Lesson - 1

Business Economics- Meaning, Nature, Scope and significance

Introduction and meaning :

(**Author : Dr. M.S. Khanchi**)

Business Economics, also called Managerial Economics, is the application of economic theory and methodology to business. Business involves decision-making. Decision making means the process of selecting one out of two or more alternative courses of action. The question of choice arises because the basic resources such as capital, land, labour and management are limited and can be employed in alternative uses. The decision-making function thus becomes one of making choice and taking decisions that will provide the most efficient means of attaining a desired end, say, profit maximation.

Different aspects of business need attention of the chief executive. He may be called upon to choose a single option among the many that may be available to him. It would he in the interest of the business to reach an optimal decision- the one that promotes the goal of the business firm. A scientific formulation of the business problem and finding its optimals solution requires that the business firm is he equipped with a rational methodology and appropriate tools.

Business economic meets these needs of the business firm. This is illustrated in the following presentation.

Economic Theory and Methodology

Decision problems in Business

Business Economic
Application of Economic
Theory and Methodology to
solving Business problems

Optimal Solution
to Business
Problems

it may be that business economics serves as a bridge between economic theory and decision-making in the context of business.

According to Mc Nair and Meriam, ñBusiness economic consists of the use of economic modes of thought to analyse business situations.ò

Siegel man has defined managerial economic (or business economic) as ñthe integration of economic theory with business practice for the purpose of facilitating decision-making and forward planning by management.ò

We may, therefore, define business economic as that discipline which deals with the application of economic theory to business management. Business economic thus lies on the borderline between economic and business management and serves as a bridge between the two disciplines.

Nature of Business Economics :

Traditional economic theory has developed along two lines; viz., normative and positive. Normative focuses on prescriptive statements, and help establish rules aimed at attaining the specified goals of business. Positive, on the other hand, focuses on description it aims at describing the manner in which the economic system operates without staffing how they should operate.

The emphasis in business economics is on normative theory. Business economic seeks to establish rules which help business firms attain their goals, which indeed is also the essence of the word normative. However, if the firms are to establish valid decision rules, they must thoroughly understand their environment. This requires the study of positive or descriptive theory. Thus, Business economics combines the essentials of the normative and positive economic theory, the emphasis being more on the former than the latter.

Scope of Business Economics :

As regards the scope of business economics, no uniformity of views exists among various authors. However, the following aspects are said to generally fall under business economics.

1. Demand Analysis and Forecasting
2. Cost and production Analysis.
3. Pricing Decisions, policies and practices.
4. Profit Management.
5. Capital Management.

These various aspects are also considered to be comprising the subject matter of business economic.

1. Demand Analysis and Forecasting :

A business firm is an economic organisation which transform productive resources into goods to be sold in the market. A major part of business decision making depends on accurate estimates of demand. A demand forecast can serve as a guide to management for maintaining and strengthening market position and enlarging profits. Demands analysis helps identify the various factors influencing the product demand and thus provides guidelines for manipulating demand.

Demand analysis and forecasting provided the essential basis for business planning and occupies a strategic place in managerial economic. The main topics covered are: Demand Determinants, Demand Distinctions and Demand Forecastmg.

2. Cost and Production Analysis :

A study of economic costs, combined with the data drawn from the firmôs accounting records, can yield significant cost estimates which are useful for management decisions. An element of cost uncertainty exists because all the factors determining costs are not known and controllable. Discovering

economic costs and the ability to measure them are the necessary steps for more effective profit planning, cost control and sound pricing practices.

Production analysis is narrower, in scope than cost analysis. Production analysis frequently proceeds in physical terms while cost analysis proceeds in monetary terms. The main topics covered under cost and production analysis are: Cost concepts and classification, Cost-output Relationships, Economics and Diseconomics of scale, Production function and Cost control.

3. Pricing Decisions, Policies and Practices :

Pricing is an important area of business economic. In fact, price is the genesis of a firms revenue and as such its success largely depends on how correctly the pricing decisions are taken. The important aspects dealt with under pricing include. Price Determination in Various Market Forms, Pricing Method, Differential Pricing, Product-line Pricing and Price Forecasting.

4. Profit Management :

Business firms are generally organised for purpose of making profits and in the long run profits earned are taken as an important measure of the firms success. If knowledge about the future were perfect, profit analysis would have been a very easy task. However, in a world of uncertainty, expectations are not always realised so that profit planning and measurement constitute a difficult area of business economic. The important aspects covered under this area are : Nature and Measurement of profit, Profit policies and Technique of Profit Planning like Break-Even Analysis.

5. Capital Management :

Among the various types business problems, the most complex and troublesome for the business manager are those relating to a firmôs capital investments. Relatively large sums are involved and the problems are so complex that their solution requires considerable time and labour. Often the decision involving capital management are taken by the top management. Briefly Capital management implies planning and control of capital

expenditure. The main topics dealt with are: Cost of capital Rate of Return and Selection of Projects.

Conclusion :

The various aspects outlined above represent major uncertainties which a business firm has to reckon with viz., demand uncertainty, cost uncertainty, price uncertainty, profit uncertainty and capital uncertainty. We can therefore, conclude that the subject matter of business economic consists of applying economic principles and concepts to dea1 with various uncertainties faced by a business firm.

Significance of Business Economics :

The significance of business economics can be discussed as under :

1. Business economic is concerned with those aspects of traditional economics which are relevant for business decision making in real life. These are adapted or modified with a view to enable the manager take better decisions. Thus, business economic accomplishes the objective of building a suitable tool kit from traditional economics.

2. It also incorporates useful ideas from other disciplines such as psychology, sociology, etc. If they are found relevant to decision making. In fact, business economics takes the help of other disciplines having a bearing on the business decisions in relation various explicit and implicit constraints subject to which resource allocation is to be optimized.

3. Business economics helps in reaching a variety of business decisions in a complicated environment. Certain examples are :

   (i) What products and services should be produced?

   (ii) What input and production technique should be used?

   (iii) How much output should be produced and at what prices it should be sold?

(iv) What are the best sizes and locations of new plants?

(v) When should equipment be replaced?

(vi) How should the available capital be allocated?

4. Business economics makes a manager a more competent model builder. It helps him appreciate the essential relationship Characterising a given situation.

5. At the level of the firm. Where its operations are conducted though known focus functional areas, such as finance, marketing, personnel and production, business economics serves as an integrating agent by coordinating the activities in these different areas.

6. Business economics takes cognizance of the interaction between the firm and society, and accomplishes the key role of an agent in achieving the its social and economic welfare goals. It has come to be realised that a business, apart from its obligations to shareholders, has certain social obligations. Business economics focuses attention on these social obligations as constraints subject to which business decisions are taken. It serves as an instrument in furthering the economic welfare of the society through socially oriented business decisions.

