X commodity shifts rightwards from D_1 to D_2, This change causes a rise in the price of X commodity from P_1 to P_2, and quantity sold increases from OX_1 to OX_2. The equilibrium point shifts from point 1 to 2.

The two commodities X and Y are substitutes. So, the increase in the demand for X commodity is accompanied by a decrease in the demand for Y commodity. The demand curve of Y commodity shifts to the left and as a result its price and quantity sold decreases. The equilibrium point shifts from 1 to 2.

The figures 12.1 (a) and 12.1 (c) show this shift in equilibrium for the two goods X & Y respectively. Figures 12.(b) & (d) show how a typical firm in Industry X & Y behaves.

The increase in price of X creates economic profits for the producers of X and losses for the producers of Y. Thus the firms are interested in diverting resources from the production of Y to the production of X. This phenomena is shown in the fig 12.1 (e), where movement takes place from point A to B on the production possibility curve. This shift explains the effect of tastes and preferences of the consumers on the decision of firms to produce the particular commodities. The increase in the price of X commodity induces the producers to shift the resources towards X to maximize profits. The output of every firm that produces X commodity increases. The sum of the increases in the output of existing firms is equal to X_1X_2 in figure 12.1(a).

Likewise, the fall in the price of y induces the firms to reduce their output. The aggregate output of Y commodity declines by Y_1Y_2 as shown in figure 12.1(c).

In the long run, the supernormal profits in industry X attract some outside firms, while many firms of industry Y exit.

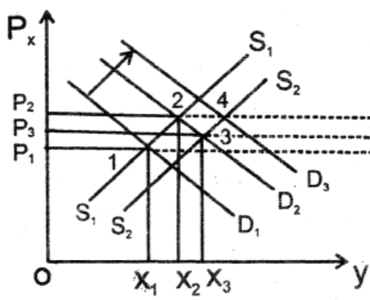


Fig. 12.1(a) : Industry X

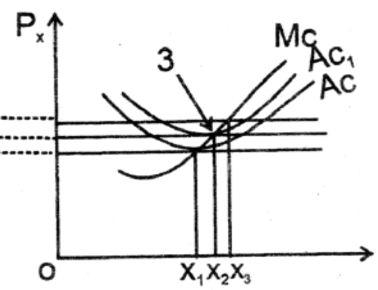


Fig. 12.1(b) : A firm in Industry X

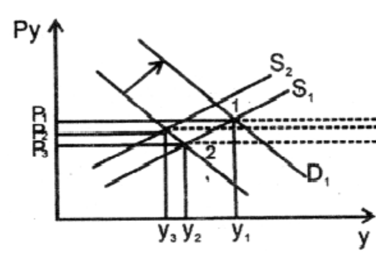


Fig. 12.1(c) : Industry Y

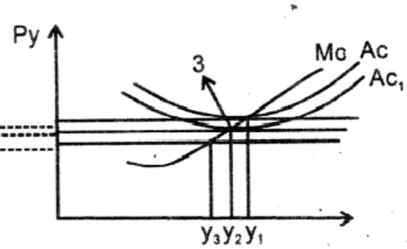


Fig. 12.1(d) : A Firm in Industry Y

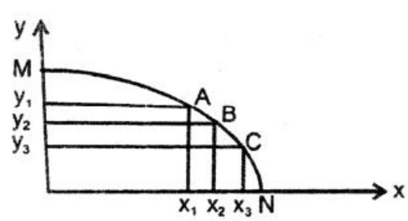


Fig. 12.1(e)

12.3.3 Impact of entry & exit on factor markets

This entry & exit affects the demand for factors of production. We shall now discuss how things change in factor market due to this. The factor markets are shown in the following figures :

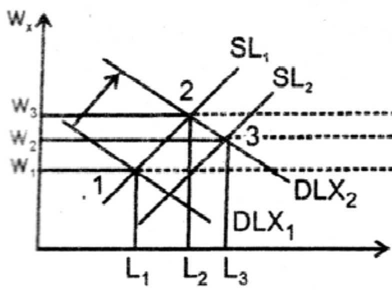


Fig. 12.2(a)
Labour Market for Industry X

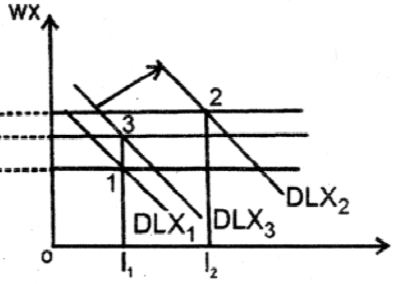


Fig. 12.2(b) Demand for labour by a firm in Industry x

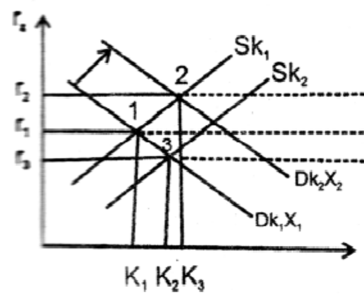


Fig. 12.2(c) : Capital Market for Industry X

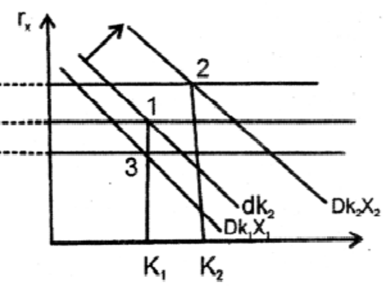


Fig. 12.2(d) : Demand for capital by a firm in Industry X

The increase of output in industry X increases the demand for labour and capital. The D_{LX_1} and D_{KX_1} show the demand curves for labour and capital respectively in the X industry. These curves shift outwards and wages and interest rise. The employment of these factors rises in the industry X. The demand for labour and capital by a single firm in industry X is shown in fig 12.2(b) & (d) respectively. The situation is industry Y is just the reverse. Here, demand curves for both the factors see a decline. The decline in amount of labour and capital employed is by L_1L_2 & K_1K_2 respectively. This is shown in figure 12.3(a) & (c) respectively.

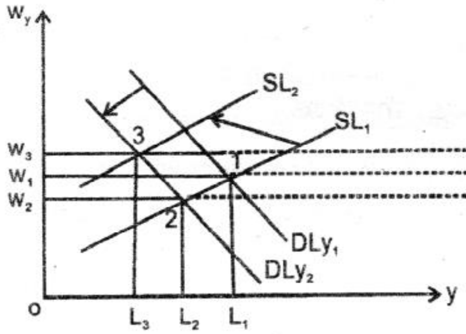


Fig. 12.3(a) : Labour Market for Industry y

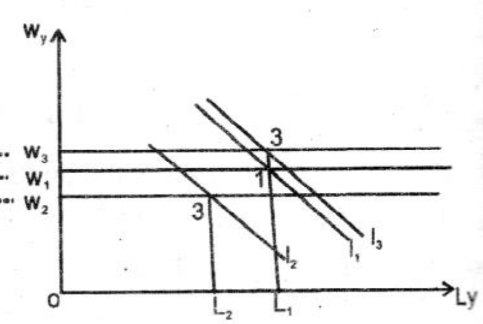


Fig. 12.3(b) : Demand for labour by a firm in Industry y

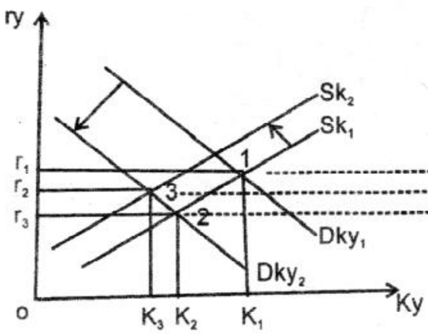


Fig. 12.3(c) : Capital Market for Industry y

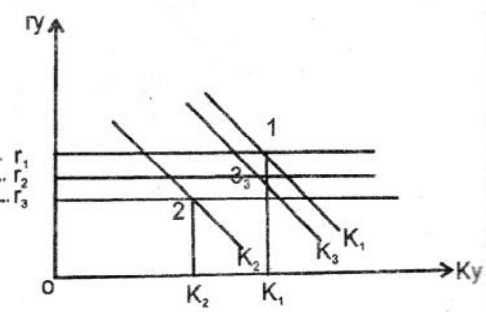


Fig. 12.3(d) : Demand for Capital by a firm in Industry y

Now, the prices of labour and capital have risen in the factor market of X industry and fallen in the factor market of Y industry. But in perfectly competitive markets, the disequilibrium is self correcting because in the long run there is perfect mobility of factors between the two markets. So, the owners of factors withdraw their services from Y industry and increase their supply in industry X, where prices of factor of production are higher as they are profit Maximisers. As a result of this mobility of factors, the supply curves of factors in industry X shift upward and the supply curves of factors in industry Y shift downward. The shifts are shown in figures 12.3 (a) & (c) for factor L&K, respectively.

But because of the assumption that X is less capital intensive than Y, the wages are higher in the new equilibrium and the price of capital (r) is lower in the new equilibrium. The demand for labour in X industry is stronger than the demand for capital. The release of labour by industry Y is smaller than the change in labour required by industry X. On the other hand, the release in capital by industry Y is larger than the need of capital by industry X. So, the overall demand for labour is higher than the initial situation and the wage level increases. The demand for capital is lower than the initial situation, so the price of capital decreases. As a result of new prices of factor, the individual firms adjust their demand for labour and capital. Finally, the firms are in equilibrium at the point 3, as shown in different figures.

Due to the entry of new firms, the supply curve of X industry shifts downward from S_1 to S_2, as shown in figure 12.1(a). The new equilibrium price of commodity X is lower than in short run, but higher than the original price. On the other hand, the exit of the firms from industry Y shifts the supply curve of industry Y upwards from S_1 to S_2, as shown in figure 12.1(c). The new equilibrium price of Y commodity is higher than the short run price but lower than the original equilibrium price. Both the industries are thus increasing cost industries.

The individual firms also adjust their output according to new prices which is shown in figures 12.1 b & d. The firms in industry X produce the output lower than the short run output but higher than the initial out put. The firms in industry Y produce the output higher than the short run but lower than initial output. The change in the prices of factors causes the upward shift in LAC curve of the firm in industry X and the downward shift in LAC curve of the firms in industry Y. So, In the long run, the firms earn just normal profits.

The above explained path towards a general equilibrium would be modified if apart from external changes, we take into account the possible induced changes caused by the adjustments in the system. We have seen that in the new equilibrium wages were higher and the price of capital is lower than initial level. If the propensity to consume of wage earners is higher than the other factor owners, then there will be further increase in the demand for X. The demand curve D_3 in figure 12.1 (a) has been drawn for this purpose. From this point we can trace the additional changes in the all other markets, starting with this change.

Check Your Progress-I

Q1. What are the main assumptions taken while discussing the path to general equilibrium ?

Q2. How do exogenous factors cause a shift in equilibrium points ?

Q3. How does the capital intensity of an industry affect the factor Prices ?

12.4 LET US SUM UP

Dear learner, in this lesson, we discussed the following :

We started from an initial equilibrium which was disturbed by an external change. We explained the interactions of markets and individual decision making in perfect product and factor markets. The reactions of firms and individuals lead to new equilibrium. It should be noted that this effect is certain only in perfectly competitive markets and continuous production function with decreasing returns to scale

12.5 KEYWORDS

Exogenous factors: The factors which originate from outside the model or system and influence the system.

Path to equilibrium : The path covered by a system, in particular the different markets like factor market, product market etc. while moving from an earlier equilibrium position to a new one, thus reaching a stage of stability.

12.6 EXAMINATION ORIENTED QUESTIONS

1. What are the assumptions of Path to general equilibrium ?
2. Discuss the Path to general equilibrium with an example.
3. Discuss how do the factor markets get affected in the path to equilibrium?
4. Discuss the dynamic process of reaching general equilibrium.

12.7 HINTS TO CYP

CYP-I

Q1. See sub-section 12.3.1

Q2. See sub-section 12.3.2

Q3. See sub-section 12.3.3

12.8 SUGGESTED READINGS

1. Koutsoyainnis, A. (1979), Modern Micro Economics, Macmillan Press LTD.
2. Weintraub, E.R (1974), General Equilibrium theory, Macmillan studies in Economics, Macmillan.

M.A. ECONOMICS LESSON NO. 13

COURSE NO. ECO 201 UNIT-II

GRAPHICAL TREATMENT OF THE TWO FACTORS, TWO COMMODITIES, TWO CONSUMERS EQUILIBRIUM

STRUCTURE

13.0 OBJECTIVES

The main objectives of this lesson are to enable you to explain the two Factors, two Commodities, two Consumers Equilibrium and to describe Graphical Treatment of the Two Factors, Two Commodities, Two Consumers General Equilibrium system.

13.1 LEARNING OUTCOMES

After going through this lesson, you should able to discuss the two Factors, two Commodities, two Consumers Equilibrium. You shall also be able to describe Graphical Treatment of the Two Factors, Two Commodities, Two Consumers General Equilibrium system.

13.2 INTRODUCTION

Dear learner, in this lesson, we shall discuss a Graphical Treatment of the Two Factors, Two Commodities, Two Consumers Equilibrium that consists of two factors of Production, two commodities and two consumers and is also known as 2×2×2 model. This lesson also discusses the three static properties of a general equilibrium which includes the efficient allocation of resources among firms, the efficient distribution of commodities between two consumers and the efficient combination of products and these properties are also called the conditions of Pareto Optimality.

13.3 A GRAPHICAL TREATMENT OF THE TWO FACTORS, TWO COMMODITIES, TWO CONSUMERS EQUILIBRIUM

A general equilibrium consisting of two factors of production, two commodities and two consumers is known as 2×2×2 model of general equilibrium. A rigorous treatment of the 2×2×2 model was given by H.F. Johnson in 1971.

13.3.1 Assumptions

In a perfectly competitive static market system, the 2×2×2 model of general equilibrium rests upon the following assumptions :-

1. There are two factors, labour and capital. These factors are identical and perfectly divisible. The quantities of these factors are exogenously given.
2. Only two commodities, X and Y are produced. The technology is given. The production functions of these two commodities are independent and represented by two maps of isoquants. These isoquants are convex to the origin, implying, Diminishing Marginal Rate of Factor Substitution along an isoquant. Each production function exhibits constant returns to

scale. Further, there are no externalities in the production of either commodity on each other. In other words, there is absence of external economies and diseconomies.

3. There are two consumers, A and B in the economy. The preferences of consumers are represented by the ordinal indifference curves which are convex to the origin, implying Diminishing Marginal Rate of Substitution between the two commodities. It is further assumed that the choices of two consumers are independent i.e. it is assumed that the externalities in consumption like Bandwagon, Snob, Veblen & other effects do not exist. Also, consumers' choices are not influenced by advertising or other activities of the firms

4. The goal of each consumer is to maximize his own satisfaction subject to his income constraint.

5. The goal of each firm is to maximize profits, subject to the technological constraint.

6. The productive factors are owned by the consumers.

7. All factors are fully employed and all incomes received by their owners are spent.

8. There are conditions of perfect competition in the commodity and factor markets.

9. Consumers and firms pursue their goals given the same set of prices (Px, Py, w, r)

In this model, general equilibrium is reached, when the commodity and factor markets are cleared at a set of equilibrium price. At the same time, each one of the two firms and two consumers are simultaneously in equilibrium.

The general equilibrium solution will require the determination of the values of the following variables:

(i) The Total quantities of the two commodities, X and Y, which are produced by firms and purchased by the consumers.

(ii) The allocation of the given quantities of capital and labour to the production of each commodity.

(iii) The quantities of two commodities that will be bought by the two consumers.

(iv) The price of the commodities and those of the factors of production.

(v) The distribution of the ownership of factors of production (KA, KB, LA, LB) between the two consumers subject to their budget constraint.

13.3.2 Static Properties of a General equilibrium

There are three static properties in a general equilibrium solution.

(i) The efficient allocation of resources among firms.

(ii) The efficient distribution of commodities between two consumers.

(iii) The efficient combination of products.

There properties are also called marginal conditions of Pareto optimality. According to Pareto, a situation is efficient it is impossible to make anyone better off without making someone worse-off.

I. The Efficient Allocation of Resources (Efficiency of production)

The equilibrium of production can be attained, it the available resources are distributed among various firms in efficient way. The firm is in equilibrium if it chooses the factor combination that minimizes cost. So the optimal combination can be determined when the $MRTS_{LK}$ is equal to the ratio of prices of the factor or the slope of isoquant is equal to the slope of iso cost line.

$MRTS_{LK} = w/r$

Slope of Isoquant = slope of isocost line

This equilibrium situation can be shown through the Edgeworth box (Fig 13.1) In the figure, the units of labour and capital are measured on the horizontal and vertical axes respectively. X_1, X_2, X_3 and X_4 are the isoquants related with X commodity with origin at O. Y_1, Y_2, Y_3 and Y_4 are the isoquants related with commodity Y with origin O_1.

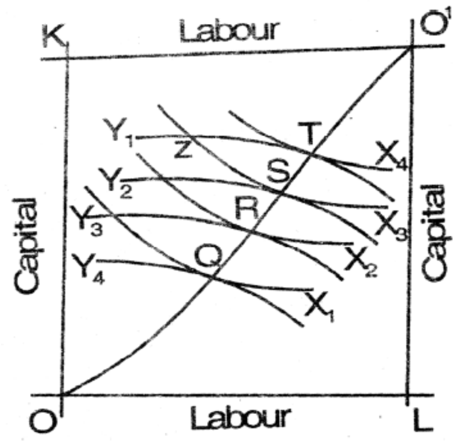


Fig. 13.1

These isoquants are tangent at Q, R, S and T points. If we join all these tangency point, oo^1 contract curve is drawn. All the points of tangency indicate the efficient allocation of resources (L and K) between the firms.

Each point of the Edge Worth box have a specific allocation of labour and capital in the production of X and Y commodity. But each point of the Edge Worth box does not represent efficient allocation of resources. For example, Point Z is not a point of efficient production. A reallocation of resources from point Z towards S or T or between S & T leads to a greater production of at least one of the commodities. At the tangency points of X and Y, the slopes of isoquants are equal.

Slope of X isoquant = slope of Y isoquant

$$MRTS^{X}_{LK} = MRTS^{Y}_{LK}$$

Since the firms in simple general equilibrium model are profit maximizers in competitive markets, the equilibrium takes place when $MRTS_{LK}$ in both the commodities X and Y become equal to the ratio of prices of the two factors.

$$MRTS^{X}_{LK} = MRTS^{Y}_{LK} = w/r$$

If the prices of two factors are given, we can determine the amounts of X and Y commodities which maximize the profits of the firms by driving the production possibility curve from Edge work contract curve . The different points on contract curve indicate such amounts of X and Y commodities that ensure efficient allocation of resources. The respective amounts of Q, R, S and T are shown in figure 13.2 through the points Q', R', S' and T'. In fig. 13.2, the points Q', R', S' & T' correspond with the points P, Q, R and S respectively on the contract curve shown in figure 13.1. The production of X & Y corresponding to points Q', R' & S' are X_1, X_2 & X_3 and Y_3, Y_2 & Y_1 respectively.

By joining these points, we can draw the production possible curve. It shows the maximum possible quantities of the two commodities that can be produced with the given resources.

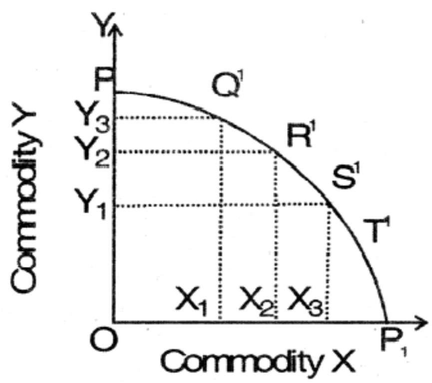


Figure 13.2

The production possibility curve is also known as the product transformation curve. The slope of this curve is measured by the marginal rate of product transformation between X and Y. It shows the quantity of Y that must be sacrificed in order to obtain an additional unit of X.

$$MRPT_{X,Y} = - dy/dx$$

Since dy/dx is negative, the $MRPT_{X,Y}$ is a positive number . The MRPT is equal to the ratio of marginal cost of X and Y. The profit maximizing firm under the conditions of perfect competition equate the price of the commodity to the Marginal Cost of X and Y. So, the slope of production possibility curve is measured through the ratio of price of two commodities

$$MRPT_{X,Y} = MC_X / MC_Y = P_X / P_Y$$

Given the price of two commodities, the general equilibrium of a maximizing firm is determined at the point of production possibility curve that has a slope equal to the ratio of price. It is shown in fig 13.3. In the figure, PP^1 is the production possibility curve. AB is the price line, the slope of price line is measured by the ratio of prices. The general equilibrium of production is determined at point E, where MRPT $= MC_X / MC_Y$. The two firms are in equilibrium producing the levels of output X_1 and Y_1. The equilibrium levels of output ensure maximum profit and efficient allocation of capital and labour.

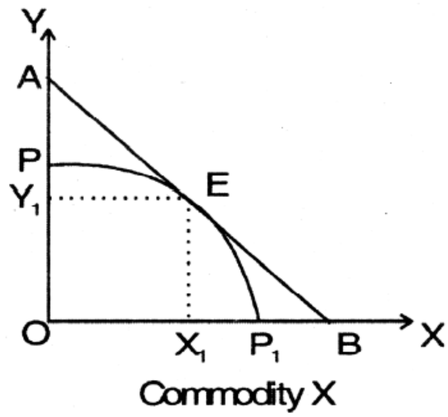


Figure 13.3

(ii) The Efficient Distribution of Commodities between Two Consumers (Equilibrium of consumption)

The general equilibrium requires that the produced commodities should be distributed between the two consumers A and B in efficient manner. Given the prices of two commodities, consumers' goal is at maximizing utility. Each consumer maximises his utility by equating the Marginal Rate of Substitution of the two commodities to the ratio of prices of the two commodities

$$MRS_{XY} = P_X / Py$$

In a perfectly competitive market, both consumers are faced with the same prices, the condition for general equilibrium of both consumers is

$$MRS^{A}_{X,Y} = MRS^{B}_{XY} = P_X / P_Y$$

MRS $_{X\ Y}$ of each consumer is measured by the slope of his indifference curve. The slope of price line represents the ratio of prices of X and Y commodities. The efficient distribution of two commodities between two consumers is explained through figure 13.4.

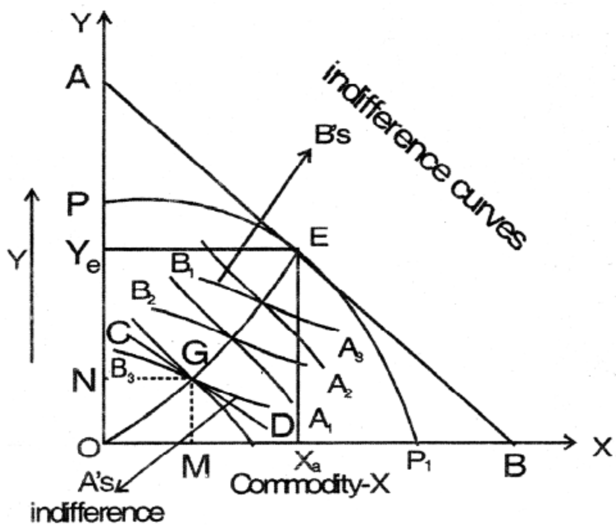


Figure 13.4

In the figure, commodity X and commodity Y are measured along the horizontal and vertical axes respectively. Given the prices, AB is the price line. PP_1 is the production possibility curve. The equilibrium of the firm is determined at E where the production possibility curve is tangent to the price line. Now we construct an Edge worth box for consumption with the precise dimensions of OX_e and OY_e. Within this box, the indifference maps of the two consumers are given. A_1, A_2 and A_3 are the indifference curves of consumer A with O as origin. B_1 , B_2 and B_3 are the indifference curves of consumer B with E as origin. The tangency points of the indifference curves of two consumers indicate that MRS $^{A}_{X,Y}$ =MRS $^{B}_{X,Y}$ at these points. The tangency points of the indifference curves of two consumers represent Pareto efficient distribution. Thus there is an infinite number of possible Pareto efficient equilibrium points for a given product mix. However, under the conditions of Perfect Competition, only one point on the contract curve is consistent with the general equilibrium system. The equilibrium of the two consumers is determined at point G where the MRS_{XY} of two consumers is equal to the price ratio of two commodities. At point G, CD, is the common price line for both consumers and it is parallel to AB. Thus, G is the point of general equilibrium of consumption Where $MRS^A = MRS^B = P_X/P_Y$. The consumer A achieves maximum satisfaction by buying OM of X and ON of Y. The Consumer B achieves maximum satisfaction by buying MX_e of X and NY_e of Y commodity.

(iii) The efficient combination of products or simultaneous equilibrium of production and consumption (Efficiency in Product Mix).

The general equilibrium of the system as a whole requires the simultaneous equilibrium of both consumption and production. The equilibrium is attained where the Marginal Rate of Product Transformation is equal to the Marginal Rate of Substitution of the two commodities between two consumers

$$MRST_{XY} = MRS^A_{XY} = MRS^B_{XY}$$

In a perfectly competitive market, the equilibrium of production takes place when the following condition is satisfied:

$$MRPT_{X,Y} = P_X/P_Y$$

The equilibrium of consumption requires the following condition

$$MRS^A = MRS^B = P_X/P_Y$$

When the above condition is fulfilled only then the plan of producer is consistent with the plan of household sector and the two sectors are in a state of equilibrium. The simultaneous equilibrium of production and consumption is shown through Figure 13.5.

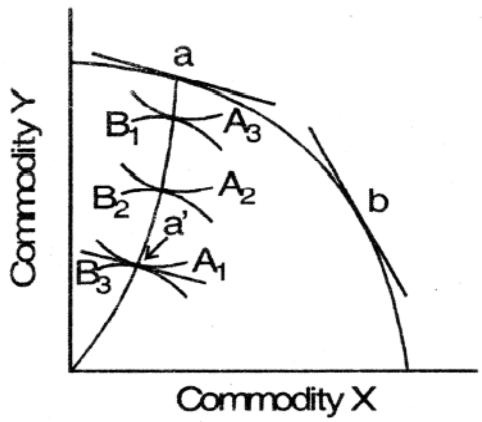


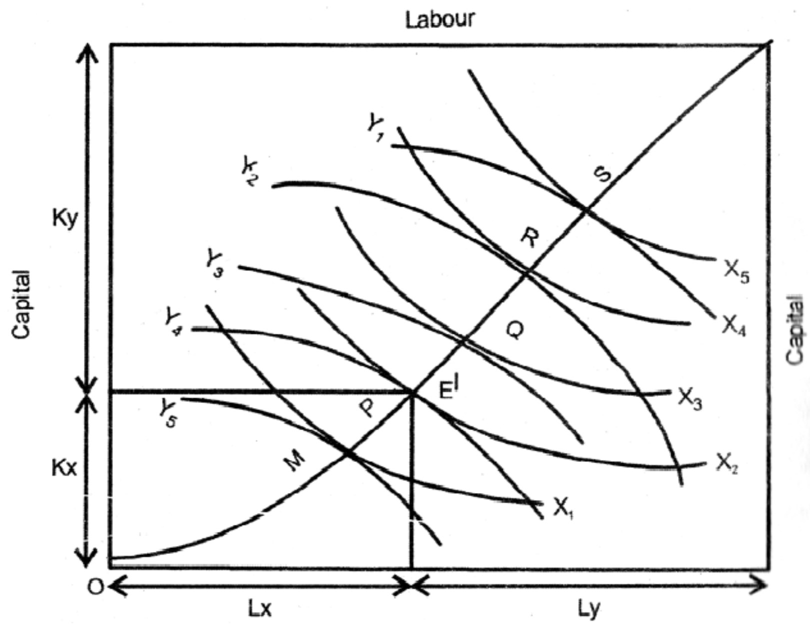
Figure 13.5

As we can observe from the figure, MRSxy = MRPTxy at point a' on the Contract Curve. This condition is also called as the efficiencyof productive substitution or optimal composition of output.

Thus, in a two factor, two commodity and two consumer model, the general equilibrium is determined, when there is optimal allocation of goods among the consumers, optimal allocation of factors among the firms and optimal composition of production. The simultaneous determination of both consumption and production under the perfectly competitive conditions signifies general equilibrium of whole system.

13.3.3 General Equilibrium and the Allocation of Resources

'In this section we shall examine the allocation of resources i.e, labour and capital between X and Y commodities in general equilibrium analysis. It is assumed that the quantities of the two factors are fixed and all units of labour are homogeneous. The allocation of fixed resources has been explained through figure 13.6. In the figure, the quantity of labour in measured on X–axis and quantity of capital is measured on the Y-axis. Earlier, we have discussed and drawn fig. 13.4. Inside the Edgework Box diagram, there are six unknowns. These are Xe, Ye, distribution among the two consumers (Xe^A, Xe^B, Ye^A, Ye^B). Here we shall discuss how the resource allocation takes place. Point E' on the contract curve corresponds to point E on the PPC (Fig. 13.4). Here resource allocation shows that Lx, Ly and Kx and Ky are the labour and capital used for X & Y respectively.



A further discussion can also lead to the knowledge of more unknowns of the model.

Check Your Progress

Answer the questions in the space given below :

Q.1. What are the assumptions of a graphical treatment of 2×2×2 model of general equilibrium?

____

____

____

____

Q.2. State the static properties of general equilibrium.

____

____

____

____

Q.3. What do you mean by equilibrium of consumption ?

____

____

____

____

13.4 LET US SUM UP

Dear learner, in this lesson, we discussed :

- A general equilibrium consisting of two factors of Production, two commodities and two consumers is known as 2×2×2 model of general equilibrium.
- There are three static properties of a general equilibrium solution.
- The efficient allocation of resources among firms.
- The efficient distribution of commodities between two consumers.
- The efficient combination of products.
- The production possibility curve is also known as the product transformation curve. The slope of this curve is measured by the Marginal Rate of Production Transformation.
- The general equilibrium of the system as a whole requires the simultaneous equilibrium of both consumption and production. The equilibrium is attained where the Marginal Rate of Production Transformation is equal to the Marginal Rate of Substitution of the two commodities between two consumers.

13.5 KEYWORDS

13.6 EXAMINATION ORIENTED QUESTIONS

1. What are the assumptions of a 2 - 2 - 2 model of General Equilibrium?
2. Discuss the equilibrium of Consumption.
3. How is simultaneous equilibrium of consumption and production determined ? Explain graphically.
4. Explain the Static Properties of the 2 - 2 - 2 model of General Equilibrium.
5. Discuss with the help of Edgeworth box diagram how efficient allocation of resources between firms takes place.
6. What do you mean by optimal composition of output? Explain diagrammatically.

13.7 HINTS TO CYP

Q.1. Refer to sub-section 13.3.1.

Q.2. Refer to sub-section 13.3.2.

Q.3. Refer to sub-section 13.3.2.

13.8 SUGGESTED READINGS

- Koutsoyiannis, A., Modern Microeconomics, 2nd edition (1997).
- Blaug, M, (976) Economic Theory in Retrospect, Irwin.
- Johnson, H.G. (1971) Two Sector Model of General Equilibrium, Aldine Press, New York.
- Stiglr, A.J (1952, The theory of Price.
- Walras Leon, (1954) Elements of pure economics, trans. Jaffe, Allen & Unwin.

M.A. ECONOMICS LESSON NO. 14

COURSE NO. ECO 201 UNIT-III

MARGINALPRODUCTIVITY THEORY OF DISTRIBUTION, PRODUCT EXHAUSTION THEOREM AND PARETO OPTIMALITY CONDITIONS.

STRUCTURE

14.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

Marginal Productivity Theory of Distribution

Product exhaustion theory

14.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to:

- Define and explain the **Marginal Productivity Theory**.
- Differentiate between various interpretations of the theory (Clark's, Marshall–Hicks' versions).
- Critically assess the assumptions and limitations of the Marginal Productivity Theory.
- Define the **Product Exhaustion Problem** and explain its significance in income distribution.
- Illustrate how the **Cobb-Douglas Production Function** addresses the Product Exhaustion Problem.
- Describe the contributions of **Wicksteed, Wicksell, Walras, Barone, Samuelson**, and **Hicks** in resolving the problem.
- Apply theoretical insights to analyze how total output is distributed among the factors of production in competitive markets.

14.2 INTRODUCTION

A theory which tries to determine the prices of the factors of production and which has been fairly held by professional economists is known as Marginal Productivity Theory of distribution. The essence of the theory is that price of a factor of production depends upon its Marginal Productivity. But the total product will be exhausted if all factors are paid reward equal to their marginal product. This version has been called a Product Exhaustion Theorem. In this lesson we shall be discussing these topics in detail.

14.3 MARGINAL PRODUCTIVITY THEORY OF DISTRIBUTION

We now turn to the question as to what determines the prices of factors of production. A theory which tries to answer this question is known as Marginal Productivity theory of distribution. It may however be pointed out that in recent years, its popularity has somewhat declined due to bitter criticisms leveled against it. The essence of this theory is that the price of a factor of production depends upon its Marginal Productivity. It also-seems to be very fair that a factor of production should get its reward according to the contribution it makes to the total output, i.e Marginal Productivity. Marginal productivity theory was first put forward to explain the determination of wages i.e reward for labour but subsequently prices of other factors of production such as land, capital etc. also were explained with Marginal Productivity.

The origin of the concept of Marginal Productivity can be traced to Ricardo and West. But both Ricardo and West applied the Marginal Productivity doctrine only to land. The concept of Marginal Productivity is implicit in the Ricardian Theory of Rent. But the idea of Marginal Productivity did not gain much popularity till the last quarter of 19th century, when it was re-discovered and popularised by "economists like J.B. Clark, Jevons, Wicksteed, Walras and later Marshall and J.R. Hicks. It may however be pointed out that there are many versions of Marginal Productivity theory. Theories of Marginal Productivity propounded by various economists differ from each other in some respects. There are in fact so many versions of marginal productivity theory that Joseph Schumpeter has gone so far as to remark that there are almost as many Marginal Productivity theories as there are economists[1]. We shall explain here the views or versions of J. B. Clark, Marshall and Hicks, as these will provide all the points at issue concerning marginal productivity theory. Since marginal productivity theory has been mainly evolved for the determination of reward for labour, we shall discuss below its application to wage determination. But it should be understood to apply equally to the rewards of other factors of production.

14.4 MARGINAL PRODUCTIVITY THEORY: CLARK'S VERSION

J. B. Clark, an American economist, developed marginal productivity theory of distribution in a number of articles and later on presented it in complete form as an explanation for the distribution of wealth in a country in his book "the distribution of wealth". In order to bring out the fundamental factors at work in the mechanics of income distribution, Clark assumed a completely static society, free from the disturbances caused by economic growth or change. In other words, he assumed constant population, a constant amount of capital and unchanging techniques of production. "By isolating of wages and other productive factors towards which all actual values in the real dynamic world around us tend at any moment of time"[2]. Besides the assumption of static economy, he has also assumed perfect competition in the factor market and prefect mobility on the part of both labour and capital.

Besides assuming that total stock of capital remains constant, Clark also supposes that the form of capital can be varied at will. In other words, physical instruments of production can be adapted to varying quantities and abilities of available labour. Further, he treats labour as a homogeneous factor by taking identical labour units and discusses how the wage rate of labour is determined.

Every rational employer or entrepreneur will try to utilize his existing amount of capital to maximize his profits. For this he will hire as many labourers as can be profitably put to work with the given amount of capital. For an individual firm or industry, marginal productivity of labour will decline as more and more workers are added to the fixed quantity of capital. He will go on hiring more and more labour units as long as the addition made to the total product by a marginal labour unit is greater than the wage rate he has to pay for it. The employer will reach equilibrium position when-the wage rate is just equal to the marginal product of labour.

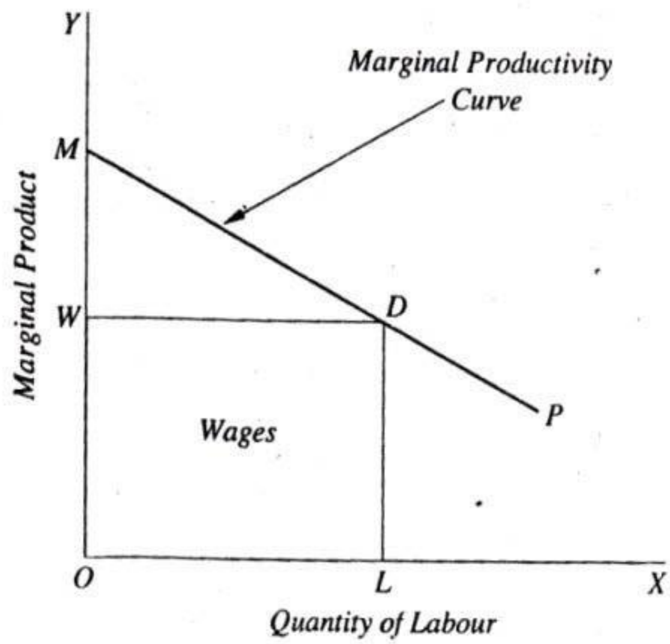


Fig 14.1

Just consider the adjoining Fig.14.1 where units of labour are represented on the X-axis and the marginal product of labour on the Y-axis. The MP curve shows the variation in marginal product of labour as more units of it are employed. If the prevailing wage rate which an employer must pay is equal to OW, then it will be profitable for the employer to go on employing additional workers until the marginal product of labour becomes equal to the prevailing wage rate OW. It will be evident from Fig. that if the prevailing wage rate is OW, then the employer will employ OL units of labour, since the marginal productivity of labour is equal to OW at OL employment of labour. He would not employ more than OL amount of labour as the marginal product of labour will fall below the wage rate OW and he would therefore be incurring losses on the employment of marginal workers beyond OL. Thus the employer would be maximizing his profits by equalizing the marginal product with the wage rate OW. Since perfect competition is assumed to be prevailing in the labour market, an individual firm or industry will have got no control over the wage rate. An individual firm or industry has therefore to determine only the number of factor units (labour in the present case) to which it has-to give employment at the prevailing (existing) wage rate. Thus, at micro level for (individual firm or industry) marginal productivity theory is the theory of employment.

A marginal product schedule or curve shows a particular wage-employment relationship. Since Clark has assumed a stationary state, he takes the total supply of labour available for employment in the whole economy as given and constant. In other words, in Clarkian analysis, aggregate supply curve of labour has been assumed to be perfectly inelastic. Given the total supply of labour in the economy, the wage rate will be determined by the marginal product of the available amount of labour assuming that, all labour units get employment. Given the aggregate amount of labour that is seeking employment, the wage rate that the labourers will secure will be equal to the addition made to the total product by the employment of the marginal unit of labour. In other words, if the total quantity of labour seeking employment is ‘n’ units, then each unit of labour will get wage which will be equal to the difference between total production when ‘n’ labour units were employed and that when n–1 labour units were employed. In other words, in the competitive labour market, the wage rate will be determined by the Marginal Product of a given quantity of labour force.

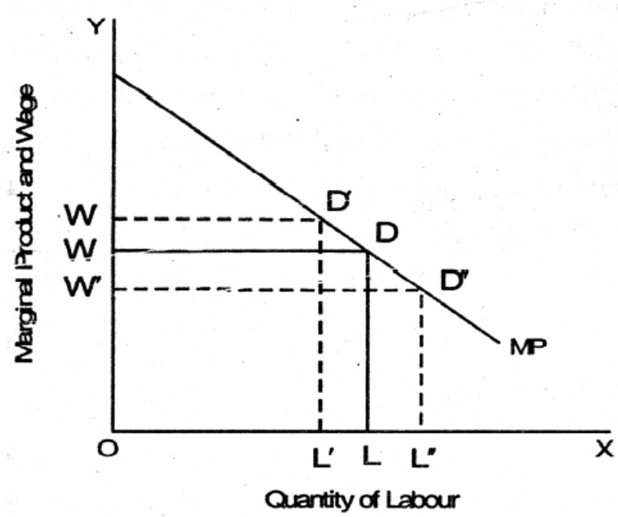


Fig. 14.2. Wage rate is determined by Marginal Productivity

If the labourers compete with each other for obtaining jobs, they will bid the wage rate down if some of them find themselves unemployed. The employers will bid the wage rate up if the prevailing wage rate is smaller than the marginal product the employers demand for labour force. This is so because at the wage rate lower than marginal product, the employers' demand for labour force will be more than the available number of labourers. Consider Fig 14.2. In this figure, MP curve represents diminishing marginal product of labour as some units of labour are employed in the economy, assuming the quantities of other factors used as unchanged. Now if the available quantity of labour force is OL in the whole economy, the marginal product of OL quantity of labour is LD. The wage rate will be determined by this marginal product LD and, therefore, equilibrium wage rate which will settle in the market will be equal to LD or OW. At a higher wage rate L'D'(=OW'), the employer will employ OL' amount of labour leaving LL' amount of labour unemployed. Unemployed workers in their attempt to get employment will bring the wage rate down to the level LD at which all are employed.

On the other hand, at a lower wage rate than LD, say L"D", the employers will demand OL" amount of labour since their profits will be maximum if they are employing OL" amount of labour at the wage rate L"D" (=OW"), but the available amount of labour is OL. Thus, at lower wage rate than LD(=OW) the demand for labour by the employers is just the same amount of labour which is actually available.

If the labour force actually available in the whole economy is OL", then the marginal productivity of labour force will be L"D" and therefore the wage rate will be equal to OW". If the actual quantity of labour force is OL' then the wage rate will be equal to OW'. Thus, given the quantity of labour in the economy wage rate is determined by marginal productivity of labour. One assumption which is implicit in the Clarkian marginal productivity theory and is applied to the economy as a whole is that of full employment. In other words, it is assumed that the existing amount of labour in the economy is fully employed. To sum up "in Clark's presentation, the marginal productivity of a given quantity of available labour determines its wage level when we consider the market as a whole. In the disaggregated picture, however, where a single employer finds the wage level determined by the forces beyond his control, the marginal product of labour determines the level of employment.

For analytical purposes, Clark's Marginal Productivity theory of distribution may be divided into the following three components :-

1. There is the premise that a rational employer will be guided by the marginal productivity of the factor in determining the number of units of that factor he has to employ. This premise has been called by Prof. Carter as the marginal productivity principle. (This is essentially based on the twin assumptions that the law of diminishing returns is working and that the employers are rational).

2. Secondly, there is the assumption of perfect competition so that market forces tend to equalize rates of return for all units of a factor.

3. Thirdly, there is the premise of long-run general equilibrium in all markets.

Given the above three assumptions, it can be said that when there is a given fixed supply of labour in the market, the level of wages will be determined by the marginal product of labour.

14.5. MARGINAL PRODUCTIVITY THEORY: MARSHALL—HICKS' VERSIONS

Alfred Marshall, who was contemporary of J.B. Clark, gave a different version of the marginal productivity theory. Marshall's version has also been called by many as the Marginal Productivity theory. Marshall, differed with those like Clark who held that wage rate (or for that matter, the price of any other factor) is determined by the marginal product of labour. Marshall said that it was wrong to regard the marginal productivity concept in the matter of wage determination as a wage theory. This is because he believed that wage rate (or any factor price) is determined by both demand for and supply of labour. Marginal productivity concept explains only the demand side of the problem. That is, given the wage rate, a national employer will employ as many units of labour as will equalize the wage rate with the marginal product of labour. At different wage rates, the employer will employ different amounts of labour units, depending upon the corresponding amount of the value of the marginal product. Thus, according to Marshall, the relationship between the wage rate and the marginal productivity of labour provides us with the supply curve of labour. For a complete theory as explanation for wage determination, the supply curve of labour has also to be introduced into the analysis. The wage rate at which the supply curve of labour cuts the demand curve of labour (governed by the marginal productivity) will be determined. No doubt, the wage rate thus determined by demand and supply will be equal to the marginal product. It is worth quoting Marshall himself on this point: "This doctrine has sometimes been put forward as a theory of wages. But there is no valid ground for any such pretensions… Demand and supply exert equally important influence on wages, neither has a claim to predominance any more than has either blade of scissors or either pier of an arch....... (but) the doctrine throws into clear light the action of one of the causes that govern wages."[3]

It is clear from the above quotation from Marshall that he considered marginal productivity principle (which he calls the marginal productivity doctrine) as one of the two forces that determine wages, the other force being the supply of labour. Marshall believed that wages would tend to be equal to the marginal product, but he emphasized several times that the wages are not determined by marginal product, since like all other marginal quantities, marginal product is determined together with the price (wage) by the interaction of demand and supply.

Furthermore, Marshall pointed out that the marginal productivity doctrine plus the competition in the labour market would in the long run tend to make wage of labour in different industries or uses equal to each other and to the marginal product (assuming of course that labour is homogenous). Another reason why Marshall rejected Clark's version of marginal productivity theory and also therefore why he declined to call the marginal productivity doctrine as a theory of wages was his opposition to use the assumptions of stationary state in his theory of distribution. As pointed out earlier, Clark's marginal productivity theory is dependent upon the rigid application of the stationary state abstraction. Through Marshall made use of the technique of stationary state abstraction in his theory of product prices, in his distribution theory he greatly qualified the techniques of stationary state by introducing some dynamic elements in it. Marshall in this theory of distribution allowed for the gradual growth of population and for the changes in the capital accumulation. Marshall pointed out that the changes in the real wages would in the long run, affect the growth of population and, therefore, influence the size of labour force or supply of labour. Further according to Marshal, the changes in the return on the capital would greatly affect the present and future level of savings and would therefore bring about changes in the capital accumulation. Clark in his theory of distribution had assumed away the changes in population (of labour force) and capital accumulation as a result of the changes in the factor rewards and thereby ignored the long-run repercussions of the changes in the factors reward.

On another point also Marshall expressed his dissatisfaction with Clark's version of marginal productivity theory and this difference was due to the fact that Marshall was more concerned with immediate short-run problems of individual employers. The relevant problem for the marginal productivity theory is to ascertain the marginal productivity of labour separately from the productive contribution of capital. This is because the employer (producer) has a joint demand for doses of combined labour and capital. Clark, as noted above, in order to find out the marginal productivity of labour separately assumed that the form of capital could be suitably adapted so that a given quantity of capital could be used with any number of labourers. Since the adaptability or variation in the form of capital can be achieved only in the long run, Marshall did not adopt this technique of separating out the marginal productivity of labour. Instead, he evolved the concept of marginal net productivity of labour. Marginal net productivity of labour is obtained by measuring the marginal product of the joint addition of labour and capital and subtracting from it the cost of capital. The concept of marginal product is essential and very useful for the employer who is to make short-run adjustments in the amount of labour used.

It may however be pointed out that later writers such as Pigou, Robertson and Hicks did not adopt the concept of marginal net productivity of labour and instead relied upon the adaptation of the form of capital in ascertaining labourer's marginal productivity separately. For the purpose of illustration, let us take the case of the factor labour and the price of its use, the wage rate, and explain how Marshall- Hicks' Marginal Productivity theory explains the determination of wages.

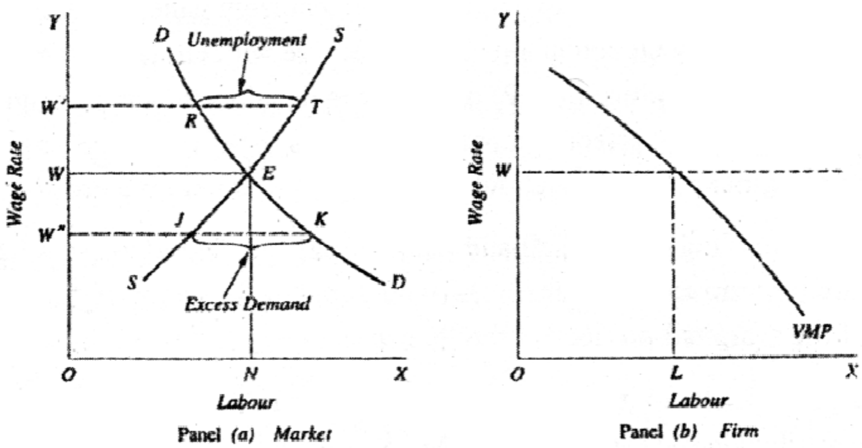


Fig. 14.3. Wage rate is determined by demand and supply and is equal to VMP of labour

It will be seen from Figure 14.3 that the demand curve for labour as derived from the marginal productivity curve is given by DD which slopes downward to the right. On the other hand, supply curve of labour is sloping upward showing that as the wage rate rises, the amount of labour supplied to a given occupation or use will increase. Even for the economy as a whole, the supply curve of labour can slope upward indicating positive response of labour to the rises in the wage rate. It will be seen from Panel (a) of fig 14.3. that demand and supply curves of labour intersect at point E and the labour market is in equilibrium with OW as the wage rate and ON quantity of labour employed. Thus, with the working of competitive forces, wage rate OW and level of employment ON are determined through the interaction of demand for and supply of labour. If somehow wage rate OW' is prevailing, quantity demanded of labour W'R at this wage rate falls short of quantity supplied W'T at this wage rate. Under such a situation all workers who want work at this wage rate will not get jobs and will therefore become involuntary unemployed. The competition among these unemployed workers will bring down the wage rate to the equilibrium level OW. On the other hand, if wage rate is lower than OW, say OW", the excess demand for labour will emerge. This excess demand for labour will lead to the competition among employers which will result in rise in the wage rate to OW.