Conclusion :

The usefulness of business economics lies in borrowing and adopting the toolkit from economic theory, incorporating relevant ideas from other disciplines to take better business decisions, serving as a catalytic agent in the process of decision making by different functional departments at the firmôs level, and finally accomplishing a social purpose by orienting business decisions towards social obligations.

Lesson - 2

Theory of Consumer's Behaviour : Utility Analysis

(Author: Dr. M.S. Khanchi)

The theory of consumerôs behaviour seeks to explain the determination of consumerôs equilibrium. Consumerôs equilibrium refers to a situation when a consumer gets maximum satisfaction out of his given resources. A consumer spends his money income on different goods and services in such a manner as to derive maximum satisfaction. Once a consumer attains equilibrium position, he would not like to deviate from it. Economic theory has approached the problem of determination of consumerôs equilibrium in two different ways: (1) Cardinal Utility Analysis and (2) Ordinal Utility Analysis Accordingly, we shall examine these two approaches to the study of consumerôs equilibrium in greater defait.

Utility Analysis or Cardinal Approach :

The Cardinal Approach to the theory of consumer behaviour is based upon the concept of utility. It assumes that utility is capable of measurement. It can be added, subtracted, multiplied, and so on.

According to this approach, utility can be measured in cardinal numbers, like 1,2,3,4 etc. Fisher has used the term ôUtilô as a measure of utility. Thus in terms of cardinal approach it can be said that one gets from a cup of tea 5 utils, from a cup of coffee 10 utils, and from a rasgulla 15 utils worth of utility.

Meaning of Utility :

The term utility in Economics is used to denote that quality in a good or service by virtue of which our wants are satisfied. In, other words utility is defined as the want satisfying power of a commodity. According to, Mrs. Robinson, ñUtility is the quality in commodities that makes individuals want to buy them.ò

According to Hibdon, ñUtility is the quality of a good to satisfy a want.ò

Features :

Utility has the following main features :

(1) ***Utility is Subjective :*** Utility is subjective because it deals with the mental satisfaction of a man. A commodity may have different utility for different persons. Cigarette has utility for a smoker but for a person who does not smoke, cigarette has no utility. Utility, therefore, is subjective.

(2) ***Utility is Relative :*** Utility of a good never remains the same. It varies with time and place. Fan has utility in the summer but not during the winter season.

(3) ***Utility and usefulness :*** A commodity having utility need not be useful. Cigarette and liquor are harmful to health, but if they satisfy the want of an addict then they have utility for him.

(4) ***Utility and Morality :*** Utility is independent of morality. Use of liquor or opium may not be proper from the moral point of views. But as these intoxicants satisfy wants of the drinkards and opiumeaters, they have utility for them.

Concepts of Utility :

There are three concepts of utility :

(1) ***Initial Utility :*** The utility derived from the first unit of a commodity is called initial utility. Utility derived from the first piece of bread is called initial utility. Thus, initial utility, is the utility obtained from the consumption of the first unit of a commodity. It is always positive.

(2) ***Total Utility :*** Total utility is the sum of utility derived from different units of a commodity consumed by a household.

According to Leftwitch, ñTotal utility refers to the entire amount of satisfaction obtained from consuming various quantities of a commodity.ò Supposing a consumer four units of apple. If the consumer gets 10 utils from

the consumption of first apple, 8 utils from second, 6 utils from third, and 4 utils from fourth apple, then the total utility will be 10+8+6+4 = 28

Accordingly, total utility can be calculated as :

$TU = MU_1 + MU_2 + MU_3 + \_\_\_\_\_ + MU_n$

or

TU = EMU

Here TU = Total utility and MU_1, MU_2, MU_3, + _____ MU_n =

Marginal Utility derived from first, second, third _____ and nth unit.

(3) ***Marginal Utility :*** Marginal Utility is the utility derived from the additional unit of a commodity consumed. The change that takes place in the total utility by the consumption of an additional unit of a commodity is called marginal utility.

According to Chapman, ñMarginal utility is the addition made to total utility by consuming one more unit of commodity. Supposing a consumer gets 10 utils from the consumption of one mango and 18 utils from two mangoes, then. the marginal utility of second .mango will be 18-10=8 utils.

Marignal utility can be measured with the help of the following formula $MU_{nth} = TU_n$ ï TU_{n-1}

Here MU_{nth} = Marginal utility of nth unit,

TU_n = Total utility of ónô units,

TU_{n-l} = Total utility of n-i units,

Marginal utility can be (i) positive, (ii) zero, or (iii) negative.

(i) ***Positive Marginal Utility :*** If by consuming additional units of a commodity, total utility goes on increasing, marginal utility will be positive.

(ii) ***Zero Marginal Utility :*** If the consumption of an additional unit of a commodity causes no change in total utility, marginal utility will be zero.

(iii) ***Negative Marginal Utility :*** If the consumption of an additional unit of a commodity causes fall in total utility, the marginal utility will be negative.

Relationship between total utility and Marginal Utility :

The relationship between total utility and marginal utility may be better understood with the help of a utility schedule and a diagram as shown below :

Table No. I

No. of units Consumed	Total Utility	Marginal Utility
0	0	-
1	10	10
2	18	18
3	24	6
4	26	2
5	26	0
6	24	-2
7	21	-3

The relationship between total utility and marginal utility can be explained with the help of the above table and diagram based thereon.

1. Total utility, initially, increases with the consumption of successive units of a commodity. Ultimately, it begins to fall.

2. Marginal Utility continuously diminishes.

3. As long as marginal utility is more than zero or positive, total utility increases, total utility is maximum when marginal utility is zero. It falls when marginal utility is negative.

4. When marginal utility is zero or total utility is maximum, a poin of saturation is obtained.

Laws of Utility Analysis :

Utility analysis consists of two important laws

1. Law of Diminishing Marginal Utility.

2. Law of Equi-Marginal Utility.

1. Law of Diminishing Marginal Utility :

Law of Diminishing Marginal Utility is an important law of utility analysis. This law is related to the satisfaction of human wants. All of us experience this law in our daily life. If you are set to buy, say, shirts at any given time, then as the number of shirts with you goes on increasing, the marginal utility from each successive shirt will go on decreasing. It is the reality of a manôs life which is referred to in economics as law of Diminishing Marginal Utility. This law is also known as Gossenôs First Law.

According to Chapman, ñThe more we have of a thing, the less we want additional increments of it or the more we want not to have additional increments of it.ò

According to Marshall, ñThe additional benefit which a person derives from a given stock of a thing diminishes with every increase in the stock that he already has.ò

According to Samuelson, ñAs the amount consumed of a good increases, the marginal utility of the goods tends to decrease.ò

In short, the law of Diminishing Marginal Utility states that, other things being equal, when we go on consuming additional units of a commodity, the marginal utility from each successive unit of that commodity goes on diminishing.

Assumptions :

Every law in subject to clause ñother things being equalò This refers to the assumption on which a law is based. It applies in this case as well. Main assumptions of this law are as follows:

1. Utility can be measured in cardinal number system such as 1,2,3 _____ etc.
2. There is no change in income of the consumer.
3. Marginal utility of money remains constant.
4. Suitable quantity of the commodity is consumed.
5. There is continuous consumption of the commodity.
6. Marginal Utility of every commodity is independent.
7. Every unit of the commodity being used is of same quality and size.
8. There is no change in the tastes, character, fashion, and habits of the consumer.
9. There is no change in the price of the commodity and its substitutes.

Explanation of the Law :

The Law of Diminishing Marginal Utility can be explained with the help of Table and Figure.

Table No.2

No. of Breads	**Marginal Utility**

1	**8**
2	**6**
3	**4**
4	**2**
5	**0point of Satiety**
6	**-2**

It is clear from the above Table that when the consumer consumes first unit of bread, he get marginal utility equal to 8. Marginal utility from the consumption of second, third and fourth bread is 6, 4 and 2 respectively. He gets zero marginal utility from the consumption of fifth bread. This is known as point of satiety for the consumer. After that he gets negative utility i.e. -2 from the consumption of sixth unit of bread. Thus, the table shows that as the consumer goes on consuming more and more units of bread, marginal utility goes on diminishing.

Pricing Decision :

A retailorôs price policy is a crucial positioning factor and must be decided in relation to its target market, its product and service assortments and its competition. This involved the decisions regarding the price lilies to be earned and overall markdown or sale policies:

Promotion Decision :

Retailers use the promotional tools - advertising, personal selling, sales promotion and public relations to reach. Customers Personal selling requires careful training of sales people in how to greet customers, meet their needs and handle their complaints.

THE FUTURE OF RETAILING :

Present scenario of retailing is that retailerôs margins are very low. They are able to survive on low margins due to remarkable capacity for thrift. In many traditional shops the family provides much of the labour. He performs several functions distribution, finance and risk taking. When there is keen competition, retailers tend to undercut each other. They compensate themselves by taking higher margins on other products, or by increasing the turnover.

WHOLESALING :

Wholesaling is the sale, and all activities directly related to the sale, of goods and services, to business and other organizations for (1) resale (2) use in producing other goods and services or (3) operating an organization.

Wholesalers buy mostly from producers and sell mostly to retailers, industrial consumers and other wholesalers.

NATURE AND IMPORTANCE OF WHOLESALING :

Here we will focus on firms engaged primarily in wholesaling. Retailers may also be occasionally be involved in wholesale transaction.

Manufacturers small or big cannot establish their own direct link with retailers or customers. It is not cost effective to them. At the other end of the distribution channel, most retailers and final users buy in small quantities and have only a limited knowledge of the market and source of supply. Thus there is gap, a wholesaling middleman can fill this gap by providing services of value to manufacturers and or to the retailers. Wholesaling brings to the total distribution system the economies of skill, scale and transactions.

Wholesaling skills are efficiently concentrated in a relatively few hands. This saves the duplication of effort that would occur if many producers had to perform wholesaling function themselves.

Economics of Scale are there because of the specialization of who leasing function that might otherwise require several small departments run by producing firms. Wholesalers typical can perform wholesaling functions more efficiently than most manufacturers can.

FUNCTION OF WHOLESALERS :

Wholesalers perform number of functions. They facilitate the task of producer and retailer by performing one or more of the following channel functions :

Selling and promoting :

Wholesalersô sales force help manufacturers reach many small customers at low cost. The wholesalers have more contacts and are often more trusted by the buyer than the distant manufacturer.

Buying and assorting :

Wholesalers can select items and build assortments needed by their customers, thereby saving the consumers much work.

Warehousing :

Wholesalers hold inventories, thereby reducing the inventory costs and risks of suppliers and customers.

Transportation :

wholesalers can provide quicker delivery to buyers because they are closer than the producers.

Financing :

Wholesalers finance their customers by giving credit and they finance their suppliers by ordering early and paying bills in time.

Risk bearing :

Wholesalers absorb risk of the manufacturers by taking title and bearing the cost of theft damage, spoilage and obsolescence. Market information:

Wholesalers give information to suppliers and customers about competitors new product and price developments.

Management services and advice :

Wholesalers often help retailers train their sale clerks, improve store layouts and displays and setup accounting and inventory control systems.

TYPES OF WHOLESALERS :

Wholesalers can be broadly divided into three broad categories Merchant wholesaler, Agent wholesaling middleman and Manufacturersô Sales facility.

Fig.9 - Types of wholesaling institutions :

Wholesaling Middleman

Merchant Wholesa1ing including	Agent Wholesalers middleman, including	Manufacturers Sales facilities including
Full Service	Manufacturers agents	Branches
Truck Jobber	Brokers	Offices
Drop Shippers		

Merchant Wholesalers :

A merchant wholesaler is independently owned business that takes title to the merchandise it handles. Merchant wholesalers include Full service, Truck jobbers, Drop Shippers.

Full service wholesalers :

Full service wholesalers provide a full set of services, such as carrying stock, using a sales force, offering credit, making deliveries and providing management assistance. They are either wholesale merchants or industrial distributors. Wholesale merchants sell mostly to retailers and provide a full range of services. Industrial distributors are merchant wholesalers that sell to producers rather than to retailers.

Truck Jobbers :

They perform a selling and delivery function. They carry a limited line of goods (such as milk, bread or snack food) that they sell for cash as they make their rounds of supermarkets, small groceries, hospitals etc.

Drop Shippers :

They operate in bulk industries such as coal and heavy equipment. They do not carry inventory or handle the product. Once an order is received, they find a producer who ships the goods directly to the customer.

Agent wholesaling middleman :

It is an independent firm that engaged primarily in wholesaling by actively negotiating the sale or purchase of products or behalf of other firms but that does not take little to the products being distributed.

Manufacturers Agents :

Agents represent buyers a seller on a more permanent basis. Manufacturersô agents represent two or more manufacturers of related lilies. They have a formal agreement with each manufacturer covering prices, territories, order handling procedures, delivery and warranties and commission rates. They know each manufacturerôs product line and use their wide contact to sell the products.