With OW as the given wage rate, it is evident from panel (b) of Figure 14.3 that firm will employ OL amount of labour because with OL labour employment, value of marginal produce (VMP) of labour is equal to the given wage rate. A firm operating under perfect competition equates the given wage rate with the value of marginal product of labour (VMP) so as to maximize profits.

Thus, according to Marshall and Hicks, wage rate and likewise other factor prices are determined by Demand for and supply of a factor and are equal to value of the marginal product of the factor.

Marginal productivity theory has been a pillar in the neo-classical theory of income distribution and even in modern Micro-economics it constitutes an important factor in the determination of factor prices.

14.6 CRITICAL EVALUATION OF MARGINAL PRODUCTIVITY THEORY

Since there exist many versions of the theory, much of the criticism of the theory may be attributed to the conflicting interpretations of the theory given by its supporters. For instance, some of the protagonists of the theory believed that with the marginal productivity theory they not only succeeded in explaining the present system of distribution of income or the way how factor prices were determined but also succeeded in showing its moral desirability. In other words, they thought that the prices of factors are not only determined by marginal productivity but that the factors ought to be paid in accordance with their marginal productivity. In their view, factor prices are not only determined by and equal to marginal products but it is also socially just and ethically desirable that the various factors be paid according to their contribution to the total national product i.e., equal to their marginal productivity.

In order to properly evaluate the marginal productivity theory, it is useful to remember that there are at least two views—Clarkian and Marshallian—of the marginal productivity theory differing mainly in their inclusiveness. Clarkian version of marginal productivity theory states that in the long run and under conditions of competitive equilibrium and fixed factor supplies, factor prices are determined by marginal productivity and price of a factor in all its various uses will tend to be equal. On the other hand, Marshallian version says that only the demand for a factor is determined by marginal product of the factor. However, even in the Marshallian version the price of a factor, in equilibrium, will be equal to its marginal product and the price of a factor in all its various uses will tend to be equal in the long run.

Clarkian theory includes the Marshallian doctrine and claims more than the latter. As we shall see below, Marshallian doctrine which we preferred to call marginal productivity principle rate than theory in our above analysis, still constitutes the heart of modern theory of factor pricing, though more inclusive Clarkian theory which is only one-sided has been abandoned. We shall discuss below in detail the various objections raised against marginal productivity theory and shall indicate which objections are valid and which ones are based upon only misinterpretation of the theory and are thus misplaced. Most of the criticisms have been leveled against the Clarkian version, but some objections have been raised on the very concept of marginal product and therefore apply equally to the Marshallian principle of marginal productivity which is a fundamental doctrine of the modern theory of factor pricing.

(1) Marginal Productivity theory is based on several unrealistic assumptions.

It has often been argued that marginal productivity theory takes too many assumptions which are quite unrealistic. Therefore, it is concluded that this theory has no validity. The theory of marginal productivity (Clarkian version) assumes a stationary state, perfect competition, perfect mobility of factors, equal-bargaining power of buyers and sellers, and perfect knowledge, which are all far away from the actual condition of the real world. World is not static. Instead, developments are continually taking place, making the actual world a dynamic one. Competition is not perfect, instead there are large imperfections in the factor markets which make any analysis of factor pricing based on the assumption of perfect competition quite useless. Bargaining powers of buyers and sellers of factor services, for instance, of employers and labourers, are not equal and thus make the exploitation of the weaker party possible. Prof. Paul Douglas, an advocate of marginal productivity theory, has given the following list of implicit assumptions made by marginal productivity theory.

(i) Employers are able to measure and predict in advance the marginal product of a factor.

(ii) There is free and complete competition among employers.

(iii) Labour knows its marginal productivity.

(iv) There is free and complete competition among workers for jobs.

(v) Capital is perfectly mobile.

(vi) Labour is perfectly mobile.

(vii) All labour is employed.

(viii) All capital is employed.

(ix) Bargaining power of labour and management are equal.

(x) Government does not interfere in the wage agreement.

At a first glance at these assumptions, one gets the impression that a theory which makes such unrealistic assumptions can hardly be useful. But Prof. Douglas has made a strong plea in defence of this theory. He has pointed out that most of the assumptions are reasonably good description of actual long-run market conditions and are thus generally valid over time.

2. Under Imperfect competition, factor rewards are not equal to Value of Marginal Product : – Another significant criticism levelled against marginal productivity theory (both Clarkian and Marshall-Hicks's versions) is that being based upon the assumption of perfect competition both in product and factor markets, it is unable to explain the determintion of factor prices under conditions of imperfect competition in the factor and product markets. As we shall discuss in the next lesson, following the development of imperfect and monopolistic competition theories by Joan Robinson and Chamberlin there emerged two concepts of marginal productivity, namely, Marginal Revenue Product (MRP) and Value of the Marginal Product (VMP). Thus, when there prevails imperfect competition in the product market (assuming perfect competition in the factor market), a factor of production would not get remuneration equal to the Value of the Marginal Product, as is generally presumed in the marginal productivity theory. Under imperfect competition in the product marker, a factor of production is remunerated according to a different principle, namely, Marginal Revenue Product (MRP) which is less than the Value of the Marginal Product (VMP). According to Joan Robinson, a factor is exploited if it is paid less than Value of its Marginal Product, whereas in Marginal Productivity theory, as it was presented in neo-classical economic thought, there was just and fair distribution of total product: every factor getting equal to its contribution in the total production. We are therefore of the view that in the context of imperfect competition in the product market, Marginal Productivity theory needs to be modified.

3. Factors do not get reward equal to MRP in conditions of monopsony. If imperfect competition or monopsony prevails in the factor market, a factor will not get the reward even equal to its Marginal Revenue Product (MRP). Under imperfect competition or monopsony or oligopsony in the factor market, the firm to be in equilibrium, will equate marginal wage of labour with the marginal revenue product of labour and, as shall be seen later, this Marginal Wage is greater than the Average Wage or the wage rate which will be paid to the labour. We shall graphically explain this in the next chapter as to how the wage rate determined under conditions of monopsony is less than even the marginal revenue product (MRP) of labour. When a factor is paid less than its marginal revenue product, Joan Robinson calls it monopolistic exploitation. We thus see that under monopsony or imperfect competition, factors do not get rewards equal to their marginal revenue proactive theory as it was presented by Marshall and Hicks, did not visualize the possibility of exploitation of labour or any other factor of production.

4. Marginal Productivity theory cannot explain rewards of the factors used in fixed proportions. Another serious shortcoming of marginal productivity theory is that it cannot explain the rewards of the factors which are used in fixed proportions. Marginal productivity takes it granted that a good degree of elasticity of substitution exists between the factors of production so that increase in one factor, keeping other factors constant, leads to the increase in the addition to the total product, that is, it has a positive marginal productivity and therefore gets positive reward for its contribution to production. But when the factors are used in fixed proportions, increase in one factor keeping the others constant, will not lead to any increase in total production at all. That is to say, in case of fixed proportions or fixed relations between the factors, marginal productivity of the factors will be zero. In view of their zero marginal productivity, according to the marginal productivity theory, their rewards or prices should also be zero. But this is quite absurd; the factors of production, even when they have fixed relations with each other, obtain positive rewards. To quote Prof. J. Pen, “If the relation between labour and capital were fully fixed-as many people think it is, the traditional distribution theory would collapse. For if the relation between L (i. e. , labour) and K (i.e. capital) is fixed the marginal product of both is zero. An addition of a unit of labour, with equipment constant, yields nothing, just as an extra amount of capital, without the addition of labour, would. In such a case the wage would also be zero, and also the interest. That is of course nonsense; in fact a wage rate and an interest rate come about even with fixed relations between L and K, but the marginal productivity theory is no longer suitable for explaining them.

5. Marginal Productivity Theory neglects the role of labour unions in influencing the wage rate. Still another serious drawback of Marginal Productivity theory is that, in its original and rigid version, trade unions or collective bargaining cannot raise the wages of labour without creating unemployment. Thus, according to this theory, trade unions are superfluous and collective bargaining by them is a futile activity. Given the downward-sloping nature of the marginal productivity curve, at the higher wage secured by the union, the employer will demand or employ less number of workers than before, so that some labour will be rendered unemployed. But, as we shall see in a later lesson, increase in wage rate by the union does not always cause unemployment. Indeed, we shall study there that under conditions of Monopsony, the increase in the wage rate by the union may be accompanied by the increase in employment rather than the creation of unemployment. As noted above, under conditions of imperfect competition in the product and factor markets, labour (or any other factor) is exploited, that is, paid less than the value of its marginal product or less than its marginal revenue product. In this context trade unions can play a useful role of removing the exploitation of labour by getting the wage rate raised to the level of Value of Marginal Product or Marginal Revenue Product.

6. Marginal productivity theory wrongly applied to the macro-analysis of wage-employment relationship. Marginal productivity theory has also been criticized for its application to macro- economic field and wrong conclusions drawn in this connection. At times of severe depression and huge unemployment during the thirties, A. C. Pigou, a famous neo-classical economist, argued on the basis of marginal productivity theory that cut in wages of labour in the whole economy would bring about increase in employment, since, given the falling marginal revenue productivity curve of labour, at a lower wage rate more labour will be employed and economy would be able to get out of depression. Thanks to J. M. Keynes who successfully challenged the above argument. According to him, what is valid in the case of a single industry or firm may not be valid for the economy as a whole. He argued for the macro view of the problem, according to which wage is not only the cost of production to a firm or industry but also an income for the labourers who constitute a majority in the society. Therefore, according to Keynes, if all-round cut is made in the wages of the working class, their incomes will decline, which in turn would bring about a fall in the aggregate demand for goods. This fall in aggregate demand will adversely affect the employment opportunities and production in the economy. Thus, cut in wages, according to Keynes, instead of removing unemployment and depression will further deepen them. We thus see that application of Marginal Productivity theory to the macro level yields quite incorrect and invalid results.

7. Positive Relationship between factor rewards and productivity ignored. Marginal productivity theory also ignores the positive interrelations between rewards of the factors and their productivity, especially between wages and the efficiency or productivity of labour. It has been pointed out that rise in wages has a favourable effect on the efficiency and productivity of labour. With higher wages, workers can afford to have better standard of living and better health which will raise their productivity and efficiency. This positive relationship between wages and labour efficiency especially holds good in case of developing countries like India where the wage rates in many industries are even below the minimum subsistence level. With wages even below the subsistence level, workers remain underfed and undernourished and, as a result, unhealthy and inefficient. If following the rise in wages, efficiency and productivity of labour improve, then it may be worthwhile from the viewpoint of employers to raise wages. It has therefore sometimes been asserted that “higher wages are economical” or there is an ‘economy of high wages’. But, as mentioned above, marginal productivity theory completely ignores this favourable effect of higher wages on productivity of labour.

Now, if the favourable effect of higher wages on labour productivity is recognized, the unique level of wage-equilibrium arrived at in the marginal productivity theory is not valid. With every rise in the wage rate, there will be a different curve of marginal productivity of labour and a different wage-employment equilibrium. Thus there are various possible positions of wage equilibrium depending upon the productivity and efficiency and there is a choice for the firm or the industry to select among them. That there is a unique wage equilibrium, as has been asserted by the strict and rigid version of marginal productivity theory, is therefore not acceptable.

8. Profit maximization assumption of marginal productivity theory criticized. Like the neo classical theory of pricing of products, marginal productivity theory of distribution which has been developed by neo- classical economists is also a marginal approach to the problem and therefore assumes that the entrepreneurs or employers seek to maximize profits. It is only if the entrepreneurs are maximizers of profits that they will equate wage with marginal product of labour. If they do not seek to maximize profits, they may employ the amount of labour at which marginal product of labour stands higher than the wage. Likewise, they may employ the number of workers at which the marginal product is lower than the wage. Just as Hall and Hitch have criticized the marginalist approach as it applies to product pricing on the ground that entrepreneurs do not maximize profits, R. A. Lester, an American economist, has criticized the marginal productivity theory of distribution on the ground that entrepreneurs do not behave as maximizers of profits and therefore wages in the real world will differ from the marginal product of labour. Lester has provided empirical evidence in support of his views. But Machlup and Pen have defended the marginal productivity theory of distribution and the profit maximization assumption on which it is based. To quote Pen, "It is not necessary for every entrepreneur to be able to find the exact point of equilibrium, some will overshoot the equality of wage and marginal product, others will remain below it. However, the trend is towards equality". In this sense the theory gives only a rough approximation of equality, but as such it is probably not bad.

9. Entrepreners are ignorant products of factors. Marginal productivity theory has also been challenged on the ground that it assumes that entrepreneurs are fully aware of their production function or, in other words, they know what are the marginal products of various factors and how they change with the expansion in their employment. In the empirical investigations when the entrepreneurs are asked whether for employing various factors of production they estimate their marginal productivities and take them into account, they flatly say 'no'. They generally reply to the economic investigator's question, "look, you are telling that I perform mental processes and calculations, that I don't perform and in fact couldn't perform if I wanted to, because I do not know the production function. Your whole distribution theory is something you have made up. It's all very ingenious but there's no rhyme or reason to it."

But many economists have tried to answer the above argument by pointing out that entrepreneurs may not be consciously calculating the marginal productivities of the various factors and taking the decisions according to these, but unconsciously or subconsciously they do behave according to the logic of marginal productivity, since they are out for maximum profits. Further, they point out that forces of competition compel them to behave in accordance with the marginal productivity doctrine. Moreover, it is pointed out that there are many entrepreneurs working independently of one another to employ the various factors. Some may employ a factor at a reward which is higher than its marginal productivity and some others may employ it at a reward which is less than its marginal productivity. But the forces of competition will ensure that these discrepancies remain within bounds and on the average reward for a factor of production is approximately in line with its marginal product.

Check Your Progress I

Answer the questions in the space provided :

Q1. Discuss Clark's version of Marginal Productivity Theory of distribution.

____

____

____

____

Q2. ExplaindiagrammaticallyMarshall-Hick’sversion.

Q3. Give a critique of M.P. Theory of distribution.

CONCLUSION

We have discussed above the various criticisms levelled against marginal productivity theory of distribution. Marginal productivity theory of distribution does not explain fully the determination of all factor prices. But marginal productivity of a factor is the most important economic factor governing the prices of factors. Other factors, such as power, social conventions, status and prestige do play a part in fixation of remunerations, but the economic factor of marginal productivity does exercise an important influence on the fixation of factor rewards.

14.7. PRODUCT EXHAUSTION PROBLEM OR ADDING-UP PROBLEM

As soon as it was propounded that the factors of production are paid equal to their marginal products, a perplexing problem cropped up over which there was a serious debate among the famous economists at that time. The perplexing problem which was posed was that if all factors were paid rewards equal to their marginal products, would the total product be just exactly exhausted? In other words, if each factor is rewarded equal to its marginal product, the total product should be disposed off without any surplus or deficit. The problem of providing that the total product will be just exhausted if all factors are paid rewards equal to their marginal products has been called “adding-up Problem” or Product Exhaustion Problem.

Let us illustrate the product exhaustion problem by assuming that only two factors, labour and capital, are required for production. Let 'a' stand for labour and b for capital. It should be remembered that marginal product of a factor can be shown as equivalent to marginal product when a certain quantity of the variable factor is used or employed. The reward for the fixed factor can then be shown as surplus (or residual income) of the total product over the marginally determined reward of the variable factor. Consider figure 14.4. where labour has been treated as a variable factor and shown on the X-axis and capital has been taken on the Y-axis as the fixed factor. If OL is the equilibrium amount of labour employed, the marginal product of labour is LM and wage rate determined according to this is OW. The total wage bill, that is , labour's share is equal to OLMW. The total product produced is, that is, the whole area under the marginal productivity curve of labour.

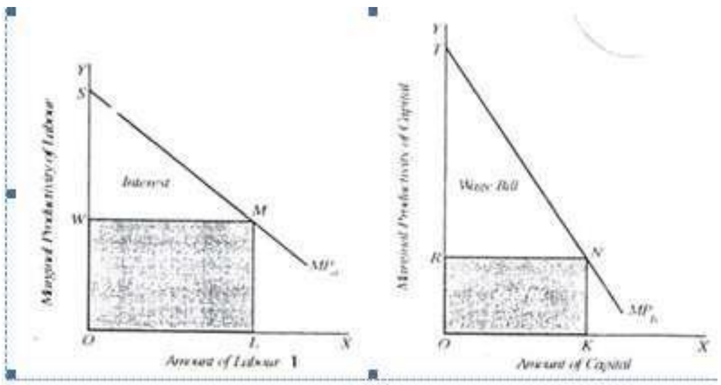


Fig 14.4 Fig 14.5.

Thus, the total interest on capital will be equal to OSML—OLMW = WMS. With OLMW as marginally determined wages and WMS as residual income of interest, the total product is exhausted. But the problem of product exhaustion is to show that interest on capital determined as residual income is in fact equal to the marginal product of capital used multiplied by the amount of capital used. To show this, we have to treat capital as a variable factor and labour as a fixed factor. This has been done in figure 14.5, by taking the same amount of labour used earlier as a fixed factor. Capital is now measured on the X-axis. If now the equilibrium amount of capital used is OK, KN is the marginal product of capital equal to which rate of interest OR is determined. Thus OKNR is the marginally determined interest income on capital. Now the residual income TRN will go to the labour as wages. Thus in Figure 14.5. the wage-bill is determined as residual income.

Now, in order to show that payment in accordance with marginal productivity to both labour and capital would exactly exhaust the total product, it is required to be proved that the area OKNR in figure 14.5 is equal to the area WMS in Figure 14.4, and area RNT in Figure 14.5. is equal to OLMW in figure 11.4. In this way we will be able to verify that income of a factor determined as marginal product is equal to income of the factor determined as residual.

It should be noted that we have not proved the product exhaustion problem: we have only illustrated the problem itself.

14.7.1 Wicksteed's Solution of Product Exhaustion Problem

Philip Wicksteed was one of the first economists who posed this problem and proved a solution for it. Wicksteed applied a mathematical proposition called Euler's Theorem to prove that the total product will be just exhausted if all the factors are paid equal to their marginal products. Let Q stand for the total output of the product, 'a' stand for the factor labour and 'b' stand for the factor capital and 'c' stand for entrepreneur. Assuming that there are only three factors employed for production. Then, the adding-up problem implies that;

$$Q = MP_a \times a + MP_b \times b + MP_c \times c$$

That is, the marginal product of factor 'a' multiplied by the amount of factor 'a' plus the marginal product of factor 'b' multiplied by the amount of factor 'b' plus the amount of factor 'c' multiplied by MP of 'c' equals the total product of the firm. Marginal products of various factors can be expressed as partial derivatives. Thus, the marginal product of labour (i.e., factor 'a') can be expressed as $\frac{\partial q}{\partial a}$ and the marginal product of capital (factor b) as $\frac{\partial q}{\partial b}$ and the marginal product of entrepreneur (factor c) as; $\frac{\partial q}{\partial c}$ then for the adding-up problem to be fulfilled, the following equations must hold good :

$$Q = a\frac{\partial q}{\partial a} + b\frac{\partial q}{\partial b} + c\frac{\partial q}{\partial c}$$

Where $a\frac{\partial q}{\partial a}$ represents share of the total product going to labour.

$b\frac{\partial q}{\partial b}$ represents share of the total product going to Capital

$c\frac{\partial q}{\partial c}$ represents share of the total product going to entrepreneur.

Now, Euler's theorem states that if Q is homogeneous function of the first degree, that is if Q = f(a, b, c), then for any given increase in the variables 'a', 'b' and 'c' by an amount 'n', the output Q will also increase by 'n'. Thus homogenous production function of the first degree or linearly homogenous function is of the following form:

$$nQ = f(na, nb, nc)$$

Now, according to Euler's Theorem, for the linearly homogenous function:

$$Q = a\frac{\partial q}{\partial a} + b\frac{\partial q}{\partial b} + c\frac{\partial q}{\partial c}$$

Thus if production function is homogenous of the first degree, then according to Euler's theorem the total product

$$Q = a\frac{\partial q}{\partial a} + b\frac{\partial q}{\partial b} + c\frac{\partial q}{\partial c}$$

Where $\frac{\partial q}{\partial a}, \frac{\partial q}{\partial b}, \frac{\partial q}{\partial c}$

are partial derivatives of the production function and therefore represent the marginal products of labour, capital, and entrepreneur respectively. It follows, therefore, that if production function is homogenous of the first degree (that is, where there are constant returns to scale) then, according to Euler's theorem, if the various factor, a, b and c are paid rewards equal to their marginal products, the total product will be just exhausted, with no surplus or deficit.

14.7.2. Cobb-Douglas Production Function and Product Exhaustion Problem

Let us demonstrate clearlythe product exhaustion problem with Euler's theorem by taking two factor Cobb-Douglas production, $Q=AL^aK^b$ where when a + b = 1 is homogeneous function of first degree (i.e. exhibits constant returns to scale).

To study the distribution of the total product between the factors, we have to determine marginal physical products of each factor by differentating the given production function with respect to labour and capital. Thus

$$MP_L = \frac{\partial Q}{\partial L} = AaL^{a-1}K^b$$

$$MP_K = \frac{\partial Q}{\partial K} = AbL^aK^{b-1}$$

Total payment to each factor is equal to the rate of reward per unit of the factor (which is taken to be equal to marginal product of a factor) multiplied by the amount of the factor employed. Thus

$$Y_L = \frac{\partial Q}{\partial L}.L = AaL^{a-1}K^b.L = aAL^aK^b \qquad .....1$$

Where Y_L stands for labour share i.e real income of labour in the total product.

$AL^aK^b = Q$, from equation (1) we get

$$Y_L = aQ$$

Similarly, share of capital (Y_k) in the total product can be written as

$$Y_k = \frac{\partial Q}{\partial K}.K = bAL^aK^{b-1}.K = bAL^a K^b$$

$$= bQ$$

Total payment made to both the factors is the sum of the payments made to them. Thus,

$$Y = Y_L + Y_k = aQ + bQ$$

$$= (a+b) Q$$

Now, if a + b = 1 (as is the given in Cobb-Douglas production function), then

$$Y = Q$$

We thus see that Euler's theorem is able to explain product exhaustion problem when production function is homogenous of the first degree. In the above way, Wicksteed, assuming constant returns to scale and applying Euler's theorem, proved the adding-up problem, that is, demonstrated that if all factor are paid equal to their marginal products, the total product will be just exactly exhausted.

Wicksteed's solution was criticized by Walras, Barone, Edgewroth and Pareto. It has been asserted by these writers that production function is not homogeneous of the first degree, that is, returns to scale are not constant in the actual world. Thus, Edgeworth satirically commented on Wicksteed's solution, "there is significance in this generalization which recalls the Youth of philosophy Justice is a perfect cube, said the ancient sage ; and rational conduct is a homogeneous function, adds the modern savant." Critics have pointed out that production function is such that it yields a U-shaped long-run average cost curve: The U- shape of the long-run average cost curve implies that up to a point, increasing returns to scale occur and after it diminishing returns to scale are found. In case a firm is still working under increasing returns to scale, then if all factors are paid equal to their marginal products, the total factor rewards would exceed the total product. On the other hand, if a firm is working under diminishing returns to scale, and all factors are paid equal to their marginal products, the total factor rewards would not fully exhaust the total product and will therefore leave a surplus. It follows that Euler's theorem does not apply and therefore the adding- up Problem does not hold good when either there are increasing returns to scale or decreasing returns to scale.

Another drawback pointed out in Wicksteed's solution is that when there is constant returns to scale ,the long-run average cost curve of the firm is a horizontal straight line, which is incompatible with perfect competition. (Under horizontal long-run average cost curve, the firm cannot have a determinate equilibrium position). But competition was essential to the marginal productivity theory and therefore to Wicksteed's solution. Thus Wicksteed's solution leads us to two contradictory things.

14.7.3 Solution of product Exhaustion Problem by Wicksell, Walras, Barone, Samuelson and Hicks :-

After Wisksteed, scholars like Wicksell, Walras and Barone, each independently, advanced more satisfactory solution to the problem that marginally determined factor reward would just exhaust the total product. These authors assumed that the typical production function was not homogenous of the first degree, but was such that yielded U-shaped long-run average cost curve. They pointed out that in the long run under perfect competition, the firm was in equilibrium at the minimum point of the long-run average cost curve. At the minimum point of the long-run average cost curve, the returns to scale are momentarily constant, that is, the operation of constant returns to scale are constant within the range of small variations of output. Thus the condition required for the marginally determined rewards to exhaust the total product, that is, the operation of constant returns to scale, was fulfilled at the minimum point of the long-run average cost curve, where a perfectly competitive firm is in long-run equilibrium. Thud in the case of perfectly long-run equilibrium, if the factors are paid reward equal to their marginal products, the total product would be just exactly exhausted.

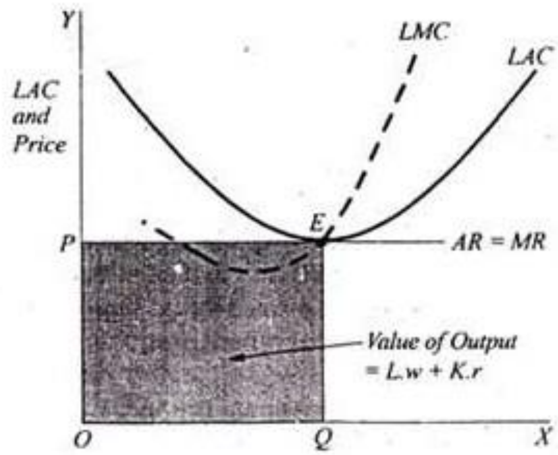


Fig. 14.6. Total product is just exhausted with equilibrium at the minimum point of LAC under perfect competition

As seen above, Wicksteed provided a solution to the product exhaustion problem by assuring that production function was linearly homogenous, that is constant returns to scale prevailed. Since all production function was linearly homogenous, the controversy remained unresolved whether or not , even under perfect competition and with the usual U—shape of long-run average cost curve with varying return to scale, product-exhaustion problem was valid. The credit for resolving this issue goes to Hicks and Samuelson, who showed that the solution to the product exhaustion problem depended crucially on the market conditions of perfect competition and not on the property of production function. As a result of free entry and exit, the equilibrium of a firm under perfect competition in the long run ensures that the firms produce at the minimum point of the long-run average cost curve LAC (see Figure above) where they make neither any economic profits, nor any losses. Now if according to marginal productivity theory, factors are paid rewards equal to the value of their marginal Products and the theory of perfect competition states that in the long run the perfectly competitive firms earn zero economic profits, it follows that under perfect competition in both the product and factor markets the level of output and employment of factors (labour and capital) would be such as to satisfy the equilibrium conditions-of factor employment and long-run competition equilibrium with zero economic profits. We show below mathematically how under perfect competition in the product and factor markets, adding up problem is solved or product exhaustion problem holds.

Mathematically, zero economic profit condition implies that the value of output equals the total cost of production. Let 'L' stands for labour, 'C' for capital, w for wage rate, r for price of capital, Q for output of a product and P for the price of the product. According to the long-run perfectly competitive equilibrium, at the minimum point E of LAC curve (see Figure 14.6.) we have

Value of output = Total Cost

$$PQ = Lw + Kr \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Now, according to marginal productivity theory, factors are rewarded equal to value of their marginal products. Thus

$$w = VMP_L = P.MPP_L \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

$$r = VMP_k = P.MPP_K \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. (3)$$

substituting the values of w and r as obtained above in (2) and (3) respectively in (1) we have

$$PQ = L.P.MPP_L + K.P.MPP_K$$

Thus, for a given price P, if the factors (labour and capital) are paid equal to their marginal physical products (MPP_L and MPP_K), the total payments to labour and capital would be equal to the total product Q and thus total product would be just exhausted.

From above it is clear that it is the property of perfect competition, namely that it ensures long-run equilibrium at the minimum point of the U-shaped long-run average cost curve with zeroeconomic profits that solves the product exhaustion problem. It may, however, be noted that at the minimum point of the U-shaped LAC curve, returns to scale are locally constant.

It follows from above that the two main solutions of the adding-up problem were offered. First, Wickstreed's solution which assumed the operation of constant returns to scale. Secondly the solution provided by Wicksell, Barone Walras, Samuelson and Hicks which assumed that the firms operated at the lowest point of the long-run average cost.

Check Your Progress-II

Answer the questions in the space provided:

Q1. What do you mean by product exhaustion problem or adding up problem?

____

____

____

____

Q2. What are the major highlights of Wicksteed's solution to product exhaustion theorem?

____

____

____

____

Q3. Is it justifiable to say that it is the property of perfect competition that solves the product exhaustion problem ?

____

____

____

____

14.8. LET US SUM UP

Dear learner, in this lesson, we explored the Marginal Productivity Theory of Distribution, understanding how factor incomes are determined by their marginal contributions, and examined different versions of the theory by Clark, Marshall, and Hicks. We also studied the Product Exhaustion Problem, which concerns whether total output can be completely distributed among factors without any surplus or shortfall, and reviewed key solutions provided by Wicksteed, the Cobb-Douglas production function, and economists like Wicksell, Walras, Barone, Samuelson, and Hicks. This lesson has helped you grasp the theoretical foundations and practical challenges in explaining how income is shared among the factors of production.

14.9 KEYWORDS

1. **Marginal Productivity**

The additional output produced by employing one more unit of a factor of production, holding other inputs constant.

2. **Marginal Revenue Product (MRP)**

The additional revenue generated from selling the output produced by one additional unit of a factor.

3. **Product Exhaustion Problem**

A theoretical issue questioning whether the total product (output) can be exactly distributed among the factors of production according to their marginal products without any leftover.

4. **Cobb-Douglas Production Function**

A specific mathematical form of production function that shows how inputs (typically labor and capital) combine to produce output, often used to demonstrate product exhaustion.

5. **Factor Pricing**

The determination of payments to the factors of production (wages for labor, rent for land, interest for capital) based on their contribution to the production process.

14.10 EXAMINATION-ORIENTED QUESTIONS

1. State and critically examine the Marginal Productivity Theory of Factor Pricing.
2. How does the Marshall-Hicks' version of the Marginal Productivity Theory of Distribution differ from Clark's version?
3. Show that when factors are paid rewards equal to their marginal products, the total product is exactly exhausted.
4. The Marginal Productivity Theory neglects the supply side of factors. Critically examine this limitation.
5. Discuss the Product Exhaustion Problem by taking the example of the Cobb-Douglas production function.

14.11 HINTS TO CHECK YOUR PROGRESS 1

Q 1 See section 14.4

Q2. See section 14.5

Q3. See section 14.6

Hints to Check Your Progress II

Q 1 SEE section 14.7

Q2. See SUB section 14.7.1

Q3. See SUB section 14.7.2

14.12 SUGGESTED READINGS

1. Ahuja, H.L. *Advanced Economic Theory*, 9th edition, 1997.
2. Koutsoyiannis, A. *Modern Microeconomics*, 2nd edition, 1979.
3. Misra, S.K. and V.K. Puri, *Advanced Microeconomic Theory*, 1st edition, 2001.

4. Schumpeter, J.A. *History of Economic Analysis*, p. 939.

5. Cartter, A.M. *Theory of Wages and Employment*, p. 518.

6. Marshall, A. *Principles of Economics*, p. 518.

7. Reder, M.W. *Studies in The Theory of Welfare Economics*, Columbia University Press, New York, 1947, p. 24.

M.A. ECONOMICS **LESSON NO. 15**

COURSE NO. ECO 201 **UNIT-III**

MARGINAL CONDITIONS OF PARETO OPTIMUM, SOCIAL WELFARE FUNCTION, COMPENSATION PRINCIPLE, SCITOVOSKY PARADOX, SCITOVOSKY DOUBLE CRITERION, THEORY OF SECOND BEST AND ARROW'S IMPOSSIBILITY THEOREM.

STRUCTURE

15.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

- Discuss the **conditions of Pareto optimality** and their relevance in Welfare Economics.
- Explain the concept and types of the **Social Welfare Function**.
- Describe the **Compensation Principle** and related welfare concepts, including critiques and paradoxes.

15.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to:

- Discuss the **conditions of Pareto optimality** and their relevance in Welfare Economics.
- Explain the concept and types of the **Social Welfare Function**.
- Describe the **Compensation Principle** and related welfare concepts, including critiques and paradoxes.

15.2 INTRODUCTION

Pareto Criterion states that an economic change which harms no one and makes some one better off indicates an increase in social welfare. But this criterion does not apply to those economic changes which harm some and benefit others. Economists like Kaldor, Hicks and scitovsky have made efforts to evaluate the change in social welfare resulting from any economic reorganization which harms somebody and benefits the other. They have put forward a criterion known as the compensation principle on the basis of which they claim to evaluate those changes in economic policy which make some individuals better off and others worse off.

The theory of second best states that due to the existence of Monopoly or imperfections in some markets or in case of externalities and public goods, all marginal conditions of Pareto optimality are not fulfilled. On the other hand, Prof. Arrow pointed that the construction of social welfare function which reflects the preferences of all individuals constituting society is an impossible task. In this lesson we shall be discussing these topics in detail.

15.3. MARGINAL CONDITIONS OF PARETO OPTIMUM

Verfredo Pareto (1848-1923) was the first to part with traditional approaches to social welfare in two important respects. First, he rejected notion of cardinal utility and its additive nature and, second, he detached Welfare Economics from the inter-personal comparisons of utilities. Pareto's concept of maximum social welfare which is based on ordinal utility and is also free from value judgment occupies a significant place in modern Welfare Economics. Pareto's optimum may not be sufficient condition for attaining maximum social welfare but it is necessary condition for it. Pareto's optimum (often called economic efficiency) is a position from which it is impossible to make anymore better off without making someone worse off by any re-allocation of resources or inputs and outputs. In the words of Pareto, we are led to a position of maximum ophelimity (welfare) as one where it is impossible to make a small change of any sort such that the ophelimities of all the individuals, except those that remain constant, are either all increased or all diminished." Thus, in the Pareto optimum position the welfare of an individual of the society cannot be increased without decreasing the welfare of another member.

The concept of Pareto efficiency can be easily understood, if we describe the situation when resource allocation is not Pareto efficient. Suppose government has undertaken the task of distributing apples and oranges among individuals and this distribution does not bother about individuals' preferences. Assume some individuals like apples and hate oranges, while others like oranges and hate apples. If government distributes apples and oranges in a way that those who like apples get oranges and those who like oranges get apples, it is obvious that such distribution of goods is not Pareto efficient. If the individuals are allowed to exchange apple and oranges, they will voluntarily do so and both groups of individuals will become better off.

Pareto argues from his criterion that competition leads the society to an optimum position but he did not give any mathematical proof of it, nor did he derive the marginal conditions to be fulfilled to achieve the optimum position. Later on, Lerner and Hicks derived the marginal conditions which must be fulfilled for the attainment of Pareto optimum. These marginal conditions are based on the following important assumptions:

1. Each individual has his own ordinal utility function and possesses a definite amount of each product and factor.
2. Production function of every firm and the state of technology is given and remains constant.
3. Goods are perfectly divisible.
4. A producer tries to produce a given output with the least-cost combination of factors.
5. Every individual wants to maximize his satisfaction.
6. Every individual purchases some quantity of all goods.
7. All factors of production are perfectly mobile.

Given the above assumptions, various marginal conditions (first-order conditions) required for the achievement of Pareto optimum or maximum social welfare are explained below :

1. The Optimum Distribution of Products among the Consumers : Efficiency in Exchange. The first condition relates to the optimum distribution of the goods among the different consumers composing a society at a particular point of time. The condition says: "The marginal rate of substitution between any two goods must be the same for every individual who consumes them both"[4]. Marginal rate of substitution of one good for another is the amount of one good necessaryto compensate for the loss of a marginal unit of another to maintain a constant level of satisfaction. So long as the marginal rate of substitution (MRS) between two goods is not equal for any two consumers, they will enter into an exchange which would increase the satisfaction of both or of one without decreasing the satisfaction of the other. This condition can be better explained with the help of the Edgeworth Box diagram. In Figure 15.1, goods X and Y, which are consumed by two individuals A and B composing a society are represented on the X and Y axes respectively. O_A and O_B are origins for A and B respectively. I_{a1}, I_{a2}, I_{a3}, and I_{b11}, I_{b2}, I_{b3}, are the indifference curves showing successively higher level of satisfaction of consumers A and B respectively. CC' is the contract curve passing through various tangency points Q, R, S of the indifference curves of A and B. The marginal rates of substitution (MRS) between the two goods for individuals A and B are equal on the various points of the contract curve CC'. Any point outside the contract curve does not represent the equality of MRS between the two goods for two individuals A and B of the society.

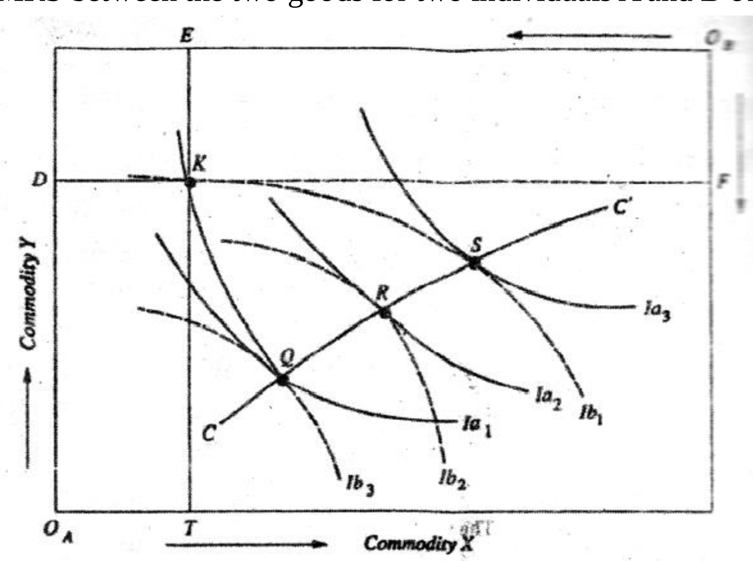


Fig. 15.1. The Optimum Distribution of Goods

Let us consider point K where indifference curves I_{a11}, and I_{b11}, of individuals A and B respectively intersect each other instead of being tangential. Therefore, at point K marginal rate of substitution between two goods X and Y (MRS_{XY}) of individual A is not equal to that of B. With the initial distribution of goods as represented by point K, it is possible to increase the satisfaction of one individual without any decrease in that of the other or to increase the satisfaction of both by redistribution of the two goods X and Y between them. A movement from K to S increases the satisfaction of A without any decrease in B's satisfaction. Similarly, a movement from K to Q increases B's satisfaction without any decrease in A's satisfaction. The movement from K to R increases the satisfaction of both because both move to their higher indifference curves. Thus, a movement from K to Q or to S or any other point on the segment SQ on the contract curve will, according to Pareto criterion, increase the level of social welfare.

From above it follows that movement from any other point off the contract curve to a point on the relevant segment of the contract curve will mean increase in social welfare. At any point off the contract curve in the Edgeworth box, the indifference curves of the two individuals will intersect which will mean that MRS_{XY} of two individuals is not the same. And, as explained above, this indicates that through exchange of some units of goods between them, they can move to some point on the contract curve where the social welfare (that is, welfare of two individuals taken together) will be higher.

Since the slope of an indifference curve represents the marginal rate of substitution (MRS_{XY}), at every point of the contract curve, which represents tangency points of the indifference curves, MRS_{XY} of two individuals are equal. Therefore points on the contact curve represent the maximum social welfare. However, a movement along the contract curve in either direction will make one individual better off and the other worse off since it will put one individual on his successively higher indifference curves and the other on his successively lower indifference curves. Thus, every point on the contract curve denotes maximum social welfare in the Paretian sense but we cannot say anything about the best of them with the help of Pareto criterion.

2. The Optimum Allocation of Factors : Efficiency in Production. The second condition for Pareto optimum requires that the available factors of production should be utilized in the production of different foods in such a manner that it is impossible to increase the output of one good without a decrease in the output of another or to increase the output of both the goods by' any re-allocation of factor of production. This situation would be achieved if the marginal technical rate of substitution between any pair of factors must be the same for any two firms producing two different products and using both the factors to produce the products.

This condition too can be explained with the help of Edgeworth Box diagram relating to production. This is depicted in Fig. 15.2. Let us assume two firms A and B producing goods X and Y by using two factors labour and capital. The available quantities of labour and capital are represented on X and Y axes respectively. Q_A ansd Q_B are the origins for firms A and B respectively I_{a1}, I_{a2}, I_{a3}, and I_{b11}, I_{b2}, I_{b3}, of firms A and B represent successively higher quantities of goods X and Y respectively which they can produce by different combinations of labour and capital. The slope of the isoquant, which is convex to the origin, represents the Marginal rate of Technical substitution between two factors. MRTS of one factor for another is the amount of one factor necessary to compensate the loss of the marginal unit of another so that the level of output remains the same. So long as the MRTS between two factors for two firms producing goods X and Y is not equal, total output can be increased by transfer of factors from one firm to another.

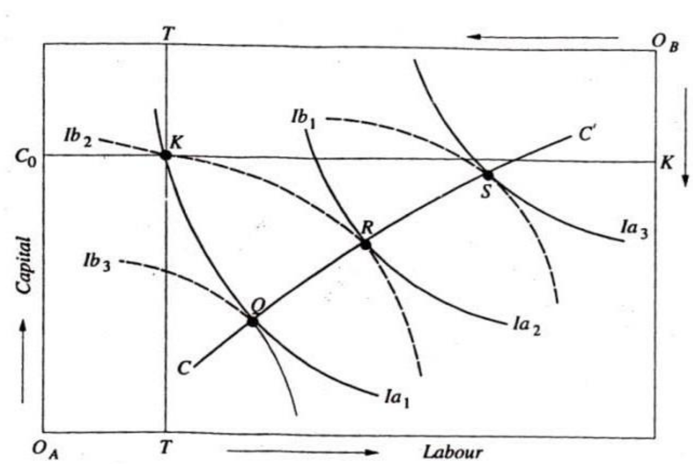


Fig. 15.2 The optimum Allocation of factors

In terms of the above diagram, any movement from K to S or to Q raises the output of one firm without any decrease in the output of the other. The total output of the two firms increases when through redistribution of factors between the two firms, a movement is made from the point K to the point Q or, S on the contract curve. Aglance at Figure 15.2. will reveal that movement from point K off the contract curve to the point R on the contract curve will raise the output of both the firms individually as well as collectively. Therefore, it follows that corresponding to a point off the contract curve there will be some points on the contract curve production at which will ensure greater total output of the two firms. As the contract curve is the locus of the tangency points of the isoquants of two firms, the Marginal Rate of Technical Substitution of the two firms is the same at every point of the contract curve CC^1. It, therefore, follows that on the contract curve at every point of which MRTS between the two factors of two firms is the same, the allocation of factor between the two firms producing X and Yrespectivelyis optimum. When the allocation of factors between the two firms is such that they are producing at a point on the contract curve, then no re-allocation of factors will increase the total output of the two firms taken together.

But it is worth mentioning that there are several points on the contract curve and each of them represents the optimum allocation of labour and capital between the two firms producing goods X and Y. But which one of them is best cannot be said on the basis of Pareto criterion because movement along the contract curve in either direction represents such factor re-allocation which increases output of one and reduces the output of another firm. That is why contract curve is also called conflict curve.

3. The Optimum Direction of Production: Efficiency in Product Mix. The third condition relates to the technical conditions of production and the state of consumer's preferences. This is also called overall condition of Pareto-optimality or economic efficiency. The fulfillment of this condition determines the optimum quantities of different commodities to be produced with given factor endowments. That is, the fulfillment of this condition ensures Pareto optimality or optimum product mix. This condition states that "the marginal rate of substitution between any pair of products for any person consuming both must be the same as the Marginal rate of transformation (for the community) between them. According to this condition, for the attainment of economic efficiency or maximum social-welfare, goods should be produced in accordance with consumer's preference. One way of proving this condition of Pareto optimality is to take a representative consumer who reflects society's preferences between the two goods. The preferences of this representative consumer reflect the preferences of the society as a whole. The second method to condition of Pareto optimality is the technique of drawing an Edgeworth Box where the indifference curves of the two consumers composing a society are drawn. We explain this condition through both these methods. First consider Fig. 15.3. in which we draw indifference curves reflecting the preferences of the society between the two commodities.

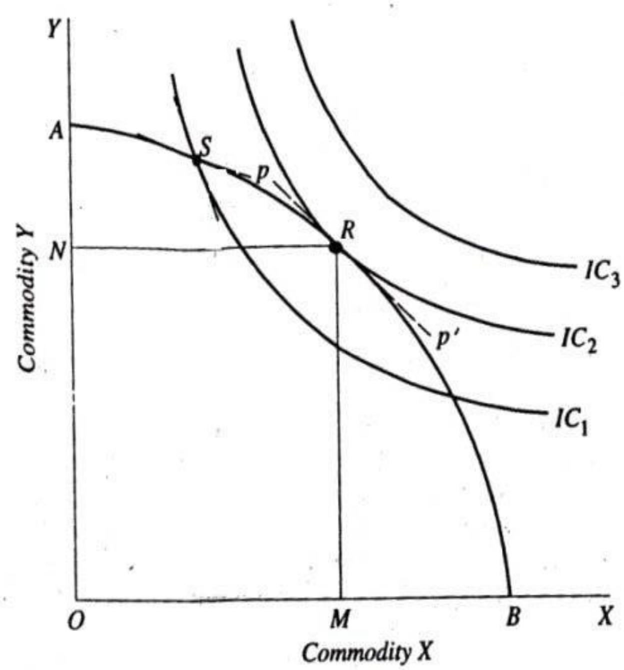


Fig 15.3.

In Fig. 15.3, commodities X and Y have been represented on the X and Y axes respectively. AB is the community's transformation curve between any pair of goods X and Y. This curve represents the maximum amount of good X that can be produced for any quantity of good Y, given the amounts of other goods that are produced and fixed supplies of available resources. Ic_1 and Ic_2 are the indifference curves of a representative consumer reflecting the preferences of the society between the two goods, the slope of which represents the marginal rate of substitution between the two goods of the society. The MRT_{XY} of the community and MRS_{XY} of the consumers of the society any equal to each other at point R at which the community's transformation curve is tangent to the indifference curve IC_2 of the consumers. Point R represents optimum composition of production in which commodities X and Y are being produced and consumed in OM and ON quantities. This is because of all the points on the community's transformation curve, point R lies at the highest possible indifference curve Ic_2 of the consumer. For instance, if a combination of goods X and Y represented by S is being produced and consumed; the consumers would be at a lower level of welfare because S lies on his lower indifference curve Ic_1 which intersects the community's transformation curve instead of being tangential to it. Tangents drawn to the indifference curve and the transformation curve at the point S intersect each other showing their slopes are not equal, that is, MRS_{XY} of theconsumers is not equal to the MRT_{XY} of the society. As a result, at point S, MRS_{XY} of the consumers is not equal to the MRT_{XY} of the community. With the situation at S there is a possibility of moving the consumers to a higher indifference curve by changing the direction or composition of production i.e. by increasing the production of X and reducing the production of Y. Thus, the optimum direction of production (or product-mix) is established at point R where community's transformation curve is tangent to the indifferent curve of a representative consumer in the society.

4. The Optimum Degree of Specialization. This is the condition necessary for determining the optimum level of output of every product by every firm. If a firm produces two goods, it has to decide the proportion of the resources to be used in the production of two goods. If the Pareto optimality is to be attained, "the Marginal Rate of Transformation between any-two products must be the same for any two firms that produce both". This condition tells to which degree any firm should specialize in order to maximize the total output of that product by all firms.

Let us first explain this condition for achievement of Pareto optimality with a numerical example Suppose the Marginal Rate of Transformation between two goods X and Y of firm A is 3 : 11 and of firm B, is 2 : 11 . That is

$$MRT_{XY} \text{ of the Firm A} = \frac{\Delta y}{\Delta x} = 3/1$$

$$MRT_{XY} \text{ of the Firm B} = \frac{\Delta y}{\Delta x} = 2/1$$

It will be recalled that Marginal Rate of Transformation A being equal to 3 implies that, it can produce 3 units of Y by sacrificing one unit of X (that is, devoting resources released from reduction in output of X by one unit) ; MRT_{XY} of firm B implies that it can produce 2 units of Y by reducing output of X by one unit and thus releasing resources. It follows that if resources by the two firms are reallocated so that firm A shifts some resources from X to Y and firm B from Y to X, the total output of two goods will increase. Thus if firm A produces one unit less of X and the resources so released are devoted to Y, it will produce 3 units of Y whereas the loss in production of X is one unit. On the other hand, if firm B produce 2 units less of Y and resources so released are allocated to the production of good X, there will be gain in one unit of X for the loss of 2 units of Y. Thus, with this rearrangement of resources in the two firms the combined output of the two firms of good Y will increase by one unit, while the combined output of X will remain the same. If the MRT_{XY} between the two goods in the two firms is the same, the re-allocation of resources between the two goods would not lead to the increase in combined output or increase in output of one good without loss of output of another: Let us explain this condition with the help of transformation or production possibility curves of the two firms.