Brokers :

A broker brings buyer and sellers together and assists in negotiations. The parties hiring them pay brokerage. They do not carry inventory, get

involved in financing or assume risk. Examples are: Food brokers, real estate brokers, insurance brokers and security brokers.

Table No. 3

Rupees	M.U. of Apples	M.U. of Bananas
1	10	8
2	8	6
3	6	4
4	4	2
5	2	1

It is clear from the Table that if the consumer, spends Rs.3 on apples and Rs.2 on bananas, the marginal utility lie gets from the last rupee on both becomes equal i.e. 6. In this way he gets maximum satisfaction. The total utility from both the commodities will be 10+8+6+8+6 = 38, which is maximum. In case the consumer spends his income in any other manner, he will act lesser total utility.

In this diagram units of money are shown on ox-axics and marginal utility on oy-axics. It indicates that if the income of the consumer is Rs. 5.00, he will spend Rs. 3.00 on apples and Rs. 2.00 on bananas, because third rupee spent on apples and second rupee spent on banana yield him equal marginal utility i.e. 8 utils. By distributing his income on apples and bananas in this manner, the consumer gets total utility of 38 utils. It will be the maximum total utility derived by the consumer out of his expenditure of Rs. 5.00. So by spending his income in this manner the consumer will get maximum satisfaction.

If the, consumer spends his income on apples and bananas in any other manner, his total utility will be less than the maximum as shown in diagram.

It is evident from the above figure that by spending one rupee less on apples the loss will be equal to ABCD and by spending one rupee more on

bananas the gain win be equal to EFGH. It is clear that (ABCD) < (EFGH), hence loss is more than gain.

Importance of the Law :

The importance of the law of equi-marginal utility can be explained as follows:

1. **Consumption :** If a consumer spends his income, as suggested by this law, on different commodities in such a way that the last unit of money spent on them yields him equal marginal utility, he will be getting maximum satisfaction out of his income.

2. **Production :** Every producer aims at earnings maximum profit. To achieve this objective he must utilize different factors of production in such a way that the marginal productivity of each factor is equal.

3. **Exchange :** Acting upon the law of equi-marginal utility, every person will go on substituting goods giving more utility for the ones giving less utility, till the marginal utility of all becomes equal. Exchange will stop at that point.

4. **Distribution :** It refers to the distribution of national income among the factors of production, i.e. land, labour, capital, etc. Distribution is done in such a way that in the long-run every factor gets its share out of national income according to its marginal productivity.

5. **Public Finance :** At the time of levying taxes, finance minister takes the help of this law. He levies taxes in such a manner that the marginal sacrifice of each tax-payer is equal. Then only it will have the lest burden on all tax-payers. To achieve this objective, a finance minister may substitute one tax for the other.

Criticism of the Law :

This law has been subjected to the following criticism.

1. **Cardinal measurement of utility is not possible :** Measurement of utility is not possible. How can a consumer say that he would get 10 utils

of utility from first apple and 8 utils, of utility from the second. Unless marginal utility is estimated, application of the law will remain dubious.

2. **Consumers are not fully rational :** The assumption that consumers are fully rational is not correct. Some consumers are idle by nature, and so to satisfy their habits and customs, they sometimes buy goods yielding less utility. Consequently, they do not get maximum satisfaction.

3. **Shortage of Goods :** If goods giving more utility are not available in the market, the consumer will have to consume goods yielding less utility.

4. **Ignorance of the consumer :** Consumer is ignorant about many things concerning consumption. Many a times, he is ignorant about the right price of the goods. He is ignorant about the less expensive substitutes that may be way available in the market. He is also ignorant about the different uses of goods. On account of this ignorance, the consumer fails to spend his income in a manner that may yield him maximum satisfaction.

5. **Influence of Fashion, Customs and Habits :** Actual expenditure of every consumer is influenced by fashion, customs, and habits. Under their influence, many a times the consumer buys more of such goods which give less utility.

   Consequently, he buys less of those goods which give more utility. Hence he fails to spend his income according to this law.

6. **Constant Income and Price :** An important assumption of the law is that the income of the consumer and the price of the goods should remain constant. Income of the consumer is limited, as such he cannot increase his satisfaction beyond a particular limit. Likewise, prices being constant, he will get only as much of satisfaction as the amount of goods that he can buy with limited income. He cannot extend his satisfaction beyond this limit.

7. **Change in the Marginal Utility of Money :** The assumption that marginal utility of money remains constant is also unrealistic. In actual

life, marginal utility of money may increase or decrease. Due to increase in the marginal utility of money, a consumer will have to rearrange his expenditure on different goods.

8. **Complementary Goods :** The law does not apply to complementary goods. It is so because complementary goods are used in a fixed proportion. By using less of one commodity, use of the other cannot be increased.

Lesson : 3

Indifference Curve Approach

(Author: Dr. M.S. Khanchi)

Indifference Curve approach was first propouned by British economist Edgeworth in 1881 in his book ñMathematical Physics.ò The concept was further developed in 1906 by Italian economist Pareto, in 1913 by British economist W .E. Johnson, and in 1915 by Russina economist Stutsky. The credit of rendering this analysis as an important tool of theory of Demand goes to Hicks and Allen. In 1934, they presented it in a scientific form in their article titled ñA Reconsideration of the Theory of Value.ò It was discussed in detail by Hicks in his book, ñValue and Capitalò.

An indifference curve is a geometrical presentation of a consumer is scale of preferences. It represents all those combinations of two goods which will provide equal satisfaction to a consumer. A consumer is indifferent towards the different combinations located on such a curve. Since each combination located on such a curve. Since each combination yields the same level of satisfaction, the total satisfaction derived from any of these combinations remains constant.

An indifference curve is a locus of all such points which shows different combinations of two commodities which yield equal satisfaction to the consumer. Since the combination represented by each point on the indifference curve yields equal satisfaction, a consumer becomes indifferent about their choice. In other words, he gives equal importance to all the combinations on a given indifference curve.

According to ferguson, ñAn indifference curve is a combination of goods, each of which yield the same level of total utility to which the consumer is indifferent.ò

According to leftwitch, ñA single indifference curve shows the different combinations of X and y that yield equal satisfaction to the consumer.ò

Indifference Schedule :

An indifference schedule refers to a schedule that indicates different combinations of two commodities which yield equal satisfaction. A consumer, therefore, gives equal importance to each of the combinations:

Supposing a consumer two goods, namely apples and oranges. The following indifference schedule indicates different combinations of apples and oranges that yield him equal satisfaction.

Table No. 1 Indifference Schedule

Combination of Apples and Oranges	Apple	Oranges
A	**1**	**0**
B	**2**	**7**
C	**3**	**5**
D	**4**	**4**

The above schedule shows that the consumer get equal satisfaction from all the four combinations, namely A, B, C and D of apples and oranges. In combination A the consumer has I apple + 10 oranges, in combination B he has 2 apples +7 oranges, in combination C he has 3 apples +5 oranges, and in Combination D he has 4 apples + 4 oranges. In order to have one more apple the consumer sacrifice, some of the oranges in such a way that there is no change in the level of his satisfaction out of, each combination.

Indifference Curve :

Indifference curve is a diagrammatic representation of indifference schedule. The indifference curve shown in figure 1 is based on Table No.1

In this diagram, quantity of apples is shown on ox-axis and that of oranges on oy-axis. IC is an indifference curve. Different points A,B,C, and D on it indicate those combinations of apples and oranges which yield equal satisfaction to the consumer.

Law of Diminishing Marginal Rate of Substitution :

The concept of indifference curve analysis is based on law of diminishing marginal rate of substitution. The law was discussed by Lei-iier, Hicks and Allen. To understand the law, it is essential to know marginal rate of substitution.

The study of indifference curve shows that when a consumer gets one more unit of X-commodity his satisfaction increases. If the consumer wants that his level of satisfaction may remain the same, that is, if he wants to remain on the same indifference curve, he will have to give up some units of y-commodity. In other words, in exchange for the satisfaction obtained from the additional unit of apple, he will have to give up that many units of changes whose satisfaction is equal to the additional satisfaction obtained from an additional apple.

Utility gained of apples = Utility lost of oranges.

According to Prof. Bilas, ñThe marginal rate of substitution of X for Y (MRSxy) is 4 @ defined as the amount of y which the consumer is just willing to give up to get one more unit of x and maintain the same level of satisfaction.ò

Explanation of the law of Diminishine Marginal Rate of Substitution :

According to this law, as a consumer gets more and more units of X, he will be wining to-give up less and less units of Y. In other words, the marginal rate of substitution of x for y will go on diminishing while the level of satisfaction of the consumer remains the same.

The law can be explained with the help of Table No.2 and Figure 2 below :

Table No. 2

Marginal Rate of Substitution

Combination	Apples (x)	Oranges (y)	MRSxy
A	**1**	**10**	**-**
B	**2**	**7**	**3:1**
C	**3**	**5**	**2:1**
D	**4**	**4**	**1:1**

Table No.2 indicates that the consumer will give up 3 oranges for getting the second apple, 2 oranges for getting the third apple and 3 orange for getting the fourth apple. In other words, marginal rate of substitution of apples for oranges goes on diminishing.

It is clear the diagram that when consumer moves from point A to point B, he give up 3 oranges to obtain one additional apple. In this situation, consumerôs marginal rate of substitution of apple for orange is 3: 1. When he moves from B to C, he gives up only 2 oranges to get one additional apple. The marginal rate of substitution of apple for orange now diminishes to 2 : 1. It is evident from this example that as the consumer increases the consumption of apples, for getting every additional unit of apple he gives up less and It less units of oranges, that is, 3: 1, 2: 1, 1: I respectively. It is called diminishing marginal rate of substitution and the law relating it is called law of diminishing marginal rate of substitution.

Assumptions :

Indifference curve approach has the following main assumptions:

1. **Rational Consumer :** It is assumed that the consumer will behave rationally. It means the consumer would like to get maximum satisfaction out of his total income.

2. **Diminishing Marginal rate of Substitution :** It means as the stock of a commodity increases with the consumer, he substitutes it for the other commodity at a diminishing rate.

3. **Ordinal Utility :** A consumer can determine his preferences on the basis of satisfaction derived from different goods or their combinations. Utility can be expressed in terms of ordinal numbers, i.e., first, second etc.

4. **Independent Scale of Preference :** It means if the income of the consumer changes or prices of goods fall or rise in the market, these changes will have no effect on the scale of preference of the consumer. It is further assumed that scale of preference of a consumer is not influenced by the scale of preference of another consumer.

5. **Non-Satiety :** A consumer does not possess any good in more than the required quantity. He does not reach the level of satiety. Consumer prefers more quantity of a good to less quantity.

6. **Consistency in Selection :** There is a consistency in consumerôs behaviour. It means that if at any given time a consumer prefers A combination of goods to B combination, then at another time he will not prefer B combination to A combination.

   A>B = B|>A

   It means if A is greater than (>) B, B cannot be greater than (>) A.

7. **Transitivity :** It means if a consumer prefers A combination to B combination, and B Combination to C Combination, he will definitely prefer A combination to C combination. Likewise; if a consumer is indifferent towards A and B and he is also indifferent towards Band C, then he will also he indifferent towards A and C.

Properties of Indifference Curves :

1. Indifference curve slopes downward from left to right, or an indifference curve has a negative slope: the downward slope of an indifference curve indicates that a consumer will have to curtail the consumption of one commodity if he wants to consume large quantity of another commodity to maintain the same level of satisfaction. If an indifference curve does not slope downwards it can either be a vertical line or horizontal line or an upward sloping curve. Consider the following shapes of an indifference curve.

In the diagram quantity of apples is shown on ox-axis and quantity of oranges of oy-axis. Let us suppose, indifference curve is a vertical line MB. Combination A on this curve represents more units of oranges with the same units of apples as compared with combination C. Consequently, A combination yields more satisfaction than C combination. So an indifference curve cannot be vertical or parallel to oy-axis.

If indifference curve is a horizontal line then H combination, will yield more satisfaction than C combination, because in H combination there are more units of apples than in C combination Consequently, an indifference curve cannot be a horizontal line or parallel to ox-axis.

If indifference curve is upward sloping like IJ, the consumer will get more satisfaction from combination A than B and C. Consequently, an indifference curve cannot be upward sloping.

If indifference curve is downward sloping, the consumer will get equal Satisfaction from A as well as B combinations, because in case of combination A if quantity of oranges is more than in combination B, then the quantity of apples is less than in combination B. Consequently, the slope of indifference curve will be downward sloping.

2. Indifference curve is convex to the point of origin: An indifference curve will ordinarily he convex to the point of origin. This property is based on the law of diminishing marginal rate of substitution.

If an indifference is not convex to the point of origin 0, it can either be a straight line or concave. But it can be proved with the help of diagram that on the basis of the assumption of the law of diminishing marginal rate of substitution both these situations are not possible.

In the first indifference curve is a downward sloping straight line. It signifies that marginal rate of substitution of apples for oranges remains constant, as shown by AB = CD = EF. Such an indifference curve can be possible only in case of perfect substitutes.