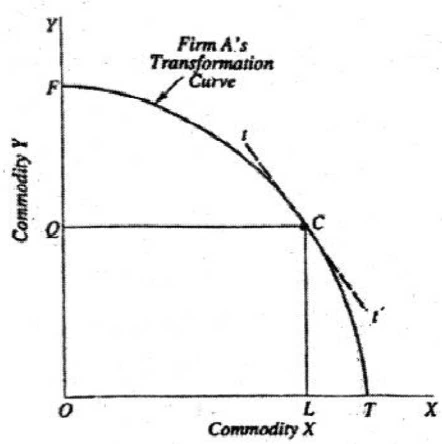


Fig. 15.4(a)

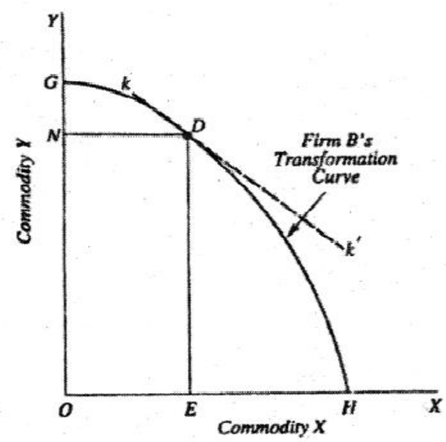


Fig. 15.4(b)

In Figure 15.4. (a) and 15.4. (b), FT and GH are the transformation curves of firms A and B respectively. They have been drawn on the assumption of increasing costs (that is, diminishing returns) conditions. A transformation curve is a locus of various combinations of two goods, which a firm can produce by fully utilizing its given resources. The slope of the transformation curve represents the Marginal Rate of Transformation (MRT) between two goods.

Suppose firm A is producing OL of X and OQ or LC of Y and firm B is producing OE of X and ED of Y. Thus the total output of both the firms is (OL+ OE) of X and (LC + ED) of Y. Now we have to prove that total output of goods X and Y will increase if the two firms specialize in accordance with their relative efficiency in the production of two goods. In Figure 15.4. (a) and (b) we have drawn tangents to the transformation curves of firms A and B at their current production at points C and D respectively. The slopes of these tangents to the transformation curves indicate the opportunity costs of commodity X in the two firms (opportunity cost = $\frac{\Delta y}{\Delta x}$ = slope).

It will be seen from the figures that the slope of tangent tt' on firm A's transformation curve is greater than the slope of tangent kk' on firm B's transformation curve. There will be gain in total output of goods X and Y if firm A reduces output of commodity X and increases the output of commodity Y and, on the contrary, firm B should reduce output of Y and devote the resources so released to the production of X until the slopes of the transformation curve in the two become equal. To sum up, Pareto-optimum is achieved from the view point of the degree of specialization when the two firms produce such combinations of goods that the slope of the transformation curve of the two firms is the same.

5. The optimum factor-Product Relationship. The fifth marginal condition of Pareto optimum relates to optimum factor - product relationship. It states "The Marginal Rate of Transformation between any factor and any product must be the same for any pair of firms using the factor and producing the product. It may be noted that the Marginal Rate of Transformation of factor into a product means how many units of a product are produced by an additional unit of a factor. Therefore, this condition simply states that the marginal product of any factor in producing a particular product must be the same in all firms producing that product. If this marginal condition is not met, 'a switching around of inputs can give us something for nothing, it can increase one or both outputs without any increase in input use.

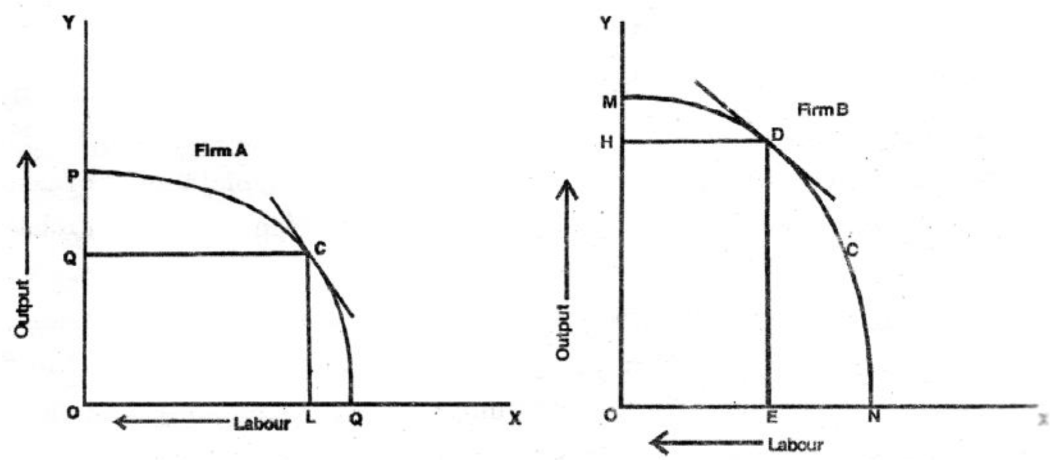


Fig.15.5 Firm A (a) *Firm B (b)*

This condition can be explained with the help of the Figures 15.5. (a) and (b) in which units of factor labour have been measured on the horizontal axis from right to left and output of a product on the vertical axis. It will be seen from these figures that as the amount of labour used increases, the total output of the product increases but at a diminishing rate. The curve QP in Figure 15.5. (a) shows as more units of labour are used in Firm A, output of the product increases but at a diminishing rate. Similarly, the curve MN in Figure 15.5 (b) shows as more labour is employed, total product increases but at a diminishing rate. It should be noted that here it is an input (labour in our figures) that is being transformed into output. The slope of the transformation curve in the present case $\frac{\Delta Q}{\Delta L}$ will therefore measure the Marginal Product of labour. It will be seen from the figures that the marginal product of labour is greater in firm A than in B (as will be known from the slopes of tangents drawn at points C and D to the transformation curves of the two firms), then total output of the product will increase if labour is shifted from firm B to firm A. The total output will be maximized and Pareto optimality attained in the use of labour in the two firms when, through re-allocation of labour between the two firms, the Marginal Rate of Transformation of labour into product (i.e. MP of labour) becomes the same.

6. The Optimum Allocation of a Factor's Time. The sixth marginal condition relates to the optimum allocation of a factor's time, particularly human labour time as between 'work for money income' and 'leisure'.

The achieve the optimum position 'the Marginal Rate of Substitution between the amount of (product) X received for aiding in its production (by a given firm) and the time spent in rendering this aid must be the same for each unit owner as the marginal rate of transformation between the time of his factor unit spent in aiding production (in this way) and the (product) X'.

Thus, according to this condition, in case of factor 'labour' the marginal rate of substitution between leisure and 'work for money income' must be equal to the marginal rate of transformation between labour time and the product. Since there is an inverse relationship between leisure and money income i.e. if a person enjoys more leisure, he receives less money income for he would be working for smaller time. Thus, an indifference curve of the factor (labourer) can be constructed by joining the various combinations of money income and leisure which give equal satisfaction to the factor owner (individual labourer). The slope of this indifference curve represents the Marginal Rate of Substitution between leisure and income. Similarly, each factor unit owner has a transformation curve between a factor unit's time and the product. If a factor works for more hours, he produces more output of any commodity. Every point on the transformation curve represents the marginal rate of transformation between factor unit's time and the product. The optimum allocation of factor's time requires that the marginal rate of substitution between leisure and income of the factor must be equal to the marginal rate of transformation between factor's time and produce. If the MRS between leisure and income is greater than the MRT between factor's time and product, the satisfaction of the individual can be increased by transfering a factor unit's time from work to leisure. Optimum position is reached when the two become equal to each other. This is illustrated in Figure 15.6.

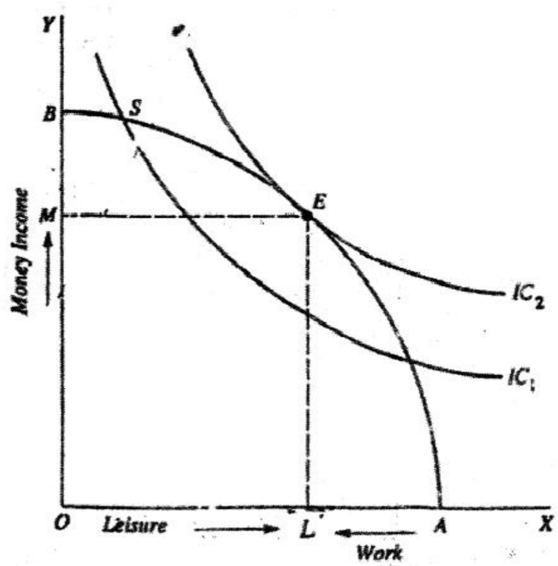


Fig. 15.6 The optimum Allocation of a factor's unit time

In Fig. 15.6, the work has been measured on the X-axis from right to left and product on the Y-axis, taking A as the zero point of the work and origin O as zero level of the product. AB is a transformation curve representing the diminishing Marginal Rate of Transformation between work and the product because of the concavity of the curve, the work (working hours) done being represented from right to the left. IC_I and IC_2 are the indifference curves of the factor owners, each representing a given level of satisfaction derived from the various combinations of income and leisure. These are also referred to as Income leisure trade off curves. Optimum position is attained at point E where the transformation curve AB is tangent to the indifference curve IC_2. Both the curves have equal slope at point E representing the equality of MRT and MRS. Obviously point S in the diagram is not an optimum position because transformation curve AB and IC_1 intersect each other instead of being tangential to each other. From point S there is the possibility of moving to a higher indifference curve such as IC_2, which is tangential to the transformation curve substituting leisure for income.

7. Inter-temporal Optimum Allocation of Money Assets. The seventh condition relates to the inter-temporal allocation over time of money or capital. Thus this condition relates to the lenders and borrowers of capital or money assets. It states that the rate of interest at which an individual is willing to lend capital must be equal to its marginal productivity to the borrower. In other words, this condition states that the marginal rate of substitution between money funds at any pair of times (say t_1 and t_2) must be the same for every pair of individuals or firms including pairs, one member of which is a firm and the other an individual. For instance, if an individual A is prepared to give Rs. 100 at any time t_1 for Rs. 110 to be received at time t_2, while an individual B is prepared to give Rs. 100 at time t_1, for more than Rs. 100 (say Rs. 115) at time t_2, then the marginal rate of substitution between money capital at these pairs of times for the two individuals is not the same. The social welfare will increase if individual A lends Rs. 100 to the individual B.

It is also worth noting that to reach the optimum position, the optimum amount of lending and borrowing for each individual and firm must take place in a given period of time. This condition can be explained with the help of Figure 15.7. In Figure 15.7, money income in the present and in the future have been shown on the X and Y axes respectively.

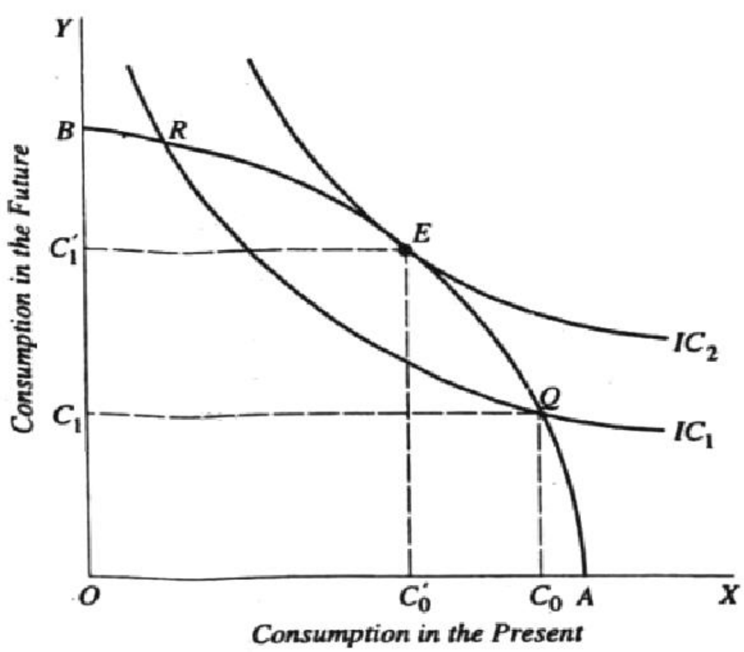


Fig. 15.7. The optimum amount of lending and borrowing

Ic_1 and Ic_2 are inter-temporal indifference curves of the lender. An inter-temporal indifference curve is the locus of various combinations of present and future money incomes which give equal satisfaction to the lender. The curve is convex to the origin which represents the diminishing Marginal Rate of Substitution of future income for present income. A lender wants to forego successively lower present income for a certain amount of future income. Similarly, AB is the inter-temporal production possibility curve of the borrower which is concave to the origin due to decreasing Marginal Productivity of capital. Point Q on the production possibility curve is the initial endowment of the individual. The inter-temporal production possibility curve AB and the inter-temporal indifference curve Ic_2 are tangential to each other at E. Therefore, E is the optimum position of the individual where Marginal Rate of Substitution of the individual lender between present income and future income is equal to the Marginal Rate of Substitution (more accurately Marginal Rate of Transformation) between present income and future income of the borrower. Since Ic_1 intersects the time production possibility curve AB at point Q, it cannot be optimum position because there is a possibility of lender moving to his higher time indifference curve such as Ic_2. With optimal position at E, the individual would consume C_0^1 in the present and C_1^1 in the future. This means that he has reduced his current consumption by $C_0C_0^1$ so that he should have greater consumption in the future. This further implies that he is lender of money and his lending will be equal to $C_0C_0^1$. It should be noted about consumption in the present that if this initial endowment point was R, his optimal consumption would have been more than his present consumption. In that case he would have been borrower. Thus, for a Pareto optimum, all the above seven marginal conditions must be satisfied

CHECK YOUR PROGRESS I

Q1. What is Pareto optimality ? What are the Marginal conditions of Pareto optimality?

____

____

____

____

Q2 With the help of Edgeworth Box Diagram, discuss how a firm improves its welfare in movement from off the contract curve to on the contract curve?

____

____

____

____

Q3 What is a contract curve? Why is it called conflict curve ? Is it possible to identify only on the basis of 'Pareto optimality' the best point on a contract curve ?

____

____

____

____

15.4 SOCIAL WELFARE FUNCTION AND THEORY OF SOCIAL CHOICE

The concept of social welfare function was propounded by A. Bergson in his article "A reformulation of Certain Aspects of Welfare Economics" in 1938. Prior to it, various concepts of social welfare had been given by different welfare theorists but they failed to provide a satisfactory solution to the problem of maximization of social welfare. Bentham talked of welfare in terms of 'the greatest happiness of the greatest number.' Neo-Classical welfare theorists discussed the problem of social welfare on the basis of cardinal measurability of utility and Interperson of utility. Analysis of Pareto optimality maximizes social welfare by satisfying various marginal conditions of production, distribution and allocation of resources among products. But unfortunately they are not fulfilled due to the existence of various externalities and imperfections in the market. Moreover, Pareto optimality analysis fails to measure the change in social welfare resulting from any change which benefits one section-of societyand harms the other. Compensation principle as given by Kaldor-Hicks-Scitovsky attempts to measure the changes in social welfare resulting from such economic changes which harm some and benefit others through hypothetical compensating payments. Compensation theorists claimed to give a value free objective criterion based on ordinary concept of utility but, as discussed earlier, these criteria are not operationally feasible.

By providing the concept of social welfare function Bergson and Samuelson have attempted to provide a new approach to Welfare Economics. They have put forward the concept of social welfare function that considers only the ordinal preference of individuals. They agree with Robbin's view that interpersonal comparison of utility involves value judgment but they assert that without making some value judgments, economists cannot evaluate the impact of changes in economic policy on social welfare. Thus, according to them, Welfare Economics cannot be separated from value judgments. According to them, Welfare Economics is essentially a normative study. But the approach to study it must be scientific despite the fact that the use of value judgment in it is unavoidable.

15.4.1. The Classical Social Welfare Function

Before making a detailed study of the Bergson-Samuelson's social welfare function, a brief explanation of some alternative welfare functions will be useful, as it will bring the different ethical judgments made in them. An important social welfare function was put forward by Bentham, Pigou and Marshall. According to them, social welfare is the sum of cardinal utilities obtained by all members of a society. In algebraic form, the classical social welfare function can be written as :

$$W = U_1 + U_2 + \ldots\ldots\ldots\ldots + U_n$$

where W denote social welfare, U_1, U_2, etc represent the cardinal utilities of the individual members of the society.

The goal of a society is to maximize social welfare, that is, the aggregate of the utilities of the individuals comprising the society. In the classical welfare function it is further assumed that the law of diminishing marginal utility applies to money income. Given this, maximum social welfare (utility) will be achieved if income is so distributed that marginal utility of income is equal for all individuals in the society. The neo-classical economists make further assumption that various individuals have the same tastes and therefore same capacity for satisfaction with the result that their utility functions are alike. Withthese assumptions, according to classical welfare function, maximization of social welfare is achieved only with equal distribution of income.

A Critique. There are many limitations of classical welfare function and is therefore not accepted by modern economists. This is for two reasons. First, classical welfare function is based on cardinal measurement of utility which has been rejected by the modern economists. In the opinion of modern economists, utility is an ordinal concept and cannot be measured cardinally. Second, classical welfare function assumes that utilities derived by individuals from consumption of goods can be compared interpersonally and can also be added up. This is also not accepted by modern economists who think that interpersonal comparison of utility is not scientific and therefore cannot be validly made. Let us elaborate on this point. Suppose Government adopts a fiscal policy which transfers income from an individual A to B so that A becomes better off and B worse off. According, to neo-classical economists, we can not only measure cardinally the gain in utility by A and loss in utility by B but also can compare the gain and loss of utility by the two persons. However, for making such interpersonal comparison of utility, value judgment is needed because there is no scientific and objective way of making such comparison. Various scholars can make different ethical judgments to evaluate whether gain in utility of A is greater or less than the loss of utility of B.

Finally, classical welfare function makes an important ethical assumption that social welfare is the sum of cardinal utilities of individuals comprising the society. This implies that same weight is given to a drunkard as a learned professor of Economics. This ethical assumption has been considered as invalid by modern economists. Therefore, policy recommendations based on classical social function, such as equal income distribution and progressive income taxation are unacceptable unless they are justified and supposed by other methods.

15.4.2. Pareto's Social Welfare Function

We have already critically examined Pareto's criterion for social welfare and the concept of Pareto optimality based on it. Here, a passing reference to Pareto's social welfare function is called for as it will help in comparing it with other social welfare functions. According to Pareto's welfare function, when one individual is made better off with no else becoming worse off social welfare increases, and, on the contrary, when one individual is made worse off without any one becoming better off, social welfare decreases. Further, maximum social welfare is attained in an organization when no one can be made better off without any one being made worse off, that is, when no one's utility can be raised without reducing the utility of others. This maximum social welfare situation is, also called Pareto optimality or economic efficiency.

It should be noted that in Pareto's welfare function, no interpersonal comparison of utility is made. Only a broad value judgment is made that it is a good thing to make one or some persons better off without making anybody worse off. Pareto's welfare function has been criticized on the ground that it is of limited operational significance because as a result of reorganization and adoption of new economic policies, some people become better off and others worse off and therefore in this case on the basis of Pareto criterion it cannot be said whether social welfare increases, decreases or remains the same. Pareto's welfare function and the concept of Pareto optimality based on it has recently come in for severe criticism by Prof. Amartya Sen. Sarcastically he writes, "A state can be Pareto optimal with some people in extreme misery and others rolling in luxury, so long as the miserable cannot be made better off without cutting into the luxury of the rich"[1]. Prof. Sen laments that Pareto's welfare function pays no attention to distributional considerations of utility or welfare.

He deplores that Pareto's welfare optimum is rooted in self-seeking behaviour of individuals whichTo quote him again, "In the small box to which Welfare Economics got confined with Pareto optimality as the only criterion of judgment and self-seeking behaviour as the only basis of economic choice, the scope of saying something interesting in Welfare Economies became exceedingly small.

15.4.3 Maximin Or Rawlsian Social Welfare Function

Another important social welfare function has been proposed by the noted philosopher John Rawls[2]. Rawls begins his welfare analysis by considering a society as being in an initial position in which no individual knows what his final utility position will be. The problem he poses is that what type of welfare criterion would be adopted by the society when it is in such an initial position where everybody has to behave under uncertainty about how the welfare criterion chosen will ultimately affect his utility or welfare. Assuming that individuals are risk averse he asserts that such a welfare criterion will be chosen that departure from perfect equality would be made only when with unequal distribution-of utilities, the worst of individual is actually better off than under equality. Rawlsian social welfare function can be written as

$$W(U_1, U_2, \ldots\ldots, U_n) = \min (U_1, U_2, \ldots\ldots, U_n)$$

Thus, this social welfare function implies that social welfare of resource allocation depends only on the worst off individual, that is, person with minimum utility.

15.4.4. Bergson-Samuelson Social Welfare Function

Social welfare function is an ordinal index of society's welfare and is a function of the utility level of all individuals constituting the society. Bergson-Samuelson Social Welfare Function can be written in the following manner :

$$W = W(U_1, U_2, \ldots\ldots, U_n)$$

Where W represents the social welfare $U_1, U_2, \ldots\ldots, U_n$ represent the ordinal utility indices of different individuals of the society. The ordinal utility index of an individual depends upon the goods and services he consumes and the magnitude and kind of the work he does and the leisure he enjoys.

Social welfare function and value judgments. The important thing to note about Bergson Samuelson social welfare function is that in its construction explicit value judgments are introduced. Value judgments determine the form of the social welfare function; with a different set of value judgment, the form of social welfare function would be different. Social welfare function is thus not unique; its form varies with different value judgments. Value judgments are essentially ethical notions which are introduced from outside Economics. The value judgments required to construct social welfare function may be obtained through democratic process through voting by individuals or it may have to be imposed on the society in a dictatorial manner. Whether the form of social welfare function is decided through democratic method of voting and majority rule or it is arrived at through consensus, the form of social welfare function depends upon the value judgments of those who promulgate them since it expresses their views regarding the effect which the utility level of each individual has on the social welfare. In the words of Prof. Scitovsky, "The social welfare function can be thought of as a function of each individual's welfare which in turn depends both on his personal well being and on his appraisal of the distribution of welfare among all members of the community".

Since the value judgments required for the formulation of social welfare function are not of the economist himself and instead they are introduced from outside Economics they are not obtained through any scientific method. It has been claimed that social welfare function has solved the basic problem of Welfare Economics, since it thinks unnecessary for the economists themselves to make value judgments concerning what is a desirable distribution of welfare as between individuals constituting the society. In other words, economist need not himself decide about what is the most desirable distribution of welfare. He can take value judgments regarding distribution as given from outside Economics.

Social Welfare Function is Individualistic. Bergson's social welfare function is called individualistic social welfare function because it reflects individual's own evaluation of his welfare rather than of anyone else. Further individual's welfare or utility depends on economic variables such as goods and services consumed and leisure enjoyed which have a direct effect on individual's own welfare. The ordinal utility level of an individual is function of services and not of others. Moreover, the utility level of an individual depends on his own evaluation of desirability or usefulness of different goods and services consumed which depends upon his tastes. An individual may derive more utility from the consumption of liquor whereas another individual may derive verynominal utility or no utility at all from it.

The social welfare function and its form depends upon the value judgments of the person or institution whom the society has authorized to decide. The authorized person or institution may be anybody, but for true value judgments regarding the social welfare he must be unbiased because changes in social welfare will depend upon his value judgments. "These judgments as to what constitutes justice and virtue in distribution may be those of the economist himself or those set up by the legislature, by some other governmental authority or by some other unspecified person or group. A social welfare function can be attained by common consensus or it may be forced upon the society by a dictator."

In modern age of democratic governments, people elect their representatives who constitute the Government. The political party which has a majority forms the Government and rules the country. The representatives of Government formed by the majority rule formulate various policies on the basis of value judgments and it is expected of it that all the policy decisions by the Government will aim at maximizing social welfare rather than maximising the welfare of an individual or a particular section of the society.

Bergson and Samuelson expressed the view that all value judgments used to construct the social welfare function must be consistent which implies that if in a given situation A is preferred to B and B is preferred to C, then A must be preferred to C. This is nothing new to students of Economics as this is the well known assumption of transitivity in social choice among various alternatives.

15.4.5. Representation of Bergson-Samuelson Social Welfare Function through social indifference curve:- We can explain the social welfare function with the help of social indifference curve or welfare frontiers. Let us assume a society of two persons.in such a case social welfare function can be reprensted with the help of social indifference curves.

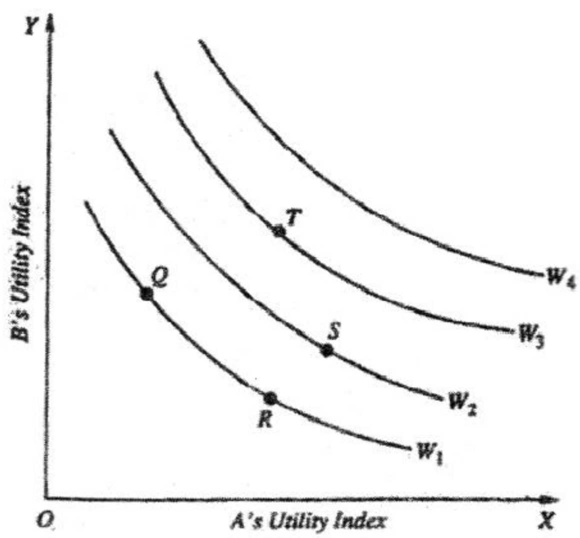


Fig. 15.8

In Fig. 15.8. the utilities of individuals A and B have been represented on the horizontal and vertical axes respectively. W_1, W_2, and W_3, are the social indifference curves representing successively higher levels of social welfare. A social indifference curve is a locus of the various combinations of utilities of Aand B which results in an equal level of social welfare. The properties of social indifference curve are just like those of individual consumer's indifference curves. Given a family of social indifference curves, the effects of a proposed change in policy on social welfare can be evaluated. In terms of Fig. 15.8, any policy change that moves the economy from Q to T is an improvement. Similarly, a movement from Q to S or from R to S also represents an improvement in social welfare and a movement from T to Q or T to S represents a decrease in social welfare. A movement along the same social indifference curve represents no change in level of social welfare. It is worth noting that the considerations of fairness and equity are incorporated into the social welfare function and are reflected in the shapes of social indifference curves. Thus, it is only on the basis of social welfare function incorporating value judgments regarding distribution of welfare among individuals that we are able to say whether or not society as whole has become better off when A's utility increases and B's utility decreases. Thus in Figure 15.8. if a change causes the two individuals A and B to move from point R to Q, it implies that individual B has become better off and A has become worse off, that is, B's utility increases while that of decreases. But since points R and Q lie on the same social indifference curve, social welfare remains the same. This means that in movement from point R to Q, loss of utility by A has been evaluated to be equal to the gain in utility by B. Thus in the construction of social welfare function curve, interpersonal comparison of utility and therefore value judgments have been made.

Similarly, if a change is effected which shifts the individual from point Q to point S in social welfare function depicted in Fig. 15.8. social welfare increases but with this change B's utility has declined and A's utility has increased. Thus, in the evaluation of social welfare, gain in utility of A has been adjudged to be of greater value as compared to the loss of utility of B, when the two move from point R to point S. Thus, it is now clear that in the construction of social welfare function curves as shown in Fig.15.8, inter-personal comparison of utility and value judgments have been made.

Analysis of Pareto optimality failed to provide unique optimum solution' which represents maximum social welfare. There are a large number of solutions which are optimum on the basis of Pareto criterion. In terms of Edgeworth box diagram, every point on the contract curve represents the optimum position. In terms of Grand Utility Possibility frontier, all points on it are Pareto optimal or economically efficient. But Pareto criterion does not tell us the best of them. Thus, Paretian analysis leaves us with a lot of indeterminacy in the choice of maximum social welfare point. Now, the significance of social welfare function is that it enables us to obtain a unique optimum position regarding social welfare. This unique optimum position is best of all the Pareto optima and therefore ensures the maximum Social welfare. By including the concept of grand utility possibility frontier along with Bergson-Samuleson social welfare function we have been able to obtain a unique optimum position or maximum social welfare position.

It should be noted that, grand utility possibility frontier is a locus of the various physically attainable utility combinations of two persons when the factor endowments, state of technology and preference orders of the individuals are given. In other words, every point on the grand utility possibility curve represents the position of Pareto optimality or economic efficiency with regard to the allocation of the products among the consumers, allocation of factors among different products and the direction of production. Thus, every point on the grand utility possibility curve represents a Pareto optimum or economic efficiency as judged criterion and as we move from one point to another on it the utility of one individual increases while that of the other falls. As explained above, social indifference curves depicting social welfare function obtained on the basis of value judgment regarding fairness or equity, in distribution of welfare.

Thus, welfare analysis by combining grand utility possibility frontier showing Pareto optimality or economic efficiency with the concept of social welfare function representing equitable distribution as obtained by value judgments made enables us to arrive at a solution for maximum social welfare which combines efficiency with equity.

Here we can discuss the concept of 'Point of Constrained Bliss'. Social Welfare is maximised at the point where the grand utility possibility frontier touches the highest possible social indifference curve or Isowelfare line. This is the point of constructed bliss or the maximum attainable social welfare.

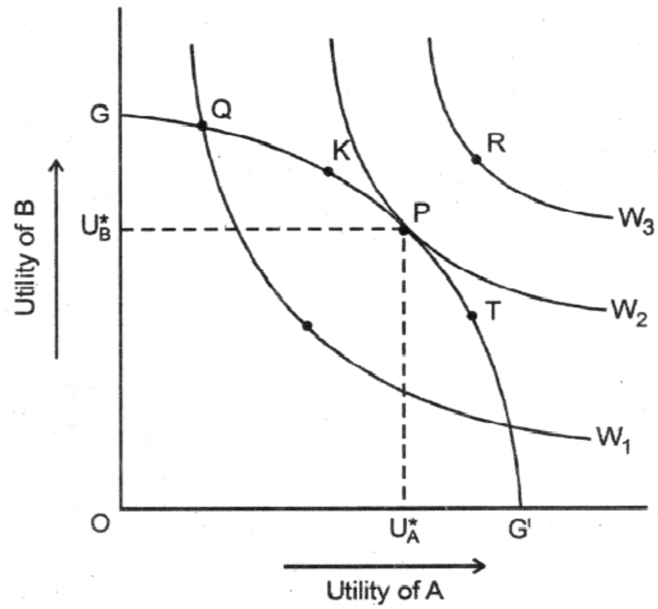


Fig. 15.9 Source : Advanced Micro Economic Theory by Misra & Puri

As we can observe from the fig. above, a point like 'S' need not be considered as it does not lie on the grand utility possibility frontier. But points like Q, K, T and P on the curve are all Pareto optimum. 'R' is unttainable. As 'P' is the point of tangency of the GUPF with the highest possibility curve, it is the point of 'Constrained Bliss'.

Check Your Progress II

Answer the questions in the space provided.

Q1 What are the limitations of classical welfare function ?

____

____

____

____

Q2 What are the main highlights of Rawl's social welfare function ?

Q3 What is the role of value judgements in Bergson-Samuleson Social Welfare function ?

Q4 Why is Bergson's Social welfare function called individualistic ?

Q5. Explain Social welfare function through social indifference curves.

15.5. WELFARE ECONOMICS : COMPENSATION PRINCIPLE

Pareto laid the foundation of the modern Welfare Economics by formulating the concept of social optimum which is based on the concept of ordinal utility and is free from interpersonal comparisons of utilities and value judgments. He aimed at formulating a value-free objective criterion designed to test whether a proposed policy change increases social welfare or not. Pareto criterion states simply that an economic change which harms no one and makes someone better off indicates an increase in social welfare. Thus, this criterion does not apply to those economic changes which harm some and benefit others. In terms of Edgeworth Box diagram, Pareto criterion fails to say as to whether or not social welfare increase as movement is made in either direction along the contract curve because it rejects the notion of interpersonal comparison of utility. Every tangency point of the two indifference curves on the contract curve represents a Pareto optimum. There is thus no unique Pareto optimum position. This criterion does not tell us about changes in the level of social welfare if one moves on the contract curve from one tangency point to another because such movement harms one and benefits the other. Thus the analysis of welfare in terms of Pareto optimality leaves a considerable amount of indeterminacy, for there are numerous Pareto optimum points on the contract curve. Economists like Kaldor, Hicks and Scitovsky have made efforts to evaluate the changes in social welfare resulting from any economic reor others. These economists have sought to remove indeterminacy in the analysis of Pareto optimality. They have put forward a criterion known as the 'compensation principle' on the basis of which they claim to evaluate those changes in economic policy or an organization which makes some individual better off and others worse off. The 'compensation principle' is based the following assumptions.

Assumptions

1. The satisfaction of an individual is independent of the others and he is the best judge of his welfare.
2. There exist no externalities of consumption and production.
3. The tastes of the individuals remain constant.

4. The problem of production and exchange can be separated from the problems of distribution. Compensation principle accepts the level of social welfare to be a function of the level of production. Thus, it ignores the effects of a change in distribution on social welfare.

5. Utility can be measured-ordinally and interpersonal comparisons of utility are not possible.

Given the above assumptions, a criterion of compensation principle can be discussed. Kaldor, Hicks and Scitovsky have claimed to formulate a value-free objective criterion of measuring the changes in social welfare with the help of the concept of compensating payments.

15.5.1 Kaldor-Hicks Welfare Criterion : Compensation Principle

Nicholas Kaldor was the first economist to give a welfare criterion based on compensating payments. Kaldor's criterion helps us to measure the welfare implications of a movement in either direction on the contract curve in terms of Edgeworth box diagram. According to Kaldor's welfare criterion, if a certain change in economic organization or policy makes some people better off and others worse off, then that change will increase social welfare if those who gain from the change could compensate the losers and still be better off than before. In the words of Prof. Baumol, "Kaldor's criterion states that a change is an improvement if those who gain evaluate their gains at a higher figure than the value which the losersset upon their losses." Thus, if any policy change benefits any one section of the society (gainers) to such an extent that he is better off even after the payment of compensation to the other section of the society (losers) out of the benefits received, then that change leads to increase in social welfare. In Kaldor's own words, "In all cases.... where a certain policy leads to an increase in physical productivity and thus in aggregate real income.... it is possible to make everybody better off without making anybody worse off. It is quite sufficient.... to show that ?even if all those who suffer as a result are fully compensated for their loss, the rest of the community will still be better off than before." Prof. J.R. Hicks supports Kaldor for employing compensation principle to evaluate the change in social welfare resulting from any economic reorganization that benefits some people and harms the others. This criterion states that "if A is made so much better by the change that he could compensate B for his loss and still have something left over, then the reorganization is unequivocal improvement. In other words, a change is an improvement if the losers in the changed situations cannot profitably bribe the gainers not to change from the original situations. Hicks has given his criterion from the losers' point of view, while Kaldor had formulated his criterion from gainers' point of view. Thus the two criteria are really the same though clothed in different words. That is why that are generally called by a single name Kaldor-Hicks criterion.

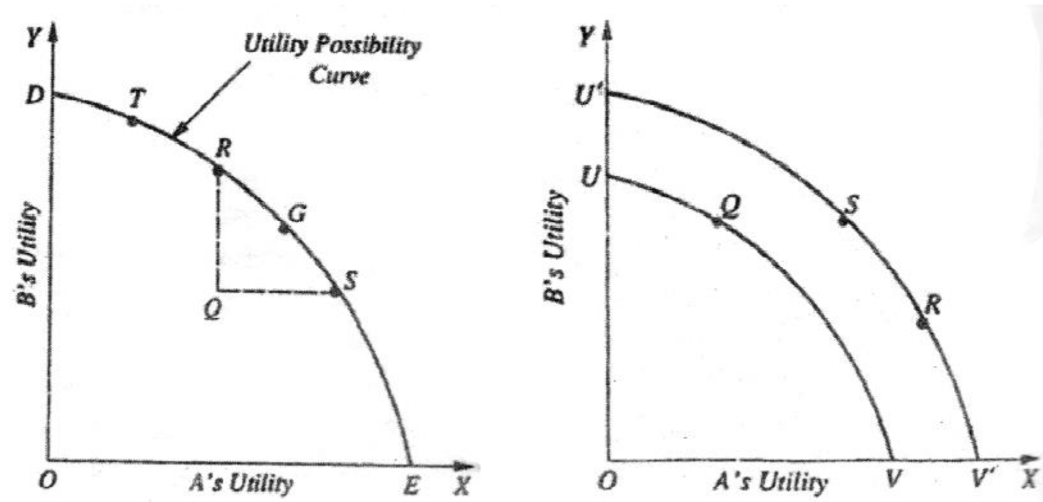


Fig 15.10 (a) Fig. 15.10 (b)

Kaldor-Hicks criterion can be explained with the help of the utility possibility curve. In figure 15.10(a), ordinal utilities of two individuals A and B are shown on X and Y axes respectively. DE is the utility possibility curve which represents the various combinations of utilities obtained by individuals A and B. As we move downward on the curve DE, utility of B decreases while that of A increases. Suppose the utilities obtained by A and B from the distribution of the income or output between them are represented by point Q inside the utility possibility curve DE. Let us assume that as a result of some change in economic policy, the two individuals move from point Q to point T on the utility possibility curve DE. As a result of this movement, utility of individual B has increased while the utility of A has declined, that is, B has been become better off and A has become worse off than before. Therefore, this movement from point Q to point T cannot be evaluated by means of Pareto criterion. Of course, points such as R, G, S or any other point on the segment RS of the utility possibility curve DE are socially preferable to point Q on the basis of Pareto criterion.

But since the movement from Q to T involves interpersonal comparison of utility, it cannot be determined whether or not social welfare increases on the basis of Pareto criterion. However, the compensation principle propounded by Kaldor-Hicks enables us to say whether or not social welfare has increased as a result of movement from Q to T. According to Kaldor-Hicks criterion, we have to see whether the individual B who gains with the movement from position Q to position T could compensate the individual A who is loser and still be better off than before. Now, it will seen from the Figure 15.10(a) that utility possibility curve DE passes through points R,G and S. This means that by mere redistribution of income between the two individuals, that is, if the individual B gives some compensation to individual A for the loss suffered, they can move from position T to position R. It is evident from the figure that at position R, individual A is as well off as at the position Q but individual B is still better off as compared to the position Q. It means that due to a policy change and consequent movement from position Q to position T, the gainer (individual B) could compensate the loser (individual A) and is still better off than at Q. Therefore, according to Kaldor-Hicks criterion, social welfare increases with the movement from position Q to position T, as from T they could move to the position R through mere redistribution of income.

It is noteworthy that, according to Kaldor-Hicks criterion, compensation may not be actually paid to judge whether or not welfare has increased. It is enough to know whether the gainer could compensate the loser for the loss in his welfare and still be better off. Whether redistribution of income (that is, payment of compensation) should be actually made following the change in policy is left for the Government to decide. If it is possible for the gainer to compensate the loser and still be better off, economists say that the social welfare has increased. It may be noticed that gainers can compensate the losers and still be better off only when the change in economic policy leads to the increase in output or real income. That is why Kaldor and Hicks claim that they have been able to distinguish between changes in output and changes in distribution. When their criterion is satisfied by any change in the situation, it means that the economy has moved to a more efficient position and as a result social welfare can be said to have increased. Now, whether redistribution of income is actually made through payment of compensation by the gainers to the losers, according to them, is different matter. Now, the implication of Kaldor-Hicks criterion becomes more clear if through redistribution, the position of the two individuals changes from T to G (see fig. 15.10 (a). It is quite manifest that at position G, both the individuals A and B are better off than at the position Q. Thus the position T to which the two individuals moved as a result of certain change in economic policy is superior to initial position Q from the viewpoint of social welfare, since from position T movement can be made, merely through redistribution of income to position G where both are better off as compared to position Q.

It may be noted that in the situation depicted in figure 15.5 (a), the change in economic policy brings about a movement from a position inside the utility possibility curve to a point on it. Now let us see what happens to social welfare if as a result of the adoption of a certain economic policy, the utility possibility curve moves outward and the two individuals move to a point on a higher utility possibility curve. It can be shown that, according to Kaldor-Hicks criterion, such a movement causes an improvement in social welfare. Consider figure 15.(b). UV is the original utility possibility curve and Q represents the position at which the two individuals are initially placed. Now suppose utility possibility curve shifts outward to the new position U´V´ where the two individuals are placed at point R. In the movement from Q on UV to R on U´V´, the utility of A has increased and that of B has declined. But position R denotes greater social welfare on the basis of Kaldor’s criterion when compared to the position Q on the utility possibility curve UV, because with U´V´ as the utility possibility curve it is possible to move, through mere redistribution of income from position R to position S where the individual B has been fully compensated for his loss of utility, the individual A is still better off as compared to position Q. To conclude, any change in the economy that moves the individuals from a position on a lower utility possibility curve to a position on a higher utility possibility curve increases social welfare.

15.5.2 Scitovsky's Paradox

Scitovsky pointed out an important limitation of Kaldor-Hicks criterion that it might lead to contradictory results. He showed that if in some situation position B is shown to be an improvement over position A on Kaldor-Hicks criterion, it may be possible that position A is also shown to be an improvement over B on the basis of the same criterion. For getting consistent result when position B has been revealed to be preferred to position A on the basis of a welfare criterion, then the position A must not be preferred to position B on the same criterion. According to Scitovsky, Kaldor-Hicks criterion involves such contradictory and inconsistent results. Since Scitovsky was the first to point out this paradoxicial result in Kaldor-Hicks criterion, it is known as 'Scitovsky Paradox".

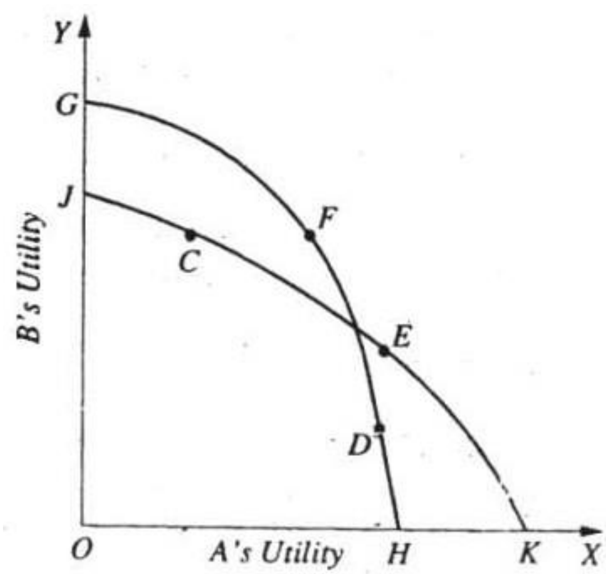


Fig. 15.11. Scitovsky Criterion

How Kaldor-Hicks criterion may lead to contradictory results in some situation is depicted in Figure 15.11. In this figure, JK and GH are the two utility possibility curves which intersect each other. Now suppose that the initial position is at point C on JK. Further suppose that due to a certain policy change, utility possibility curve changes and takes the position GH and the two individuals find themselves at position D. Position D is superior to position C on the basis of

Kaldor-Hicks criterion because from position D movement can be made through mere redistribution to position F where individual B has been fully compensated but individual A is still better off as compared to the original position C. Thus movement from position C to position D satisfies Kaldor-Hicks Criterion. But, as been pointed out by Scitovsky, reverse movement from position D on the new utility possibility curve GH to the position C on the old utility possibility curve JK also represents an improvement on Kaldor-Hicks criterion, that is, C is socially better than D on Kaldor-Hicks criterion.

This is because from the position C movement can be made by mere redistribution of income to position E on the utility possibility curve JK on which position C lies and which also passes through the position E. And, as will be observed from fig.15.11, at position E that while A is as well off as at position D, individual B is still better off than at D. We thus see that the movement from position C to the position D due to a policy change is passed by the Kaldor-Hicks criterion. This implies that D is socially better than C on this criterion and C is socially better than D on the same criterion. So Kaldor-Hicks criterion leads us to contradiction and inconsistent results. It is Worth being noted that these contradiction results are obtained by Kaldor-Hicks criterion when following a policy change new utility possibility curve intersects the former utility possibility curve. After bringing out the possibility of contradictory results in Kaldor-Hicks criterion Scitovksy formulated his own criterion which is generally known as Scitovsky's Double Criterion.

15.5.3. Scitovsky's Double Criterion Of Welfare

To rule out the possibility of contradictory results in Kaldor-Hicks test, Scitovsky formulated a double criterion. Hicks test and also the fulfillment of the reversal test. It means that a change is an improvement if the gainers in the changed situation are able to persuade the losers to accept the change and simultaneously losers are not able to persuade the gainers to remain in the original situation. Scitovsky double criterion can also be explained with the help of utility possibility curve.

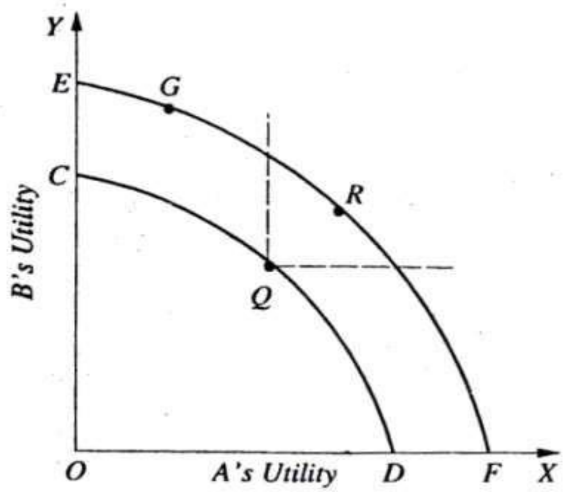


Fig. 15.12. Scitovsky's Double criterion

In Figure 15.12, CD and EF are the two utility possibility curves which do not intersect each other at any point. Suppose there is a change from position Q on the utility possibility curve CD to point G on EF as a result of the adoption of a new economic policy. Such a movement is an improvement on Kaldor-Hicks criterion because G lies on the utility possibility curve EF passing through point R. From the position G, movement can be made to the position R simply by redistributing the income between the two individuals. And R is better than Q because the utility of both the individuals is greater at R as compared to the position Q. Thus the Kaldor-Hicks criterion is satisfied and therefore change from Q to G will increase social welfare. Now, let us see, what happens to the reversal test. It must also be satisfied, if the scitovsky double test is to be fulfilled. That is a movement from the position G back to the original position Q must not be passed by Kaldor-Hicks criterion if Scitovsky's reversal test is to be satisfied. It is evident from figure 15.12 that from position G we cannot move to any other position on the utility possibility curve CD merely through redistribution of income which is sociallybetter than G (that is, which raises utility of either A or B, the Utility of the other remaining constant or which raise the utility of both). We thus see that while movement from position Q to G is passed by Kaldor-Hicks criterion, reverse movement from position G to Position Q is not passed as per Kaldor-Hicks criterion. Hence, in Figure 15.12. the movement from the position Q to G satisfied scitovsky's criterion. Thus when the two utility possibility curves are non-intersecting and change involves movement from a position on a lower utility possibility curve to a position on a higher utility possibility curve, the change raises social welfare on the basis of Kaldor-Hicks-Scitovsky criterion. This happens only when a change brings about increase in aggregate output or real income.

15.5.4 A Critique Of The Compensation Principle

The compensation principle as developed by Kaldor, Hicks and Scitovsky, has been a topic of much discussion in Welfare Economics since 1939. Prof Kaldor was the first to give a criterion to judge the changes in social welfare when an economic change benefits some people and harms the others. Later Hicks also supported this criterion in 1940, though he put it in different words; Scitovsky tried to improve upon the Kaldor-Hicks criterion by formulating his own double criterion. These welfare economists have claimed that they have succeeded in developing a welfare criterion based on ordinal concept of utility and also which is free from any value judgments. But compensation principle has been bitterly criticized by the various welfare economists.

First, Little has pointed out that Kaldor did not formulate a new welfare criterion at all because he assumed welfare to be a function of increase in production or efficiency irrespective of the changes in distribution. Thus, according to Little, Kaldor has given only a definition of increase in wealth or increase in efficiency. Kaldor himself has interpreted the compensation principle in this sense as he says that, "when the production of wealth goes up, some income distribution could be found which makes some people better off, and no one worse off than before". However, as desired income distribution via compensation is only hypothetical, therefore, according to Little, it is not a welfare test but a definition of 'economic efficiency' in terms of over-compensation.

Second, compensation principle is not free from value judgments as is claimed by its propounders. It involves implicit value judgments. Prof. Baumol and Little are of the opinion that the contention of Prof. Kaldor that the changes which enable the gainers to compensate the losers and still be better off are good changes, is itself a value judgment. According to Little, to say that a policy which meets the Kaldor-Hicks criterion increases the output or "efficiency" of society is in effect to recommend it. According to him, Kaldor and Hicks have coined a definition of "efficiency" whose implicit ethical implications or value judgments will hardly find favour with many people. Compensation is after all only hypothetical; it is consistent with making the poor yet pooler. Thus according to little, if the value judgments implicit in Kaldor-Hicks criterion are made explicit, then, the claim of Kaldor and Hicks that they have discovered a criterion of detecting increases in wealth, production or efficiency free from value judgments is hardly acceptable.