If indifference curve is concave to the point of origin, it signifies that marginal rate of substitution of apples for oranges is increasing. It would mean that as the quantity of apples is increasing, its importance is also increasing, which it does not happen in real life.

If indifference curve is convex to the point of origin 0, it signifies that marginal rate of substitution of apples for oranges is diminishing. It means as the consumer gets more and more apples he parts with less and less units of oranges. This situation conforms to real life. Consequently, indifference curve is convex to the point of origin.

3. Two Indifference Curves never cut each other: Each indifference curve represents different levels of satisfaction, so their intersection is ruled out.

In this diagram two indifference curves IC_1 and IC_2 have been shown intersecting each other at point A, but it is not possible points A and C on indifference curve IC_1 represent combinations yielding equal satisfaction, that is, A = C Likewise points A and B on indifference curve IC_2 represent combinations yielding equal satisfaction, that is, A = B. It implies that satisfaction from B combinations equal to satisfaction from C combination, but it is not possible because in B combination quantity of oranges is more than in C combination, although quantity of apples in both combinations is equal.

4. Higher Indifference Curves represent more satisfaction

In this diagram IC_2 is higher than IC_1. Point B on IC_2 represents more units of apples and oranges than point A on IC_1 curve. Hence point B on IC_2 will give more satisfaction than point A on IC_1. It is evident, therefore, that higher the indifference curve, greater the satisfaction it represents.

5. Indifference Curve touches neither x-axis nor y-axis;

In case an indifference curve touches either axis it means that the consumer wants only one commodity and his demand for the second commodity is zero. An indifference curve may touch oy-axis if it represents

money instead of a commodity. In this diagram IC touches oy-axis at point M. It means the consumer has in his possession OM quantity of money and does not want any unit of apples. At point N consumer likes to have a combination of OQ units of apples and OP units of money. This combination will yield him same satisfaction as by keeping OM units of money.

6. Indifference curves need not be parallel to each other:

Indifference curves mayor may not be parallel to each other. It all depends on the marginal rate of substitution on two curves shown in the indifference map. If marginal rate of substitution of different points on two curves diminishes at constant rate, then these curves will be parallel to each other, otherwise they will not be parallel.

7. Indifference curves become complex in case of more than two commodities: When a consumer desires to have combinations of more than two commodities, say, three commodities, we will have to draw three dimensional indifference curves which are quite complex. If the consumer wants a combination comprising of more than three goods, such a combination cannot be expressed in the form of a diagram. In that case, we will have to take the help of algebra.

Some Exceptional Shapes of Indifference Curves :

Some exceptional shapes of indifference curves are as follows:

1. Straight Line Indifference Curve :

If two goods are perfect substitutes of each other then their indifference curve may be a straight line with negative slope. It is so because the marginal rate of substitution of such goods remains constant. Supposing, Brook Bond and Lipton tea are perfect substitutes of each other. If in place of 1 kg. of Brook Bond tea the consumer buys 1 Kg. of Lipton tea his total satisfaction remains unchanged. As such, indifference curve for such kind of goods will not be convex to the origin, rather it will be a straight line. Marginal rate of substitution (MRS) of such good goods is always equal to one.

2. Right angled Indifference Curves: Marginal rate of substitution (MRS) of perfectly complementary goods is zero.

For example, a consumer will buy right and left shoes in a fixed ratio as shown in diagram. It is clear that IC_1 and IC_2 are right angel curves, meaning thereby that if the consumer buys one piece of each of right and left-shoes, he will be on point A of IC_1. In case he buys 2 pieces of left shoe and only 1 piece of right shoe, he will be at C of the same IC1. It means, his satisfaction will remain the same. But if he also buys one more piece of right-shoe, his satisfaction will definitely increase and he will move to point B of higher indifference curve IC_2. Thus, perfectly complementary goods have indifference curves of the shape of right angle. Marginal rate of substitution in the case of such goods is zero (MRSxy = O).

Price Line or Budget Line :

Study of price line is essential to have the knowledge of consumer equilibrium through indifference curve analysis. It is also known as Budge line, consumption possibility line, or line of attainable combinations.

A price line represents all possible combinations of two goods, that consumer can purchase with his given income at the given prices of two goods.

Explanation :

Supposing a consumer has an income of Rs.4.00 to be spent on apples and oranges. Price of orange is Re. 0.50 per orange and that of apples Re. 1.0 per apple. With his given income and given prices of apples and oranges, the different combinations that a consumer can get of these two goods are show in Table and Figure below:

Table No. 3

Income (Rs.)	Apples (Re. 1.00)	Oranges (Re. 0.50)
4	0	8
4	1	6
4	2	4
4	3	2
4	4	0

It is clear from the table that if the consumer wants to buy oranges only then he can get a maximum 8 oranges with his entire income of Rupees four. On the other hand, if the consumer wants to buy apples only, then he can get a maximum 4 apples with his entire income of Rupees four. Within these two extreme limits, the other possible combinations that a consumer can get are 1 apple +6 oranges, 2 apples + 4 oranges, 3 apples +2 oranges.

In this diagram different combinations of two goods have been shown by AB Line. It is called Price Line. It is presumed that the consumer spends his entire income on the consumption of these two goods, so AB price line is the

limit line of the consumer. Slope of the price line refers to the price ratio of two goods, apples and oranges, that is,

Slope of price Line =

$$\frac{Px}{Py}$$

(Here Px = price of apples and Py = price of oranges)

Shifting of the Price Line :

Position and slope of the price line depends upon two factors : (1) Income of the consumer and (2) price of the two goods that the consumer wants. Price line may change due to these two reasons :

1. Due to change in Income: If prices of the two goods remain unchanged, then with an increase in income, the price line will shift to the right, and with a decrease in income it will shift to the left its slope remaining unchanged.

Figure No. 12 indicates that when income of the consumer was Rs.4.00 he could buy those combinations of apples and oranges as were represented by price line AB. With increase in income, price line shifts to the right as shown by the line CD. Likewise, if income decreases, price line will shift to the left, as shown by EF line, its slope remaining the same.

2. Due to change in price of one commodity

If income of the consumer and price of one commodity remains unchanged, but the price of other commodity changes, the slope of price line will also undergo a change. One end of the price line will remain at its place, but the other end touching the axis of that commodity whose price has changed to will shift forward from its original place if the price has fallen or shift backward if the price has risen. It is clear from figure that when the price of apple falls, slope of the price line will change from AB to AC.

Consumer's Equilibrium - Indifference Curve Analysis

According to the ordinal approach, a consumer has a given scale of preference for different combinations of two goods. By just comparing the levels of satisfaction, he can derive maximum satisfaction out of a given money income.

Consumerôs equilibrium refers to a situation in which a consumer with given income and given prices purchases such a combination of goods and services as gives him maximum satisfaction and he is not willing to make any change in it.

Assumptions:

1. Consumer is rational and so maximises his satisfaction from the purchase of two goods.
2. Consumerôs income is constant.
3. Prices of the goods are constant.
4. Consumer knows the price of all things.
5. Consumer can spend his income in small quantities.
6. Goods are divisible.
7. There is perfect competition in the market.
8. Consumer is fully aware of the indifference map.

Conditions of Consumer's Equilibrium :

There are two main conditions of consumerôs equilibrium;

(i) Price line should be tangent to the indifference curve, i.e. MRSxy = Px / Py

(ii) Indifference curve should be convex to the point of origin.

(iii) Price line should be tangent to indifference curve:

In this diagram AB is the price line. IC_1, IC_2 and IC_3 are indifference curves. A consumer can buy those combinations which are not only on price line AB but also coincide with the highest indifference curve which is IC_2 in this case. The consumer will be in equilibrium at combination D (2 apples + 4 oranges) because at this point price line AB is tangent to the indifference curve IC_2. At equilibrium point D, slope of indifference curve and price line coincide. Slope of indifference curve is indicative of marginal rate of substitution of good x for good y (MRSxy) and slope of price line is indicative of the ratio of price of good x (Px) and price of good y (Py). In case of equilibrium:

Slope of indifference curve = slope of price line

$$\frac{Px}{Py}$$

MRSxy = Py

(ii) Indifference Curve must be convex to the origin: It means that marginal rate of substitution of good x for good y should be diminishing. If at the point *of* equilibrium, indifference curve is concave and not convex to the origin, it will not be a position of permanent equilibrium.

In this diagram AB is the price line. IC is the indifference curve. At point E, price line AB is tangent to indifference curve, but point E is not a permanent equilibrium point because at this point, marginal rate of substitution is increasing instead of diminishing. At point E, indifference curve is concave to its point of origin 0 and so it is a violation of second condition of equilibrium. Hence the consumer is in equilibrium at point E IC curve. At point E, Price line AB is tangent to IC, curve, which is convex to the point of origin.

Income, Substitution and Price Effect :

Consumerôs equilibrium is affected by change in his income, change in the price of substitutes, and change in the price of good consumed. These changes are known as (1) Income effect (2) Substitution effect and (3) Price effect, respectively.

1. Income Effect

The income effect is the effect on the consumption of two goods caused by change in income, if prices of goods remain constant.

The income effect may be defined as the effect on the purchases of the consumer caused by change in income, if price remains constant. Income effect indicates that, other things being equal, increase in income increases the satisfaction of the consumer. As a result, equilibrium point shifts upward to the right. On the contrary, decrease in income decrease the satisfaction of the consumer and his equilibrium point shifts downwards to the left.

In this diagram consumerôs initial equilibrium is at point E on price line AB. When his income increases, his equilibrium point shifts to the right i.e. E, on price line C-D. With decrease in his income, his equilibrium point shifts to the left i.e. E 2 on price line EF. Locus of all these equilibrium points is called income consumption curve. It starts from the point of origin 0 meaning thereby

that when the income of the consumer is zero, his consumption of apples and oranges will also be zero.

Income Consumption Curve:

As shown in figure No. 16, effect of change in income is reflected in Income Consumption Curve (ICC). This curve is a locus of tangency points of price lines and indifference curves.

Income consumption curve refers to the effect of change in income on the equilibrium of the consumer.

Slope of income consumption curve is positive in case of normal goods, but it is negative in case of inferior goods.

(i) Positive Slope

Income consumption curve is positive in case of normal goods. In other words, consumption of both normal goods (x and y) increases with increase in income. As shown in the diagram, income consumption curve (ICC) of normal goods slopes upwards from left to right signifying that more of both the goods will be bought when income increases. ICC curve indicates that expenditure on both the goods will increase in almost the same ratio. ICC, curve indicates a higher proportionate increase in the expenditure on good - x, and ICC2 curve indicates a higher proportionate increase in expenditure on good - y.

(ii) Negative Slope:

Income effect of inferior goods is negative. It means inferior goods are brought in less quantity when income of the consumer increases.

Suppose x-good is inferior and y-good is normal. Price line AB, drawn on the basis of given income of the consumer and given prices of the two commodities, touches indifference curve IC_1 at point E which is the point of consumerôs equilibrium. As the income of the consumer goes on increasing, price line goes on shifting to the right as CD and GH touching IC_2 and IC_3 at points E_1 and E_2, respectively. Consequently, the quantity of good x falls from OM to OM_1 and OM_2. In this way, increase in the income of the consumer is followed by decrease in the quantity demanded for inferior good-x by MMI and M_1M_2 respectively. This decline in quantity demanded reflects negative income effect. By joining together different equilibrium points E, E_1 and E_2 one gets income consumption curve which slopes backward to the left. It indicates negative income effect.

2. Substitution Effect :

If with the change in the prices of goods the money income of the consumer changes in such a way that his real income remains constant, the consumer will substitute cheaper good for the dearer ones. Consequently, it will effect the quantity purchased of both the goods. Iô his effect is known as substitution effect.

Substitution effect shows the change in the quantity of the goods purchased due to change in the relative prices alone while real income remains constant.

Supposing the income of the consumer is Rs.4.00 which he spends on the purchase of oranges and apples. Price of oranges is 50 paise per orange and that of apples Re. 1.00 per apple. With this income he buys 4 oranges and 2 apples and finds himself in an equilibrium.

In this, diagram AB is the price line and IC_1 is the original indifference curve. Consumer is in equilibrium at point E. He is getting ON units of oranges and OM units of apples. Supposing apples become cheaper. Consequently, AB Price line will shift towards the right on ox-axis as AC and be tangent to higher indifference curve IC_2 at Point D which will be the new equilibrium point of the consumer. Now his real income will be more than before. If the real income of

the consumer should remain the same as before, we will have to take away some of his money income. Now his price line will be GH which will be parallel to price line AC. The new price line GH is tangent to indifference curve IC_1 at point F which will be the new point of equilibrium. He will substitute MQ apples for NP oranges. In this way, consumerôs marginal rate of substitution of apples for oranges will be MQ / NP. This substitution of relatively cheaper good for dearer ones is called substitution effect. Thus movement from equilibrium point E to equilibrium point F on the same indifference curve IC_1 indicates the substitution effect.

3. Price - Effect :

Price effect means change in the consumption of goods .when the price of either of the two goods changes, while the price of the other good and the income of the consumer remain constant.