Third likewise, Baumol is also of the view that Kaldor-Hicks criterion is based upon unacceptable implicit value judgments. "By using a criterion involving potential money compensations, they set up a cealed interpersonal comparison on a money basis". If an individual A evaluates his gain from a change worth Rs. 500 whereas another individual B evaluates his loss due to that economic change at Rs. 75, we cannot conclude that social welfare has increased; for if the loser is poor and the gainer a rich one, it may be possible that loss of satisfaction of the poor from Rs. 75 is far greater than the addition to the satisfaction of the rich by Rs. 500 because the marginal significance of one rupee to a poor is far greater than that of the rich. Thus without actual compensation, the change would mean a major loss of welfare to the poor individual B and a trivial gain of welfare to the rich individual A even if it passes the Kaldor criterion with flying colours. To quote Baumol again, "The Kaldor and Scitovsky criteria have thus ducked the basic problem of the interpersonal comparison required to evaluate a policy change which harms X but aids Y. They duck it by saying implicitly that the recommendation should be based on X's and Y's relative willingness and ability to pay for what they want".

Fourth, Kaldor-Hicks have claimed that through compensation principle they have been able to separate a production change from the distribution change bywhich it is accompanied. For instance, as a result of a policy change output of coca-cola increases and that of Whisky decreases. Now, if individual X prefers Cola but Y prefers whisky, the question whether there has occurred any increase in production is inseparably connected with the distribution of these beverages between X and Y. In many cases, therefore, it is difficult to say whether or not production has increased without considering how the output or real income is being distributed.

Moreover, Kaldor and Hicks think that the level of production is the main determinant of social welfare and the distribution is a secondary one. But this is quite untenable. This is because welfare has dual aspect : absolute and relative. People are dissatisfied not only because they are poor but more because other people are very rich. If all the people would have been poor they would not have been dissatisfied much. Thus, a lower total output equitably distributed ensures greater social welfare than larger output, inequitably distributed.

Fifth, Gorman is of the opinion that Scitovsky's double criterion can avoid contradictory results only when the choice is to be made from only two positions. This criterion does not help us if the choice of a position is to be made from more than two possible positions. Moreover, Prof. Samuelson refers to a more basic drawback in Scitovsky's double criterion. Scitovsky's criterion is satisfied if utility possibility curve in one situation lies above the utility possibility curve of the other situation in the neighbourhood of both actual observed points. But Samuelson asserts that "the complete test for an increase in welfare requires the utility possibility curve after the change to be nowhere inside the corresponding utility possibility curve before the change and at least somewhere outside it. In other words, every welfare position attainable before the change must be attainable after together with at least one position that was not attainable before."

Sixth, Prof. Baumol, Little and Arrow point out another major flaw in compensation principle that it does not envisage social welfare. This principle proves the social desirability of change in the social state on the basis of the criterion that gainers could compensate the losers and still be better off than before. These critics are of the opinion that policy changes which would increase social welfare when accompanied by actual compensation need not lead to improvement in social welfare if compensation is not actually made. Dr. Rothenberg has given a very good example to illustrate this. He supposes an initial social state in which a firm adopts a new invention and as a result the cost of production of the firm is reduced but it throws the competitors out of industry and the workers become unemployed. Let us suppose that the gainer firm from the invention can compensate losers out of its increased income and still be better off. If the compensation is not actually made in the changed situation, social welfare will decrease as the welfare loss suffered by the workers rendered unemployed will be very large indeed. As a matter of fact, there is no guarantee that compensation will be actually made in such cases. Thus, so long as compensation is hypothetical, it might make the rich richer and the poor poorer and therefore reduce social welfare.

It follows from above that a basic flaw in Kaldor-Hicks compensation principle is that it refers to potential welfare rather than actual welfare since it does not envisage that compensation should be actually made. In the absence of actual opinion one cannot say- whether or not actual social welfare has increased as a result of a certain policy change unless one is prepared to make some value judgments. Therefore, making value judgments, especially that concerns distribution of income or welfare, is quite indispensable in Welfare Economics. And economists should not avoid making those value judgments which are widely accepted by the people.

It may also be noted that if compensation is actually made then Kaldor-Hicks criterion is quite unnecessary, for in that case only Pareto criterion will be sufficient to judge the effect of a policy change on social welfare.

Seventh, compensation principle does not take into account the external effects on consumption and production. The exponents of compensation principle are of the opinion that an individual's welfare depends solely upon his own level of production and consumption and is not affected by the production and consumption activities of the others. But this is not a realistic assumption because a person's level of satisfaction (or dissatisfaction) depends to a large extent upon the consumption of goods and service by other persons. A person is more satisfied if his relative economic position in the society is improved. Thus, if an economic change leaves a person where he was before but makes some other individuals better off, he will not feel as well off as in the original situation, that is, his level of welfare will fall. The gains by some individuals from a policy change have usually unfavourable external effects on the welfare position of those whose position is said to have remained unchanged.

Thus compensation principle has been criticized on several grounds, especially for its unacceptable implicit value Judgments. Because of the above limitations of compensation criteria of Kaldor, Hicks and Scitovsky, some economists, especially Bergson, Samuelson and Arrow have propounded a concept of social welfare function, which incorporates explicit value judgments, to evaluate the welfare implications of policy changes and also to find out a unique social optimum.

Check Your Progress– III

Answer the questions in the space provided :

Q1. What are the main assumptions of Kaldor-Hicks compensation principle ?

____

____

____

____

Q2. Discuss with the help of utility possibility curve how is compensation principle an improvement over Pareto's criterion.

____

____

____

____

Q3. What is Scitovsky's paradox ?

____

____

Q4. What is the double criterion of welfare ?

15.6 THE THEORY OF SECOND BEST

We have explained earlier that satisfaction of Pareto criterion and the marginal conditions of Pareto-optimality lead to economic efficiency or maximum social welfare. Applying the Pareto criterion and marginal conditions we found that movement from a point inside the production possibility frontier to a point on it will increase social welfare. We have also discussed above that due to the existence of monopoly or imperfections in some markets or in case of externalities and public goods all marginal conditions of Pareto optimality are not fulfilled.

Now a pertinent question has been raised by Profs. Lipsey and Lancaster in their theory of second best. According to them, it is not possible to satisfy all conditions of Pareto efficiency (or it is not possible to satisfy these conditions in all markets) and consequently best solution, that is, maximum social welfare situation is unattainable. In their theory, they discuss whether or not efforts should be made to achieve the second best position by satisfying the remaining marginal conditions of Pareto optimum or satisfying these conditions in the remaining markets if in some market they are not fulfilled. In their theory of second best they assert that this second best solution will not lead to increase in social welfare. Let us give a concrete example. Suppose a Monopoly exists in one market and the government takes step to make this market competitive. It would appear that social welfare will increase (as price and marginal cost will be equated in this market) irrespective of whether in some other markets competition cannot be enforced and therefore in them Pareto optimal condition, cannot be satisfied. According to the second best theory, this will not lead to the increase in social welfare, as second best solution is not desirable.

The formal proof of the second best theorem is bit complicated and difficult. Therefore, we will present their argument with its graphic representation. Consider Figure 15.13. where production possibility frontier PP? has been drawn on which all points are Pareto efficient. According to Lipsey and Lancaster, it is sometimes better to move inside to the production possibility curve to achieve a higher level of social welfare in case all marginal conditions are not satisfied. To demonstrate

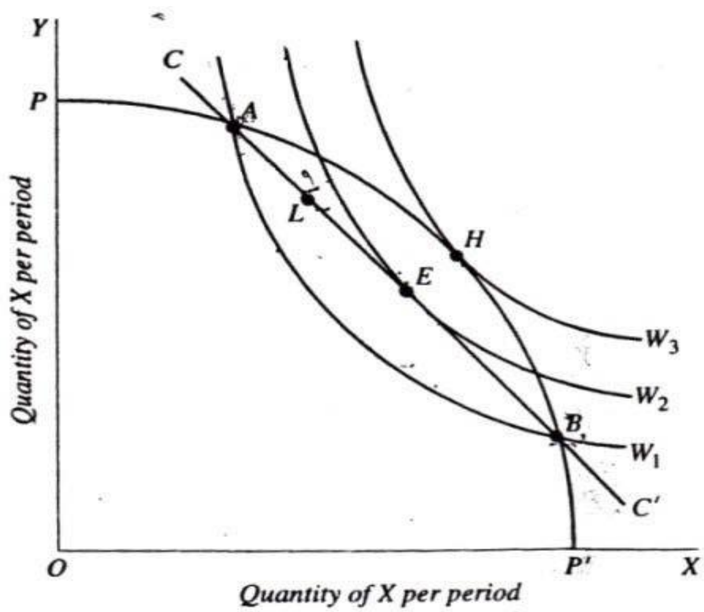


Fig. 15.13. Graphic Demonstration of the theory of second best

this, social welfare curves (also called community indifference curves) have also been drawn in Figure 15.13. like the indifference curves of the individuals. These social welfare curves represent combinations of two products X and Y which yield the same level of welfare to the society. Further, the higher level of a social welfare curve, the higher the level of social welfare. It will be seen from Figure 15.13. that point H at which social welfare curve is tangent to the production possibility curve PP´ shows the maximum social welfare point of satisfaction or satisfaction of all the marginal condition of Pareto-optimality. Now, suppose due to the existence of Monopoly in the markets for two goods, the socially best point H is unattainable. Further suppose that due to the existence of monopolies, only combinations lying on the line CC´ are attainable. Suppose the economy is at present at point L on the attainable line CC´. Now, if Pareto optimality is to be achieved, we can move from a point L which is inside the production possibility curve to point A or B, on production possibility curve PP´, which are also on the attainable line CC´. However, as will be seen from the figure, moving to point, A or B on the production possibility curve would put us on a lower social welfare curve W_1.

Another possibility is to move from point L, to the point E which is interior to the production possibility curve PP´ (which is therefore Pareto inefficient but yields a higher level of welfare as indicated by social welfare curve W_2) as compared to the points A or B. Thus, the theory of second best asserts that when one of the marginal conditions for Pareto-optimality is not satisfied (in the present example, the existence of monopoly does not satisfy the condition of optimum direction of production, namely, MRS_{XY} is not equal to MRT_{XY}), it may be better to violate other marginal conditions of Pareto optimality to achieve maximum possible social welfare.

The theory of the second best has been used to question the desirability of advocating competitive pricing in some particular markets when it is known that Pareto conditions do not hold in other markets. However, the supporters of such piecemeal policies in making some markets more competitive for achieving Pareto optimality argue that as the markets in question are unrelated, it does not matter that for attaining maximum social welfare, in other markets, conditions for Pareto efficiency are not fulfilled. Advocates of the latter policy argue that when markets are relatively independent or unrelated, Marginal Cost pricing increases social welfare. Thus, "the theory of second best applies most strongly when markets are closely related : that is, they either produce complementary goods like bread and butter, or one market is an intermediate supplier of another as in the case of tyre makers supplying automobile producers".

Can Taxation Improve Efficiency ?

Now, a pertinent question is whether taxation is bad because it has undesirable effect in terms of loss of efficiency and excess burden it creates. Our answer is in the negative for two reasons. First, economic efficiency is not the only goal of society. For example, tax on heavy (petrol consuming) cars causes inefficiency as it will raise their prices and reduce their quantity bought and sold. But the purpose or goal of levying tax on these big cars is to conserve energy, that is, reduction in the consumption of petrol. If a society is willing to tolerate some economic inefficiency to achieve the goal of conserving energy, then it is worthwhile to impose tax on big cars using large quantities of petrol.

Taxes and Pollution

But a more important beneficial effect of a tax on social welfare is when it is imposed on the activities which cause pollution of environment. The production of an industry which pollutes air or water causes external diseconomies which harm others but industries do not pay for them. In this case, the Marginal Social cost (MSC) of production differs from Private Marginal Cost (PMC). How in this case imposition of tax on the production of a polluting industry will improve efficiency is illustrated in fig. 15.14.

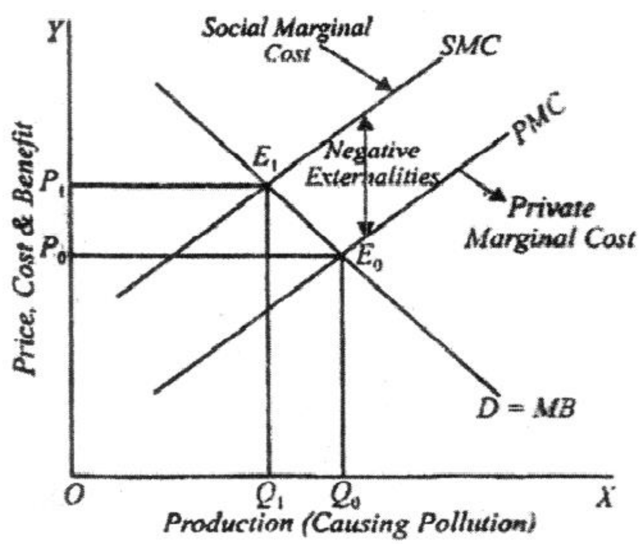


Fig. 15.14.

DD is the demand curve of the product of an industry which reflects marginal benefits (MB) to the society, as it represents how much consumers are willing to pay for different units of output of the product. The curve PMC is the supply curve of the industry based on private cost when it does not pay for the harms it does to others by way of polluting air and water. When negative externalities are added to the private marginal cost, we get social marginal cost curve (SMC). Therefore, SMC is the supply curve incorporating external diseconomies.

Initially, prior to the imposition of tax on the production of polluting industry, quantity of a product produced is Q_0 and price is P_0 at which demand curve and supply curve MC intersect. Now, with the imposition of tax per unit of output produced equal to the negative externalities, it causes the supply curve to shift to SMC. It will be seen from Fig. 15.14. that this supply curve SMC intersects the demand curve DD at point E_1. As a result, price rises to P_1 and production of polluting industry falls to Q_1. In the new equilibrium at output Q_1 and price P_1, marginal benefit (MB) from production of commodity is equal to social marginal cost (SMC). Thus, imposition of tax has achieved allocative efficiency. To conclude, taxes on industries or services which create external diseconomies may improve economic efficiency rather than cause efficiency losses.

15.7 ARROW'S THEORY OF SOCIAL CHOICE

Bergson and Samuelson made significant contributions to Welfare Economics by introducing explicit value judgments in the form of social welfare function. However, Bergson and Samuelson did not deal with the question as to how to get these value judgments or what these value judgments could be for constructing a social welfare function. It was this problem left untouched by Samuelson and Bergson, which was explored by Arrow in his path breaking work "Social choice and individual value". Prof. Arrow pointed out that the construction of social welfare function which reflects the preferences of all individuals constituting a society is an impossible task. The main contention of Arrow is that it is very difficult to set up reasonable democratic procedures for the aggregation of individual preference into a social preference for making a social choice. Arrow has proved a general theorem according to which it is impossible to construct a social ordering which will in some way reflect the individual ordering of all the members of the society.

In the original version of Bergson-Samuelson social welfare function, welfare of an individual was thought to depend solely upon the goods and service consumed by him alone and not by others. Hence, an individual ordering of an alternative 'social states' reflected his tastes and not of others. On the other hand, Arrow has pointed out that individual ordering of social states does not depend exclusively upon the commodities consumed but also on the amounts of various types of collectives such as municipal service, parks, sanitation, errection of statues of famous men etc. He has argued that welfare result of collective activity cannot be evaluated by an individual solely on the basis of his consumption. Instead, individual ordering of social states will depend on his own consumption as well as on the consumption of others in a society. Individual ordering of alternative social states reflects his value judgments which are also called simply value by Arrow. According to Arrow, it is ordering of social states according to the values of individuals as distinct from the individual tastes which should be determined for construction of a valid social welfare function.

15.7.1 Arrow's Conditions of Social Choice

We have seen earlier that the value judgments of a superman or a dictator about social welfare may not be valid due to various types of biases in human mind. As a result, a superman's or dictator's value judgment or value do not truly reflect the social choice. Arrow was the first welfare economist who attempted to lay down reasonable necessary conditions for achieving the social ordering which reflects the desire or the ordering of all individuals of the society. There are many ways in which social choice can be derived. Choice may be made by a dictator or though customs and traditions or by some spiritual or religious head as was done in traditional society or by individuals comprising a society through voting. The problem of social choice is easiest in a dictatorial rule in which all the social choices are made by the dictator and all the individuals comprising the society are compelled to accept it. Similarly, in a traditional society various religious and spiritual rules or customs make the problem of the social choice very easy. No individual can disregard the social choice made by a religious and spiritual head.

But the problem of making a social choice based on individual ordering becomes difficult in a democratic society in which every individual is free to have his own individual ordering of various social states. Now, a pertinent question is as to whether the social choice can consistently be derived from individual orderings. Prof. Arrow has laid down certain necessary conditions which social choice must satisfy in order to reflect the individual orderings. He has laid down the following five conditions which must be met for an acceptable social welfare function. In fact, these conditions reflect the value judgment of Arrow himself.

Condition 1:- Transitivity or consistency. The first condition mentioned by Arrow is that social choices must be consistent or transitive. Transitivity of the social choice implies that if an alternative 'A' is socially preferred to alternative B and alternative 'B' is strictly preferred to alternative 'C', then alternative 'C' will not be socially preferred to alternative 'A'. If alternative 'C' is found to be socially preferred to A, then the condition of transitivity would be violated and the choice would be inconsistent. It may be mentioned that the question of transitivity arises only when the social ordering has the properties of convexity. By convexity, we mean that the various alternatives must be related to each other by preference or by indifference. Thus two alternatives are said to be related or connected if for any pair of alternatives 'A' and 'B', either A is preferred to B or B to A or there is indifference between the two. Thus the condition of transitivity must be found in the social choice because it has been considered by Arrow as a condition for consistent social choice.

Condition 2: Responsiveness to Individual Preferences. The second condition is that social ordering must depict responsiveness to individual preferences. It states that social ranking must respond positively to the individual ranking. This means that the social choices must change in the same direction as the choice of the individuals constituting the society. To quote Arrow, "The social welfare function is such that the social ordering responds positively to alteration in individual value or at least not negatively. "This implies that social choice reflects the values of different individuals of the society and it changes as the individual values change.

Suppose an alternative 'A' is socially preferred to B on the basis of a set of individual orderings. If change occurs in the ordering of individuals so that some individuals still prefer alternative 'A' more strongly than before and no one's preference for it declines, then 'A' must remain socially preferred to B. It is worth mention that this condition would be violated "if there were some individuals against whom society discriminates in the sense that when their desire for some alternative increases relative to other alternatives, social desirability of that alternative is reduced."

Condition 3: The Condition of Non-imposition. The third condition is of 'non-imposition. This states that social choices must not be imposed independently of individual preference. For instance, it implies that if no individual in the society prefers alternative 'B' to alterative 'A' and any one or few other individuals in the society prefer alternative A to alternative B, then society must prefer A to B. This condition implies that the choice of an alternative by the society must satisfy Pareto criterion. This also implies that the soc.ial choice must not be determined by anyone outside the community.

Condition 4: The condition of Non-Dictatorship. The fourth condition is related to existence of non-dictatorship. It states that social choices must not be dictated by any one individual in the community. For instance, 'A' must not be socially preferred to 'B' only because any one individual in the society prefers 'A' to 'B' irrespective of the preferences of other individuals. If this condition is violated, then the individual whose preferences are regarded as social preferences will in fact be a dictator. This condition implies that the social choices must be determined by the democratic method of voting by all individuals rather than dictatorial one of imposition of his will by an individual.

Condition 5: Independence of Irrelevant Alternatives. The fifth condition is of independence of irrelevant alternatives. According to this, social ranking of any two alternatives is determined exclusively by individual ranking of these two alternatives alone and should not be affected at all by individual preferences with respect to other alternatives. In other words, a most preferred alternative out of a given series of available alternatives must be independent of (that is, not affected by) other alternatives, which are not available. Suppose three alternatives, A, B and C are available and society prefers A to B and B to C. If C were no longer available, then this condition implies that it must not be the case that society then prefers B to A. Thus, the social preference of A over B depends only on individual preferences of just these two alternatives, A and B and not on any other alternative which is not immediatelyrelevant.

The above five conditions of Arrow reflect his value judgments and they seem to be quite reasonable set of conditions for making social choices in a free democratic society. However, Arrow has shown that it is impossible to make social choices without violating at least one of the above five conditions. In other words, it is not possible to construct a social welfare function on the basis of individual values that satisfy all the above conditions. Arrow has demonstrated through his general impossibility theorem that when the choice is between more than two alternatives then the individual's voting or expression of preferences for them would lead to inconsistent or contradictory results so that no valid social choice can be made by majority rule. We will explain below Arrow's impossibility theorem.

15.7.2 Arrow's Impossibility theorem

After discussing the social choice and individual values, Arrow proved his famous impossibility theorem in terms of abstract Mathematics and symbolic logic. From this, he concluded that it is not possible to pass from the individual preference to the social preference so as to construct a social welfare function. According to Arrow's theorem, if we exclude the possibility of interpersonal comparison of utility then the only method of passing from individual tastes to social preferences which will be satisfactory and which will be defined for a wide range of sets of individual ordering are either imposed or dictatorial.

The democratic procedure for reaching a social choice or group decisions is the expression of preferences by individuals through free voting. Social choice will be determined by the majority rule. But Arrow has demonstrated through his impossibility theorem mentioned above that consistent social choices cannot be made without violating the consistency or transitivity conditions. The social choice on the basis of majority rule may be inconsistent even if individual preferences are consistent. Arrow first consider a simple case of two alternative social states and proves that in this case group decision or social choice through a majority rule yield a social choice which can satisfy all the five conditions. But when there are more than two alternatives, majority rule fails to yield a social choice without violating at least one of the five conditions. Thus, in case of more than two alternatives, social welfare function based on individual preference cannot be constructed.

Let us illustrate the proof of the impossibility theorem with help of table 1 given below.

Table 1 : Ranking of alternatives by individual and social choice

	X	Y	Z
A	3	2	1
B	1	3	2
C	2	1	3

In this table, three individuals A, B, and C who constitute the society have been shown to have voted for three alternative social states X,Y and Z, by writing 3, against the most preferred alternative, writing 2 for the next preferred alternative and 1 for the least preferred alternative. A glance at the table will reveal that individual A prefers X to Y,Y to Z, therefore X to Z. Individual B prefers Y to Z, Z to X and therefore Y to X. And individual C prefers Z to X, X to Y and therefore Z to Y. It is clear that two individuals A and B prefer Y to Z and also individual B and C prefer Z to X. Thus, the majority (two of the three individuals) prefers X to Y, also Y to Z as well as Z to X. Thus, we see that majority rule leads to inconsistent social choices because on the one hand X has been preferred to Z by the majority and on the other hand Z has also been preferred to X by the majority which is quite contradictory or inconsistent. Prof. Arrow, therefore, concludes that it is impossible to derive a social ordering of different conceivable alternative social states on the basis of the individual ordering of those sociaI states without violating at least one of the value judgments as expressed in the five conditions. This is in essence his impossibility theorem.

We have shown above how Arrow has proved that it is impossible to construct a social welfare function based on all individual preferences without involving inconsistent or non-contradictory social ranking. It is worth mentioning that in reaching this conclusion, Arrow has assumed that only ranking i.e. ordinary preferences of individuals be considered which means that no weight should be given to the intensity of preferences. For, it is very difficult to measure the intensity of feelings of different individuals with regard to various social states such as the errection of a social overhead capital such as bridge, tunnel sanitation work or park etc. It should be further noted that if weights are assigned to individual preferences i.e. if the intensity of individual preference are considered, then consistent and non- contradictory social choice is possible. Thus, according to Boumol, “in deciding whether to allocate labour to the production of some drug needed to treat a rare but dangerous disease or to the manufacture of Scrabble sets, we may recognize that, compassion aside, more people will want the Scrabble sets than the medicine. Yet on the crudest sort of interpersonal comparison of benefits, we may decide that the public as a whole will gain more from the production of the medicine because its potential users feel much more strongly about their preference than do the others.

CHECK YOUR PROGRESS IV

Q1 Discuss how imposition of tax on industries that create negative externalities improves economic efficiency.

____

____

____

____

Q2. What are Arrow’s conditions of Social Choice ? Why is it impossible to satisfy all these simultaneously ?

____

____

____

____

Q3. What do you understand by Arrow's impossibility Theorem ?

____

____

____

____

15.8. LET US SUM UP

Dear learner, in this lesson, we discussed the following:

- Pareto Criterion states that an economic change which harms no one and makes someone better off indicates an increase in social welfare.
- Economists like Kaldor, Hicks, and Scitovsky have made efforts to evaluate the change in social welfare resulting from any economic reorganization which harms some people and benefits others. They have put forward a criterion known as the compensation principle, on the basis of which they claim to evaluate those changes in economic policy which make some individuals better off and others worse off.
- The concept of social welfare function was propounded by A. Bergson in his article *"A Reformulation of Certain Aspects of Welfare Economics"* in 1938. Prior to this, various concepts of social welfare had been given by different welfare theorists, but they failed to provide a satisfactory solution to the problem of maximization of social welfare.
- The goal of a society is to maximize social welfare, that is, the aggregate of the utilities of the individuals comprising the society. In the classical welfare function, it is further assumed that the law of diminishing marginal utility applies to money income. Given this, maximum social welfare (utility) will be achieved if income is so distributed that the marginal utility of income is equal for all individuals in the society.
- John Rawls begins his welfare analysis by considering a society as being in an "initial position", in which no individual knows what his final

utility position will be. The problem he poses is: what type of welfare criterion would be adopted by the society when it is in such an initial position, where everybody has to behave under uncertainty about how the welfare criterion chosen will ultimately affect his utility or welfare.

- Pareto laid the foundation of modern welfare economics by formulating the concept of social optimum, which is based on the concept of ordinal utility and is free from interpersonal comparison of utilities and value judgments.

- Nicholas Kaldor was the first economist to give a welfare criterion based on compensating payments. Kaldor's criterion helps us to measure the welfare implications of a movement in either direction on the contract curve in terms of the Edgeworth box diagram.

- Scitovsky pointed out an important limitation of the Kaldor-Hicks criterion— that it might lead to contradictory results. According to him, the Kaldor-Hicks criterion involves such contradictory and inconsistent results. Since Scitovsky was the first to point out this paradoxical result in the Kaldor-Hicks criterion, it is known as the "Scitovsky Paradox."

- To rule out the possibility of contradictory results in the Kaldor-Hicks test, Scitovsky formulated a double criterion, which requires the fulfillment of both the Kaldor-Hicks test and the reversal test.

- Profs. Lipsey and Lancaster, in their theory of second best, state that if it is not possible to satisfy all conditions of Pareto efficiency (or if these conditions cannot be satisfied in all markets), then the best solution—that is, maximum social welfare—is unattainable. So the question arises whether efforts should be made to achieve the second-best position by satisfying the remaining marginal conditions of Pareto optimum in the remaining markets. In their theory, they assert that this second-best solution will not necessarily lead to an increase in social welfare.

- Arrow, in his path-breaking work *"Social Choice and Individual Values,"* pointed out that the construction of a social welfare function that reflects the preferences of all individuals in society is an impossible task.
- According to Arrow's theorem, if we exclude the possibility of interpersonal comparison of utility, then the only method of passing from individual tastes to social preferences that is satisfactory and defined for a wide range of individual orderings must be either imposed or dictatorial.

15.9. KEYWORDS

1. **Pareto Optimum / Pareto Efficiency :** A state of allocation where no individual can be made better off without making someone else worse off. It is used as a benchmark for economic efficiency.
2. **Compensation Principle :** A criterion introduced by Kaldor and Hicks which states that an economic change increases social welfare if the gainers from the change could hypothetically compensate the losers and still be better off.
3. **Kaldor-Hicks Criterion :** A refinement of the compensation principle suggesting that a policy is welfare-improving if those who gain could, in theory, compensate those who lose—whether or not compensation actually occurs.
4. **Scitovsky Paradox :** Asituation where a move from Ato B is considered an improvement under the Kaldor-Hicks criterion, but a reverse move from B to A also qualifies as an improvement. This paradox led to the refinement called the **Scitovsky Double Criterion**.

   5. Social Welfare Function (SWF) : A function that ranks different allocations or distributions of resources in terms of the social welfare they produce, based on individual utilities. Proposed formally by Abram Bergson.

6. **Theory of Second Best :** Introduced by Lipsey and Lancaster, it states that if a Pareto optimal condition cannot be satisfied in one market,

satisfying the remaining conditions does not necessarily lead to the second-best welfare outcome.

15.10 EXAMINATION-ORIENTED QUESTIONS

1. What is the compensation principle of judging social welfare brought about by the adoption of an economic policy?

2. State Kaldor-Hicks criterion of judging social welfare.

   Explain with the help of the Utility Possibility Curve.

   How does it judge the social welfare impact of an economic policy which makes some persons better off and others worse off?

3. Show how the Kaldor-Hicks criterion helps in solving the problem of indeterminacy involved in Pareto's analysis of maximization of social welfare.

4. Does the Kaldor-Hicks-Scitovsky criteria give us a sufficient basis for ordering states from the viewpoint of social welfare?

   Give reasons for your answer.

5. In what way is the Scitovsky criterion an improvement over the Kaldor-Hicks criterion?

   Do you agree with the view that Scitovsky criterion does not remove all the weaknesses of the Kaldor-Hicks criterion? Discuss.

6. Explain the theory of second best.

7. Discuss the various conditions of social choice given by Arrow.

5.11 HINTS TO CHECK YOUR PROGRESS

Check Your Progress I

- Q1 – See Section 15.0
- Q2 – See Section 15.3
- Q3 – See Section 15.3

Check Your Progress II

- Q1 – See Subsection 15.4.1
- Q2 – See Subsection 15.4.3
- Q3 – See Subsection 15.4.4
- Q4 – See Subsection 15.4.4
- Q5 – See Subsection 15.4.5

Check Your Progress III

- Q1 – See Subsection 15.5.1
- Q2 – See Subsection 15.5.1
- Q3 – See Subsection 15.5.2
- Q4 – See Subsection 15.5.3

Check Your Progress IV

- Q1 – See Section 15.6
- Q2 – See Subsection 15.7.1
- Q3 – See Subsection 15.7.3

15.12 SUGGESTED READINGS

- **Amartya Sen**, *On Ethics and Economics,* Oxford University Press, Delhi, 1990, p. 34.
- **J. Rawls**, *A Theory of Social Justice,* Harvard University Press, 197.
- **Scitovsky**, "The State of Welfare Economics", printed in *Papers on Welfare and Growth*, George Allen and Unwin Ltd., 1964, p. 184.

M.A. ECONOMICS **LESSON NO. 16**

COURSE NO. ECO 201 **UNIT-III**

EXTERNALITIES, MARKET FAILURE AND COASE THEOREM

STRUCTURE

16.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

- Discuss the concept of **externalities** and their types (external economies and diseconomies).
- Explain the causes and consequences of **market failure**, particularly in the presence of externalities and public goods.
- Explore the **Coase Theorem** as a potential solution to externality-related market failures through private negotiation.

16.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to:

- Explain **externalities**—both external economies and diseconomies—and their impact on **Pareto Optimality**.
- Distinguish between **external economies and diseconomies in production** and **consumption**.
- Describe the concept of **market failure** and how it results from externalities.
- Analyze how **monopoly** can act as a barrier to achieving Pareto optimal resource allocation.
- Examine the role of **public goods** in contributing to market failure.
- Describe the **Coase Theorem** and its implications for resolving externality problems through private negotiation.

16.2 INTRODUCTION

Dear learner, externalities are a pervasive and significant phenomenon in modern societies. The term externalities refers to the economic effects which occur from the production or the use of goods to other parties or economic units. On the other hand, a market determined allocation of wide spread resources is the basis for the widespread belief that for a "free enterprise", market economy is socially desirable. But, unfortunately, the conditions under which a market allocation is efficient are often not met, and market fails in many ways to achieve efficient allocation of resources. In this lesson, we shall discuss about externalities & market failure.

16.3 EXTERNALITIES (EXTERNAL ECONOMIES AND DISECONOMIES) AND PARETO OPTIMALITY

The existence of externalities is an important factor which prevents the achievement of Pareto optimality (or maximum social welfare or economic efficiency) even when perfect competition prevails. Externalities refer to the beneficial and detrimental effects of an economic unit (a firm, a consumer or an industry) on others. The beneficial externalities created by a consumer or a firm for others are known as external economies and detrimental externalities imposed on the others by a productive firm or a consumer are known as external diseconomies. To be more precise, when an economic unit creates benefits for others for which he does not receive any payment, there exist external economies. On the other hand, external diseconomies occur when an economics unit inflicts costs on others for which he is not required to pay. It is noteworthy that the term externalities covers both the external economies and external diseconomies. When for a productive firm there exist external economies, that is, beneficial external effects, then the private marginal cost of the firm will be higher than the social marginal costs, since the firm will not take into account benefits external to it (i.e. benefits created for others). And the market price fixed on the basis of private marginal cost will not reflect the social marginal cost which will be lower when external economies occur. Similarly, when with the expansion of a firm external diseconomies occur, then the private marginal cost will be lower than the social cost, since the firm will not take into account the disadvantages caused to others by its activity. Thus, when external diseconomies occur, price fixed on the basis of private marginal cost will be lower than that determined on the basis of social cost.

It follows from above that in the absence of externalities, all costs incurred and all benefits received by producers and consumers will be reflected in market prices and that there will not be any divergence between private and social costs (or benefits). But, when externalities (external economies and diseconomies) occur, market prices determined on the basis of private costs and benefits will not truly reflect social costs and therefore divergence is caused between private and social costs (or benefits).

It follows from above that externalities arise primarily due to the fact that the effects regarding costs, output, employment, labour skills, technological capabilities of the activities of a producer or consumer on others or society as a whole are not reflected in market prices and therefore market prices do not truly reflect social costs. We shall give below a few examples of externalities, that is, external economies and diseconomies in production and consumption. The basic idea behind the belief that a competitive price system is optimal is based on the fact that a producer benefits himself only by benefiting the society because he makes available certain goods and services to the society. In other words, by promoting his own interests he promotes the interests of society as well. But there are so many cases in production and consumption "when members of the economy do things which benefit others in such a way that they can receive no payment in return or where their actions are detrimental to others and involve no commensurate cost to themselves."[1] Thus, due to externalities there arises the divergence between social and private costs and between social and private benefits.

16.3.1 External Economies and Diseconomies in Production

As the firm expands its scale of production, it becomes possible for the firm to produce a unit of product at a relatively lower cost due to internal economies of large-scale production. On the other hand, external economies occur when the expansion of a firm's output creates benefits, part of which goes to others. A firm may create external benefits for others in two ways: (a) By expanding its productionthe firm may render a direct service to others such as training the laborers by its manpower training programme and thus benefit the other firms by making available skilled and trained labourers when they have to pay no cost or only nominal cost. (b) By expanding its production, a firm may make the supply of some inputs cheaper for all the firms in the industry. For example, an expansion in the production of an engineering firm may increase the demand for steel. And if the steel production is subject to internal economies of large-scale production, then the expansion of steel industry following the increase in its demand will lower its cost and price. Another example of external economies is provided by the construction of a bridge or a highway which reduces transport cost and increases the land values in the neighbouring areas. Still another example of external economies of production is provided by the pumping of water from a mine. If a firm pumps out water from its mine 'A', it will lower the cost of pumping water from a mine 'B' owned by another firm. Similarly, bees of producers of honey create benefits for the owners of nearby orange groves, for the bees help Pollinate oranges in the groves. On the other hand, orange groves create external economies for the honey producers since the orange groves provide nectar for the bees producing honey. In all these cases, a firm incurs costs for its own production but the benefits arising out of it are also reaped by others who pay no price for them.

Let us explain some external diseconomies of production. There are a good number of external diseconomies which may be created by the productive activity of a firm. The pollution of air by the factories through emitting smoke and the wastes of factories poured into streams or ocean create health hazard for men, especially those who live in the surrounding areas. For these external harms caused to the other members of the society, the firms are not required to pay any price. A factory owner pays nothing to the residents of the neighbouring colony who happen to be the victims of pollution by the factory. Another example of external diseconomies is provided by a firm or industry which has to keep more trucks on the road to do its business. This will overcrowd the road which will increase the transportation costs of other firms or industry which has to carry its own goods by trucks. The expanding firm or industry does not pay any price for the higher transport cost incurred by others.

16.3.2 External Economics and Diseconomies in Consumption

External economies in consumption arise when the consumption of a person creates beneficial effects on others. Many examples of external economies can be given. For example, the satisfaction of a telephone owner increases with the increase in the number of telephone owners because he can now contact a larger number of persons on telephone. Likewise, if a person maintains a beautiful garden or lawn, he not only raises his own satisfaction but also that of others, especially his neighbours who also enjoy the look of his garden or lawn. Similarly, when a person maintains his car in such a way that it is quite safe to drive it and also does not emit any smoke, it will also improve others' safety and health and therefore welfare. In this category, we may also include the expenditure incurred by the parents on educating their children. This will not only benefit them and their children but also other members of the society This is because the education makes a person civilized and better citizen and therefore whoever comes in contact with him, derives satisfaction from him. Thus, in the presence of external economies in consumption the social utility exceeds the private utility and therefore divergence between social and private benefits is caused when external economies in consumption prevail.

On the contrary, external diseconomies of consumption occur when a person's consumption creates unfavourable impact on other consumers. A good example of it is provided by the conspicuous consumption of a person who through demonstration effect causes a lot of dissatisfaction to his friends and neighbourers who now feel themselves inferior to him. Likewise, when a person purchases candy bars for his children it will make his neighbouring children unhappy because their parents cannot afford to buy these for them. Likewise, loud music played by your neighbour may disturb you and cause a lot of dissatisfaction. Also falls in this category the purchasing of a new Maruti SX4 car by your friend because now your Maruti 800 in your own eyes becomes old. More examples of such external diseconomies of consumption can be given.

The existence of external economies and diseconomies explained above plays a significant role in determining the activities of production and consumption in the economy. Apertinent question is how these externalities can lead to the misallocationof resources and thereby act as an obstacle to the attainment of Pareto optimality or maximum social welfare. When externalities in production and consumption prevail and as a result divergence is caused between private and social costs and between private and social benefits, the economy guided by the market prices alone, even when perfect competition prevails, will fail to achieve optimum allocation of resources (or, in other words, maximum social welfare). When external economies of production occur, private marginal cost will be greater than the social marginal cost and when external diseconomies in production are present, private marginal cost will be lower than social marginal cost.

Under these circumstances therefore a firm which creates external benefits for others will not produce its product to the extent social interest requires. This is because equating price with the private marginal cost, which is higher than the social marginal cost will result in under-production of the product. Thus in this case of the existence of external economies in production, output of the product determined on the basis of private marginal cost, will be less than the socially optimal level of output.

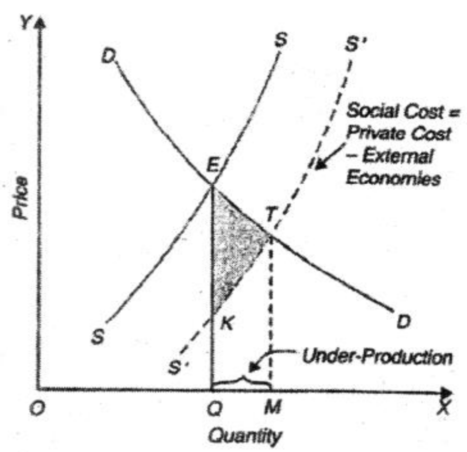


Fig. 16.1 a.

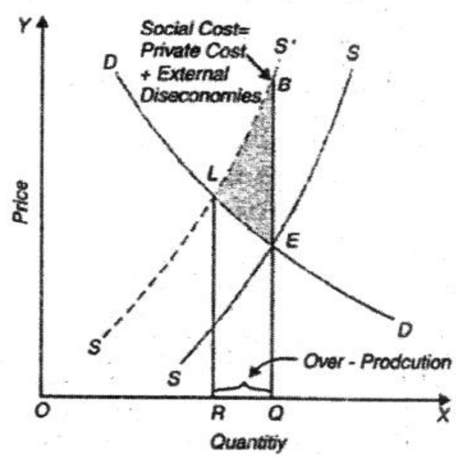


fig 16.1 b.

This is illustrated in Figure 16.1. where SS represents the supply curve for the product of the industry which has been obtained by summing up the private marginal cost curves of firms. Due to the existence of external economies, social cost will be smaller than the private costs. Therefore the supply curve S´S´ (dotted) of the product reflecting social cost will be lower than the supply curve SS based on private marginal costs. The supply curve reflecting social cost is lower because it takes into account external economies generated by the production in the industry, while private cost does not take into account these external economies. It will be seen from Figure 16.1. (a) that the given demand curve DD and the supply curve SS, based upon the private cost of production intersect at point E and thus determine OQ as the actual amount of the product produced. But the socially optimum output is OM at which the supply curve S´S´ reflecting social cost intersects the given demand curve. It is thus evident that the product is being produced in smaller quantity than the socially optimum output OM. Thus the existence of external economies result in under-production and loss of social welfare equal to the area ETK.

On the other hand, when there exist external diseconomies in production, private marginal cost will be lower than the social marginal cost, since the former will not take into account costs or harms imposed on others. Therefore, when external Diseconomies are present, equating price with marginal cost will result in over-production of the product, that is, more than socially optimum output will be produced. This is illustrated in Figure 16.1 (b). It will be seen that the supply curve SS based on private marginal costs intersects the demand curve at point E and thus determine OQ amount of output. Supply curve S´S´ (dotted) which takes into account external diseconomies and therefore curve reflects social cost lies at a higher level and intersects the demand curve at point L and therefore socially optimum output will be OR. Thus, it follows that when external diseconomies are present, equating price with private marginal cost will result in over-production of the product, that is, more than socially optimum output will be produced and will cause loss of social welfare equal to area ELB.

Likewise, when there are external economies in consumption, then the demand curve for the product determined on the basis of private marginal utility will be lower than that based on social marginal utility, for the former will fail to reflect the external economies in consumption being generated. Therefore, in this case too output determined on the basis of private marginal utility and demand will result in lower output than the socially optimum fuel. On the other hand, when there exist external diseconomies in consumption the private marginal utility will be higher than the social marginal utility, since the former will not take into account the external diseconomies. As a result, when external diseconomies in consumption are present, the output determined on the basis of private marginal utility (benefit) will be more than the socially optimum level.

16.4 MARKET FAILURES

We have already discussed Pareto efficiency. But under some circumstances, the market system cannot lead to this optimum situation of Pareto efficiency (i.e. the state of maximum social welfare). These circumstances due to which market fails to achieve economic efficiency or maximum social welfare have been called market failures. There are three main causes of market failure. They are :

1. The existence of monopoly or imperfect competition;
2. The presence of externalities, i.e., external economies and diseconomies in production and consumption; and
3. The consumption of public goods.

We shall explain below these obstacles at some length.

16.4.1 Monopoly as an Obstacle to the attainment of Pareto Optimality

An important complaint against monopoly (and as a matter of fact against all forms of imperfect competition) is that it causes misallocation of productive resources and thus hinders the achievement of maximum social welfare. The crucial condition required for Pareto optimality is that Marginal Rate of Transformation (MRT) of the community between any two commodities should be equal to the marginal rate of substitution (MRS) between these commodities of every consumer. We saw above that Perfect Competition satisfies this condition of Pareto optimality. But under Monopoly (or any other form of Monopolistic Competition) the Marginal Rate of Transformation of the community between two commodities is not equal to the Marginal Rate of Substitution between them of the consumers. Consequently, Monopoly does not ensure optimum allocation of resources and serves as an obstacle to the attainment of maximum social welfare. The reason for this is that am he restricts output and charges higher price than Marginal Cost. Assume that there are two commodities X and Y and further that the commodity X is being produced under conditions of Monopoly whereas the commodity Y is being produced under conditions of Perfect Competition. Since the commodity X is being produced under conditions of Monopoly, the price (P_X) of commodity X will be greater than Marginal, Cost. (MC_X) of its production. i.e. $P_X > MC_X$. But since the commodity Y is being produced under condition of Perfect Competition, price (P_Y) of commodity Y will be equal to its marginal cost (MC_Y) of production. That is, $P_Y = MC_Y$.

It therefore follows that:

$$\frac{MC_X}{MCy} < \frac{Px}{Py}$$

Since the ratio of Marginal Costs of two commodities $\left(\frac{MC_X}{MCy}\right)$ represents the marginal Rate of Transformation (MRT_{XY}) between them, therefore,

$$MRT_{XY} < \frac{Px}{Py} \quad \ldots\ldots\ldots\ldots..(1)$$

But the consumers, in order to be in equilibrium, will equate their Marginal Rate of Substitution between two commodities (MRS_{XY}) with the price ratio of the two commodities $\frac{Px}{Py}$. This is because each individual consumer will take the prices of commodities as given and constant for him. Thus, for consumers,

$$MRS_{XY} = \frac{Px}{Py} \quad \ldots\ldots..(2)$$

From (1) and (2) above it follows that under conditions of Monopoly in the production of X

$$MRT_{XY} < MRS_{XY}$$

Or $$MRS_{XY} > MRT_{XY}$$

That is, when Monopoly exists in the production of a commodity, Marginal Rate of Substitution between commodities will be greater than the Marginal Rate of transformation. In other words, consumers would like the commodity under Monopoly production to be produced more but monopolists would not be producing the desired quantity of the commodity and will, therefore, be causing loss of satisfaction and misallocation of resources.

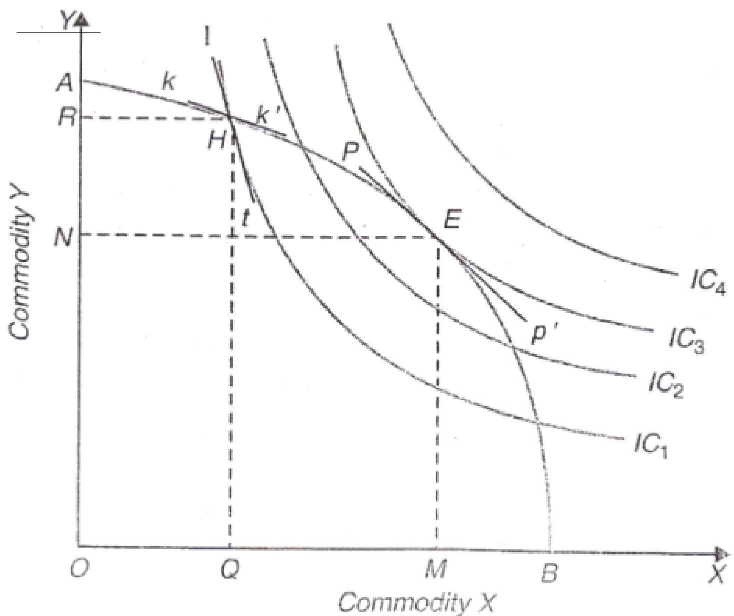


Fig. 16.2. Monopoly as an obstacle to maximization of social welfare

That monopoly causes loss of social welfare and misallocation of resources will become very clear by considering Figure 16.2. It will be seen from this figure that the transformation curve of the community AB is tangent to the community indifference curve IC_3 at point E. Therefore, at point E, Marginal Rate of Transformation (MRT_{XY}) of the community between two commodities is equal to the Marginal Rate of Substitution (MRS_{XY}) of the community. Thus E represents maximum possible level of social welfare and the combination of two commodities being produced (i.e. OM of X and ON of Y) represents optimum allocation of resources. But when the commodity X is being produced under condition of Monopoly, the equilibrium will not be at point E but instead it will be at point. This is because, under Monopoly, producers would be equating Marginal Rate of Transformation

(MRT) or ratio of marginal costs $\left(\frac{MC_X}{MCy}\right)$ with the ratio of marginal revenues $\left(\frac{MR_X}{MRy}\right)$ and not with the ratio of prices of two goods $\left(\frac{Px}{Py}\right)$. Since consumers would be equating Marginal Rate of Substitution (MRS_{XY}) with the price ratio of two goods, the Marginal Rate of Transformation in the equilibrium position at point H will not be equal to the Marginal Rate of Substitution. This is quite obvious from Figure 16.2. where at point H, transformation curve AB and consumers' indifference curve IC_1 rare intersecting each other. This implies that slopes of transformation curve AB at point H, which indicates Marginal Rate of Transformation, and the slope of consumers' indifference curve IC_1 which indicates Marginal Rate of Substitution, will not be the same. It will be observed from Figure 16.2. that at point H, Marginal Rate of Substitution between two goods (MRS_{XY}) is greater than the marginal rate of transformation (MRT_{XY}) between them as tangent II drawn to point H on indifference curve IC_1 is steeper than the-tangent kk' drawn to point H on transformation curve AB. This means that consumers' preference is that good X should be produced more but because of the existence of Monopoly in the production of commodity X, it is not produced to their desired quantity. As a result, the level of satisfaction or welfare PM of the consuming community is at a lower level than possible under the given production conditions. Consuming community's satisfaction will be greater at point E which lies on indifference IC_3 but under conditions of monopoly in the production of X, equilibrium is at point H which lies on the lower indifference curve IC_1. Thus Monopoly causes loss of satisfaction or welfare. This loss of satisfaction or welfare is due on the fact that Monopoly is not optimally allocating its resources to the production of commodity X according to the consumers' preferences. Given the transformation curve and consumingcommunity's indifference map, the optimal production pattern is represented by the point E where OM amount of commodity X and ON amount of commodity Y are being produced. But under conditions of Monopoly in the production of X, the equilibrium is established at point H where smaller quantity QQ of commodity X and large quantity OR of commodity Y are being produced. Thus Monopoly has caused misallocation of resources.

16.4.2 Presence of Externalities

Already discussed in section 16.3.

16.4.3 Public Goods and Market Failure

The existence of public goods provides us another important source of market failure. It should be noted that public goods are not necessarily produced by the public sector. It is due to the possession of certain properties that some goods are called public goods and has nothing to do with whether they are produced-in the public sector or private sector. Two essential characteristics of public goods are that they are non-rival and non exclusive in consumption. Let us explain these characteristics of public goods in some detail.

Non-Rivalry In Consumption

In order to explain what are non-rival goods, it is better to know first what are rival goods. A rival good is one, when one unit of which is consumed by an individual, that very unit cannot be consumed by another. For example, if Rekha consumes an apple, any other person, say Karishma, cannot eat the same very apple. Of course, Karishma can get another apple for her consumption from the market by paying a price for it. Similarly, if Amit drinks Pepsi cola, Bela cannot drinks the same very pack of Pepsi cola, that is, two individuals cannot consume the same very Pepsi Cola; its consumption by one individual excludes others to consume it. Thus goods like apples, Pepsi cola, shirts, machines and several such other goods, the consumption of which reduces their availability for other persons, are called rival goods. Rival goods cannot be public goods; they are private goods.

On the other hand, public goods are non-rival in consumption. National defense, parks, television signals, flood control project, pollution control project, lighthouse in the sea are some examples of public goods. Thus all persons of a nation can enjoy (consume) equally the security provided by the national defense system. All persons of a city can benefit from the television signals and enjoy the programme telecast. The enjoyment provided by a park, if there is free access to it, can be obtained by all who visit it. National defence, parks, television signals and such other goods are non-rival goods, as their consumption by one individual does not exclude its consumption by others. That is, the consumption of a non rival good by an individual does not reduce its amount available for others to consume. To conclude, public goods are non-rival.

Non- Excludability

The other essential characteristic of a public good is non-excludability in distribution of their consumption benefits. This non-exclusive nature of a public good implies that it is difficult, if not impossible, to exclude those from consuming them who are not willing to pay for them. In case of private rival goods such as shirts, cars, Pepsi cola, apples, those who do not pay for them can be easily prevented from consuming them or receiving benefits from them because the producer or seller simply does not provide them these goods, if they have not paid their price. On the contrary, in case of public goods, either it is not possible or it is very costly to prevent those people who do not pay for these goods. We will explain later that the feature of non-excludability of public goods accounts for the failure of market in case of these goods to ensure Pareto efficiency.

For example, national defence is a public good and is provided to all members of a society and its benefits are available to all equally, irrespective of whether some people pay taxes for it or not. It is difficult, if not impossible, to exclude those people from receiving benefits of security provided by national defence system who do not pay for it. Likewise, if a lighthouse is constructed in a sea, it provides light for all the ships whether any one of them pays for it or not and it is not possible to prevent those who do not pay from receiving light from the light house. This inability to exclude those who do not pay for receiving benefit also applies in case of other public goods such as television signals, pollution control project to provide clear air, flood control projects, parks etc.

Check Your Progress–I

Q.1. Write a note on externalities in production and consumption.

Q 2 Discuss how, in the presence of externalities, socially optimum output is not produced.

Q 3 What do you understand by market failures ? What are the possible cause for market failures?

Q4 Discuss how market failure takes place in case of public goods.

16.4.3.1 Free-Rider's Problem And Public Goods

It is easy to show how non-excludability of a public good can lead to market failure, that is, failure of market to achieve Pareto efficiency. As explained above, non-excludability of public goods arises because producers are not able to prevent those from consuming or enjoying benefits from these goods who do not pay their share of cost. There is a problem called a free rider's problem which states that because people cannot be excluded from consuming public goods or enjoying benefits from them, there is incentive for persons in these situations to free ride and try to enjoy benefits from reduced pollution, parks, television signals, light house without paying for them. These persons want to get something for nothing and rely on others to make purchases of public goods whose benefits they will also get automatically.

Due to this free-rider problem or inability of producers of public goods to prevent those who do not pay for receiving benefits from them that a profit-maximizing firm will either not produce a public good or produce too little of it. This creates economic inefficiency or Pareto non-optimality. Let us take an example of this free-rider's problem in case of public goods leading to economic inefficiency. Suppose the construction of a dam to check floods which cause a lot of damage to a city is required. This dam, when built, will protect equally all people of the city from the damages due to floods. However, some people of the city would not like to pay for the dam with the hope that others would pay for it and they, because of non-excludability would also enjoy its benefits. But in view of this incentive to free ride, adequate revenue to cover costs of building the dam cannot be provided and, therefore, no private entrepreneur would consider it worthwhile to construct the dam to control floods. Similarly, the production of other public goods such as lighthouse, television signals, pollution control projects would not be extended to the socially desirable level in view of the non-excludability and incentive to free ride.

16.4.3.2 Public Goods and Pareto Efficiency

Before explaining further how the free-rider's problem results in less than socially optimal production of public goods, it is important to understand how Pareto optimal level of production of public goods is determined. Because public goods are non-rival in consumption, some modifications are required in formulation of Pareto optimality conditions. To illustrate the conditions of Pareto optimality in case of public goods, we take the case of a society composed of two persons A and B and the public good is the pollution control project aimed to clean air which, if produced, would benefit both of them. But the two persons may not perceive to receive the same amount of marginal benefits from this pollution control measure. In other words, they may have different evaluation of the marginal benefits of pollution control measure. Each person will place some value on the pollution control. The marginal benefits they obtain or values they place on the different quantities of pollution abatement are depicted in Figure 16.3.

Due to differences in tastes or perceptions of two individuals, the curvesshowing their marginal benefits from the pollution-free air are different. The curves MB_A and MB_B depict the marginal benefits obtained by individuals A and B respectively from the varying quantities of pollution-free air.

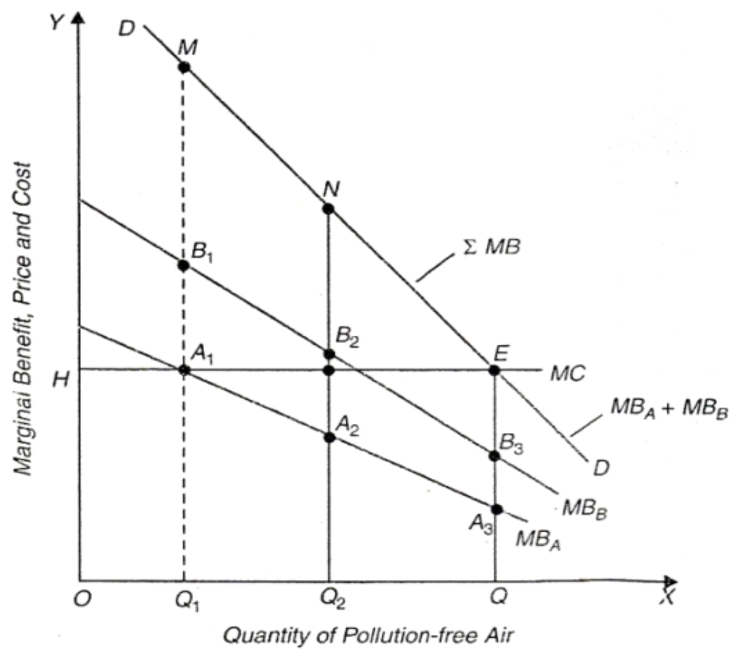


Fig. 16.3. Public good, pareto optimum and market failure

The marginal benefit curve can also be interpreted as the price which the individuals are willing to pay for the different quantities of pollution-free air. Thus, it will be seen from the marginal benefit curves that individual A will be willing to pay price Q_1A_1 for OQ_1 quantity of pollution-free air, whereas individual B is willing to pay Q_1B_1 for the same OQ_1 quantity of the pollution-free air. Similarly, for OQ_2 quantity of the pollution-free air, the individual A is willing to pay price equal to Q_2A_2 and individual B is willing to pay price equal to Q_2B_2. Therefore, the marginal benefit curves can be interpreted as the demand curves of the individual for pollution free air. It should also be noted that the marginal benefits or the price which the individuals are willing to pay depend on the values they place on the different quantities of the pollution-free air.

In order to determine the Pareto-optimal quantity of pollution-free air we need the total market demand curve or the aggregate marginal benefit curve of the individuals comprising the society. Market demand curve for a public good cannot be obtained in the way market demand curve is obtained in case of private goods. Since a private good is rival in consumption, the market demand curve of it is obtained by adding up sideways (i.e. horizontal addition) of the demand curves (i.e. marginal benefit curves) of the two individuals. But, as explained above, public goods are non-rival in consumption, that is, in case of public goods same units of output can be consumed by various people at the same time. Therefore, different individuals can pay for the same units of a public good at the same time. Thus, a pollution central project renders the air of a town free of pollution to some degree from which everybody in the town is benefited and should pay for it. Again consider Figure 16.3, where it will be seen that individual A is prepared to pay price Q_1A_1 for OQ_1 quantity of pollution-free air and the individual B is prepared to pay Q_1B_1 for same OQ_1 quantity of pollution-free air which he enjoys or consumes at the same time as individual A. Thus, for OQ_1 quantity of clean air the total price which the two individuals are willing to pay equals $Q_1A_1 + Q_1B_1 = Q_1M$. Similarly, for the same OQ_2 quantity of pollution-free air, individual A is prepared to pay price equal to Q_2A_2, and individual B is prepared to pay price equal to Q_2B_2.Thus, the total price which the individuals together are willing to pay for the same OQ_2 quantity of the good is equal to the sum of these two prices, i.e $Q_2A_2 + Q_2B_2 = Q_2N$. It therefore follows that in case of a public good, market demand curve is derived by summing up vertically the demand curves of the individuals because each individual consumes the same units of the good at the same time.

Having now obtained the market demand curve of a public good we can now show what will be the Pareto-efficient output of a public good. In this connection, it should be noted that a society has to bear the costs it incurs on labour and materials to produce pollution-free air. Pareto-efficient output is determined at a level at which price the individuals together are willing to pay for the good (that is, aggregate benefit), equals marginal cost of production. Suppose the marginal cost of production is equal to OH per unit and remains constant. In Figure 16.3, with OH as the constant marginal cost, MC is the marginal cost curve. It will be observed from the figure that price which the two individuals together are willing to pay equals marginal cost (OH) at OQ quantity of pollution-free air. As said above, price which the persons are willing to pay indicates the aggregate marginal benefit. Thus, aggregate marginal benefit and marginal cost of production are equal at OQ level of output of pollution free air. lf resources are allocated to the pollution control project to the extent that OQ quantity of pollution-free air is produced at which price (marginal benefit) equals marginal cost incurred, the social welfare (i.-e., the aggregate benefits of the two individuals) will be maximum. Thus OQ is Pareto-efficient level of output of the public good.

But a private firm will produce Pareto optimum output OQ only if each individual pays a price equal to the marginal benefit. At OQ output marginal benefit of pollution-free air is QA_3 for individual A and QB_3 for individual B. If both are willing to pay prices equal to these marginal benefits, the aggregate price per unit which they together will pay for Pareto-efficient quantity OQ amounts to OH or $QE = QA_3 + QB_3$.ln this way total revenue collected by the private firm will cover the cost of pollution control project which cleans the air and therefore will be worthwhile for the private firm to undertake the pollution control project.

But, as explained above, due to inability of the producer of a public good to exclude those who do not pay of output cannot be covered by a private producer. Therefore, in this situation too little or even none of the public good will be produced though the marginal benefits of additional units (i.e.. the value the individuals place on these additional units) exceeds the social marginal costs of producing these units. Thus, private production and functioning of market in case of public good do not lead to Pareto efficiency in the provision of public goods.

It may be further noted that in case of two individuals composing a society, there may not be much problem for an individual trying to be a free-rider because of his being constantly watched and pressured, but in the real world a society consists of many persons. There is incentive to the persons for misrepresenting the values they place (i.e. the benefits, they receive) on public goods such as national defence, pollution control project, flood control programme, television signals, and apparently claiming that they have little interest in the provision of these public goods. Since a large number of individuals are involved, each one is likely to think that his not paying for the public good will not make much difference to the overall revenue and the public good will be produced and he will be able to enjoy its benefits without making any contribution. This, of course, would be true if one individual tries to be a free-rider. But, as is likely to be the case, if many individuals and may be even all of them thinking in a similar way try to be free riders, then as explained above in Figure 16.3 enough revenue cannot be collected to cover the cost of production of a public good. In this situation there will be no production of a public good at all, at least its Pareto optimal quantity would not be produced. Thus the production of highly important and useful public goods such as national defence, pollution control project, flood control project, television signals may not be undertaken at all if we rely on private sector and market. This is a glaring example of the market failure.

An interesting way of explaining market failure to achieve Pareto efficiency in case of public goods is to emphasise that marginal cost of allowing a person to consume the public good is zero once it is produced, even if it is possible to prevent him from consuming the commodity. Thus, parks, television signals, flood control projects etc. having been produced, the cost of letting additional consumers to consume these goods or their services is zero. For example, within good limits a visitor to a public park who has not paid is not to affect the enjoyment of the park by those who have paid for it, and it costs the society or a private producer nothing for this additional person visiting the park and enjoying (consuming) it. In fact he would be made better off and no one would be worse off because no more resources of the society are used when the additional person is allowed to enjoy the park (i.e. marginal cost is zero). Now, if the marginal cost of permitting additional persons to consume the good is zero, then Pareto efficiency requires that price of the public good should be zero. But the total cost of production of public goods is not zero; to produce public good is indeed very expensive. To meet these total costs of production private producer sets a positive price to cover cost. Consequently, price set will be higher than marginal cost and less than Pareto-optimal quantity will be consumed. Thus production of a public good by the private sector does not lead to Pareto optimality in allocation of the good. In other words, market fails to achieve Pareto Efficiency.

16.5 THE COASE THEOREM

Suppose that parties affected by an externality can negotiate costlessly with one another. Coase pointed out that in such a case, an efficient outcome will result, no matter how the initial property rights to the externality-creating variables are assigned.

Coase's idea is quite simple. Suppose that the polluting firm is given the right to pollute, therefore the polluted agent will have an incentive to pay the former to restrict pollution. The firm will accept this offer so long as the gain from this outweighs the loss in profit from cutting down production. The polluting firm's loss in profit is given by the marginal profit schedule: p-c´ (x). The polluted firm's gain from restricting pollution comes from the MEC schedule: e´(x). The polluted firm can go on paying the polluter to restrict production so long as e´(x) > p-c´(x). The firms will stop at the point where e´(x) = p-c´(x). But this is precisely the condition for socially optimal amount of pollution to take place.

Example. Suppose that firm 1 is a monopoly and faces an inverse demand curve $p = 11 - x$. Total cost is $C = 2x$ and $MEC = 2x$.

x	Profit = $px - 2x$	Marginal Profit	MEC
1	8		
		6	2
2	14		
		4	4
3	18		
		2	6
4	20		
		0	8
5	20		

The marginal figures are given between two x figures to emphasize their marginal nature. The monopoly by itself will produce 5 (or 4) units. If the polluted firm wants it to reduce output (and pollution) by one unit, it will be prepared to pay up to 8, whereas the loss in profit will be 0. Hence, the firms should be able to reach agreement over this reduction. This process continues until $x = 3$. For a further one unit reduction in x, the maximum the polluted can pay is 4, which is exactly equal to the marginal loss in profit for firm 1. We can expect bargaining to stop here, and the socially optimal amount of X to be produced.

Suppose on the other hand, that the polluted firm is given the right to clean air. Then rather than producing nothing, it is in the interest of the polluting firm to offer the polluted firm a sum of money in return for the right to pollute (or to prevent the latter from going to court). For every extra unit of pollution, the firm will weigh the marginal gain in profit it can make and be prepared to offer a maximum amount based on this. The polluted firm will go on accepting so long as this payment adequately compensates it for the extra cost of pollution. The end result will be the socially optimal level of pollution.

Of course, while the efficiency result is the same in both cases, the distributional implications are quite different. In one instance, the polluted-firm has to pay the polluter, in the other the polluter has to pay. The Coase theorem also depends crucially on the absence of any transaction cost, i.e. any cost of bargaining. If bargaining costs are substantial, then the gains from negotiation may be neutralized entirely, and the socially optimal outcome will not be reached.

Identification of transaction costs can alert us to situations where in the absence of government intervention, the market will not be able to reach the desired outcome, even if property rights are established. These costs include among other things, (a) costs of discovering who is it one must deal with, (b) forming associations of affected agents to conduct bargaining, (c) conducting negotiations, (d) drawing up the contract, (e) undertaking inspection to ensure that the conditions of the contract are being made, and the like.

The Coase theorem has an interesting implication. When there is a negative externality, we have seen that the socially optimal output x_S is less than the privately optimal output x_p. This means that when output is reduced from x_p to x_s, the gain to the agent that is suffering from pollution must be greater than the loss to the polluting party. Hence there will be a tendency for the agents to merge, internalize the externality and share appropriately the increase in the total gain.

Suppose that there is a chemical dyes' firm and a fishing firm that is adversely affected by the operations of the former. Let 'm' be the output of the chemical dyes' firm which (for simplicity) sells its output in a perfectly competitive market at the price of Rs 5. Its cost function is $C(m) = m^2$. The chemical dyes' firm maximizes profit $\prod m = 5m - m^2$, by setting $5 = 2m$, so that $m^* = 5/2$. Its profit then is Rs 25/4.

Now suppose that the fishing firm is also perfectly competitive and faces a price of Rs 4. Its cost function is $C(f) = f^2 + mf$ Since $m = m^* = 5/2$,the fishing firm's profit function is $\prod f = 4f - f^2 - 2.5f$. It maximizes profit by setting $f = 3/4$. lts profit is then 9/16.

Instead, if both firms were jointly owned, then the owner would be interested in maximizing the joint profit

$$\Pi = \Pi m + \Pi f = 5m + 4f - m^2 - f^2 - mf$$

The first order conditions are $\frac{\partial \Pi}{\partial m} = \frac{\partial \Pi}{\partial f} = 0$ and the solutions then are f=1 and m = 2. We can see that the joint profit x = 7. This is higher than the aggregate profit of 25/+9/16 =6.8 that the firms make when they take decisions independently.

Hence, if the firms merge, total profit will increase and the extra profit can be shared between them. Therefore, when a negative externality exists, the goal of profit maximization itself will encourage firms to internalize the externalities.

CHECK YOUR PROGRESS–II

Q1 What is free riding ?

Q2. Why is free riding responsible for production of output less than socially optimum ?

Q3 What is coase theoram ? How does it address the issue of negative externalities ?

16.6. LET US SUM UP

Dear learner, in this lesson, we discussed the following:

- **Externalities** refer to the beneficial and detrimental effects of an economic unit (a firm, a consumer or an industry) on others. The beneficial externalities created by a consumer or a firm for others are known as **external economies**, and detrimental externalities imposed on the others by a productive firm or a consumer are known as **external diseconomies**.
- Under some circumstances, the **market system cannot lead to optimum situation of Pareto efficiency** (i.e., the state of maximum social welfare). These circumstances due to which market fails to achieve economic efficiency or maximum social welfare have been called **market failures**.
- **Public goods** are non-rival in consumption. Examples include **national**

defence, parks, television signals, flood control project, pollution control project, lighthouse in the sea.

- The other essential characteristic of a public good is **non-excludability** in distribution of their consumption benefits. This **non-exclusive** nature of a public good implies that it is difficult, if not impossible, to exclude those from consuming them who are not willing to pay for them.

16.7. KEYWORDS

1. **Externalities**

   Effects (positive or negative) of a firm's or individual's actions on others that are not reflected in market prices.

   - *Positive externalities* = External economies
   - *Negative externalities* = External diseconomies

2. **External Economies**

   Beneficial effects or spillovers produced by a firm or consumer that benefit others, without receiving direct compensation.

3. **External Diseconomies**

   Detrimental or harmful effects produced by a firm or consumer that impose costs on others, without paying for the damage caused.

4. **Market Failure**

   A situation where the market system fails to allocate resources efficiently, leading to a loss in social welfare. Common causes include externalities, public goods, monopoly power, and information asymmetry.

5. **Public Goods**

   Goods that are **non-rival** (one person's consumption does not reduce availability for others) and **non-excludable** (people cannot be excluded from using them even if they don't pay).

   Examples: National defence, parks, flood control, lighthouse.

6. **Non-Excludability**

   A characteristic of public goods where it is not possible to prevent someone from using the good, even if they haven't paid for it.

7. **Pareto Efficiency**

   An allocation where no one can be made better off without making someone else worse off. It represents an optimal distribution of resources.

8. **Optimum Allocation of Resources**

   A condition where resources are distributed in a way that maximizes total utility or welfare in the economy, often associated with Pareto efficiency.

9. **Coase Theorem**

   A proposition that if property rights are clearly defined and transaction costs are negligible, private parties can solve the problem of externalities among themselves without government intervention.

10. **Monopoly and Misallocation**

    In a monopoly, resources may be misallocated because output is restricted and prices are higher than marginal costs, leading to inefficiency and loss of social welfare.

16.8. EXAMINATION-ORIENTED QUESTIONS

1. What is meant by optimum allocation of resources? Show that monopoly leads to misallocation of resources.

2. What are the major sources of competitive market failure? Explain briefly in each case why the competitive market does not always operate efficiently.

3. Write a note on the role of externalities and public goods in causing market failure.

16.9 HINTS TO CHECK YOUR PROGRESS

CHECK YOUR PROGRESS I

Q1 – See Section 16.3

Q2 – See sub-Section 16.3.2

Q3 – See Section 16.4

Q4 – See Section 16.4.2

Check Your Progress II

Q1 – See Subsection 16.4.3.1

Q2 – See Subsection 16.4.3.2

Q3 – See section 16.5

16.10 SUGGESTED READINGS

1. W.J. Baumol and W.E. Oates, *The Theory of Environmental Policy*, 2nd Edition, Cambridge University Press, 1988.

2. Robert S. Pindyck and Daniel L. Rubinfeld, *Microeconomics*, 9th Edition, Pearson Education, 2017.

3. Ronald Coase, "The Problem of Social Cost," *Journal of Law and Economics*, Vol. 3, 1960, pp. 1-44.

4. William J. Baumol, *Economic Theory and Operations Analysis*, 4th Edition, Prentice Hall, 1978.

M.A. ECONOMICS LESSON NO. 17

COURSE NO. ECO 201 UNIT-III

POLLUTION VOUCHERS, COMMON PROPERTY RESOURCES, TRAGEDY OF THE COMMONS AND PRIVATE OWNERSHIP OF PROPERTY

STRUCTURE

17.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

- Discuss pollution vouchers, common property resources, property rights, and the tragedy of the commons.

17.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to:

- Explain how pollution vouchers work.
- Identify issues with common property resources.
- Discuss the role of property rights in resource management.
- Describe the tragedy of the commons and suggest solutions.

17.2 INTRODUCTION

Dear learner, in the earlier lesson, we have discussed about externalities, public goods and market failure as well as free riding that takes place in such cases. In this lesson, we shall be discussing about pollution vouchers and tragedy of the commons.

17.3 POLLUTION VOUCHERS

We know that "A negative externality means that when a private transaction imposes cost on society that are not paid by the private transactors."[1] A solution to the problem means ensuring that the externality is internalised by the one responsible for it. One way is to impose a pollution tax by the govt. which can force the polluter to internalise the externality by paying for it. The govt. can also use a command approach, whereby it dictates how much waste a particular company, say copper mining operation can generate.

However, these methods have not proved so effective. So the most popular of these is the marketable pollution permits or pollution vouchers, in which the govt. allows a market in which the polluters are given the right to pollute. The govt issues permits or vouchers to the owners of these polluting firms on which it is mentioned as to how much each is allowed to pollute. For example, if in an area, the target pollution level is 400 billion particulates per day and the govt. issues 400 such vouchers, it means each party can pollute up to 1 billion particulate per permit on daily basis. If a firm purchases 30 of these permits, it can emit pollutants up to 30 billion daily. But suppose, it adopts a cleaner technology and uses less than the no. of permits issued, it can also sell these to other firms.

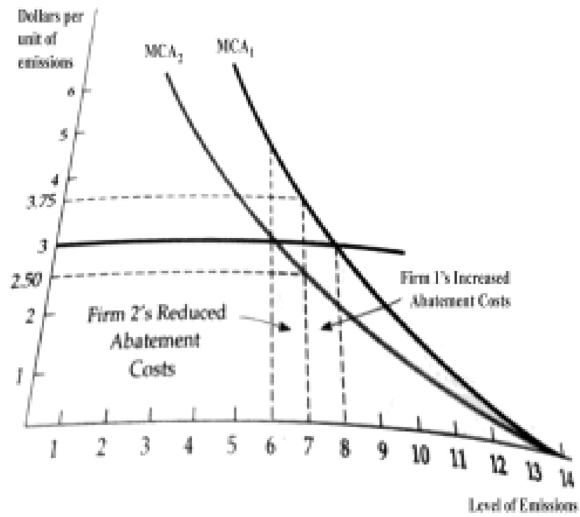


Fig. 17.1

We have drawn above the figure explaining the transferable or marketable emission permits just discussed. The X-axis shows the level of emissions allowed and the Y-axis, dollars per unit of emissions or price for purchasing these permits MCA_1 and MCA_2 are the Marginal cost of Abating (emission)curves of firm 1 and 2 respectively. The MCA curve means or measures the additional cost to the firm of installing pollution control equipment. The MCA curve is downward sloping, as the Marginal cost of reducing emissions is low when reduction has been slight, and high when it is MCA_1 substantial. MCA_1 has been drawn at a higher level than MCA_2, as firms differ with regard to abatement costs. Initially, each firm was generating 14 units of emissions. Suppose, the govt. aims to reduce emissions by 14 units at the least possible cost. It is possible if firm Areduces emissions by 6 units and B does it by 8 units. So the Marginal cost of Abatement to both is \$3. But suppose, now, the govt. permits both of them to emit only upto 7 units. This increases abatement cost (paid for buying a permit for one unit of emission) to firm 1 to \$3.75, whereas for firm 2, it reduces to \$2.50. In this case, the target of reducing emissions by 14 units is not cost effective, as the abatement cost is not the same for the two firms. Here the 'transferable emission permits' option can be used, whereby firm 2 can sell its permit to firm 1 for a price between\$ 2.50 and \$3.75.

Hence in this way, a market for externalities can be created, where firms reducing emissions the most (because they have low abatement cost) can sell permits to those who emit more (because abatement cost is high) ; and in this way, target emissions set by the govt. can be met at the minimum cost.

17.4 COMMON PROPERTY RESOURCES, PROPERTY RIGHTS AND TRAGEDY OF THE COMMONS

Market failure takes place when certain situations arise, including the one in which there are no well defined private property rights. A private property right is the right to claim ownership of an item-i.e. to do what we wish with that item. It is said to be well defined if there is a clear owner and if the right is recognized and enforced by society.

Having understood that, let us now try to answer whether anyone owns the right to the ocean, fish in the sea or airspace, in general , the 'commons'? Take the example of airspace. Suppose you live in an area, the airspace of which is being polluted due to emissions by the car driver. Though the driver is generating negative externality in your area, you can't stop him from doing that. The reason is that neither you own the airspace, nor does the driver. Had you owned it, you would have charged him a price as penalty; on the contrary you would have to pay him for not polluting, had he owned the airspace.

Like airspace, there are so many examples of commons i.e. those resources, which are not owned by anyone. These tend to be overused or utilized, as there is no payment for using these. Externalities created continue to be external, as nobody tells you to internalise them. Some other examples include animal populations, mineral exploration etc. These are together referred to as 'common property resources'. A resource to which anyone can have free access is a common property resource.

'The tragedy of the commons' is the name given to the problem created when something is owned commonly, rather than privately. Crawfish fishing in Louisiana, a popular restaurant item there , has been overfished over the years. Its harvest , which was just over one million pounds in 1950, increased to over 30 million pounds, before falling far below the efficient level.

Let us take an example to understand how tragedy of the commons occurs and how to mitigate the problem with private ownership of rights.

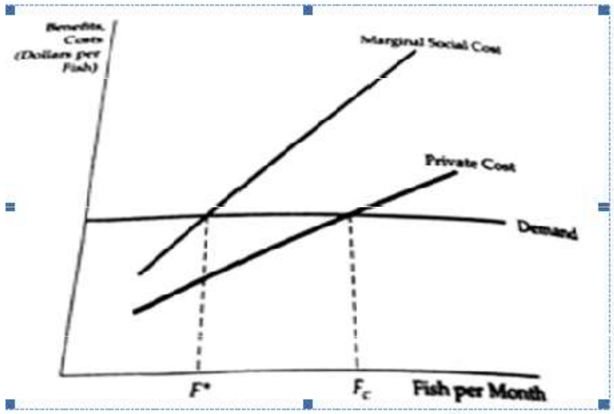


Fig. 17.2

Let there be a large lake with trout, to which an unlimited number of fishermen have access. Initially, each fisherman fishes up to the point at which Marginal revenue is equal to the (private) cost. The price (demand) line is given. As fishermen fish more and more, to that extent depletion of stock occurs and less opportunities are available for others to fish and earn revenue. But since it is a common property resource, there is no enforcement on the fisherman to behave responsibly. Here F* is the efficient point of fishing, as here Marginal benefit equals MSC. But F* is ignored by the fishermen as long as each fisherman's Marginal benefit (revenue) exceeds (private) cost. So fishing takes place up to point Fc. But here overfishing occurs. To find a solution to this problem, let one person own the right to this resource, who sets a fee for the use of this resource. The fee is equal to the marginal cost of depleting the stock of fish. It occurs at point F*. This forces the fishermen not to exceed F*.

Hence private ownership of property rights provides a solution to tragedy of the commons. For example, in their book (Textbook of Economics), William Boyes and Michael Melvis (2009 edition, p. 670) raise a question on cows and chickens vs. Endangered species. Many species like Amazon rainforests Big leaf Mahogany, Hawksbill sea turtles, Toothfish, Amazon parrots are among the ten most endangered species, because of their commercial value. So the difference between these 'animals and plants' and 'cows and chickens' is that the latter are privately owned. Whosoever owns them wants the population to flourish, so that more and more gains can accrue to him. In recent years, nations with more species on the verge of extinction have taken steps, including setting up of national parks, providing private property rights to individuals to own elephants (e.g. Botswana, Zimbabwe, South Africa), private ownership of forest land (Sweden) etc.

Another point needs to be added here. Private ownership has implications not only for natural resources, but also in other fields. We tend to spend on renovation of a house that is owned by us. That common ownership fails to incentivise the people to produce and consume is evident from the failure of communism. It has been said about China, that before 1990, people could not own homes. They had to live in govt. made and owned apartments. There were no renovations, as nobody bothered to take care of a building to which he had no right. But situation improved with private ownership, as substantial improvement with regard to beautification and cleanliness took place.

CHECK YOUR PROGRESS – I

Q1 Write a note on the concept of pollution vouchers (marketable pollution permits) and state how they help internalize the negative externality of pollution. Illustrate your answer with an example.

____

____

____

____

Q2. What is meant by the "tragedy of the commons"? Explain how private property rights can help mitigate this problem. Illustrate your answer with an example.

____

____

____

____

17.5. LET US SUM UP

In this lesson, we discussed the concepts of **pollution vouchers**, **common property resources**, and the **tragedy of the commons**. We learned how pollution vouchers serve as market-based instruments to control pollution, how the lack of well-defined property rights leads to overuse of shared resources, and how the tragedy of the commons illustrates the challenges of managing common property without regulation.

17.6 KEYWORDS

- **Negative Externalities :** Unintended harmful effects of an economic activity on third parties.
- **Transferable Emission Permits :** Marketable permits that allow holders to emit a certain amount of pollution, which can be bought and sold.
- **Emission Abatement Costs :** Costs incurred by firms to reduce pollution emissions.
- **Common Property Resource :** Resources owned collectively or not owned at all, leading to potential overuse.
- **Tragedy of the Commons :** The overexploitation of common resources due to individual self-interest.
- **Private Property Ownership :** Exclusive ownership rights that can help in resource management.

17.7 EXAMINATION-ORIENTED QUESTIONS

1. What are negative externalities?
2. How do transferable emission permits help solve the problem of negative externality?
3. Do you agree that a market for externalities can exist with pollution vouchers? Explain.
4. What are emission abatement costs?
5. What is a common property resource? Why do we tend to overutilise it?

6. What are commons? What is the tragedy of the commons?

7. How can the tragedy of the commons be mitigated?

8. What is private property ownership? Can it help overcome the tragedy of the commons? Explain diagrammatically.

9. Show how emission permits get traded between low emission producers and high emission producers. Draw a graph to explain it.

10. Why does free access to a common property resource generate inefficient outcomes? How can it be corrected?

11. Why do governments create national parks?

12. What is overfishing? Why does it take place?

13. Suppose a smoker creates negative externalities for non-smokers. How would the outcome change in case :

    (i) Smokers own the airspace,
    (ii) Non-smokers own the airspace,
    (iii) Nobody owns the airspace.

17.8 HINTS TO CHECK YOUR PROGRESS I

Q 1 See Section 17.3

Q 2 See Section 17.4

17.9 SUGGESTED READINGS

- Pindyck, Robert S., and Daniel L. Rubinfeld, *Microeconomics* (5th Edition)
- Boyes, William, and Michael Melvin, *Textbook of Economics* (6th Edition, Indian adaptation)

End Notes

1. Boyes and Melvin, *Textbook of Economics*, p. 662

M.A. ECONOMICS **LESSON NO. 18**

COURSE NO. ECO 201 **UNIT-III**

PUBLIC GOODS, OPTIMAL AND PRIVATE PROVISION FOR A PUBLIC GOOD, FREE RIDER PROBLEM, CLARK- GROVES TAX AND ITS PROBLEMS.

STRUCTURE

18.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

- Explain the concept of public goods.
- Discuss about the optional private provision of public goods.
- State the free rider problem and its implications.
- Describe the Clarke-Groves tax mechanism as a solution to public goods provision issues.

18.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to:

- Define and explain public goods and their characteristics.
- Describe how optional private provision of public goods works.
- Identify the free rider problem and why it causes under-provision of public goods.
- Explain the Clarke-Groves tax and how it incentivizes truthful revelation of preferences.

18.2 INTRODUCTION

Dear learner, public goods are those goods which are non-rival in consumption, Many public goods are provided by the government. For example, streets and sidewalks are provided by local municipalities. There are a certain number and quality of streets in a town and can use it, as it is available for all. National defence is another good example; there is one level of national defence provided for all the inhabitants of a country. Each citizen may value it differently—some may want more; some may want less—but they are all provided with the same amount. On the other hand, the free riding shows how non-exclucludiability of public goods can lead to the market failure, that is, failure of market to achieve pareto efficiency. In this lesson, we shall discuss these topics.

18.3 PUBLIC GOODS

In an earlier lesson, we argued that for certain kinds of externalities, it was not difficult to eliminate the inefficiencies. In case of a consumption externality between two people, for example, all one had to do was to ensure that initial property rights were clearly specified. People could then trade the right to generate the externality in the normal way. In case of production externalities, the provision of profit signals helped sort out the property rights in the most efficient way. In the case of common property, assigning property rights to someone would eliminate the inefficiency.

Unfortunately, not all externalities can be handled in that manner. As soon as there are more than two economic agents involved, things become much more difficult. Suppose, for example, that instead of the two roommates, one of whom is smoker and the other non smoker, we have three roommates–one smoker and two non smokers. Then the amount of smoke would be a negative externality for both of the non smokers.

Let us suppose that property right are well defined –say the non-smokers have the right to demand clean air. Just as before, although they have the right to clean air, they also have the right to trade some of that clean air away in return for appropriate compensation. But now there is a problem involved–the non-smokers have to agree among themselves how much smoke should be allowed and what the compensation should be. Perhaps one of the non-smokers is much more sensitive than the other, or one of them is much richer than the other. They may have very different preferences and resources and yet they both have to reach some kind of agreement to allow for an efficient allocation of smoke.

Instead of roommates, we can think of inhabitants of a whole country. How much pollution should be allowed in the country? If you think that reaching an agreement is difficult with only three roommates, imagine what it is like with millions of people!

The smoke externality with three people good—a good that must be provided in the same amount to all the affected consumers. In this case the amount of smoke generated will be the same for all consumers—each person may value it differently, but they all have to face the same amount.

Many public goods are provided by the government. For example, streets and sidewalks are provided by local municipalities. There are a certain number and quality of streets in town, anyone has that number available to use. National defense is another good example; there is one level of national defense provided for all the inhabitants of ac country. Each citizen may value it differently—some may want more, some may want less—but they are all provided with the same amount.

Public goods are an example of a particular kind of consumption externality: everyone must consume the same amount of the good. They are a particularly troublesome kind of externality, for the decentralization driven market solutions that economists are fond of don't work very well in allocating public goods. People can't purchase different amounts of public defense; somehow they have to decide on a common amount.

The first issue to examine is what the ideal amount of the public good should be. Then we'll discuss some ways that might be used to make social decisions about public goods.

18.3.1 When to provide a public good?

Let us start with a simple example. Suppose that there are two roommates, 1 and 2. They are trying to decide whether or not to purchase a TV. Given the size of their apartment, the TV will go in the living room, where both roommates will be able to watch it. Thus it will be a public good, rather than a private good. The question is, is it worth it for them to acquire the TV.? Let's use w_1 and w_2 to denote each person's initial wealth, g_1 and g_2 to denote each person's contribution to the TV, and x_1 and x_2 to denote each person's money left over to spend on private consumption. The budget constraints are given by :

$$x_1 + g_1 = w_1$$

$$x_2 + g_2 = w_2$$

We also suppose that the TV costs 'c' dollars, so that in order to purchase it, the sum of these two contributions must be a least 'c'.

i.e. $g_1 + g_2 > c.$

This equation summarizes the technology available to provide the public good: the roommates can acquire the TV if together they pay the cost 'c'.

The utility function of person 1 will depend on his or her private consumption,x_1, and the availability of the TV –the public good. We will write person 1's utility function as $u_1(x_1,G)$, where G will either be 0, indicating no TV, or 1, indicating that a TV is present. Consumption has a subscript to indicate that the good is consumed by person 1 or 2, but the public good has no subscript. It is consumed by both people. Of course, it isn't really consumed in the sense of being used up; rather, it is the services of the TV that are consumed by the two roommates.

Here we can talk of reservation price. It is that price, r_1, such that person 1 is just indifferent between paying r_1 and having the TV available, and not having the TV at all. If person 1 pays the reservation price and gets the TV, he will have $w_1—r_1$ available for private consumption. If he doesn't get TV, he will have w_1 available for private consumption. If he is just indifferent between these two lternatives, we must have $u_1(w_1—r_1, 1) = u_1(w_1, 0)$

This equation defines the reservation price for person 1—the maximum amount that he would be willing to pay to have the TV present. A similar equation defines the reservation price for person 2. Note that in general the function that defines the reservation price of each person will depend on that person's wealth. The maximum amount that an individual will be willing to pay will depend to some degree on how much that individual is able to pay.

Recall that an allocation is Pareto efficient if there is no way to make both people better off. An allocation is Pareto inefficient if there is some way to make both people better off; in the case, we say that a Pareto improvement is possible. In the TV problem there are only two sorts of allocations that are of interest. One is an allocation where the TV is not provided. This allocation takes the simple $(w_1, w_2, 0)$ that is, each person spends his wealth only on his private consumption.

The other kind of allocation is of the form $(x_1, x_2, 1)$, where

$$x_1 = w_1 - g_1$$

$$x_2 = w_2 - g_2$$

These two equations come from rewriting the budget constraints. They say that each individual's private consumption is determined by the wealth that has been left over after making his contribution to the public good.

Under what conditions should the TV be provided? That is, when is there a payment scheme (q_1, q_2) such that both people will be better off having the TV and playing their share than not having the TV? In the languages of Economics, when will it be a Pareto improvement to provide the TV?

It will be a Pareto improvement to provide the allocation $(x_1, x_2, 1)$, if both people would be better off having the TV provided than not having it provided.

This means

$$u_1(w_1,0) < u_1(x_1,1)$$

$$\& \quad u_2(w_2,0) < u_2(x_2,1)$$

Now we shall use the definition of the reservation prices r_1 and r_2 and the budget constraint to write :

$$u_1(w_1—r_1, 1) = u_1(w_1,0) < u_1(x_1,1) = u_1(w_1—g_1, \quad 1)$$

$$\& \quad u_2(w_2—r_2, 1) = u_2(w_2,0) < u_2(x_2,1) = u_2(w_2—g_2, \quad 1)$$

looking at the left and right hand side of these inequalities, and remembering that more private consumption must increase utility, we can conclude that

$$w_1—r_1 < w_1—g_1$$

$\&$ $\quad w_2—r_2 < w_2—g_2$ which in turn implies that

$$r_1 < g_1$$

$$\& \quad r_2 < g_2$$

This is a condition that must be satisfied if an allocation $(w_1, w_2, 0)$ is Pareto inefficient: it must be that the contribution that each person is making to TV is less than his willingness to pay for the TV. If a consumer can acquire the good for less than the maximum that he would be willing to pay, then the acquisition would be to his benefits. Thus the condition that the reservation price exceeds the cost share simply says that a Pareto improvement will result when each roommate can acquire the services of TV for less than the maximum that he would be willing to pay for it. This is clearly a necessary condition for purchase of the TV to be a Pareto improvement. If each roommate's willingness to pay exceeds his cost share, then the sum of the willingness of both to pay must be greater than the cost of the TV :

$$r_1 + r_2 > g_1 + g_2 = c$$

This condition is a sufficient condition for it to be a Pareto improvement to provide the TV. If the condition is satisfied, then there will be some payment plan such that both people will be made better off by providing the public good. If $r_1 + r_2 > c$, then the total amount that the roommates will be willing to pay is at least as large as the cost of purchase, so they can easily find a payment plan (g_1, g_2) such that $r_1 \geq g_1$, $r_2 \geq g_2$, and $g_1 + g_2 = c$. This condition is so simple that you might wonder why we went through all the detail in deriving it.

Here are some important points :

First, it is important to note that the condition describing when provision of the public good will be a Pareto improvement only depends on each agent's willingness to pay and on the total cost. If the sum of the reservation prices exceeds the cost of the TV, then there will always exist a payment scheme such that both people will be better off having the public good than not having it.

Second, whether or not it is Pareto efficient to provide the public good will, in general, depend on the initial distribution of wealth (w_1, w_2). Thus is true because, in general, the reservation price r_1 and r_2 will depend on the distribution of wealth. It is perfectly possible that for some distributions of wealth, $r_1 + r_2 > c$, and for other distributions of wealth, $r_1 + r_2 < c$.

To see how this can be, imagine a situation where one roommate really loves the TV and the other roommate is nearly indifferent about acquiring it. Then if the TV loving roommate had all of the wealth, he would be willing to pay more than the cost of the TV all by himself. Thus it would be a Pareto improvement to provide the TV. But if the indifferent roommate had all of the wealth, then the TV lover wouldn't have much money to contribute toward the TV, and it would be Pareto efficient not to provide the TV.

Thus, in general, whether or not the public good should be provided will depend on the distribution of wealth. But in specific cases the provision of the public good may be independent of the distribution of wealth. For example, suppose that the preferences of the two roommates were quasilinear. This means that the utility functions take the form

$$u_1(x_1,G) = x_1 + v_1(G)$$

&

$$u_2(x_2,G) = x_2 + v_2(G)$$

where G will be 0 or 1, depending on whether or not the public good is available. For simplicity, suppose that $U_1(0) = U_2(0) = 0$. This says that no TV provides zero utility from watching TV.

In this case the definitions of the reservation price become

$$u_1 (w_1 - r_1, 1) = w_1 - r_1 + v_1(1) = u_1(w_1,0) = w_1$$

$$u_2 (w_2 - r_2, 1) = w_2 - r_2 + v_2(1) = u_2(w_2,0) = w_2$$

which implies that the reservation prices are given by

$$r_1 = v_1(1)$$

$$r_2 = v_2(1)$$

Thus the reservation prices are independent of the amount of wealth, and hence the optimal provision of the public good will be independent of wealth, or at least some range of wealths.

18.3.2 Optimal Provision of a Public Good

In this section, we shall discuss the optimal level of a public good whose amount can be varied continuously. Consider an economy with only two goods, one private and the other public. Suppose that there are two agents, whose initial wealth levels are given by w_1 and w_2. Their respective contributions to the public good are given by g_1 and g_2, and let x_1 and x_2 denote the consumption of the private good of each person. Let G measure the amount of the public good (in rupees) and c(G) be its cost. If G amount of the public good is provided, then the two individuals have to spend Rs c(G).

The two agents face the constraint that their total initial wealth cannot exceed their total expenditures on the private good and the public good :

$$x_1 + x_2 + c(G) = w_1 + w_2$$

We consider a Pareto-efficient provision of the public good. The provision is Pareto-efficient if agent 1's utility is maximized given the utility level of agent 2. We must remember that both agents consume the same amount of the public good. The problem can then be written as

Max $u_1 (x_1 G)$,

$x_1, x_2,\ G$

subject to $u_1(x_1, G), = u_2{}^*$ and $x_1+x_2+c(G) = w_1+w_2$

let us from the Langrangean

$L = u_1(x_1, G)+\lambda_1\{u_2{}^*-u_2 (x_2,G)\} \ +\lambda_2\{w_1+w_2–x_1-x_2-c(G)\}$

The first order condition are

1. $\partial L/\partial x_1 = \partial u_1 /\partial x_1 -\lambda_2 =0$
2. $\partial L/\partial x_2 = \lambda \partial u_2 /\partial x_2 -\lambda_2 = 0$
3. $\partial L/\partial G= \partial u_1 /\partial G–\lambda_1 \partial u_2 /\partial G–\lambda_2 C'(G) = 0$
4. $\partial L/\partial\lambda_1 = u_2{}^*-u_2 (x_2,G) = 0$
5. $\partial L/\partial\lambda_2 = w_1+w_2-x_1–x_2–c (G) = 0$

From (1), we get (2) $\lambda_2 = \partial u_1 /\partial x_1$ and from (1) and (2), eliminating λ_2, we get $\lambda_1 = –(\partial u_1 /\partial x_1)/(\partial u_2 /\partial x_2)$ Using these values in (3), we get the condition $(\partial u_1 /\partial G)/(\partial u_1 /\partial x_1) + (\partial u_2 /\partial G)/(\partial u_2 /\partial x_2) = c' (G)$

This condition for the optimal provision of a public good can be written more properly as :

$MRS_{G1}+MRS_{G2}=MC_G$,

i.e. the sum of the marginal rates of substitution between the private good and the public good for the two individuals must equal the marginal cost of providing the public good.

If the efficiency condition is violated, we can show that at least one of the agents can be made better off and nobody made worse off. Suppose, for example, that the sum of the MRSs is less than the marginal cost. Let MC = 1, MRS_{G1} = 1/2 and MRS_{G2} = l/3. Then agent 1 would be willing to accept half a rupee more of the private good for the loss of Re 1 of the public good and agent 2 would be willing to accept 1/3 more rupees of the private good for the loss of Re 1 of the public good. Suppose we reduce the amount of the public good by Re 1, then we can compensate the two agents by giving them Rs 5/6, and still have Rs 1/6 left to distribute to the two individuals and make them better off. Thus if the sum of the MRSs is less than the MC, less of the public good and more of the private good should be provided.

Another way of interpreting the Pareto-efficiency condition is to think of the MRS as measuring the marginal willingness to pay for an extra unit of the public good. Then the efficiency condition simply says that at the margin, the sum of the willingness to pay of both must be equal to the cost of providing an extra unit of the public good.

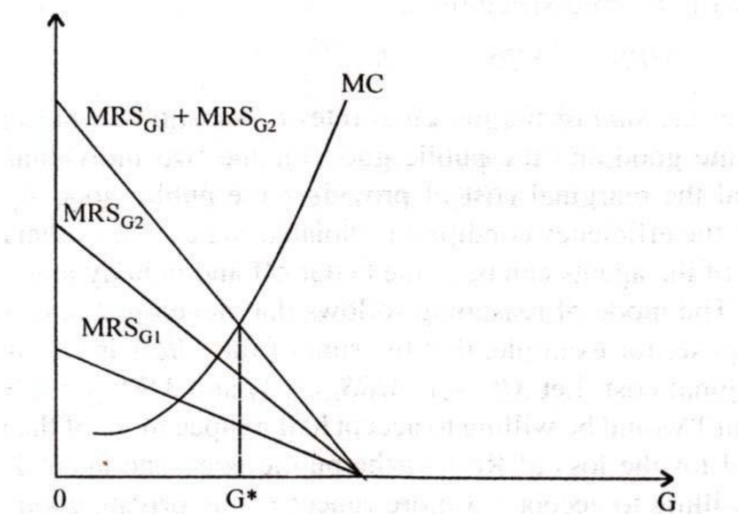


Fig. 18.1

The efficiency condition for a private good is that the MC should be equal to the MRS of each person separately. People can Consume different amounts of the private good, but they must all value it the same at the margin for efficiency. In the case of a public good, all individuals have to consume the same amount of the good, but they can value it differently at the margin.

The public good efficiency condition is illustrated in Fig. 18.1. The MRS curves are added vertically because both agents must consume the same amount of the public good. The efficient provision of the public good, G*, is obtained at the point of intersection of the MC and the $MRS_{G1} + MRS_{G2}$ curve.

18.3.3 Private Provision Of The Public Good

We have seen above that acquiring the TV will be Pareto efficient for the two roommates if the sum of their willingness to pay exceeds the cost of providing the public good. This answers the question about efficient allocation of the good, but it does not necessarily follow that they will actually decide to acquire the TV. Whether they actually decide to acquire the TV depends on the particular method they adopt to make joint decisions.

If the two roommates cooperate and truthfully reveal how much they value the TV, then it should not be difficult for them to agree on whether or not they should buy the TV. But under some circumstances, they may not have incentives to tell the truth about their values.

For example, suppose that each person valued the TV equally, and that each person's reservation price was greater than the cost, so that $r_1 > c$ and $r_2 > c$. Then person 1 might think that if he said he had 0 value for the TV, the other person would acquire it anyway. But person 2 could reason the same way! One can imagine other situations where both people would refuse to contribute in the hopes that the other person would go out and unilaterally purchase the TV.

In this kind of situation, economists say that the people are attempting to free ride on each other: each person hopes that the other person will purchase the public good on his own. Since each person will have full use of the service of the TV if it is acquired, each person has an incentive to try to pay as little as possible toward the provision of the TV.

18.4 THE FREE RIDER PROBLEM

Now that we know what the Pareto efficient allocations of public goods are; we can turn our attention to asking how to get there. In the case of private goods with no externalities we saw that the market mechanism will generate an efficient allocation. Will the market work in the case of public goods?

We can think of each person as having some endowment of a private good, w_1. Each person can spend some fraction of this private good on his own private consumption, or he or she can contribute some of it to purchase the public good.

Let's use x_1 for l's private consumption, and let g_1 denote the amount of the public good he buys, and similarly for person 2. Suppose for simplicity that $c(G) = G$, which implies that the marginal cost of providing a unit of the public good is constant at 1. The total amount of the public good provided will be $G = g_1 + g_2$. Since each person cares about the total amount of the public good provided, the utility function of person i will have the form $u_i(x_i , g_1 + g_2) = u_i(x_i, G)$.

In order for person 1 to decide how much he should contribute to the public good, he has to have some forecast of how much person 2 will contribute. The simplest thing to do here is to adopt the Nash equilibrium model. Suppose that person 2 will make some contribution g_2. We assume that person 2 also makes a guess about person 1's contribution, and we look for an equilibrium where each person is making an optimal contribution given the other person's behaviour. Thus person 1's maximization problem takes the form

$$\max_{x_1, g_1} u_1(x_1, g_1 + g_2)$$

such that $\quad x_1 + g_1 = w_1$

This is just like an ordinary consumer maximization problem. The optimization condition is therefore the same: if both the people purchase both the goods, the Marginal Rate of Substitution between the public and the private goods should be 1 for each consumer:

$$[\ MRS_1] = 1 \ [\ MRS_2] = 1.$$

However, we have to be careful here. It is true that if person 2 purchases any amount of the public good, he will purchase it until the Marginal Rate of Substitution equals one. But it can easily happen that person 2 decides that the amount already contributed by person 1 is sufficient and that it would therefore be unnecessary for him to contribute anything toward the public good at all.

Formally, we are assuming that the individuals can only make positive contributions to the public good-they can put money into the collection plate, but they can't take money out. Thus there is an extra constraint on each person's . But then it may well be that one person decides that the amount provided by the other is just fine and would prefer to make no contribution at all.

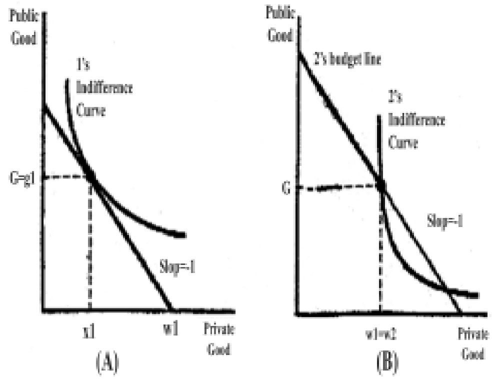


Fig 18.2

A case like this is depicted in Figure 18.2. Here we have illustrated each person's private consumption on the horizontal axis and his or her public consumption on the vertical axis. The "endowment" of each person consists of his or her wealth, w_i, along with the amount of the public good contribution of the other person-since this is how much of the public good will be available if the person in question decides not to contribute. Figure 18.2. shows a case where person 1 is the only contributor to the public good, so that g_1 = G. If person 1 contributes G units to the public good, then person 2's endowment will consist of her private wealth w_2 and the amount of the public good G-since person 2 gets to consume the public good whether or not he or she contributes to it. Since person 2 cannot reduce the amount of the public good, but can only increase it, her budget constraint is the bold line in Figure 18.2(B). Given the shape of 2's indifference curve, it is optimal from her point of view to free ride on 1's contribution and simply consume her endowment, as depicted.

This is an example where person 2 is free riding on person 1's contribution to the public good. Since a public good is a good that everyone must consume in the same amount, the provision of a public good by any one person will tend to reduce the other peoples' provision: Thus in general there will be too little of the public good supplied in a voluntary equilibrium, relative to efficient provision of the public good.

18.5 CLARK GROVES TAX AND ITS PROBLEMS

It is convenient to consider a numerical example to see just how the Clarke tax works. Suppose that we have three roommates who have to decide whether or not to acquire a TV that costs $300. They agree in advance that if they jointly decide to get the TV, then they will each contribute $100 towards the cost. Persons A and B are willing to pay $50 each to have the TV present, while person C is willing to pay $250. This information is summarized in Table 1.

Table 1 : Example of the Clarke Tax

Person	Cost share	Value	Net value	Clarke tax
A	100	50	−50	0
B	100	50	−50	0
C	100	250	150	100

Note that the TV provides a positive net value only to person C. Thus if the roommates voted on whether or not to purchase the TV, a majority would be opposed. Nevertheless, it is Pareto efficient to provide the TV since the sum of the values ($350) exceeds the cost ($300).

Let us consider how the Clarke tax works in this example. Consider person. The sum of the net values excluding person A is 100, and person A's net value is -50. Thus person A is not pivotal. Since person A is made worse off in the net value by the provision of the public good, he might have a temptation to exaggerate his bid downward. In order to ensure that the public good is not provided, A would have to bid -100 or below. But if he did this, then A would become pivotal, and he would have to pay a Clarke tax equal to the amount the other two people bid: -50 + 150 = 100. Thus reducing his bid saves him $50 in net value, but costs him $100 in taxes, leaving him with a net loss of $50.

The same thing goes for person B. What about person C? In the example person C is pivotal—without his bid the public good would not be supplied, and with his bid the good will be supplied. He receives a net value from the public good of $150, but pays a $100 tax, leaving him with a total value of his actions, of $50. Would it be worth it for him to increase his bid above his true value?. No, because that doesn't change any of his payoffs. Would it be worth it to reduce his bid? No, because that lowers the chance that the public good willbe supplied and doesn't change the amount of tax he has topay. Thus it is in the interest of each of the parties to truthfully reveal his or her net value of the public good. Honesty is the best policy—at least in a situation involving a Clarke Tax[1].

Problems With The Clark Tax : Despite the nice features of the Clarke tax it does have some problems. The first problem is that it only works with quasilinear preferences. This is because we can't have the amount that you have to pay influence your demand for the public good. It is important that there is a unique optimal level of the public good.

The second problem is that the Clarke tax doesn't really generate a Pareto efficient outcome. The level of the public good will be optimal, but the private consumption could be greater. This is because of the tax collection. Remember that in order to have the correct incentives, the pivotal people must actually pay some taxes that reflect the harm that they do to the other people. And these taxes cannot go to anybody else involved in the decision process, since that might affect their decisions. The taxes have to disappear from the system. And that's the problem if the taxes actually have to be paid, the private consumption will end up being lower than it could be otherwise, and therefore be Pareto inefficient.

However, the taxes have to be paid only if someone is pivotal. If there are many people involved in the decision, the probability that any one person is pivotal may not be very large; thus the tax collections might typically be expected to be rather small.

The final problem concerns the equity and efficiency tradeoff inherent in the Clarke tax. Since the payment scheme must be fixed in advance, there will generally be situations where some people will be made worse off by providing the public good, even though the Pareto efficient amount of the public good will be provided. To say that it is Pareto preferred to provide the public good is to say that there is some payment scheme for which everyone is better off having the public good provided than not having it. But this doesn't mean that for an arbitrary payment scheme everyone will be better off.

CHECK YOUR PROGRESS – I

Q1 Why is it difficult to allocate public goods efficiently when many people are involved?

____

____

____

____

Q2 Why do individuals tend to free ride on others' contributions when it comes to public goods, and what is the result of this behavior on the provision of public goods?

____

____

____

____

18.6 LET US SUM UP

Dear learner, in this lesson, we learned :

- **Public goods** are non-rival and often provided by the government, like streets and sidewalks.
- The **optimal provision** of a public good occurs when the total of individuals' willingness to pay (marginal rate of substitution) equals the **marginal cost** of providing it.
- The **free rider problem** happens when people use a good without paying, which leads to **market failure** and less than optimal production.

18.7 KEYWORDS

- **Public Goods :** Goods that are non-rival and non-excludable.
- **Non-Rivalry :** One person's use does not reduce availability for others.
- **Non-Excludability :** People cannot be prevented from using the good.
- **Free Rider Problem :** When someone benefits from a good without contributing to its cost.
- **Clarke-Groves Tax :** A tax system designed to encourage truthful reporting of preferences for public goods.

18.8 EXAMINATION ORIENTED QUESTIONS

1. What are public goods? Why is private production of them limited?
2. What are the key characteristics of public goods?
3. Explain the free rider problem. How does it affect public goods production?
4. Describe the optimal provision and private provision of a public good.
5. What is the Clarke-Groves tax? What problems are associated with it?

18.9 HINTS TO CHECK YOUR PROGRESS

Q1. See Section 18.3

Q2. See Section 18.4

18.10 SUGGESTED READINGS

- *Advanced Microeconomics* by Hal Varian
- Tideman, N. & Tullock, G., "A New and Superior Process for Making Social **Choices," *Journal of Political Economy*, 1976.**
- Pindyck R.S., D. Rubinfeld and P.L. Mehta (2012), Microeconomics, Pearson Education, South Asia.
- Salvatore, D. (2010), Microeconomics; Theory and Applications, Oxford University Press, New Delhi
- Cooper, R., & John, A. (2011). Economics–theory through applications, Saylor Foundation. [1]
- Shapiro, D., MacDonald, D., & Greenlaw, S. A. (2022). *Principles of Microeconomics 3e*.[2]

END NOTES

Note 1: Type of license/Conditions of use: CC BY-NC-SA (Attribution-Non commercial-ShareAlike) .Scan QR code 3.2 to access **Cooper, R., & John, A. (2011). Economics–theory through applications.**

Note 2 . This book is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). "Access for free at openstax.org." **Scan QR code 2.1 to access the e book by Shapiro, D., MacDonald, D., & Greenlaw, S. A. (2022). *Principles of Microeconomics 3e*.**

QR code 1

QR code 2

M.A. ECONOMICS LESSON NO. 19

COURSE NO. ECO 201 UNIT-III

EXPECTED UTILITY THEORY, VON- NEUMAN-MORGENSTERN METHOD, THE ST. PETERSBURG PARADOX AND BERNOULI HYPOTHESIS

STRUCTURE

19.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

- Discuss the Von Neumann-Morgenstern method of utility analysis.
- Examine the St. Petersburg Paradox.
- Explore Bernoulli's hypothesis related to decision-making under risk.

19.1 LEARNING OUTCOMES

After going through this lesson, you should be able to:

- Explain the concept of Expected Utility Theory.
- Describe the Neumann-Morgenstern method of constructing a utility index under risky conditions.
- Discuss the assumptions involved in constructing a Neumann-Morgenstern utility index.
- Illustrate the steps to construct a Neumann-Morgenstern utility index.
- Distinguish between the Neumann-Morgenstern utility measure and the neoclassical cardinal utility.
- Analyze the St. Petersburg Paradox and interpret Bernoulli's hypothesis.

19.2 INTRODUCTION

Dear learner, Neumann and Morgenstern in their now famous work, "Theory of Games and economic Behaviour" gave a method of cardinal measure of expected utility from win and prizes. On the basis of such a cardinal utility index called N-M index, rational decisions are made by the individuals in case of risky situations. Thus, Neumann-Morgenstern method seeks to assign a utility number, or in other words, construct a N-M utility index of the marginal utility of money which a person gets from extra amounts of money income. On the other hand, St. Petersburg paradox refers to the problem why most people are unwilling to participate in a fair game or bet. In this lesson, we shall discuss N-M index, St. Petersburg index and Bernoullian hypothesis.

19.3 THE EXPECTED UTILITY THEORY

The expected utility theory, developed by John von Neumann and Oskar Morgenstern, explicitly uses probabilities to help reach decisions in uncertain situations. In addition to states of nature, acts and their consequences, we now introduce the following:

A probability function expressing the beliefs of the individual as to the likelihood of nature closing each and every state.

An elementary utility function v(.) measuring the desirability of the different possible consequences. The function v is called the von Neumann-Morgenstern (vNM) utility function.

The expected utility rule says that the individual will choose that act x which maximizes the expected utility function

$U(x) = \sum_s psv$ [c (x, s] Ev(c), where E is the expectations operator.

This means that the utility of an act in an uncertain environment is equal to the mathematical expectation of the elementary utilities of associated consequences.

DISCUSSION

In an uncertain environment, whenever all act is decided upon, the consequence is not certain, various consequences may follow with different probabilities. Therefore, every act is like a gamble or a lottery. A lottery may be defined more formally as follows :

Suppose that a probability p_i can be associated with the state of nature s_i. That is, the probability that state of nature s_i will occur is p_i. Suppose an act a_i is chosen. This has consequences/pay-offs $c(a_i, 1)$, $c(a_i, 2)$,....., with associated probabilities p_1, p_2,……….. When we couple these consequences with their associated probabilities, we have a lottery:

$L(a_i) = [\{p_1, c(a_i, s_1)\}, \{p_2, c(a_i, s_2)\}, \ldots\ldots]$

Therefore, in an uncertain situation, the individual in essence has to choose between lotteries. If we make a number of assumptions about the way individuals rank lotteries (i.e. about their preferences relating to lotteries), then, just as in utility theory under certainty, there will exist a continuous and non-decreasing function U(.) that will represent these rankings.

Some additional assumptions put additional structure on U and enable us to express it as the expectation of the elementary utilities of the consequences.

Example: The vNM utility function of an individual is $v = \sqrt{m}$, and her initial wealth is 36. Will she accept a gamble in which she wins 13 with a probability 2/ 3 and lose 11 with probability 1/3?

The expected utility of the gamble is given by U= (2/3) + $\sqrt{36+13}$ (1/ 3) = $\sqrt{36}$ (2/3)7 + (1/3)5 = 19/3 > $\sqrt{36}$ = 6. Hence the gamble should be accepted.

19.4 NEUMANN-MORGENSTERN METHOD OF CONSTRUCTING UTILITY INDEX UNDER RISKY CONDITION

Making use of Bernoulli's idea that under risky and uncertain prospects as in betting, gambling, purchasing lottery tickets etc a rational individual will go by the expected utility rather than expected money values, Neumann and Morgenstern in their now famous work, "Theory of Games and economics Behaviour" gave a method of cardinal measure of expected utility from win and prizes. On the basis of such a cardinal utility index called N-M index, rational decisions are made by the individuals in case of risky situations. Thus, Neumann-Morgenstern method seeks to assign a utility number, or in other words, construct a N-M utility index of the marginal utility of money which a person gets from extra amounts of money income. The choices by an individual under risky and uncertain situations depend on the N-M utility index (i. e., expected numerical utilities) and changes in it with the changes in money income.

19.4.1. Assumptions Of Constructing N-M Utility Index

Before explaining the Neumann-Morgenstern method of measuring utility from money or the construction of N-M utility index, it will be better to describe the assumptions on which the method is based. Firstly, it is assumed that the individual possesses a scale of preference that is quite comprehensive and complete. This is similar to the assumption of indifference curve analysis of demand that the individual knows fully his indifference map depicting his scale of preferences. But unlike the indifference curve analysis of demand, the question here is the choices of "events". The events refer to the amounts of money some of which are "certain" and others uncertain, monetary amounts with probabilities or odds attached to them. Secondly, it is assumed that the individual can always say whether he prefers one event to another or he is indifferent between the two. This means that he can make probability calculations and on their basis can make comparison between the alternative events. For instance, he can compare the event of receiving Rs. 5,000 for sure, or Rs. 10,000 with 60 - 40 odds or any other probability, and can say whether he prefers one to the other or is indifferent between the two. Lastly, it is assumed that individual's choices are consistent.

19.4.2. Constructing Neumann-Morgenstern Utility Index

The first step in the Neumann-Morgenstern method of measurement of utility of money is to state the numerical probabilities of uncertain events of acquiring additional money. Suppose the purchase of a lottery ticket is under consideration of an individual. Let the prize be Rs. 5,000 which he will get if he wins and suppose if he loses, he will get the consolation prize of Rs.10. Further, suppose that odds are 60:40, that is, the probability of his winning is 0.6 and the probability of his losing (and therefore getting the consolation prize) is 0.4. With this the expected monetary value (generally called standard actuarial evaluation) of lottery ticket;

$= \square(W)+(1-\square)F$

where □ stands for probability of winning and W the monetary amount of the first prize and (1 - □) the probability of his losing and therefore of getting the consolation prize of the monetary value of F.

In our above numerical example, expected monetary value of lottery ticket

$= 0.6\ (5000) + (0.4)\ 10$

$= 3,000 + 4 \quad =3,004$

But the objective of Neumann-Morgenstern method is to evaluate the utility of a certain sum of money. For this, tool of certainty equivalents used. A certainty equivalent is the sure sum of money. An individual is indifferent between certainty equivalent and the probable amount of money with a certain probability in an uncertain and risky situation such as in a lottery ticket or gamble. For finding out the certainty equivalent, we have to interrogate the individual and to know his mind.

Take the above example of lottery ticket of the first prize of Rs. 5,000 with a probability of 0.6 of winning and 0.4 of losing (getting consolation prize of Rs. 10). We ask the individual X as to with how much sure sum of money and the above lottery ticket with the given prize and probabilities he would be indifferent. Suppose the individual reveals that he is indifferent between Rs. 3,000 and the above lottery ticket with the aforesaid prize and probabilities. Then, in this individual's case, certainty equivalent is Rs. 3,000. With this we can calculate the marginal utility of Rs. 3.000.

Since the individual is indifferent between the aforesaid lottery ticket and the sure sum of Rs 3,000, the expected utility of the lottery ticket will be equal to the utility of Rs. 3,000.

Therefore,

Utility of Rs. 3,000 = Expected utility of the lottery ticket

Rs. 3,000 = $\square.U(W)+(1-\square)U(F)$

= 0.6 x 500 utils + 0.4 x 1 util

= 300 utils + 0.4 utils

= 300.4 utils

Thus, for the individual, utility index of Rs. 3,000 is 300.4 utils. Hence:

Amount	Rs. 10	Rs. 5.000	Rs. 3000
Expected utility	1 util	500 utils	300.4 utils

Likewise, by using the above procedure utility index of other sums of money, by taking the examples of other lottery tickets and gambles with other probabilities and different amounts of prizes can be calculated. In this way a series of N – M utility indices can be converted into a schedule. prepared at different levels of income, which then can be converted into a curve.

19.4.3. Neumann-Morgenstern Utility Measure is Not The Same As Neoclassical CardinalMeasurement

It should be noted that Neumann-Morgenstern method of measuring utility cardinally is not the same thing as neoclassical cardinal measurement of utility. The neoclassicals sought to measure the strength of psychic feelings of individuals towards goods and service. In fact the word "cardinal" in these two types of analyses means two entirely different things. The Neumann-Morgenstern method does not measure the individual's quantities of satisfaction or pleasure from goods and service: it is intended to measure the utility of money with respect to predicting how an individual will make choices in riskyand uncertain situations. The Neumann- Morgenstern index is constructed by asking the consumers to choose among risky alternatives to know about individual's attitude toward gambling and not the evolution of amount of pleasure or satisfaction he obtains from the outcome of his choice. Thus, according to Professor Borumol, "the N-M utility index is cardinal in this very specific sense—it is intended to be used for making predictions. It is employed to predict which of two lottery tickets (or which of the two other risky alternatives) a person will prefer. We are given this individual's ranking of the alternative prizes offered by the lottery tickets and the odds on each prize. From this we wish to be able to infer by numerical calculation, and without actually asking the person, which lottery ticket he will choose.

Further, he writes, "It is not the purpose of the Neumann-Morgenstern utility index to set up any sort of measure of introspective pleasure intensity. Such a measure of strength of feelings" is totally unnecessary in the theory of game for which the N-M theory was constructed. Rather, the utility measure was set-up for purposes of calculation, or rather of prediction (in the subtler sense of the word) to permit the theorist to determine in the absence of the player which of several risky propositions theplayerwill prefer". Thuswe see that in Neumann-Morgenstern method, an individual's risky alternatives and lotteries come prior to constructing utility index. In this analysis it is not said that the individual prefers alternative A to alternative B because A gives him greater utility then B. Instead, because individual prefers A to B , the higher utility number is assigned to it.

19.5 THE ST PETERSBURG PARADOX AND BERNOULLI'S HYPOTHESIS

As said above, Daniel Bernoulli evinced great interest in the problem known as St. Petersburg paradox and tried to resolve this. St. Petersburg paradox refers to the problem why rational people are unwilling to participate in a fair game or bet. For example, offer of participating in a gamble in which a person has even chance (that is, 50-50 odds) of winning or losing Rs. 1000 is a fair game. To put in mathematical terms, a gamble whose expected value is zero, or more generally, the game in which the fee for the right to play is equal to its expected value is a fair one. Thus, according to St. Petersburg in an uncertain game a rational individual will not make a fair bet or, in other words, will not play the fair game.

Daniel Bernoulli[2] provided a convincing explanation of the said behaviour of a rational individual. According to him, a rational individual will take decisions under risky and uncertain situations on the basis of expected utility rather than expected monetary value. He further contended that marginal utility of money to the individual declines as he has more of it. Since the individual behaves on the basis of expected utility from the extra money if he wins a game and the marginal utility of money to him declines as he has extra money, a rational individual will not 'play the game', that is, will not make a bet. It is in this way that Bernoulli resolved 'St. Petersburg paradox'.

A graphic illustration will make clear Bernoulli's solution to the paradox. Consider Figure 19.1a. in which on the X-axis, the quantity of money (thousands of rupees) and on the Y-axis, marginal utility of money (rupees) to an individual are measured. Suppose an individual has 20 thousands of rupees with him and can make a bet at even odds (i.e.,50-50 chance) of winning or losing rupees one thousand. If he wins the bet, money with him will rise to 21 thousand (20 + 1) rupees. If as a result of an increase in money with him, his expected marginal utility of money declines, then the expected marginal utility of extra one thousand rupees to him which is depicted by the rectangle CDFE is less than the extra marginal utility of the previous one thousand (i.e.,20th thousand) rupees which is measured by the rectangle ABDC. In other words. The gain in utility in case of his winning the bet is less than the loss of utility in case of his losing the bet, though the gain and loss is the same in terms of monetary amount (i.e., Rs. one thousand). Thus, given the diminishing marginal utility of money the expected gain in utility is less than the expected loss of utility from one thousand rupees involved in the bet, a rational individual will therefore not make a bet with 50-50 odds.

It may, be further noted that a rational individual would be unwilling to bet or gamble even at favourable odds if his marginal utility of money declines very rapidly. For example, if a person is offered a bet; if he wins the bet. He will be given Rs. 1500 and if he loses it, he has to pay Rs. 1000, then as in case of rapid decline in marginal utility of money, he may be unwilling to agree to make the bet. Consider Figure.19.1.b where the individual has Rs. 20,000 at present. In case he wins the bet, his monetary gain will be Rs. 1500 which will raise the money income to Rs. 21,500 and gain in his total utility will be given by the area ABCD and if he loses the bet, his income falls by Rs. 1000 to Rs. 19,000 and as a result he suffers a loss in total utility equal to the area EFAB. It will be observed from Figure 19.b. that despite a smaller loss in money terms, the loss in terms of total utility is greater than the gain in total utility despite a greater increase in money in case he wins the bet. This happened due to the rapid decline in marginal utility of money as individual's money increases.

It may be pointed out that in our discussion above about the individual's betting it is assumed that individual derives no pleasure from gambling, that is, he does not enjoy gambling for its own-sake. This is another way of saying that the individual behaves rationally in the sense that he will behave on the basis of expected gains and losses of utility from winning and losing money through gambling.

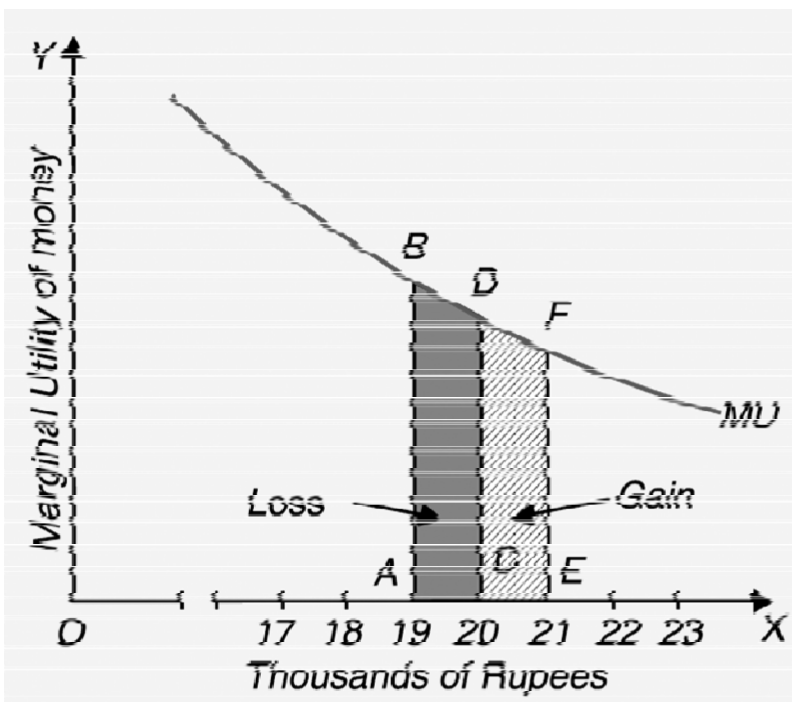


Fig. 19.a. Bernouli's hypothesis unwillingness to participate in fair game

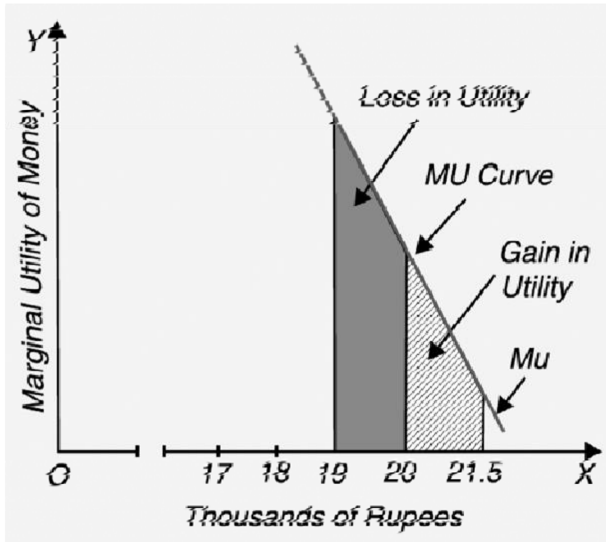


Fig. 19.b. Unwillingness to participate at favourable odds when MU of money declines rapidly

Bernoulli's hypothesis that individual's decision to participate in a gamble or not, depends on his expected utility rather than expected money value of the game is of crucial significance in any discussion of individual's behaviour under riskyand uncertain situations. So long as there is no upper bound on the utility function, the prize in a gamble can be appropriately adjusted so that the paradox is regenerated. Further Bernoulli's main point that an individual considers expected utility from the extra money rather than monetary value of the gain itself has found wide acceptance among economists.

However, a major drawback of Bernoulli's expected utility hypothesis is that it assumes cardinally measurable utility which economists today find difficult to believe. As will be explained later, J. von Neumann and O. Morgenstern adopted an entirely new approach to assigning numerical values to the utilities obtained from extra money by the individuals behaving in risky or uncertain situations, such as in case of gambling and insurance and they based their method of constructing utility index (which is arrived at in a different way from the cardinal measurement of utility by neoclassical economists) on the expected utility hypothesis of Bernoulli. They showed that we can analyse the choice by an individual under risky and uncertain situation on the basis of N-M utilityindex.

CHECK YOUR PROGRESS – I

Q1: What is the St. Petersburg Paradox?

____

____

____

____

Q2 What is a certainty equivalent in the Neumann-Morgenstern method?

____

____

____

____

Q3: According to Bernoulli, why do rational individuals avoid fair gambles?

____

____

____

____

19.6 LET US SUM UP

Dear learner, in this lesson, we explored how individuals make decisions under risk and uncertainty using expected utility theory. The key points covered are:

- **Neumann and Morgenstern**, in their seminal work *"Theory of Games and Economic Behavior"*, introduced a method to derive a **cardinal measure of expected utility** based on monetary outcomes. This is known as the **Neumann-Morgenstern (N-M) utility index**.

- The N-M method assigns utility values to various outcomes to help rational individuals make decisions in risky situations by comparing **expected utilities** rather than expected monetary values.
- The **St. Petersburg Paradox** demonstrates why individuals are reluctant to participate in a fair gamble with infinite expected monetary value.
- **Daniel Bernoulli** resolved the paradox by suggesting that rational individuals base decisions on **expected utility**, which diminishes with increasing wealth, rather than on expected monetary gain.

19.7 KEYWORDS

Expected Utility Theory: A theory used to model decision-making under risk, where individuals choose between risky or uncertain options by comparing the expected utility of outcomes rather than their expected monetary value.

- **Neumann-Morgenstern Utility Index**: A cardinal utility measure proposed by John von Neumann and Oskar Morgenstern, used to assign numerical utility values to outcomes to enable rational decision-making under uncertainty.
- **Cardinal Utility**: A type of utility measurement where the difference between utility values has meaning, allowing comparison of the strength of preferences.
- **St. Petersburg Paradox**: A classical problem in probability and economics showing that a gamble with infinite expected monetary value may not be appealing to most people, questioning the use of expected value as a decision rule.
- **Expected Monetary Value (EMV)**: The average of all possible monetary outcomes of a gamble, weighted by their probabilities.
- **Daniel Bernoulli's Resolution**: Bernoulli suggested that people evaluate risky choices based on expected utility, not expected monetary gain, and that utility increases at a decreasing rate with wealth (diminishing marginal utility).

- **Diminishing Marginal Utility of Wealth**: The concept that each additional unit of wealth provides less added satisfaction (utility) than the previous one.

19.8 EXAMINATION-ORIENTED QUESTIONS

1. What is meant by risk, and how does it differ from uncertainty?
2. What is a fair gamble? Why do most individuals refuse to play a fair gamble?
3. What is the St. Petersburg Paradox? How did Bernoulli resolve it?
4. What is the Neumann-Morgenstern concept of utility index under risky conditions? How is it constructed?
5. Under what circumstances do rational people avoid fair games?
6. What is meant by favourable odds? Explain with the help of a diagram.

19.9. HINTS TO CHECK YOUR PROGRESS

Q1 – See Section 19.4

Q2 – See Subsection 19.4.2

Q3 – See Section 19.5

19.10. SUGGESTED READINGS

- Ahuja, H.L. *Advanced Economic Theory: Microeconomic Analysis*, S. Chand & Company Ltd.
- Henderson, J., and R. Quandt. *Microeconomic Theory: A Mathematical Approach*, McGraw-Hill.
- Hader, J. *Mathematical Theory of Economic Behavior*, Addison-Wesley.
- Koutsoyiannis, A. *Modern Microeconomics*, Macmillan.
- Mas-Colell, A., Whinston, M.D., & Green, J.R. *Microeconomic Theory*, Oxford University Press.

- Salvatore, D. *Microeconomics*, Oxford University Press.
- Sen, A. *Microeconomics*, Oxford University Press.
- Varian, H. *Intermediate Microeconomics*, East-West Press.

ENDNOTES

1. John Von Neumann and Oskar Morgenstern, *Theory of Games and Economic Behavior*, 2nd Edition, Princeton University Press, 1947.
2. D. Bernoulli, *Exposition of a New Theory on the Measurement of Risk*, *Econometrica*, January 1954, pp. 23–36.

M.A. ECONOMICS **LESSON NO. 20**

COURSE NO. ECO 201 **UNIT-IV**

ALLAIS PARADOX, ATTITUDE TOWARD RISK, RISK LOVER AND GAMBLING, RISK AVERSION AND INSURANCE

STRUCTURE

20.0. OBJECTIVES

The main objectives of this lesson are to enable you to understand:

- Define Allais Paradox
- Discuss different attitudes toward risk

- Explain the concept of a risk lover and the role of gambling
- Describe Risk aversion and the importance of insurance

20.1. LEARNING OUTCOMES

After going through this lesson, you shall be able to:

- Discuss Allais Paradox and its implications for decision-making under risk
- Differentiate between risk-averse, risk-neutral, and risk-loving behaviors
- Analyze why individuals engage in gambling despite the odds
- Explain the rationale behind purchasing insurance and its link to risk aversion

20.2 INTRODUCTION

Dear learner, according to M. Allais, under situations of uncertainty, people tend to make choices inconsistently. Psychologists Daniel Kahneman and Amos Tversky explain this phenomenon by a constant factor has a larger impact when the outcome was initially certain than when it was merely probable'. In this lesson, we shall discuss this and related concepts.

20.3 THE ALLAIS PARADOX

The French economist M. Allais has shown that under situations of uncertainty, people tend to make choices inconsistently. To illustrate, we consider two situations of choice. Situation 1. The choice is between (A) a sure win of Rs 30 and (A') an 80 per cent chance to win Rs 45, and 20 per cent chance to win O. In this situation, most people tend to choose A, even though A'has an expected payoff of Rs 36. However, the choice is not surprising if the individual is risk-averse.

Situation 2. (B) A 25 per cent chance to win Rs 30 versus (B') a 20 per cent chance to win Rs 45. Here, most people tend to choose B'. This, again, is not surprising, because the expected payoff with B' is Rs 9 and that with B is Rs 7.50. Moreover, both are situations of risk.

However, the most popular pair of choices (A and B') taken together contradict the assumption of expected utility maximization. To show this, we assume that the Von-Neumann Morgenstem (VNM) utility function is v(.) and the initial wealth is w.

[Note : v(.) is an elementary utility function measuring the desirability of the different possible consequences.]

Then the choice of A over A' implies

1. $v(w + 30) > 0.8v(w + 45) + 0.2v(w)$.

The choice of B', on the other hand implies that

$O.2v(w + 45) + 0.8v(w) > O.25v(w + 30) + 0.75v(w)$, that is

2. $O.25v(w + 30) < O.2v(w + 45) + 0.05v(w)$.

Dividing both sides of (2) by 0.25, we get

$2v(w + 30) < 0.8v(w + 45) + O.2v(w)$,

which contradicts (1). Hence there is an inconsistency in the choices under the two situations. Psychologists Daniel Kahneman and Amos Tversky explain this phenomenon by arguing that 'a reduction in the probability of an outcome by a constant factorhasa larger impact whentheoutcome wasinitially certain than when it was merely probable'. Thus in the first pair of alternatives, the movement from A to A'represented a 20 per cent reduction in the chances of winning (from 100 per cent to 80 per cent), and so did the movement from B to B' (25 per cent to 20 per cent). But because the first movement was from an initially certain outcome, individuals disliked it much more.[1]

20.4 ATTITUDE TOWARDS RISK

We analyzed the problem of investment choice to describe how people make a choice among alternative investments involving different degrees of risk involved. But the underlying principles of making a choice in risky and uncertain situation, namely, expected return and the degree of risk involved apply equally well to other choices. In this section we focus on examining individual's choices in the face of risk. In the various earlier theories of consumer's behaviour we saw that in making choices among commodity bundles when there is no risk and uncertainty, the consumer maximizes his utility. We will analyse below how an individual maximizes his expected utility when risk or uncertainty is present.

People's preferences towards risk greatly differ. Most individuals generally prefer the less risky situation (that is, the situation with less variability in outcomes or rewards). In other words, most individuals seek to minimize risk and are called risk averters or risk averse. However, some individuals prefer risk and are therefore called risk-seekers or risk lovers. Some other individuals are indifferent towards risk and are called risk-neutral. But it is important to note that these different preferences towards risk depend on whether for an individual Marginal Utility of money diminishes or increases or remains constant. As shall be explained below, for a risk averse individual, Marginal Utility of money diminishes as he has more money, while, for a risk-seeker Marginal Utility of money increases as money with him increases. In case of a risk-neutral individual, Marginal Utility of money remains constant as he has more money.

Risk Averter. To explain the attitude towards risk, we will consider a single composite commodity, namely, money income. An individual's money income represents the market basket of goods that he can buy. It is assumed that the individual knows the probabilities of making or gaining money income in different situations. But the outcomes or payoffs are measured in terms of utility[1] rather than rupees.

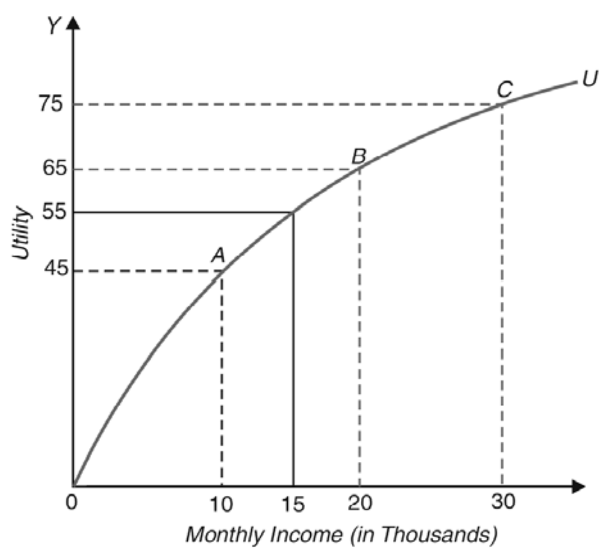


Fig. 20.1 : Utility Function of a Risk Averter

In Fig. 20.1 we have drawn a curve OU showing utility function of money income of an individual who is risk-averse. It will be seen from this figure that the slope of total utility function OU decreases as the money income of the individual increases. Note that we measure money income on the X-axis and utility on the y-axis. It will be seen from Fig. 20.1 that as money income of the individual increases from 10 to 20 thousand rupee, his total utility increases from 45 units to 65 (that is, by 20 units) and when his money increases from 20 thousand to 30 thousand rupees his total utility increases from 65 to 75 units (that is, by 10 units).

Thus in this concave utility function depicted in Fig. 20.1 marginal utility of money of an individual decreases as his money income decreases and therefore it represent the case of risk-averse individual. Suppose the individual is currently employed on a fixed monthly salary basis of Rs. 15000. There is no uncertainty about the income from his present job on the fixed salary basis and hence no risk.

Now, suppose that the individual is told to consider joining a new job of a salesman on a commission basis. The new job involves risk because his income in this case is not certain. This is because if he proves to be a successful salesman, his income may increase to Rs. 30 thousand per month but if he does not happen to be a good salesman his income may go down to Rs. 10 thousand per month. Suppose in this new job there is 50-50 chance of either earning Rs. 30 thousands or Rs. 10 thousands (that is, each has a probability of 0.5) When there is uncertainty, the individual does not know the actual utility form taking a particular action.

But given the probabilities of alternative outcomes, we can calculate the expected utility. Whether the individual will choose the new risky job or retain the present salaried job with a certain income can be known by comparing the expected utility from the new risky job with the utility of the current job. It will be seen firm the utility function curve OU in figure 20.1 that the utility of the money income of Rs 15,000 with certainty is 55. Further, in case of new risky job if he is proved to be a successful salesman and his income increases to Rs. 30 thousand, his utility from Rs. 30 thousands is 75, and if he fails as a good salesman, his income falls to Rs. 10 thousands which yields him utility of 45. (Note that in the risky job, the expected income is 20,000 which is given by or E(X) = 0.5 x 10,000 + 0.5 x 30,000 = Rs. 20,000). Given that the probability of success or failure as a salesman is 0.5, the expected utility of the new job is given by

$$E(U) = \square U(10,000) + 1 - \square U(30,000)$$

$$E(U) = 0.5\ U(10,000) + 0.5\ \ U(30,000)$$

$$=0.5x45+0.5x75$$

$$= 22.5 + 37.5$$

$$= 60.0$$

[Note : 45 is the utility on point A & 75 on point C of OU Curve]

Thus utilitiy with the present job with a fixed salary of Rs. 15,000 with no uncertainty is 55, whereas the expected utility of the new job or salesman on commission basis is 60. Though the individual is risk-averse as revealed by the nature of his utility function of money income, but since the expected utility of the risky job is greater than the utility of the present job with a certain income he will choose the risky job.

Let us now slightly change the data. Suppose that if the individual in his new job proves to be successful and earns Rs. 30,000, double the present assured income of Rs. 15,000, but if he fails in his new risky lob of a salesman on commission basis, his income falls to zero, then the expected utility of the risky job is given by

$$E(U) = 0.5\ U(O) + 0.5\ \ U(30,000)$$

$$=0+0.5x75$$

$$= 37.5$$

Now the expected utility from the new risky job is less than the utility of 55 from the present job with an assured income of Rs. 15,000 (Note that in the risky job also, expected income is Rs. 15,000 [E(x) = 0.5 x 0 + 0.5 x 30,000 = 15000]. Note again that in Figure 20.1 we are considering the choice of a risk averse individual for whom Marginal Utility of money declines as he has more of it. We are now in a position to provide a precise definition of risk-averse individual. Precisely speaking, a person who prefers a certain given income to a risky job with the same expected income is called risk averter or risk-averse. Risk aversion is the most common attitude toward risk.

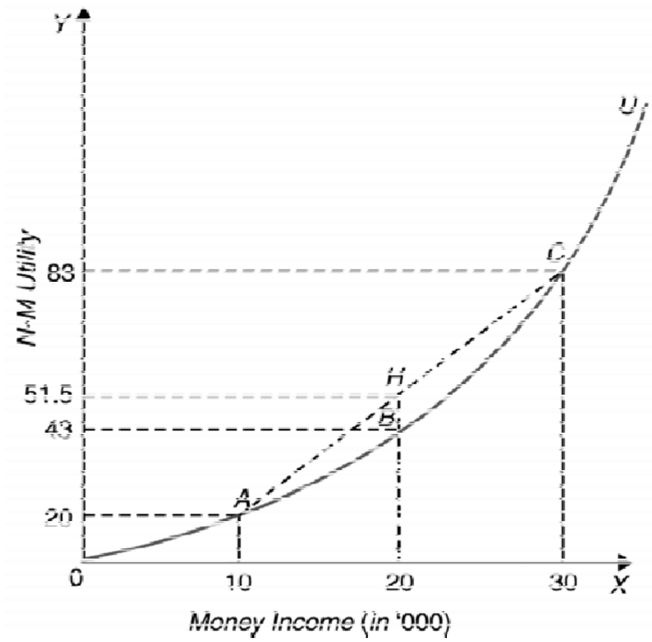


Fig. 20.2. Utility Function of a Risk-Seeker

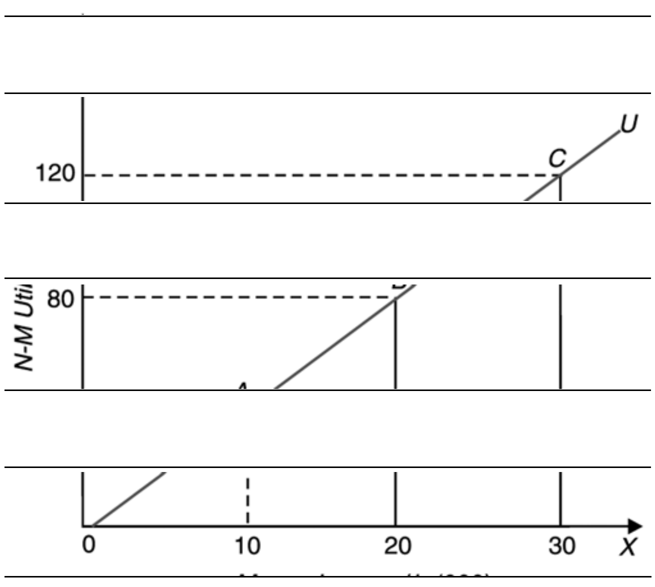


Fig. 20.3. Utility Function of a Risk-Neutral

Risk Lover On the other hand, a person is risk-preferer or risk-loving who prefers a risky outcome with the same expected income as a certain income. In case of a risk-loving individual, Marginal Utility of income to the individual increases as his money income increases as shown by the convex total utility function curve OU in Fig. 20.2. Suppose this risk-loving individual has a present job with a certain income of Rs. 20 thousands. It will be seen from the figure that the utility of Rs. 20 thousands is 43 units to this individual. Now, if he is offered a risky job with his income of Rs. 30 thousands if he happens to be highly efficient and Rs 10 thousands if he happens to be not so efficient in the new job with the equal probability of 0.5 in these two jobs, then the expected utility from the new job is given by

$$E(U) = 0.5\ U(10{,}000) + 0.5\ U(30{,}000)$$

It will be seen from Fig. 20.2 that the utility of Rs. 10 thousands to this individual is Rs. 20 while utility of Rs. 30 thousands to him is 83. Therefore,

$$\begin{aligned} E(U) &= 0.5\,(20) + 0.5\,(83) \\ &= 10 + 41.5 \\ &= 51.5 \end{aligned}$$

Since the expected utility from the new risky job is 51.5 which is greater than the utility of 43 from the present job with a certain income of Rs. 20 thousands, the risk-loving individual will prefer the new risky job even though the expected income in the new risky job is also Rs. 20,000 as (0.5 x 10,000) + 0.5 (30,000) = Rs. 20,000).

As mentioned above, most of the individuals are risk averse but there is a good deal of evidence of people who are risk seekers. It is risk-loving individuals who indulge in gambling, buy lotteries, engage in criminal activities such as robberies, big frauds even at risk of getting heavy punishment if caught.

Risk-Neutral. A person is called risk neutral, if he is indifferent between a certain given income and an uncertain income with the same expected value. An individual will be risk neutral if his Marginal Utility of money income remains constant with the increase in his money. The total utility function of a risk neutral person is shown in Fig. 20.3. It will be seen from this figure that utility of a certain income of Rs. 20 thousands is 80. Now, in a risky job when income increases to Rs. 30 thousands if he proves to be a successful salesman, the utility of Rs. 30 thousands is 120 units. On the other hand, if in a new risky job, he proves to be a bad salesman, his income goes down to Rs. 10,000 whose utility to the individual is 40 units. We assume that there is equal probability of high and low income in the new risky job. Note that expected value of income in the new job with an uncertain income is 20,000 as (0.5 x 10,000 + 0.5 (30,000)= 20,000.

The expected utility of the new risky job is given by

$$
\begin{aligned}
E(U) &= 0.5\ U(10{,}000) + 0.5\ U(30{,}000). \\
&= 0.5\ (40) + 0.5\ (120) \\
&= 20+60 \\
&= 80
\end{aligned}
$$

It is seen from above that in case of risk-neutral person, Expected utility of an uncertain income with the same expected value (Rs. 20,000 in the present case), is equal to utility of an assured or a certain income. That is, risk-neutral person is indifferent between them.

20.5 RISK AVERSION AND FAIR BETS

People differ greatly in their attitudes towards risk. In Bernoulli's hypothesis we have seen that a person whose Marginal Utility of money declines will refuse to accept a fair gamble. A fair gamble or gamble is one in which the expected value of income from a gamble is equal to the same amount of income with certainty. That person who refuses a fair bet is said to be risk averse. Thus, the risk averter is one who prefers a given income with certainty to a risky gamble with the same expected value of income. Risk aversion is the most common attitude towards risk. It is because of the attitude of risk aversion that many people insure against various kinds of risk such as burning down of a house, sudden illness of a severe nature, car accident and also prefer jobs or occupations with stable income to jobs or occupations with uncertain income.

This attitude of risk aversion can be explained with Neumann-Morgenstern method of measuring expected utility. It may be noted that marginal utility of income of a risk-averter diminishes as his income increases. In Figure 20.4, Neumann-Morgenstern utility function curve U(I) has been drawn. It will be seen from this figure that N-M utility curve starts from the origin and has a positive slope throughout indicating that the individual prefers more income to less. Further, the N-M utility curve shown in Figure 20.4 is concave which shows the marginal utility of income of a person diminishes as his income increases. Therefore, the utility curve in Figure 20.4 represents the case of a risk averter or the attitude of risk aversion. With Rs. 2,000 income, the person's utility is 50 which rises to 70 when his income increases to Rs. 3,000. As his income further increases to Rs. 4,000, his utility rises to 75.

Now suppose the person's current income is Rs. 3,000 and he is offered a fair gamble in which he has a 50-50 chance of winning or losing Rs. 1,000. Thus, the probability of his winning is l/2 or 0.5. If he wins the game, his income will rise to Rs. 4,000 and if he loses the gamble, his income will fall to Rs. 2,000. The expected money value of his income in this situation of uncertain outcome is given by:

$$E(V) = \frac{1}{2} \times 4000 + \frac{1}{2} \times 2000 = \text{Rs. } 3000$$

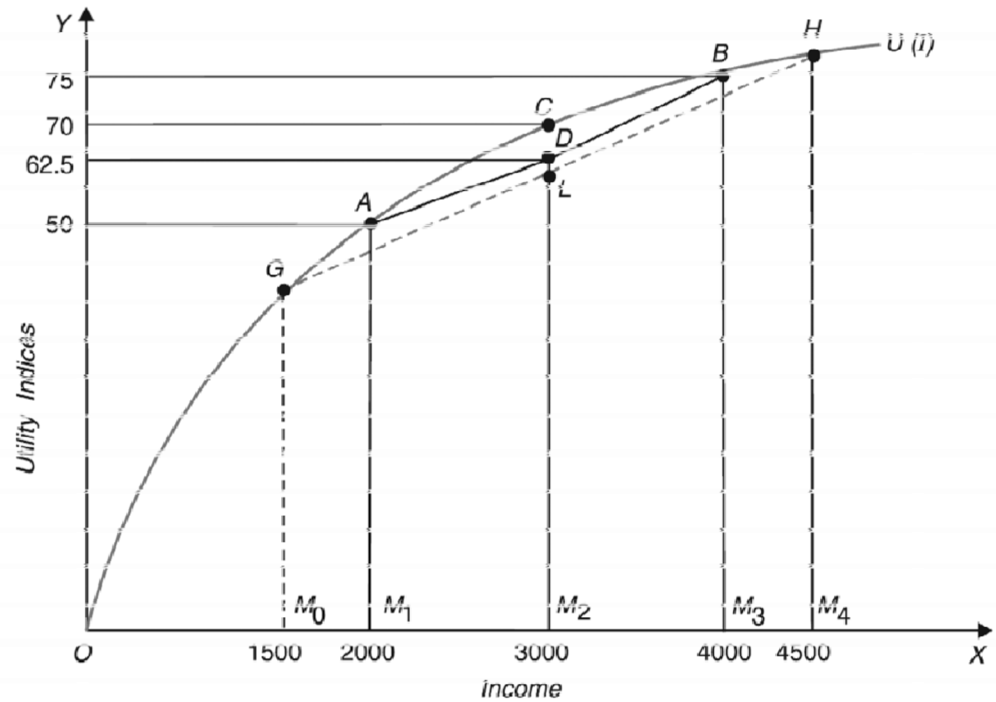


Fig. 20.4

If he rejects the gamble he will have the present income (i.e., Rs. 3,000) with certainty. Though the expected value of his uncertain income prospect is equal to his income with certainty, a risk averter will not accept the gamble. This is because as he acts on the basis of expected utility of his income in the uncertain situation (that is, Rs. 4,000 if he wins and Rs. 2,000 if he loses) can be obtained as under:

Expected Utility (EU) = □U(Rs. 4000) + 1-□U (Rs. 2000)

As will be seen from Figure 20.4 the utility of the person from Rs. 4.000 is 75 (point B on the utility curve and utility from 2000 is 50 (point A in Figure 20.4) the expected utility from this uncertain prospect will be :

$$E(U) = \frac{1}{2}(75) + \frac{1}{2}(50)$$
$$= 37.5 + 25 = 62.5$$

In the N-M utility curve U(I) in Figure 20.4 the expected utility can be found by joining point A (corresponding to Rs. 2,000) and point B (corresponding to Rs. 4,000) by a straight line segment AB and then reading a point on it corresponding to the expected value of the gamble Rs. 3,000, the expected value of the utility is M_2D (:62.5) which is less than M_2C or Rs. 70 which is the utility of income of Rs. 3,000 with-certainty. Therefore, the person will refuse to accept the gamble (that is, he will not gamble). It should be carefully noted that his rejection of gamble is due to diminishing marginal utility of money income for him. The gain in utility from Rs. 1,000 in case he wins is less than the loss in utility from Rs. 1000 if he loses the gamble. That is why his expected utility from the uncertain income prospect has been found to be lower than the utility he obtains from the same income with certainty.

It follows from above that in case marginal utility of money income diminishes a person will avoid fair gambles. Such a person is called risk averter as he prefers an income with certainty (i.e., whose variability or risk is zero) to the gamble with the same expected value (where variability or risk is greater than zero). Let us illustrate it with another example. Suppose to our person with a certain income of

Rs. 3,000, two fair gambles are offered. First, a 50:50 chance of winning or losing Rs. 1000 as before and the second a 50:50 chance of winning or losing Rs.1,500 with the even chance of winning and losing. The expected value of income in the second gamble will be 1/2(1500) + 1/2 (4500) = Rs. 3000. On the N- M utility curve U(I) in Figure 20.4, we draw a straight line segment GH joining point G (corresponding to income of Rs. 1500) and H corresponding to income of Rs. 4500). It will be seen from this straight-line segment GH that the expected utility from the expected money value of Rs. 3,000 from the second gamble is M_2L which is less than M_2D of the first gamble.

Thus the person will prefer the first gamble which has lower variability to the second gamble which has a higher degree of variability of outcome. It should be remembered that risk in this connection is measured by the degree of variability of outcome. In the first gamble, the degree of variability of outcome is less and therefore the risk is less and in the second gamble, the degree of variability is greater which makes it more risky. And in case of income with certainty there is no variability of outcome and therefore involves no risk at all. A risk-averse person therefore prefers the income with certainty to any gamble with the same expected money value as the income with certainty.

20.6 RISK PREFERENCE AND GAMBLING

In sharp contrast to risk-averse individual, the risk lover or risk seeker will play a gamble. As stated above, a risk lover prefers uncertain outcome having the same expected value of income to the equivalent income with certainty. In case of risk preferring or loving individual, Marginal Utility of money increases as his income increases. In Fig. 20.5, the N-M utility curve of a risk lover has been drawn that is convex. The convexity of the utility curve implies that marginal utility of money income increases as his money income increases. Now suppose that present income of the individual is 25 thousand rupees and is offered a gamble with 50- 50 chance of winning or losing Rs. 15 thousand rupees. Therefore, if he wins this money income will rise to Rs. 40,000 and if he loses, his money income will fall to Rs. 10 thousand rupee. The expected money value of the gamble is given by.

$$E(V) = 0.5 \times 40{,}000 + 0.5 \times 10{,}000$$
$$= 20{,}000 + 5{,}000$$
$$= 25{,}000$$

In order to find out the expected utility of 25 thousand rupees we draw a chord AB which connects the utility M_1A at income of Rs. 10 thousand and utility M_3B at income level of Rs 40 thousand corresponding to the expected value of income of Rs. 25 thousand, the expected utility is M_2D where point D lies on the chord AB corresponding to income of Rs. 25 thousand. It will be seen from Fig. 20.5 that expected utility M_2D of expected income of Rs. 25 thousand is greater than the utility M_2C of certain income of Rs. 25 thousand.

Since his expected utility from gamble is greater than that from his income with certainty, he will accept the gamble. Further, it can be known that risk lover, a person whose marginal utility of money increases with increase in his income will prefer a gamble having greater risk, that is, the gamble which has a larger variability of outcome to a gamble with less risk (that is, with smaller variability of outcome).

An application : Farmer's Gambling Against nature

We will now provide an interesting and useful application of choice under risk and uncertainty to the case of a farmer who in the absence of irrigation has to face an uncertain situation regarding outcome of his use of fertilizers which yields high output only in the presence of water, the availability of which, in the presence of irrigation, depends on the rain. If the farmer fertilizes his land and nature is kind and there is good and timely rainfall, he will reap a bumper crop. On the country, if it does not rain, his output (and therefore his income) will be low because the use of fertilizers in the absence of water is likely to burn his seed and damage his soil. If the farmer does not use fertilizers, he will have a moderate but a certain income. Thus the farmer for deciding whether or not to fertilize the land is playing a game against nature.

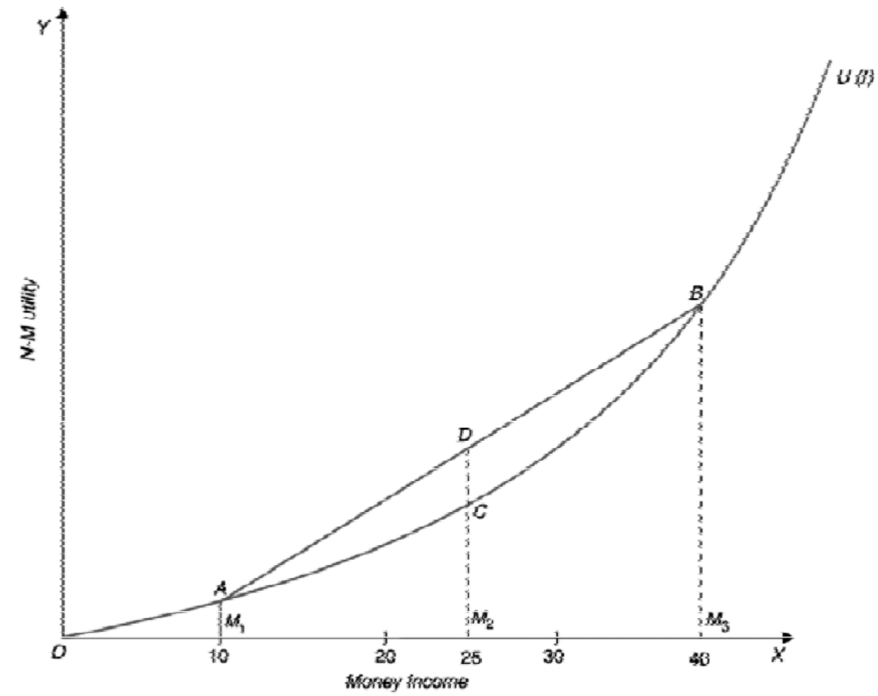


Fig. 20.5 : N-M Utility curve of a risk lover and his preference for gambling

Let us illustrate it with a numerical example. Suppose if our farmer uses fertilizer his income rises to Rs. 5000 but only if there is a good rainfall. And if it does not rain, his output or income will be small because in the absence of water, the use of fertilizers may destroy much of his seed and also cause damage to his soil. Let his output or income in this case comes about to be Rs. 2500. Suppose in the case of non-use of fertilizers, he has a moderate output worth Rs. 3750 regardless of whether it rains or not. What choice a farmer will make under this uncertain situation depends on his attitude towards risk. To know this we have constructed a pay off matrix of the farmer which is given in Table 20.1. It is further assumed that the farmer expects that there is 50:50 chance of rainfall.

Table 20.1 : Farmer's payoff matrix of fertilizing his land

Choice	Outcomes	
	Rain	No Rain
Using fertilizers	Rs. 5000	Rs. 2500
Not using fertilizer	Rs. 3750	Rs. 3750
Probability of a rainfall	Rs. 1/2 or 0.5	Rs. 1/2 or 0.5

The choice of not using fertilizers yields a certain income of .Rs. 3750. The choice by the farmer in favour of using fertilizers yields expected value equal to 1/2 (5000) + 1/2 (2500) = Rs. 3750. Thus the choice facing the farmer is between a certain income of Rs. 3750 and a gamble of using fertilizers with an expected value of Rs. 3750. This is quite a fair game. But whether or not the farmer will choose to fertilize his land depends upon his attitude towards risk. His choice is between a certain income of Rs. 3750 from not using fertilizers and a gamble of using the fertilizers the expected money value of which also equals Rs. 3750. If the farmer is risk averter, he will choose not to fertilize the land, that is, he will not gamble. On the other hand, if a farmer is risk lover he will fertilize the land that is, he will gamble. Of course, if he is risk neutral, he will be indifferent between the two choices.

20.7 RISK AVERSION AND INSURANCE

Our foregoing analysis shows why most people buy insurance when they are faced with a risky and uncertain situation. As mentioned above, most people are risk averters and therefore they buy insurance to avoid risk. Now, an important question is how much money or premium a risk-averse individual will pay, to the insurance company to avoid risk and uncertainty facing him. Suppose the individual buys a house which yields him income of Rs. 30 thousands per month. But if the house catches fire and due to the damage caused, his income from it falls to Rs. 10 thousands Per month and thus he suffers a loss of income. For the sake of simplifying analysis suppose there is 50 percent chance of the house catching fire. Then the expected value of income in this risky and uncertain situation is

$$\begin{aligned} E(X) &= 0.5 \text{ X } 30{,}000 + 0.5 \text{ X } 10{,}000 \\ &= 15{,}000 + 5{,}000 \\ &= 20{,}000 \end{aligned}$$

It is important to note that expected income of Rs. 20,000 is the weighted average of the two uncertain alternatives (30 thousands and 10 thousands) using their probabilities as weights. Different probabilities of the occurring of these incomes (30 and 10 thousands) would yield different expected income. Further note that the expected income is not the actual income that a person would get; it is weighted average of the two uncertain outcomes.

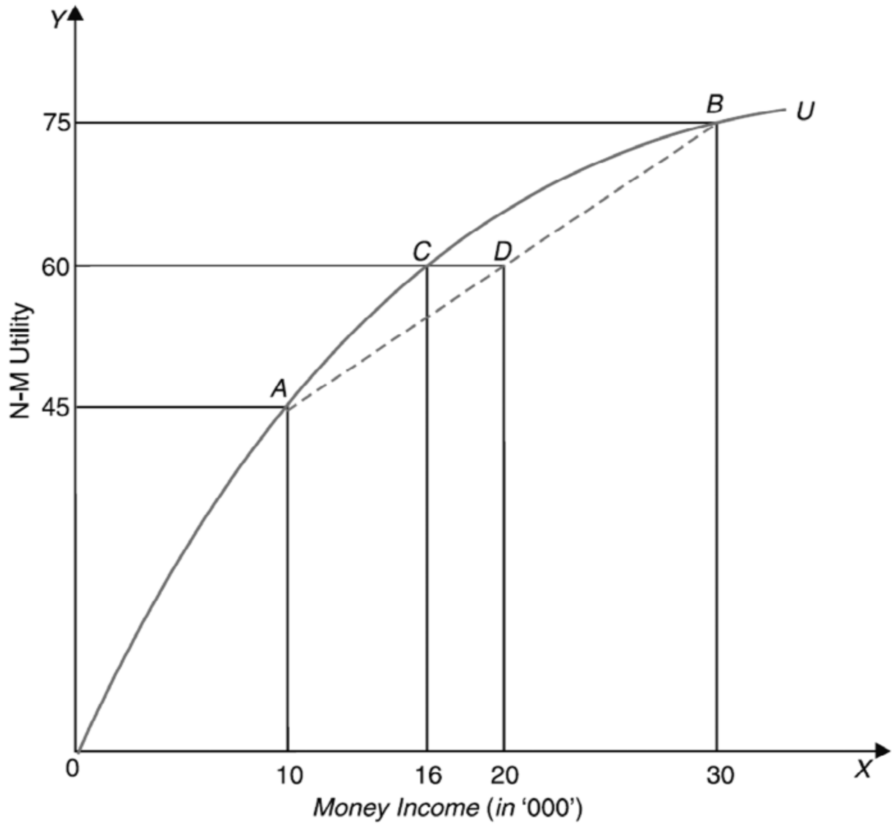


Fig. 20.6. Insurance and risk premium

The utility function OU with a diminishing marginal utility of money income of a risk- averse individual is shown in Fig. 20.6. With money income of Rs. 30 thousands, his utility is 75 and with his lower income of 10 thousands his utility is 45. Given that there is probability of 0.5 for each outcome, expected utility :

E(U) = 0.5 X U(30,000) + 0.5 X U(10,000)

= 0.5 X 75 + 0.5 X 45

= 37.5 + 22.5

= 60

It will be seen from Fig. 20.6 that we have drawn a straight line AB joining the utilities of 75 and 45. It is on this straight line or chord AB that the amount of expected utility will be corresponding to the expected value of income in the present risky and uncertain situation. It will be seen from Fig.20.6 that on this straight line AB and corresponding to the expected Value of income of Rs. 20 thousands, the expected utility is 60 which corresponds to point D on the straight line AB. But it will be seen from the individual's utility function OU, that utility of 60 is equal to that of an assured and certain income of Rs. 16 thousands. Thus the individual with an expected uncertain income of Rs. 20 thousands will be willing to forego Rs. 4 thousands (or DC to get a certain or guaranteed income of Rs. 16 thousands as the expected utility of uncertain expected income of Rs. 20 thousands is equal to the utility of a certain income of Rs. 16 thousands. This means that if the individual gives up Rs. 4 thousands (20-16 = 4) from his uncertain expected income he will get the same utility of 60 as with a certain income of Rs. 16 thousands. Rs. 4 thousands equal to distance DC is called the risk premium. Therefore, the risk premium is the amount of money that a risk-averse individual will be willing to pay to avoid the risk. By paying the risk premium the individual can insure himself against a large loss from a fire and to get an assured or certain income.

It is clear from above why people buy insurance for fire, accident, ill health and even life.

CHECK YOUR PROGRESS I

Answer the questions in the space provided:

Q1. What inconsistency in individual choices is illustrated by the Allais Paradox, and how does it challenge the expected utility theory?

____

____

____

____

Q2. Who is considered a risk-averse individual, and how does diminishing marginal utility explain his behavior toward risky choices?

____

____

20.8 LET US SUM UP

In this lesson, we discussed:

- According to M. Allais, under situations of uncertainty, people tend to make choices inconsistently.
- Psychologists Daniel Kahneman and Amos Tversky explain this phenomenon by arguing that *"a reduction in the probability of an outcome by a constant factor has a larger impact when the outcome was initially certain than when it was merely probable."*
- People make choices among alternative investments involving different degrees of risk.
- While making choices among commodity bundles when there is no risk or uncertainty, the consumer maximizes his utility.
- People's preferences toward risk greatly differ. Most individuals generally prefer less risky situations (i.e., situations with less variability in outcomes or rewards). In other words, most individuals seek to minimize risk and are called **risk averters** or **risk averse**.
- However, some individuals prefer risk and are therefore called **risk seekers** or **risk lovers**.
- Some individuals are indifferent toward risk and are called **risk neutral**.
- It is important to note that these different preferences toward risk depend on whether, for an individual, the marginal utility of money diminishes, increases, or remains constant.
- A **fair gamble** is one in which the expected value of income from the gamble is equal to the same amount of income received with certainty.
- Some individuals indulge in gambling; if their expected utility from the

gamble is greater than that from their income with certainty, they will accept the gamble.

- A **risk lover** is a person whose marginal utility of money increases with an increase in income. Such a person will prefer a gamble with greater risk—that is, a gamble with a larger variability of outcome—to a gamble with less risk (smaller variability of outcome).

- Most people buyinsurance when faced with riskyand uncertain situations. They are **risk averters** and therefore buy insurance to avoid risk.

20.9 KEYWORDS

1. **Risk Averse :** A person who prefers certainty over a gamble with a higher or equal expected value; characterized by a concave utility function.

2. **Risk Lover (Risk Seeker) :** An individual who prefers risk and chooses a gamble with higher variability even if the expected value is the same; represented by a convex utility function.

3. **Risk Neutral :** A person who is indifferent between a certain income and a gamble with the same expected value; has a linear utility function.

4. **Expected Utility :** The weighted average of utilities across all possible outcomes of a risky choice, where weights are the probabilities of each outcome.

5. **Fair Gamble**: **:** A gamble in which the expected monetary value is equal to the value of a certain outcome.

20.10 EXAMINATION-ORIENTED QUESTIONS

1. Discuss the Allais paradox.
2. Explain the attitude of a risk-averter, risk lover, and risk-neutral individual.
3. Why do people buy insurance? Explain.
4. What is meant when it is said that a person is risk averse? Why are most people risk averse, while some are risk lovers?

5. Why do some people indulge in gambling, even though it may bring heavy losses to them? Illustrate graphically.

6. Explain graphically the case of insurance.\

7. Explain, through a utility curve, the different attitudes toward risk.

20.11 HINTS TO CHECK YOUR PROGRESS

Q1 : See Section 20.3

Q2 : See Section 20.4

20.12 SUGGESTED READINGS

- Ahuja, H.L., *Advanced Economic Theory: Microeconomic Analysis*, S. Chand & Company Ltd.
- Misra, S.K. and V.K. Puri, *Advanced Microeconomic Theory*, First Edition (2001).
- Henderson, J. and Quandt, R., *Microeconomic Theory: A Mathematical Approach*, McGraw-Hill.
- Hader, J., *Mathematical Theory of Economic Behaviour*, Addison-Wesley.
- Koutsoyiannis, A., *Modern Microeconomics*, Macmillan.
- Mas-Colell, A., Whinston, M.D. & Green, J.R., *Microeconomic Theory*, Oxford University Press.
- Salvatore, D., *Microeconomics*, Oxford University Press.
- Sen, A., *Microeconomics*, Oxford University Press.
- Varian, H., *Intermediate Microeconomics*, East West Press.

ENDNOTES

- Source: Tversky, A. and Kahneman, D., *"The Framing of Decisions and the Psychology of Choice"*, *Science*, Vol. 211.
- Sen, Anindya, *Microeconomics*, Fourth Impression.

M.A. ECONOMICS	LESSON NO. 21
COURSE NO. ECO 201	UNIT-IV

THE FRIEDMAN-SAVAGE HYPOTHESIS AND MARKOWITZ HYPOTHESIS

STRUCTURE

21.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

- Discuss the **Friedman-Savage Hypothesis**.
- Describe the **Markowitz Hypothesis**.

21.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to:

- Discuss **Friedman-Savage Hypothesis** and its implications.
- Describe **Markowitz Hypothesis** with supporting illustrations.
- Interpret individual choices involving **risk and uncertainty** through both hypotheses.
- Compare and contrast the assumptions and conclusions of both models.
- Apply these concepts in analyzing **risk-related economic behavior**.

21.2 INTRODUCTION

Dear learner, Milton Friedman and L.J. Savage in their well-known article put forward a hypothesis that explains why the same persons buy insurance and also engage in gambling. In buying insurance they seek to avoid risk and in engaging in gambling they take risk. In other words, people behave both as risk averters (i.e., when they buy insurance), and also as risk lovers when they gamble by buying lottery tickets or bet at horse races. This seemingly contradictory behaviour on the part of the people could not be explained with Bernoullian hypothesis of diminishing marginal utility of money. On the other hand Markowitz has criticized Friedman-Savage hypothesis for its assertion that very poor and very rich will be unwilling to take risk and therefore will not indulge in gambling etc. and it is only the persons in the middle income range that will take risks. According to his hypothesis, small increases in money income yield rising marginal utility of money, but large increases in income bring about diminishing marginal utility of money, the individual will be willing to accept small fair bets but will be reluctant to take risks involving large amounts of money. In this lesson, we will discuss both the hypotheses.

21.3. THE FRIEDMAN SAVAGE HYPOTHESIS

We studied above that, according to Bernoulli, Marginal Utility of money diminishes for most of the people and therefore they are unwilling to make fair bets. But if the Marginal Utility of money always diminishes for all the people, then the widespread practice of gambling found among the people all over the world cannot be explained . Moreover, why some people actually prefer to make choices under very risky situations such as in a gamble or a race? Are these people irrational and thoughtless? Do they gamble for the sake of fun and pleasure? Watson provides an answer to this. According to him, "To point to the entertainment and pleasure that many people find in gambling is not enough, nor does it suffice to dismiss gambling and other decisions under risk as "irrational". Though the world abounds with people who are thoughtless and scatter-brained in their decisions, much gambling is done with cold and careful calculations. Remember, too, that gambling has flourished for centuries and in many cultures, whatever its morals and legality might be, gambling is not an aberration in the behaviour of a part of the population."[1]

Neumann-Morgenstern provided the method of measuring numerically the Marginal Utility of money. Based on Neumann-Morgenstern cardinal utility analysis, Milton Friedman and L. J. Savage in their well-known article put forward a hypothesis that explains why the same persons buy insurance and also engage in gambling. In buying insurance, they seek to avoid risk, and in engaging in gambling, they take risk. In other words, people behave both as risk averters (i.e., when they buy insurance), and also as risk lovers when they gamble by buying lottery tickets or bet at horse races. This seemingly contradictory behaviour on the part of the people could not be explained with Bernoullian hypothesis of diminishing marginal utility of money.

Friedman and Savage abandoned Bernoullian hypothesis of diminishing marginal utility of money for all ranges of income and instead adopted another hypothesis. According to Friedman Savage hypothesis, for most people, Marginal Utility of money income diminishes up to a certain level of money income, it increases from that middle level to a certain higher level of money income and thereafter at very high levels of income it again diminishes. With this hypothesis and using Neumann- Morgenstern Utility Curve, Friedman and Savage[2] explain both types of behaviour of buying insurance to avoid risk and of indulging in gambling and thereby to take risks. Friedman-Savage hypothesis is depicted in Figure 21.1. where the Neumann-Morgenstern Utility Curve having both the concave and convex portions has been drawn. It will be observed from this figure that the section of this N-M utility curve U(I) up to point K is concave and in the middle portion from E to F is convex, and beyond point F, that is, at higher levels of income, it again becomes concave. This implies that at the lower levels of income marginal utility of income of the person diminishes; and between middle incomes corresponding to point E and F, marginal utility of income is increasing and beyond B, Marginal Utility of income again diminishes with further increases in income.

Let us now explain how Friedman and Savage explain the behaviour of the same person as risk averter (as when he buys insurance) and risk lover (when he buys a lottery ticket). Suppose the person's present income is 35 thousand rupees which lies at the concave part of the N-M utility curve corresponding to point C where marginal utility of income is diminishing. Suppose the individual feels that he has a 50:50 chance of losing the sum of money income equal to 30. This means that there is probability of 0.5 or ½ that he may have income of 35 and probability of 0.5 that his income may fall to 5 thousand rupees. As explained in the preceding section, the person with diminishing marginal utility of money will try to avoid this uncertain prospect (which is a kind of gamble) and will therefore buy insurance to eliminate risk and thereby have an income with certainty giving the insurance premium.

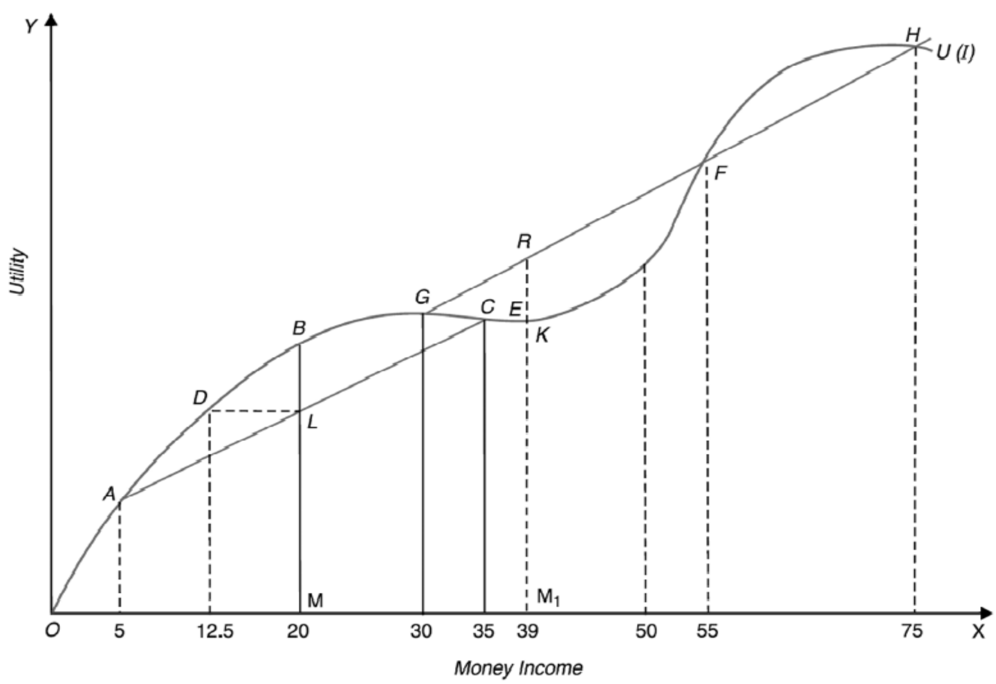


Fig. 21.1 Friedman-savage hypothesis

Note that his expected value of income in this case is

$$E(X) = 0.5 \times 35 + 0.5\ (5)$$

$$= 17.5 + 2.5 = 20 \text{ thousands}$$

As can be seen from Figure 21.1, that expected utility from Rs. 20 thousand is ML lying at chord AC which is less than utility MB of income of Rs. 20 thousand with certainty. To avoid risk, he can pay insurance premium LD or 7.5 thousand to have certain income of Rs. 12.5 thousands.

Now, assume that the same person with the present income of Rs. 35 thousands is considering to buy a lottery ticket that offers him chance of winning a large sum of money, say 40 thousand rupees, the first prize of the lottery and thereby raising his income to Rs. 75 thousand.

lf he does not win, his income will fall to Rs. 30 thousands; 5 thousands being the price of the lottery ticket. Now, if there is 20 per cent chance of winning the lottery; Its expected value will also be equal to 0.2 x 75 + 0.8 (30) = 15 + 24 = 39 and as it will be seen from Figure 21.1 that the expected utility from the sum of Rs. 39 thousands is M_1R (note that point R lies on the straight line segment GF which is greater than utility M_1K which he gets from income of Rs. 39 thousands with certainty. Thus, the person would purchase the lottery ticket, that is, he will gamble.

Now, suppose the individual's income is Rs. 50 thousands which lies in the middle income segment where the marginal utility of money income is increasing. With Rs. 50 thousands income the individual will be willing to buy lottery tickets, indulge in gambling or undertake risky investment since the gain in utility from extra money will be much greater (marginal utility of money income is rising) than the loss of utility from the small payment for a lottery ticket or from equal monetary loss in a gamble.

A person with an income beyond 55 or in the segment FH enjoys quite high income and therefore marginal utility of money to him is declining. As a result of this he would be unwilling to take risk either in a gamble or in undertaking risky investment except at very favourable odds.

Friedman-Savage think that the N-M utility curve of money indicates the behaviour or attitude of people towards risk in different socio-economic groups. They of course admit that there are many differences between the persons within a same socio-economic group: some have great preference for gambling and others are unwilling to take any risk at all. Even then Friedman and Savage think the curve describes the propensities of broad classes. The middle income group with increasing marginal utility of money are those, they argue, who are eager to take risks to improve their economic conditions. The expectation of more money means much to this group of persons: if their efforts succeed, they will lift themselves up into the next upper socio-economic class. These persons want not just more consumer goods; they have ambition to look up in the social scale. They want to rise, to change their life styles. No wonder that marginal utility of money increases for them.

21.4. THE MARKOWITZ HYPOTHESIS

Both the Bernoullian hypothesis and Friedman-Savage hypothesis assume that marginal utility of money depends on the absolute level of income. When Neumann- Morgenstern utility index is constructed for an individual for various levels of income, his choice regarding risky situations can be determined. Professor Markowitz has criticized Friedman-Savage hypothesis for its assertion that very poor and very rich will be unwilling to take risk and therefore will not indulge in gambling etc. and it is only the persons in the middle income range that will take risks. According to Prof. Markowitz, Marginal Utility of money income relates only to the change in the level of money income rather than to the absolute level of income. With this, Markowitz has tried to explain why most people buy insurance and also indulge in gambling. Markowitz takes the deviations from the present income of the individual as determining the attitude of the people towards risk. According to Markowitz hypothesis, with small increases in income from the present level, the marginal utility of money income increase but large increase in income yield diminishing marginal utility of money. On the other hand, with small decreases in income yield from the present level marginal utility of money increase, but large decrease in income yield diminishing marginal utility of money. Markowitz hypothesis is illustrated through the curve of MU of money in Fig. 21.2 which has been derived from the N—M utility curve.

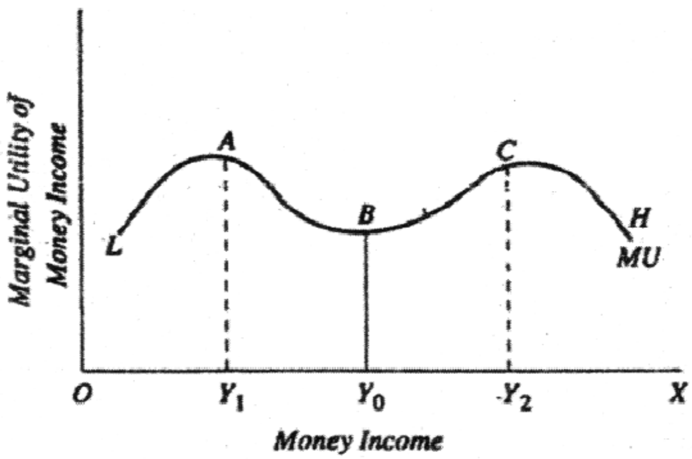


Fig. 21.2 The Markowitz hypothesis

In Markowitz hypothesis, this MU curve has three inflexion points, A, B and C and four segment BC, CH, BA, and AL as against two inflexion points and three segments in MU curve of Friedman-Savage hypothesis. Y_0 in figure 21.2 is the present level of income. As the person's income increase from Y_0 to Y_2, marginal utility of money increases along BC. However, increase in income beyond Y_2 yield diminishing marginal utility of money. On the other hand, small decreases in money income from Y_0 to Y_1, yield increase in marginal utility of money but large decreases in income, that is, to the left of Y_1 bring about diminishing marginal utility of money.

Since according to Markowitz hypothesis, small increases in money income yield rising marginal utility of money, but large increases in income bring about diminishing marginal utility of money, the individual will be willing to accept small fair bets but will be reluctant to take risks involving large amounts of money. Likewise, since small decrease in income yield rising marginal utility of money and large ones diminishing MU, the individual will be eager to insure against small losses but will adopt a "devil-may-care-attitude" to the risky situation involving very large losses.

CHECK YOUR PROGRESS I

Answer the questions in the space provided:

Q1. Explain the Friedman-Savage Hypothesis and discuss how it accounts for the behavior of individuals who both buy insurance and indulge in gambling.

Q2. Explain the Markowitz Hypothesis and discuss how it differs from the Friedman-Savage Hypothesis in explaining individuals' attitudes towards risk.

21.5 LET US SUM UP

Dear learner, in this lesson, we discussed:

Milton Friedman and L.J. Savage in their well-known article put forward a hypothesis that explains why the same persons buy insurance and also engage in gambling.

In buying insurance they seek to avoid risk and in engaging in gambling they take risk. In other words, people behave both as risk averters (i.e., when they buy insurance), and also as risk lovers when they gamble by buying lottery tickets or bet at horse races.

This seemingly contradictory behaviour on the part of the people could not be explained with Bernoullian hypothesis of diminishing marginal utility of money.

According to Friedman-Savage hypothesis, for most people marginal utility of money income diminishes up to a certain level of money income, it increases from that middle level to a certain higher level of money income and thereafter at very high levels of income it again diminishes.

Both the Bernoullian hypothesis and Friedman-Savage hypothesis assume that marginal utility of money depends on the absolute level of income.

Markowitz has criticized Friedman-Savage hypothesis for its assertion that very poor and very rich will be unwilling to take risk and therefore will not indulge in gambling etc. and it is only the persons in the middle income range that will take risks.

According to Markowitz hypothesis, small increases in money income yield rising marginal utility of money, but large increases in income bring about diminishing marginal utility of money, the individual will be willing to accept small fair bets but will be reluctant to take risks involving large amounts of money.

Likewise, small decrease in income yield rising marginal utility of money and large ones diminishing MU, the individual will be eager to insure against small losses but will adopt a "devil-may-care-attitude" to the risky situation involving very large losses.

21.6 KEYWORDS

- **Marginal Utility (MU) :** The additional satisfaction or utility a person receives from consuming an additional unit of a good or service.
- **Risk Averter :** An individual who prefers to avoid risk and tends to buy insurance to protect against uncertain losses.
- **Risk Lover :** An individual who enjoys taking risks and may engage in gambling or speculative activities.
- **Bernoullian Hypothesis :** The hypothesis that individuals make decisions under uncertainty by maximizing expected utility, where marginal utility of money diminishes as income increases.
- **Friedman-Savage Hypothesis :** A theory explaining that marginal utility of money income decreases up to a certain level, then increases for a range of income, and finally decreases again at very high income levels, accounting for both risk-averse and risk-loving behavior in the same individual.

- **Markowitz Hypothesis :** A critique of Friedman-Savage hypothesis stating that small income changes produce rising marginal utility while large income changes cause diminishing marginal utility, explaining willingness to accept small bets but reluctance for large ones.

21.7 EXAMINATION ORIENTED QUESTIONS

- Discuss in detail the Friedman-Savage Hypothesis.
- State and explain the Markowitz Hypothesis.
- Explain on what grounds has Markowitz criticised Friedman-Savage hypothesis.
- How does Friedman-Savage hypothesis justify the same individuals taking and avoiding risk at the same time?

21.8 HINTS TO CHECK YOUR PROGRESS

Q1 see section 21.3

Q2 see section 21.4

21.9 SUGGESTED READINGS

Koutsoyannis, A : *Modern Microeconomics*, Macmillan

Mas-Colell, A., M.D. Whinston & J.R. Greene : *Microeconomics Theory*, OUP

Salvatore, D : *Microeconomics*, OUP

Sen, A : *Microeconomics*, OUP

Varian, H : *Intermediate Microeconomics*, East West Press

Ahuja, H.L. — *Advanced Economic Theory, Microeconomic Analysis*, S. Chand & Company Ltd.

Handerson J and R Quandt — *Microeconomics Theory, A Mathematical Approach*, McGraw-Hill

Hader, J. — *Mathematical Theory of Economic Behaviour*, Addison-Wesley

Endnotes

D.S. Watson, *Price Theory and its uses*, p.128

M. Friedman and L.J. Savage, "The utility Analysis of choice involving risk", *Journal of Political Economy*, Aug, 1948, pp. 279-304

M.A. ECONOMICS	LESSON NO. 22
COURSE NO. ECO 201	UNIT-IV

INVESTOR CHOICE PROBLEM AND DEALING WITH ASYMMETRICAL INFORMATION

STRUCTURE

22.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

- Discuss the investor choice problem.
- Describe the asymmetrical information in markets.

22.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to:

- Explain the choice of an investment portfolio.
- Discuss the concept of asymmetrical information.
- Describe how information problems affect markets.

22.2 INTRODUCTION

Dear learner, at the time of taking investment decision, the investor knows a set of possible outcomes in terms of returns from various assets and the probability of each but is unaware of what particular outcome will occur. Asymmetric information occurs when the workers who sell their labour services know their ability and efficiency, while the firms who hire them are not well informed about it.

22.3. CHOICE OF AN INVESTMENT PORTFOLIO: THE INVESTOR'S PROBLEM

The theory of choice under risks and uncertainty is also applicable in case of an investor who has to invest his saving in various types of assets having degrees of risks to get optimum return from them. For instance, if an investor does not want to bear risk shares from the stock market whose value and dividend can vary a good deal. From these shares he can get much higher return if the stock market goes well or his return may be verylow if the stock market is gripped by depression. Obviously he faces a choice problem of combining the assets with assured fixed returns such as fixed deposits in banks, debentures of reputed companies with some to arrive at an optimum portfolio of investment.

It may be noted that at the time of taking investment decision, the investor knows a set of possible outcomes in terms of returns from various assets and the probability of each but is unaware of what particular outcome will occur. It may be noted that a risky asset must have a higher expected return than the risk free assets because a risk averse investor will always invest in a risk free asset if expected return from a risky asset does not exceed the return from the former. In economic theory the degree of risk involved in an asset is measured by the variance or standard deviation of the variability of the return from an asset. Similarly, if a person has invested his money in a number of assets, the standard deviation of the variability of return from this investment portfolio as a whole will indicate the degree of risk involved.

Let us take an example of a person who invests his savings in two assets. In order to know how an investor will allocate his money in various types of assets:

Fixed Deposits and (2) Equity shares of a company. Suppose an investor invests a fraction of his savings in equity shares of a company, then the fraction of it will be invested in a risk-free investment. Let the expected return from investment in shares be denoted by R_m and the return from a risk-free in investment by R_f. If the investor buys shares, then $R_m > R_f$ otherwise he would only invest in fixed deposits and would not buy risky shares. The expected return from his total portfolio Rp, is the weighted average of the expected returns on the two types of assets. Thus,

$$R_p = bR_m + (1—b)R_f \qquad \text{……………………. (i)}$$

Where

R_p = the average return from total investment portfolio

R_m = the return from the shares which represent risky investment

R_f = the return from a risk-free investment

b = the fraction of investment made in shares.

The above equation (i) measuring the expected return from an investment portfolio can be rewritten as

$$R_p = bR_m + R_f - R_f b$$

$$R_p = R_f + b(R_m—R_f) \qquad \text{……………………… (ii)}$$

Measuring Risk of investment in an Asset: As regards the riskiness of an investment portfolio, it can be shown by a little mathematical manipulation that the standard deviation of the portfolio (with one risky asset and the other risk free asset) as a whole is given by the proportion of investment made in risky asset multiplied by the standard deviation of the return from that asset. Thus,

$$\sigma p = b \ \sigma m$$

or $$\frac{\sigma p}{\sigma m} = b \qquad \text{.............................(iii)}$$

where

σp = standard deviation of the total investment portfolio

σm = standard deviation of the risky asset b

= the fraction of investment in risky asset.

Budget Line for the Investor : Like a consumer who has to spend his income on good for his consumption, investor also faces a budget line which describes the trade off between risk and return. In order to know what fraction of his investment the person will choose to make in risky asset, that is, 'b', the budget line measuring the trade off between risk and return of the various investment portfolios has to be drawn.

We know from equation (iii) that b = $\frac{\sigma p}{\sigma m}$. Now, putting this value of b in equation (ii) given the expected rate of return from the portfolio we have

$$R_p = R_f + \frac{\sigma p}{\sigma m}(R_m - R_f)$$

Rearranging, we have

$$R_p = R_f + \frac{(Rm - Rf)\ \sigma p}{\sigma m} \qquad \text{..................... (iv)}$$

The above equation (iii) is the budget equation of the investor measuring trade off between risk and rate of return and it is the equation of a straight line because R_m, R_f σ_m, and the slope $\frac{(Rm - Rf)}{\sigma m}$ are constants. The slope of the budget line $\frac{(Rm - Rf)}{\sigma m}$ is the price of risk because it shows how much extra risk an investor must bear if he has to have expected return. Figure 22.1, presents this budget line which slopes upward starting from the point Rf which represents a rate of return on a risk-free asset. The upward-sloping budget line implies that if the investor wants a higher expected return (Rp) investment portfolio, the standard deviation of the return (i.e., the degree of risk) increases. Aglance at the budget line will reveal that if he invests all his savings in a risk- free asset, namely Fixed deposits in a bank, he will have a fixed Rf return. On the other hand, if the investor devotes all his saving to the risky asset of shares, he earns an expected return Rm but have to bear more risk measured by standard deviation ∫m. But ordinarily an investor will opt for a mixed portfolio devoting a fraction of his money in each of the two assets.

Optimum Investment Portfolio : Choosing Between Risk and Return: Which of the investment portfolio, the investor will choose depends on his attitude towards risk. In Figure 22.1 we have drawn three indifference curves reflecting attitude towards risk of our investor. As we saw in a previous chapter on indifference curve, riskiness is ‘bad’ or undesirable and therefore more of it yields less satisfaction and therefore as we move rightward indicating greater risk or standard deviation of the return, the investor should receive higher expected return to keep him equally satisfied. Therefore, indifference curves between riskiness and rate of return slope upward to the right.

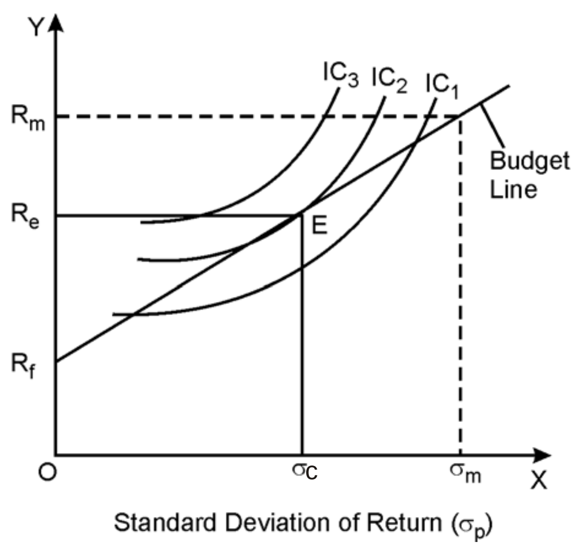


Fig. 22.1. Choice of an Investment Portfolio

It will be seen from Figure 22.1 that the budget line RfL, the slope of which $\frac{(Rm - Rf)}{\sigma m}$ depicts the price of risk and reflects the actual market situation regarding earnings or return from various types of asserts, is tangent to the indifference curve IC2 (which reflect his attitude toward risk) at point E which is the maximum attained satisfaction-maximizing investment portfolio. According to this optimum investment portfolio point E, the investor has OR_e as the return from his investment portfolio and bears the risk equal to σc.

It may be noted that how much additional risk an investor will be prepared to bear to earn a higher expected return depends on how risk averse he is. The difference in attitude towards risk will make their indifference curves between risk and return of different slopes and concavity. Figure 22.2, depicts the choices of two different investors A and B. From their indifference curves, it will be observed that investor A is more risk averse and in equilibrium chooses an investment with portfolio with lower return R_a and lower risk σ_a, whereas investor B chooses an investment with higher return R_B and greater risk σ_b.

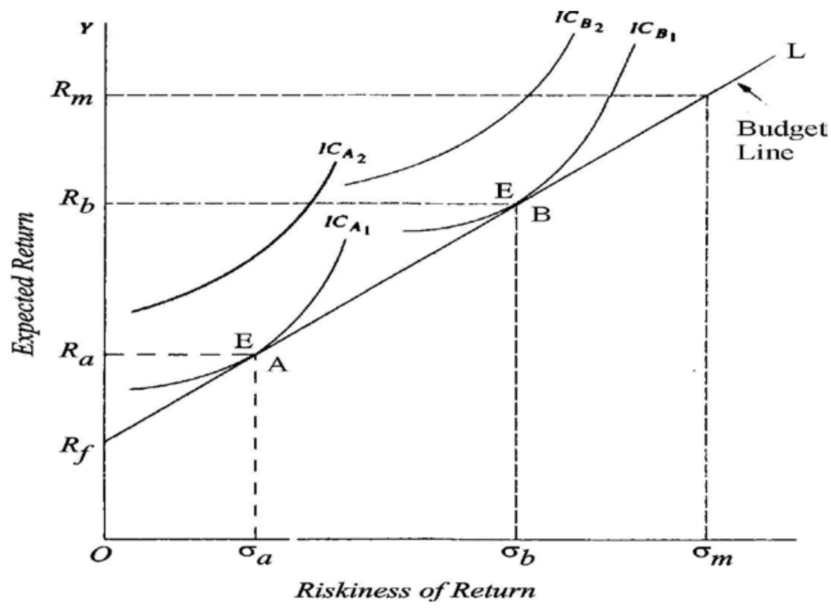


Fig. 22.2 : Choice of Two different investors

22.4 INFORMATION PROBLEM AND MARKETS WITH ASYMMETRIC INFORMATION

In the perfect competition model it is assumed that market participants, namely, the consumers of a product and its producers (i.e., firms) had perfect information about the price, availability and quality of the products. Thus under it, the consumers know about what products are available in the market and at what prices and where they are available. Besides, they know fully well the quality of the products available, that is, their various attributes and characteristics. In other words, they have full knowledge of the opportunity set available to them. The consumers also know their preferences of various goods and can tell at what rate they are willing to exchange one good for another, for example, wheat for cloth or oranges for apples. Further, students who consume education service and paying for it know fully well how much they will benefit from it.

In the perfect competition, the firms, the producers of goods, are fully informed about the most efficient technique available for producing goods, the productivity of workers hired by them. They know the efficiency with which workers are actually doing their jobs. Besides, firms are perfectly informed about the quality of other inputs and the prices at which they are available from various sources. Finally, the firms know about the prices at which they can sell not only today but in the future too. Indeed in perfect competition model they fully know the demand and supply curves of their products produced by them.

22.5 THE INFORMATION PROBLEM

However, in the real world, the households and the firms do not have full information about the price, quality and availability of the product. In fact, they have only imperfect information about opportunity set available to them. Perfect competition model often misleads us in the situations characterized by imperfect information. A leading example of imperfect information is asymmetric information about the quality of used goods, for example, used cars in the market. By asymmetric information we mean one party in the market for used cars does not know (i.e. has no perfect information) about the Quality of the product being sold. In case of used cars while the sellers know about the true quality of their products, the buyers do not know the quality of the used cars which may turn out to be lemons (that is, defective pieces). Similarly, asymmetric information also occurs while the workers who sell their labour services know their ability and efficiency, the firms who hire them are not well informed about it.

Economists such as Joseph E. Stiglitz of the University of California and Akerlof have incorporated imperfect information of the households and firms in their models of market equilibrium. They emphasize certain economic phenomena such as lemon problem, hiring of suitable persons for the jobs can be explained well only by incorporating imperfect information in economic models of functioning of markets. Stiglitz and Walsh write, "Prices and market provide the basic of the economy's incentive system. But there are some information problems that markets do not handle or do not handle well. And the imperfect information sometimes inhabits the ability of the market to perform the tasks they perform so well when information is complete.[1]

According to Stiglitz, for bringing Economics closer to the real world one should appreciate the problem of information faced by the firms (i.e. producers and employees) and the buyers of the products. In this context the example of insistence by employers about the possession of college graduate degree for certain jobs is worth mentioning. The possession of college graduate degree conveys information to the employers about their quality or productivity of the person applying for particular jobs. Thus Stiglitz and Walsh write, "College graduates may receive a higher income than high school graduates, not only because they have learned things in college that make them productive but also because their college degree conveys valuable information to employers. Employers cannot easily learn in an interview which applicants for a job will be productive. College graduates are, on average, more productive workers.... It may simply have enabled firms to sort more easily students who are productive from the less productive."[2]

Price system in a competitive economy provides an efficient solution to the information problem. Prices play a significant role in bringing about coordination in productive activities of the firms. It is the prices of products that convey information to the firms about the economic scarcity of the products. The high prices of certain products convey information to the firms what people prefer and adjust their productive activity accordingly. The relative prices of goods tell the firms at what rate the consumers are willing to exchange various goods, that is, their trade-offs. They are guided by the prices in deciding about what goods are to be produced and in what quantities. Thus prices serve as a signal to the firms of marginal benefits of producing extra units of various goods. And this is all that the firms need to know about for deciding about their productive activity.

Likewise, prices of inputs convey information to the firms about the extent of their scarcity. On the basis of market prices of inputs they decide what inputs and in what quantities to be used for producing various commodities. They try to economise on the relatively more scarce inputs and use relatively more cheap inputs. But, as pointed out by Stiglitz, there are some information problems that markets do not handle well. And imperfect information sometimes inhibits the ability of markets to perform the task they perform well when information is complete.

Therefore, when there is imperfect information, for getting correct information buyers and firms are willing to pay for it. Information can be considered as a good like any other good and like other goods it has a price. Investors spend a lot of money on newspapers, business magazines that provide them information about shares and bonds of companies, other investment opportunities such as mutual funds. By subscribing to them people pay a price for the information they provide. The growth of internet has reduced the costs of getting various types of valuable information which was earlier difficult to obtain.

However, despite the growth of information technology, market for information is far from perfect. Two factors are responsible for it. First, information is not exactly like any other good. While in case of other goods buyers can have a look at them and know its components and attributes. But, correct information is not easy to get when you buy, it. Advertisements in the newspapers generally give misleading information to the people. Some writers in business magazines give biased information about the stocks and mutual funds of various companies, the quality of various cars, etc. Of course, the sellers of information either through advertisements in newspapers, or television and electronic media appear to say

'Trust us we are giving you correct information'. However, the people and the firms are generally sceptical about the correctness of this information conveyed to them.

Even if a consumer or a firm buys certain information which it thinks will provide benefit and is worth paying for it, information it gets is far from perfect. Getting information is quite costly. Consumers as well as firms can be victims of inaccurate or imperfect information. A profit-maximizing firm hires workers as long as Marginal Revenue Product (MRP) of workers exceeds the wage rate. How can a firm judge the productivity of workers who apply for a job? Likewise, the consumers are often misled by advertisements informing them by the firms that their oil can blacken their hair naturally or can grow new hair and remove their baldness. Their claims turn out to be untrue.

It follows from above that imperfect information is a fact of life.

CHECK YOUR PROGRESS I

Answer the questions in the space provided:

Q1: What is the main objective of an investor while selecting an investment portfolio?

____

____

____

____

Q2: What do we mean by "asymmetric information" in financial markets?

____

____

____

____

Q3: How does lack of information affect the decision-making of buyers and sellers?

____

____

____

____

22.6 LET US SUM UP

Dear learner, in this lesson, we discussed:

- The theory of choice under risks and uncertainty as it applies to investors who allocate their savings across various types of assets with different levels of risk to optimize returns.
- While making investment decisions, investors are aware of possible outcomes and their associated probabilities, but they do not know which specific outcome will occur.
- In the perfect competition model, it is assumed that market participants—both consumers and producers—have perfect information about prices, availability, and product quality.
- According to Stiglitz, understanding the real-world economics requires acknowledging the information problems faced by firms (producers and employees) and buyers.
- The price system in a competitive economy helps solve the information problem by coordinating productive activities and conveying scarcity through prices.
- In markets with imperfect information, individuals and firms are willing to pay to acquire correct and useful information. Information is treated as an economic good with a price.

- Investors often spend significantly on newspapers, business magazines, and financial platforms to gain information about stocks, bonds, mutual funds, and other investment opportunities.

22.7 KEYWORDS

1. **Investment Portfolio** – A collection of financial assets such as stocks, bonds, mutual funds, and other investments held by an investor.
2. **Risk** – The possibility of loss or uncertainty in the return on an investment.
3. **Asymmetric Information** – A situation where one party in a transaction has more or better information than the other party.
4. **Perfect Competition** – A market structure in which all participants have equal and perfect knowledge about prices, products, and market conditions.
5. **Expected Return** – The weighted average of all possible returns from an investment, based on their probabilities.
6. **Information Cost** – The expenditure incurred to obtain accurate and useful information about products, investments, or market conditions.

22.8 EXAMINATION ORIENTED QUESTIONS

1. Explain the choice of an investment portfolio.
2. Discuss the information problem and markets with asymmetric information.

22.9 HINTS TO CHECK YOUR PROGRESS

- Q1: See Section 22.3
- Q2: See Section 22.4
- Q3: See Section 22.5

22.10 SUGGESTED READINGS

1. Ahuja, H.L. *Advanced Economic Theory, Microeconomic Analysis,* S. Chand & Company Ltd.
2. Handerson J and R Quandt: *Microeconomics Theory, A Mathematical Approach*, McGraw-Hill.
3. Hader, J. *Mathematical Theory of Economic Behaviour*, Addison-Wesley.
4. Koutsoyannis, A: *Modern Microeconomics*, Macmillan.
5. Mas-Colell A, M.D. Whinston & J.R. Greene: *Microeconomic Theory*, OUP.
6. Salvatore, D: *Microeconomics,* OUP.
7. Sen, A: *Microeconomics*, OUP.
8. Varian, H: *Intermediate Microeconomics,* East West Press.
9. Pindyck, Robert. S and Daniel l. Rubinfeld, Microeconomics (fifth edition).
10. Boyes, William and Michael Melvin, Textbook of Economics (sixth edition, Indian adaptation)
11. Sloman, John, Economics (Third edition).
12. Pindyck R.S., D. Rubinfeld and P.L. Mehta (2012), Microeconomics, Pearson Education, South Asia.
13. Salvatore, D. (2010), Microeconomics; Theory and Applications, Oxford University Press, New Delhi
14. Cooper, R., & John, A. (2011). Economics–theory through applications, Saylor Foundation. [3]
15. Shapiro, D., MacDonald, D., & Greenlaw, S. A. (2022). *Principles of Microeconomics 3e*.[4]

ENDNOTES

1. Joseph E. Stiglitz and Carl E. Walsh, *Economics*, 3rd Edition, 2002, Norton and Co.

2. Stiglitz and Walsh, op. cit. p.234

3. Type of license/Conditions of use: CC BY-NC-SA (Attribution-Non commercial-ShareAlike) .Scan QR code 1 to access **Cooper, R., & John, A. (2011). Economics–theory through applications.**

4. This book is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). "Access for free at openstax.org." **Scan QR code 2 to access the e book by Shapiro, D., MacDonald, D., & Greenlaw, S. A. (2022). *Principles of Microeconomics 3e*.**

QR code 1

QR code 2

THE MARKET FOR LEMONS, AND QUALITY CHOICE, ADVERSE SELECTION, MORAL HAZARD, SIGNALLING AND INCENTIVE SYSTEM

STRUCTURE

23.0 OBJECTIVES

The main objectives of this lesson are to enable you to:

- Define the concept of *the market for lemons* and how it illustrates problems of quality choice and adverse selection.
- Describe the idea of *moral hazard* and its implications for market behavior.
- Discuss the role of *signalling* as a solution to information asymmetry.
- Explain how *incentive systems* can align interests and improve outcomes in markets with imperfect information.

23.1 LEARNING OUTCOMES

After going through this lesson, you shall be able to:

- Explain how information asymmetry affects markets, especially in the context of quality uncertainty.
- Describe adverse selection using the example of the market for lemons.
- Discuss the concept of moral hazard and how it arises in various economic contexts.
- Explain how signalling and incentive structures can mitigate problems caused by imperfect information.

23.2 INTRODUCTION

Dear learner, so far in our study of markets we have assumed that sellers and buyers were perfectly and equally informed about the quality of goods being sold in the market. This assumption is valid if it is either quite easy to verify the quality of a goods or it is not costly to ascertain which goods are of high quality and which goods are of low quality. In view of the known difference in qualities of various goods, the different prices of the goods will then reflect the quality difference between them.

When it is not easy to know about the quality of goods or it is costly to get information about them and buyer and seller of goods are not equally informed about the quality of goods, then there is asymmetric information which means the market situation when the buyer and seller have a different information while making a transaction.

The various examples of asymmetric information can be given. An important example of an asymmetric information is of the market for used cars. In this case the sellers are better informed than the buyers about quality of the goods that are sold. The theory of lemons was first developed in connection with the sale of used cars. Some used cars are of bad or low quality (i.e. they are lemons) and others are of good quality. While the sellers know whether the cars they are selling are lemons or of good quality. The other examples of markets with asymmetric information are provision of insurance service and labour market. The customer is better informed than the insurance company about the probability of his getting ill. In case of labour market, different workers have different productivities and it is difficult for the employer to know the productivity of the workers and employees. The problem about the asymmetric information is that it leads to market failure, that is, failure to achieve Pareto efficiency. We shall discuss the markets with asymmetric information and the problems it gives rise to.

23.3 THE MARKET FOR LEMONS

The important example of market with asymmetric information is the market for lemons. In the market for lemons, the buyers and sellers have different information about the quality of goods being bought and sold. The word lemon is used to describe a defective or low quality product.[1] Let us consider the market for used cars. We assume that the used cars are of different quality; some of them are of good quality while others are just lemons. The bad quality cars (i.e. lemons) quite often break down and require a lot of repairs. However, while the sellers of used cars fully know the quality of their used cars, the buyers are not certain about their quality due to lack of information. Therefore, the market for used cars is an important example of asymmetric information. The price of the used car in the market depends on the average quality of the used car offered for sale. This is because the buyers being uninformed about the quality of the used car will not be willing to pay more than what an average quality used car is worth. This means that the owners of bad quality cars will get price for their used cars more than what they are worth. This has an important consequence. Since the owners of good quality used cars will not be able to get the price for their better quality cars since the price determined in the market for used cars will be equal only to what average quality cars are worth. As a result, the owner of goods quality used cars will withdraw their cars from the market. This will reduce the number of used cars available for sale in the market. With this the average quality of used cars offered for sale will also go down and so also the price of the used cars. This will cause some more car owners whose used cars all relatively of goods quality to go out of the market for used cars. Due to this process of withdrawing from the market for used cars, lemons are left for sale in the market. This phenomenon is called adverse selection as due to asymmetric information on the part of sellers and buyers the bad quality products drive out the good quality products from the market.

Graphic Illustration Of Lemons Problem

The lemon problem is illustrated in Figure 23.1. It is assumed that two types of used cars-high quality cars and low quality cars are available for sale in the market. To begin with, we assume that both the sellers and buyers know which cars are high quality cars and which are low quality cars. This means there will be two markets, one for high quality cars and the other for low-quality cars. In panel (a) of this figure S_H is the supply curve of high quality used cars and D_H is the demand curve for them. It will be seen from Figure 23.1 that the supply curve S_H of the high-quality used cars (panel (a)) lies at a higher level than supply curve S_L of the low-quality used cars. This is because sellers of high-quality used cars will be willing to supply their product at a relatively higher price as compared to the low quality product. Similarly, demand curve of high-quality used car D, is higher than that of low-quality used cars.

It will be seen from Fig. 23.1 that P_H is the price of high-quality used cars and Q_H is the quantity of high-quality used car bought and sold. On the other hand, in panel (b) the price of used cars determined is P_L and the quantity transacted is Q_L. It will be seen from Figure 23.1 the $P_H > P_L$ and $Q_H = Q_L$ ..

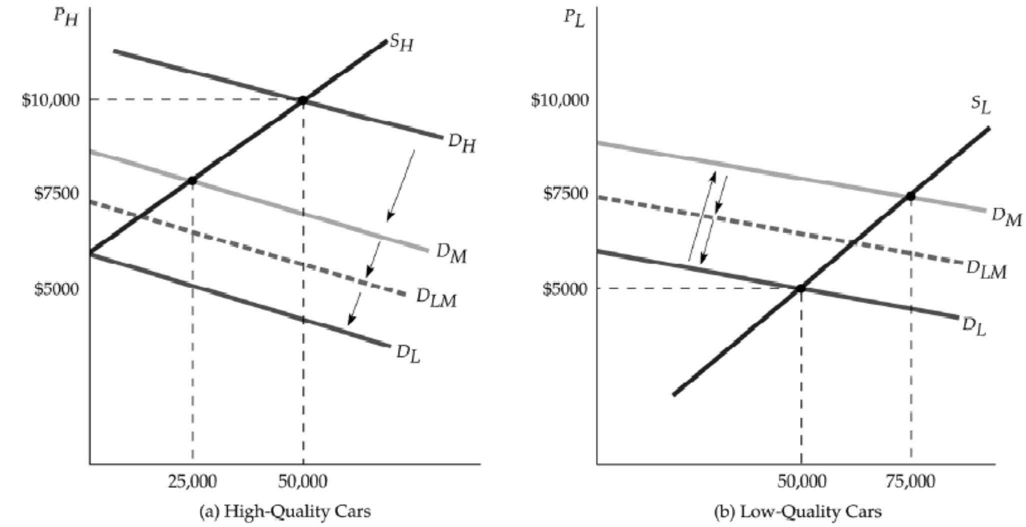


Fig. 23.1

Since sellers of the used cars have better information about the quality of their products than the buyers of them, there is the problem of asymmetric information. In the first instance the buyers might think that there are 50 per cent chances of their getting high-quality cars. Therefore, when deciding to buy the used cars, the buyers would consider all cars to be of medium quality represented by D_M curve in both the panels of Figure 23.1. The demand curve D_M is lower than the demand curve D_H in panel (a) but higher than D_L in panel (b). This means that now greater quantity Q'_L of the low-quality used cars will be transacted than the high-quality used cars whose equilibrium quantity sold has now declined to Q'_H. Thus, in the market for used cars the quantity of low quality cars has increased. As a result the buyers of the used cars will think that the most of the used cars are of low quality, they would be willing to pay lower price for them. This will cause the demand curve for used cars to shift below in both the panels, say to D_{LM} representing demand for lower medium used cars. Given their supply curves, and the demand curve D_{LM}, in the composition of the used cars, quantity of low-quality cars will further increase. As a result, the buyers of used cars, will soon expect that theaverage quality of used cars available for sale will be lower than before causing a downward shift in the demand curve for them. This process of downward shifting of the demand curves in both the panels will continue and average price will continue falling until, given their supply curves, only the low-quality cars (i.e. lemons) are sold in the market and the high quality cars are completely driven out of the market. It will be seen that at the price P_L of the low-quality car (i.e. lemon) no high-quality used cars will be offered for sale in the market.

23.4 ASYMMETRIC INFORMATION AND THE MARKET FAILURE

The example of sale of used cars brings out how asymmetric information leads to market failure, that is, failure to achieve Pareto efficiency. This is due to the externality between the seller of high-quality and low-quality used cars. When some individuals try to sell their low-quality cars, they affect the buyers' perceptions of the quality of the average car available for sale in the market. This causes reduction in the price that they are willing to pay for the used cars available in the market which hurts the individuals who want to sell their goods or high-quality cars. Because of asymmetric information, that is, the buyers cannot easily determine the quality of user cars; they can do so only after they have purchased them and used them for a while. As seen above, the asymmetric information creates the lemons problem and drives good cars out of the market and thereby, harms the welfare of the owners of good quality cars who want to sell their cars. Thus, externality created by asymmetric information prevents the achievement of Pareto efficiency and creates market failure.

23.5 MEASURES ADOPTED TO SOLVE THE PROBLEM OFADVERSE SELECTION

The lemons problem arises because the sellers of good used cars are not able to inform accurately to the buyer about the quality of their cars. To solve this problem, some institutes have been developed to deal with the problem. The first institution that has come about existence is of used cars dealers who provide guarantees for the cars they sell. The well-reputed dealers who want to keep their goodwill in the market give assurance for the quality of car brought through them. Besides, there are automobiles service centres that check the quality of used cars for a certain fee. The buyers of used cars can thus get information about the quality of the used cars from these testing centres.

Some other devices that have developed to solve the problem of quality uncertainty are the use of brand names, the practice of providing franchisees (such as Apollo Hospitals, franchisees of NIIT etc.) chain stores etc. These brands names and franchisees of some reputed companies also provide assurance about the good quality of goods or service being provided. As we shall see below with regard to insurance markets, the device of the co insurance and the deductibles have been adopted to tackle the problem of moral hazard that arises due to asymmetric information.

23.5.1 The Insurance Market and Adverse Selection

After Akerlof who first analyzed the lemons problem created by asymmetric information in the market for used cars, his analysis has been extended to the markets for insurance, financial credit and labour employment. These other markets are also characterized by asymmetric information. In this section we will study how Akerlof's analysis has been applied to the insurance market.

In the case of insurance, the insurance company has less information about the state of health of the individuals who want to get their life insurance or want insurance for sickness or in case of general insurance about the accident proneness of their cars, etc. let us take the case of insurance against illness. Suppose there are two groups of individual who want to insure for the risk of illness. One group of individuals belongs to high risk group. There are H number of individual who a constitute the high risk group. The second group is low risk individuals and there is L number of them. The possibility of their becoming ill and therefore becoming entitled to claim insurance money is P_H and P_L respectively. The individuals belonging to both these groups get themselves insured and the insurance company is not able to distinguish between them. While insurance company does not know the true state of their health, the individuals that is, the buyers of insurance, are fully informed about the likelihood of their becoming ill. Thus, this is a case of asymmetric information. The weighted average probability of their becoming ill is

$$\bar{p} = P_H \frac{H}{H+L} + P_L \frac{L}{H+L}$$

where P stands for the weighted average probability of illness. Obviously, $P_H > P > P_L$. Suppose the cost of insurance is C. If the insurance company is to recover at least the cost of insuring them against illness, then the insurance premium to cover both the high-risk and low-risk individuals which we denote I is given by

$$I \geq C\bar{P}$$

As the individuals know their risk of becoming ill, as is being assumed here, while the individuals with low risk (P_L) may not be willing to buy the insurance, given the above insurance premium, the individuals with high risk will be very much eager to buy it. This is because > P_L the insurance premium will be much greater than CP_L in case of low-risk individuals. Therefore, the low-risk individuals are likely to drop out and in that case to recover the insurance cost of high-risk individuals, the insurance companies will have to raise the rate of insurance premium. As a result, only the high risk individuals with low risk of illness will go without insurance. Thus, as in case of used cars, the problem of adverse selection also arises in case of insurance market. In fact, the term adverse selection was first used in case of health insurance because the insurance companies do not get unbiased selection of individuals who buy insurance policy and it is only the high-risk individuals who purchase it. In such a situation the insurance claims will mostly be made by the individuals with high risk and as a result the insurance company who charges premium rate on the basis of average risk of both kinds of individuals with high-risk and low risk of illness will go bankrupt.

However, the various measures have been proposed to overcome the problem of adverse selection. First, the insurance company can offer the low-risk individuals to get coinsurance (under the coinsurance scheme the individuals share a portion of the loss with the insurance company) or get some deductibility scheme (that is, get some amount deducted from the claim for the loss suffered). The insurance company can offer various individuals to make a choice among different schemes with different rates of coinsurance, the amounts of deductibles, length of contract etc. From the choices made the insurance company can know the risk characteristics of various individuals willing to buy insurance and adjust the insurance premium accordingly.

From above it follows that some schemes have been devised to tackle the problem of adverse selection in connection with health insurance. But one thing is certain from above that there cannot be a single insurance premium policy. There can be at least two insurance policies, one for high-risk individuals who get complete insurance and the other for low-risk individuals who get insurance with a large deductible. In this case the low-risk individuals get only partial insurance. However, given the fact that there is not only two groups of individuals with either high risk or low risk but a continuum of individuals with several probabilities, there can be no equilibrium in the insurance market that ensures market efficiency.[2]

It may be emphasized that the problem of adverse selection in health insurance arises because insurance company cannot fix their rates of premium on the average incidence of illness in the population. Therefore, to solve the problem of adverse selection, compulsory purchase plan of insurance by a group of people containing both high-risk and low-risk individuals and an insurance company fixing its premium rate on the basis of average incidence of health problem has been adopted in some cases. In such a situation it is claimed that each of the individuals comprising compulsory purchase group can be made better off. The high-risk individuals are better off because they buy insurance at a rate that is relatively lower than the risk they face and the low-risk individuals can purchase insurance at a rate less than the rate that would be prevailing when only high-risk people purchased it. Such compulsory insurance plans with some variation have been extensively adopted. For example, all teachers of a college or university can be brought under the compulsory health insurance plan with the insurance company fixing its premium rate on the basis of average risk of health problem faced by the teachers. Similarly, all employees of a factory can be asked to participate in the compulsory insurance plan. But the scheme of compulsory purchase plan is repugnant to most of the economists who lay stress on 'free choice' as a means of achieving Pareto efficiency. But a compulsory purchase plan of insurance restricting choice of individuals ensures Pareto efficiency by tackling the adverse consequences of the existence of externality between the low-risk and high-risk individuals. In this connection, it may be noted that some employers offer health insurance plan to their employees as a part of the package of fringe benefits. As a set of all employees with low-risk and high-risk of illness participate in the health plan the insurance company can fix its premium rate on the basis of average risk and also the problem of adverse selection is eliminated. But in this case cost of insurance plan is borne by the employer and the employees get it as an incentive to work for the employer.

23.6 THE PROBLEM OF MORAL HAZARD

Another problem that often arises in the insurance market is that of moral hazard. The moral hazard refers to a person or firm's behaviour which maychange after buying insurance so that it increases the probability of theft, fire, illness or other accident. This is because when a person buys insurance, the loss from an accident, fire or getting ill is shifted from the person to the insurance company Therefore, when insured against risk, a person takes less care or takes fewer precautions to avoid accident, fire, illness or any other disaster. As a result, the probability of the accident, illness or any other disaster which is insured increases. And when the accident, illness or any other insured disaster occurs, the person or firm inflates the loss suffered. This behaviour of the insured person raises the cost of insurance of the insurance company. For example, a person gets his car insured for theft and accident. Knowing that the loss occurred as a result of an accident willbe borne by the insurance company, he will drive the carmore carelesslyafter buying insurance and thus increasing the probability of accident taking place.

Take another example, say of medical insurances involving moral hazard. After buying medical insurance, individuals usually spend less on health care which increases the probability of their getting ill. Besides, they are likely to spend more on treatment when fall ill than if they had no insurance. This behaviour on the part of individual increases the amount payable by the insurance companies on account of mediclaim.

Similarly, in case of fire insurance, the firms are likely to take fewer precautions such as installation of fire detector system and thereby increasing the probability of a fire and in addition, the firms inflate the loss of property damaged if fire actually occurs. In fact, the probability of fire rises very much, if the property is insured for an amount greater than its true value.

It is worth nothing that if the care taken by an individual who buys insurance is observable, then there is no problem. In case care is observable, the insurance company can fix its premium on the basis of the amount of care taken by the person or firm. Thus, in real life, the insurance companies usually charge different rates from firms which have a fire sprinkler system in their building; or to charge smokers higher premium rates than non-smokers for health insurance. In these cases, the firms try to discriminate among those who buy insurance depending on the care or choices they make that affect the probability of damage.

But the problem of moral hazard arises because the insurance companies cannot observe all the relevant actions regarding care the individuals take after buying insurance. If the problem of moral hazard is not reduced, it could lead to the very high insurance costs and premium rates which would defeat the very purpose of insurance. The purpose of insurance is to distribute the given risks of a large monetary loss among many individuals or firms participating in an insurance programme. But if the facility of available insurance increases the total risks and claimed losses, then insurance does not prove to be efficient and it may not be even possible to provide insurance.

Insurance companies have tried to reduce or overcome the problem of moral hazard. One method generally adopted by the insurance companies is specifying the precautions that individuals or firms buying insurance must take to be eligible for making a claim. For example, the insurance company might require that individual must get medical check-up annually to continue enjoying health insurance. Similarly, the insurance companies may charge higher premium from drivers involved in accidents, may require the installation of fire detector system by the firms as a condition for providing insurance to them. By specifying these conditions, the insurance company attempts to reduce the probability of illness, accident or fire and thereby reduce the possible claims on it.

Another method often adopted by the insurance companies to reduce the problem of moral hazard is that of coinsurance. In coinsurance the individual or firm shares a good portion of a potential loss with the insurance company. That is, in case of coinsurance, the insurance company insures only a part of the possible loss or value of property insured. The idea behind such a proposal is that when the individual or firm shares a good part of potential loss with the insurance company, the individual or firm will take more care and precautions to avoid losses.

A similar proposal to overcome the problem of moral hazard is the provision made in the insurance policy that includes a large "deductible" under which the insured individual or firm has to pay a part of the claim. This also ensures that the insured individual will have incentive to take some amount of care.

23.6.1 Moral Hazard and Allocative Inefficiency

Moral hazard is not only a problem for insurance companies. It also affects the ability of the market to achieve efficient resource allocation. This is illustrated in Figure 23.2. where we have shown the demand curve driving in kilometers per week. The demand curve D is sloping downward to the right. This is because at the higher cost of driving, the individuals switch over to other means of transportation. To begin with, let us suppose the cost of driving per kilometer includes the cost of insurance and the insurance company can measure correctly the number of kilometers driven. Thus in this case the problem of moral hazard does not arise. The owner-driver knows the more driving by him will increase his insurance premium and therefore his total cost of driving will increase. For example, as will be seen from Figure 20.b when the cost of car driving is Rs. 15 per kilometer, (including Rs. 5 per kilometer of insurance cost), the driver will do car driving 200 kilometers per week. Since there is no problem of moral hazard, marginal utility of car driving for 200 kilometers is equal to marginal cost of driving (Rs. 15) per kilometer. 200 kilometers car driving is socially efficient level.

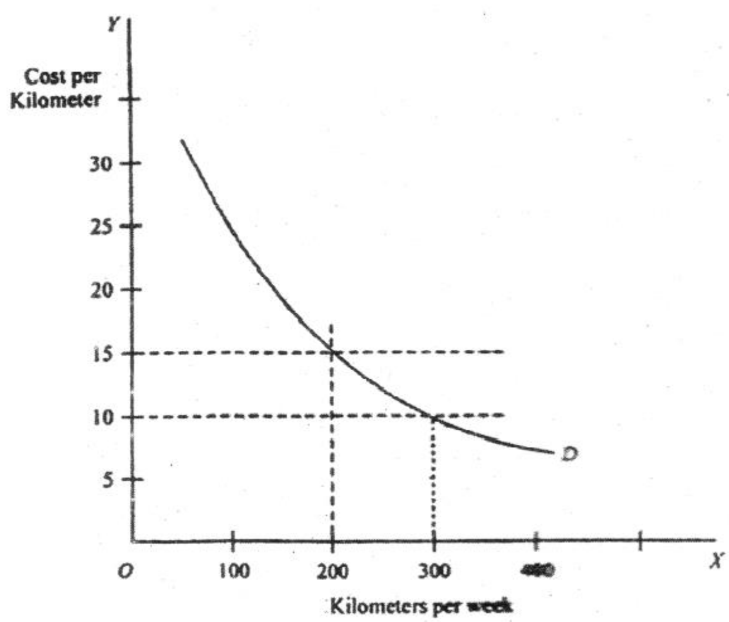


Fig. 23.2. Moral hazard and allocative efficiency

However, the problem of moral hazard arises because it is very difficult for an insurance company to monitor driving habits of individuals and further that insurance premium does not depend on the number of kilometers driven. Once a car owner buys a car insurance policy he has to pay a given insurance premium for a week, month or a year. Therefore, the owner-driver thinks that any additional accident cost will be borne by the insurance company and further that since the insurance premium does not vary with the number of kilometrs driven, he will have therefore incentive to drive car more kilometers than without insurance. Without the insurance premium of Rs. 5 per kilometer, the cost of driving per kilometre will now be Rs. 10 (that is, Rs. 15 - 5 = 10). It will be seen that from Figure 20.b that at the cost of driving Rs. 10 per kilometre, the owner-driver will drive 300 kilometers per week which is more than the socially efficient number of 200 kilometers. Thus it is clear that the problem of moral hazard leads to socially inefficient kilometers of driving activity.

23.7. MARKET SIGNALLING

In the earlier section of this chapter we have explained how asymmetric information leads to the lemons problem. Since sellers know more about the quality of their products than the buyers, the buyers assume that products offered for sale are of low quality. This results in fall in price of the product and ultimately only low-quality products are sold in the market, the good quality products having been driven out of the market. This is called the adverse selection which is the consequence of asymmetric information. Similarly, in case of insurance market we have seen that since the insurance companies cannot know who are high-risk individuals and who are low-risk individuals in order to provide insurance at an appropriate rate of insurance premium the result is again of adverse selection, that is, only high-risk individuals buy insurance at high premiums. In the extreme case such high rates of insurance premiums may come about that no one is willing to buy insurance and the insurance market even ceases to exist. Similarly, as shall be explained at length below, the problem of adverse selection occurs in case of employment of workers by the employers who cannot distinguish between high-productivity and low-productivity workers while deciding to employ them.

The problem ofadverse selectioncaused byasymmetric information has been sought to be resolved through market signalling. It has been proposed that if sellers of high-qualityproducts, low-risk individuals ormore productive workers cansomehow inform or send signals to the potential customers about their superior quality, the problem of adverse selection can be mostly overcome. A firm can send signals that indicates the high quality of its products to its potential buyers by adopting brand names, and byoffering guarantees and warranties. The firm whose products are of low quality cannot offer guarantees and warranties since it will be very costly to do so.

In case of employment of workers education serves as a good signalling device regarding the productivity of workers. Michael Spence showed that in some markets the sellers could send signals to the potential buyers conveying them about the better quality of their products. In the context of labour market, he suggested that education served as a good signal in a labour market with asymmetric information. Suppose a firm plans to hire some new workers. The new workers (i.e. sellers of their labour) know about their quality, that is; productivity of the labour they provide, whereas the buyers do not know how hard they work, how skilful they are. That is, the employer firms do not know who are more productive workers and who are less productive workers. The firms will be able to know about their productivity only after they have been hired and have been working for some time. At the time of hiring them the firms are quite uncertain about their productivity. This lack of information may lead to the adverse selection explained earlier in the context of market for used cars and insurance market.

A pertinent question is why don't firms first hire workers, observe them for some time to know how well they work and terminate the services of those workers who have low productivity and retain only those who are actually more productive. However, this is not a practical proposition because labour laws in many countries do not permit the firing of workers who have worked for more than 6 months in a firm. Besides, firms have to invest a lot of money on providing on the job training to workers. If they, are fired at a later time, the resources spent on their training will go waste. Therefore, it is much better for the firms if they know about the potential productivity of the workers before hiring them.

Spence has claimed that education of workers is a good signal that can be used by the employers to distinguish between high-productivity workers and low-productivity workers. A person's education level can be easily measured by the number of years of schools, the degrees obtained, the college or university where he studied, the grade or marks obtained etc. The education undoubtedly can raise one's productivity.

But as has been stressed by Spence in his model of signalling that even if education does not improve one's productivity, it will still be a useful signal of productivity because it is easier and less costly for more productive persons to obtain a higher level of education as compared to the low-productivity persons. More productive persons are more intelligent highly motivated, more hard-working and are therefore likely to obtain a higher level of education which can be used as a signal of their higher productivity to enable them to get highly lucrative jobs. For all these reasons terms are also right in thinking that education is a signal of productivity.

Spence in his important contribution regarding education as a signal emphasized that education is a good signal if cost of acquiring it is less by a more productive worker as compared to the low productive worker. We explain below his model of signalling in some detail.

Spence Model Of Signalling.

Suppose P_H represents the productivity of high-quality workers and they are H in number and P_H represents the productivity of low-quality of workers whose number is L. The total numbers of workers seeking jobs are therefore H + L. If

P_H workers are paid wages equal to their marginal productivity P_L and L workers are paid wages equal to their marginal productivity Q_1, then there will be efficient use of labour. But, since the employers are unable to distinguish between high- productivity P. workers and low productivity workers it is not possible to achieve Pareto efficiency in labour use. The weighted average productivity of workers which we represent by p is given by

$$P = P_H \frac{H}{H+L} + P_L \frac{L}{H+L}$$

$$P_H > \bar{P} > P_L$$

In view of the fact that it is not possible for the employers to distinguish between high and low productivity workers, they may pay each worker equal to their average marginal productivity as obtained above. Paying workers equal to their average marginal productivity rather than their own marginal productivity would violate the conditions of achieving Pareto efficiency. Let us suppose as Spence does that though education does not cause increase in productivity of workers, it costs less for more productive workers (H) to acquire education as compared to the less productive workers. Using education as a signal of their higher marginal productivity, the educated workers are able to get jobs and are paid high wages. On the other hand, low-quality workers are unable to get education because it is either not easy for them to get education or is very costly for them to acquire education which can serve as signal to get employment.

Spence showed that a number of conditions must be fulfilled if education is to serve as a good signalling device in the job market and equilibrium is achieved that ensures market efficiency in the sense that more productive workers in fact get employment and are paid equal to the value of their marginal products.

The conditions stipulated by Spence in his signalling model can be stated as under. Let C, the cost of education for high-productivity workers and C. the cost of education of the low-productivity workers.

The first essential condition of Spence's model of education as a successful signaling device for higher productivity is $C_H < C_L$ that is, it costs less for high-productivity workers to acquire education.

The second condition is : $C_L < P_H - P_L$ that is, It costs more for low-productivity workers to acquire education for using as a necessary signal then the differential in productivity, and therefore the differences in wages paid to the two types of workers.

$C_L < P_H - P_L$, that is it casts less to high productive workers to invest in education so as to use it as a signal than the differential in productivity, that is perceived to result from possessing that signal.

Those having acquired education to use it as a signal are paid a wage $W_H = P_H$ and those who have not acquired educational signal are paid wage $W_L = P_L$. the following equilibrium condition follows from the above conditions:

$$C_H < (P_H\text{-}P_L) < C_L \qquad \ldots\ldots\ldots\ldots\ldots\ldots \text{(i)}$$

According to the above equilibrium condition, since $C_H < (P_H\text{-}P_L)$, high productive workers find it worthwhile to invest in acquiring education as a signal and get job a the wage rate equal to their high productivity, P_H, and, on the other hand, since $(P_H\text{-}P_L) < C_L$ or $C_L > (P_H\text{-}P_L)$, the low productivity workers do not find it worst while to invest in acquiring education as a signal for perceived right productivity and are satisfied with accepting the low wage rate equal to P_L.

The above equilibrium condition (i) is called separating equilibrium. It shows how the two groups of workers separate themselves for acquiring education as a signal, to get jobs. However, the equilibrium condition (i) is a necessary but not sufficient condition for the success of education as a signalling device. For example, suppose that the equilibrium conditions (i) is satisfied but all workers before they invest in acquiring education are paid a wage equal to their average productive.

Because, as stated above, $C_L > P_H - P_L$, it does not pay the low-productivity workers to invest in acquiring education to serve as signal to get high-paid jobs. However, even through $C_H < P_H - P_L$ but at the same time $C_H > P_H - P$, then even high-productivity workers would not find it sufficient to induce them to invest in acquiring education. Hence in view of being presently paid wage equal to the average productivity, $\bar{p}$ the wage differential ($P_H - \bar{p}$) is considered by the high-productivity workers being not worthwhile to invest in acquiring education as a signal to earn wage rate equal to P_H. Under these circumstances, both the low-productivity workers and high- productivity workers would have no incentive to acquire education as signal. Hence the equilibrium condition (i) derived above is a necessary but a sufficient condition for the success of education as a signal.

Spence's model of treating education as a mere signalling device has been criticized on the ground that it represents a social waste of resources. Based on the assumption that acquiring education does not lead to the increase in true productivityand merelyservesasa signalofperceived productivity,highproductivityworkers are inducedto spend a lot of time (even years) simply to invest in acquiring education to act as successful signal of high perceived productivity. Thus, commenting on Spence's model, Griffiths and Wallwrite, "The efficiency gain is thatofmoreappropriateallocation ofhigh-productivityemployeesaftertheeducativesignalis acquired:theefficiencylossis the resources used up by society in providing some individuals with educational opportunities merelyto serve as a non-productive signalling device and; of course, the opportunitycost of work foregone by high- productivity workers during the educational process.

23.8 INCENTIVES

We turn now to a slightly different topic, the study of incentive Systems. As it turns out, our investigation of this topic will naturally involve asymmetric information. But it is useful to start with the case of full information.

The central question in the design of incentive systems is "How can I get someone to do something for me?" Let's pose this question in a specific context. Suppose that you own a plot of land but you are unable to work on the land yourself. So you try to hire someone to do the farming for you. What sort of compensation system should you set up?

One plan might involve paying the worker a lump sum fee independent of how much he produces. But then he would have little incentive to work. In general a good incentive plan will make the payment of the worker depend in some way on the output he produces. The problem of incentive design to determine exactly how sensitive the payment should be to the produced output.

Let x be the amount of "effort" that the worker expends, and let y = f(x) be the amount of output produced; for simplicitywe suppose that the price of output is 1 so that y also measures the value of the output. Let s(y) be the amount that you pay the worker if he produces y dollars worth of output. Presumably you would like to choose the function s(y) to maximize your profits y-s(y). what are the constraints that you face? In order to answer this question we have to look at things from the worker's perspective.

We assume that the worker finds effort costly, and write c(x) for the cost of effort x. We assume that this cost function has the usual shape both total and marginal costs increase as effort increases. The utility of the worker who chooses effort level r is then simplys(y)-c(x) = s(f(x))-c(x). The worker may have other alternative available that give him some utility u. This could come from working at other jobs or from not working at all. All that is relevant for the design of the incentivescheme is that the utility that the worker gets from this job must be at least as great as the utility he could get elsewhere. This gives us the participation constraint:

$$s(f(x))-c(x) \geq u$$

Given this constraint we can determine how much output we can get from the worker. You want to induce the worker to choose an effort level c that yields you the greatest surplus given the constraint that the worker is willing to work for you:

$$\text{Max}_X\ f(x) - s(f(x))$$

Such that $s(f(x))-c(x) \geq \bar{u}$

In general, you will want the worker to choose x to just satisfy the constraint so that s(f(x))-c(x) = $\bar{u}$. Substituting this into the objective function we have the unconstrained maximization problem

Maxx f(x))-c(x) - $\bar{u}$

But it is easy to solve this problem! Just choose x^* so that the marginal product equals the marginal cost:

$MP(x^*)= MC(x^*)$

Any choice of $x^*$where the marginal benefit is not equal to the marginal cost cannot maximize profits.

This tells us what level of effort the owner wants to achieved now we have to ask what he has to pay the worker to achieve that effort. That is, what does the function s(y) have to look like to induce the worker to choose to make x^* the optimal choice?

Suppose that you decide that you want to induce the worker to put in x^* amount of effort. Then you must make it in his interest to do so; that is, you must design your incentive scheme s(y) so that the utility from choosing to work x^* is larger than the utility of worker any other amount x. This gives us the constraint

$s(f(x^*))-c(x^*) \geq s(f(x))-c(x)$ for all x

This constraint is called the incentive compatibility constraint. It simply says that the utility to the worker from choosing x* must be greater than the utility of any other choice of effort.

So we have two conditions that the incentive scheme must satisfy: first, it must give total utility to the worker of $\bar{u}$, and second, it must make the marginal product of effort equal to the marginal cost of effort at the effort level x*. There are several ways to do this.

Rent. The landowner could simply rent the rand to the worker for some price R, so that the worker gets all the output he produces after he pays the owner R. For this scheme

$s(f(x))= f(x) -R$

If the worker maximizes s(f(x))-c(x)= f(x) -R -c(x), he will choose the effort level where MP(x*) = MC(x*), what the owner wants. The rental rate R is determined from the participation condition. Since the total utility to the worker must be ? we have

$$f(x^*) - c(x^*) - R = \bar{u}$$

which says $R = f(x^*) - c(x^*) - \bar{u}$

Wage labor. In this scheme the landowner pays the worker a constant wage per unit of effort along with a lump sum K. This means that the incentive payment takes the form uu

$$S(x) = wx + K$$

The wage rate w is equal to the marginal product of the worker at the optimal choicer x^* $MP(x^*)$. The constant K chosen to just make the worker indifferent between working for the landowner and is, it is chosen to satisfy the participation constraint.

The problem of maximizing f(x) - c(x)then becomes

$$\max_X wx + K - c(x),$$

which means that the worker will choose o so as to set his marginal cost equal to the wage: w = MC (x). Since the wage is $MP(x^*)$, this means that the optimal choice of the worker will be x^* such that $MP(x^*) = MC(x^*)$ which is just what the firm wants.

Take-it-or-leave-it. In this scheme the landowner pays the worker B* if he works x* and zero otherwise. The amount B* is determined by the participation constraint $B^* - c(x^*) = \bar{u}$, so $B^* = \bar{u} - c(x^*)$. If the worker chooses any level of effort $x \neq x^*$, he gets a utility of -c(x). If he chooses x*, he gets a utility of z. Hence the optimal choice for the worker is to set x= x*.

Each of these schemes is equivalent as far as the analysis goes: each one gives the worker a utility of $\bar{u}$, and each one gives the worker an incentive to work the optimal amount x*. At this level of generality there is no reason to choose between them.

If all of these schemes are optimal, what could a nonoptimal scheme look like? Here is an example.

sharecropping. In sharecropping the worker and the landowner each get some fixed percentage of the output. Suppose that the worker's share takes the form $s(x) = \langle f(x) + F$, where F is some constant and <1. This is not an efficient scheme for the problem under consideration. It is easy to see why. The worker's maximization problem is

$$\max_x \langle f(x) + F - c(x),$$

which means that he would choose a level of effort ? where

$$\langle MP(\square) = MC(\square).$$

Such an effort level clearly cannot satisfy the efficiency condition that $MP(x) = MC(x)$.

Such an effort level clearly cannot satisfy the efficiency condition that $MP(x) = MC(x)$.

Here is a way to summarize this analysis. In order to design an efficient incentive scheme it is necessary to ensure that the person who makes the effort decision is the residual claimant to the output. The way the owner can make himself as well off as possible is to make sure that he gets the worker to produce the optimal amount of output. This is the output level where the marginal product bf the worker's extra effort equals the marginal cost of putting forth that effort. It follows that the incentive scheme must provide a marginal benefit to the worker equal to his marginal product.[3]

CHECK YOUR PROGRESS I

Answer the questions in the space provided:

Q1: How does asymmetric information lead to market failure?

Q2: What are the key measures used to reduce adverse selection in markets?

____

____

____

____

Q3: How does moral hazard affect the behavior of insured individuals?

____

____

____

____

23.5 LET US SUM UP

Dear learner, in this lesson, we discussed the following :

- In the market for lemons, the buyers and sellers have different information about the quality of goods being bought and sold. The word lemon is used to describe a defective or low quality product.
- The lemons problem arises because the sellers of good used cars are not able to inform accurately to the buyer about the quality of their cars.
- In the case of insurance, the insurance company has less information about the state of health of the individuals who want to get their life insurance or want insurance for sickness or in case of general insurance about the accident proneness of their cars, etc.
- Another problem that often arises in the insurance market is that of moral hazard. The moral hazard refers to a person or firm's behaviour which may change after buying insurance so that it increases the probability of theft, fire, illness or other accident.

- This is because when a person buys insurance, the loss from an. accident, fire or getting ill is shifted from the person to the insurance company, Therefore, when insured against risk, a person takes less care or take fewer precautions to avoid accident, fire, illness or any other disaster.
- Moral hazard is not only a problem for insurance companies. It also affects the ability of the market to achieve efficient resource allocation.

  The problem of adverse selection caused by asymmetric information has been sought to be resolved through market signalling.
- In order to design an efficient incentive scheme it is necessary to ensure that the person who makes the effort decision is the residual claimant to the output.

23.6 KEYWORDS

1. **Lemon** – A defective or substandard product, often used to refer to poor-quality used cars.
2. **Asymmetric Information** – Asituation in which one party in a transaction has more or better information than the other.
3. **Adverse Selection** – A market process in which buyers or sellers with higher risks are more likely to participate, due to information asymmetry.
4. **Moral Hazard** – The tendency of individuals or firms to take greater risks once they are protected from the consequences, such as through insurance.
5. **Market Signalling** – Actions taken by informed parties (e.g., sellers) to reveal information to uninformed parties (e.g., buyers), like offering warranties.
6. **Incentive System** – A structure designed to align the interests of individuals with desired outcomes, often by making them bear the consequences of their actions.

23.7 EXAMINATION ORIENTED QUESTIONS

1. What is asymmetric information?
2. Explain why lemons in the used cars market drive out good quality used cars.
3. Discuss the role of signalling and incentives in overcoming information problems in markets.
4. Explain the concept of the 'Market for Lemons'. How does asymmetric information affect the quality of goods traded in a market?
5. How does asymmetric information lead to market failure? Discuss with relevant examples.
6. What measures are adopted to solve the problem of adverse selection in the market?
7. Explain adverse selection in the insurance market with suitable illustrations.
8. Define moral hazard. How does it lead to allocative inefficiency in markets?
9. What is market signalling? How can it help overcome the problem of asymmetric information?

23.8 HINTS TO CHECK YOUR PROGRESS

Q1. See section 23.4

Q2. See section 23.5

Q3. See section 23.6

23.9 SUGGESTED READINGS

1. Ahuja, H.L. *Advanced Economic Theory, Microeconomic Analysis*, S. Chand & Company Ltd.

2. Handerson J and R Quandt: *Microeconomics Theory, A Mathematical Approach*, McGraw-Hill.

3. Hader, J. *Mathematical Theory of Economic Behaviour*, Addison-Wesley.

4. Koutsoyannis, A: *Modern Microeconomics*, Macmillan.

5. Mas-Colell A, M.D. Whinston & J.R. Greene: *Microeconomic Theory*, OUP.

6. Salvatore, D: *Microeconomics*, OUP.

7. Sen, A: *Microeconomics*, OUP.

8. Varian, H: *Intermediate Microeconomics*, East West Press.

9. Pindyck, Robert. S and Daniel l. Rubinfeld, Microeconomics (fifth edition).

10. Boyes, William and Michael Melvin, Textbook of Economics (sixth edition, Indian adaptation)

11. Sloman, John, Economics (Third edition).

12. Pindyck R.S., D. Rubinfeld and P.L. Mehta (2012), Microeconomics, Pearson Education, South Asia.

13. Salvatore, D. (2010), Microeconomics; Theory and Applications, Oxford University Press, New Delhi

14. Cooper, R., & John, A. (2011). Economics–theory through applications, Saylor Foundation. [4]

15. Shapiro, D., MacDonald, D., & Greenlaw, S. A. (2022). *Principles of Microeconomics 3e*.[5]

ENDNOTES

1. George Akerlof, "The Market for Lemons," *The Quarterly Journal of Economics*, Aug. 10, pp. 488–550.

2. M. Rothschild and J. Stiglitz, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics*, Nov. 1976.

3. This lesson draws extensively from *Intermediate Microeconomics* by Hal R. Varian.

4. Type of license/Conditions of use: CC BY-NC-SA (Attribution-Non commercial-ShareAlike) .Scan QR code 1 to access **Cooper, R., & John, A. (2011). Economics–theory through applications.**

5. This book is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). “Access for free at openstax.org.” **Scan QR code 2 to access the e book by Shapiro, D., MacDonald, D., & Greenlaw, S. A. (2022).** ***Principles of Microeconomics 3e*.**

QR code 1

QR code 2