Supposing IC is the original indifference curve and AB the original price line and consumer is in equilibrium at point E. As the price of apple falls, new price line will be AD which touches higher indifference curve IC_1 ant point E_1. It means fall in price of any good will increase the satisfaction of the consumer.

On the contrary, if the price of apples rises, the new price line will be AN which will touch the lower indifference curve at point E_2, the new equilibrium point. It means arise in price will reduce the level of satisfaction of the consumer.

By joining together different equilibrium points E_2, E, E_1, one gets the price consumption curve (PCC). The price consumption curve for commodity X is the locus of points of consumerôs equilibrium when the price of only X varies, the price of Y and income of the consumer remaining constant.

Price Effect is the Suni of Substitution Effect and Income Effect :

When the price of a commodity changes, it has two effects: (i) There is change in the real income of the consumer leading to change in his consumption. It is called income effect; (ii) Secondly, due to change in relative prices, the consumer substitutes relatively cheaper goods for the dearer ones. It is called substitution effect. The combination of this income and substitution effect is called price effect. Thus, Price Effect = Income Effect + Substitution Effect.

Supposing AB is the original price line and IC the original indifference curve. Consumer is in equilibrium at point E. When the price of apple falls, the new price line shift from AB to AC. The new price line touches higher indifference curve IC at point E, which is the new equilibrium point. Movement from E_1 signifies the Price Effect.

Fall in price of apples means increase in the real income of the consumer. If the monetary income of the consumer is reduced to such an extent that the real income remains the same as before, in that case the new price line will be PH and new equilibrium point E 2ô The movement from E toE_2 reflects the Substitution Effect. If due to fall in price of apples, the money income of the consumer is not reduced, the consumer will move from equilibrium point E_2 to E_1. Thus movement from E_2 to E_1, shows the Income Effect.

Due to fall in the price of apples a consumer buys more of apples, it is called price effect. Consumer buys MT units of apples. Of these, he buys MN units on account of substitution effect and NT units on account of income effect. It means, with regard to demand for apples:

Price Effect = OM to OT i.e. MT

Substitution Effect = OM to ON i.e. MN

Income Effect = MN to MT i.e. NT

Thus MT = MN + NT

Price Effect = Substitution Effect + Income Effect

Critism :

Robertson, Armstrong, Knight etc. have criticised indifference curve analysis on account of the following.

1. **Unrealistic assumption :** Indifference curve analysis is based on the assumption that a consumer has complete knowledge regarding the preference of two goods. In reality, he cannot take quick decisions in real life in respect of different combinations.

2. **Complex analysis :** Indifference curve analysis can explain easily that behaviour of the consumer which is restricted to the combination of only two goods. If the consumer wants combinations of more than two goods, then indifference curve analysis becomes highly complex.

3. **Imaginary :** Indifference curve analysis is based on imaginary combinations. A consumer does not decide always like a computer as to which of the combinations of two goods he would prefer.

4. **Assumption of Convexity :** This theory does not explain why an indifference curve is convex to the point of origin. In real life, it is not necessary that all goods should have diminishing marginal rate of substitution.

5. **Unrealistic combinations :** When we consider different Combinations of two goods, sometimes we come across such funny combinations that have no meaning for the consumer. For instance, there is a combination of 10 shirts + 2 pairs of shoes. If in the subsequent combinations shirts are given up to get more pairs of shoes then we way arrive at a combination representing 2 shirts + 10 pairs of shoes, which is rediculous.

6. **Impractical :** Indifference curve analysis is based on the unrealistic assumption that goods are homogenous. ôThis assumption holds good only under perfect competition, which is more 9 theoretical concept. In real life, monopolistic and oligopolistic conditions are found more prevelents.

However, compared to utility analysis, indifference curve analysis is an improved technique of consumerôs behaviour.

Revealed Preference Theory

Revealed Preference Theory was put forward by Somvelson in 1938 and it is based on the actual market behaviour of consumer. Accordingly, it is a behaviouristic approach. On the other hand, utility and Indifference Curve Approaches are the psychological and interospective approaches.

Assumptions :

Revealed Preference Theory involves the following assumptions :

1. There is no change in the taste of the consumers.
2. The choice of the consumer for a particular combination reveals his preference.
3. The consumer chooses only one combination on a given price income line.
4. Assumption of consistency in the consumerôs behaviour. That is if combination B canôt be preferred to A is another situation.

   A > B, then B |> A
5. Assumption of transitivity in consumerôs behaviour: Transitivity means that if combination A is preferred to B and, B>C, then A must be preferred to C. This assumption is necessary for Revealed Preference Theory.
6. Assumption of the concept of ordinal utility, i.e., Revealed preference Theory regards utility to be merely comparable.

Theme of Revealed Preference Theory :

The basic essence or theme of Revealed Preference Theorem is shown below in the diagram.

In this diagram, AB is the price line which represents given Price Income situation. The consumer can buy or choose any combination on this price line and below this price line within OAB. Suppose the consumer chooses combination Q. It means consumer has revealed his preference for Combination Q to all other combinations on AB price line and below it. It means combination is revealed preferred to all other combinations on and below AB price line, such as E, G, D, and F. In other words, all combinations with in OAB are revealed inferior to combination Q. In this situation consumer is purchasing ox of x and oy of y. However, a combination within right angle KQT such as Z is artainly superior to Q.

Implications Significance of Revealed

Preference Theory :

Revealed Preference Theory is significant because :

1. It can be used to explain convexity of Indifference Curve.
2. it can be used to explain Demand Theory or fundamental theorem of consumption or Slutsky Theorem.

I. Revealed Preference Theory can be used to derive the convexity of indifference curve.

This proved as under :

In this diagram under given price Income situation, original price line is AB on which consumer selects a combination Q. Q is revealed preferred to all other combinations on AB line and within OAB. A point like 2 within KQT is superior to Q. We have to prove the convexity of indifference curve through point Q. The indifference curve cannot enter the area OAB as this area is

inferior to Q. The indifference curve cannot enter the area KQT because it is superior to Q. The area between QT and QB and between QK and QA is the zone of ignorance. We can reduce the zone of ignorance with reference to, new line such as A, B and on this line suppose Q is selected. It means Q, is revealed preferred to all other combinations on A_1, B_1 and below it. But since Q_1 is inferior to Q, Q_1, BB_1 is also inferior to Q. This area is out off from the zone of ignorance. Similarly, we can repeat the same process on the upper side. The superior area can also be discovered from the zone of ignorance by drawing another line such as yy which passes through Q on this line. Combination G is superior to Q as G was not available when Q was selected. In the same way we can draw more such points and we can draw an offer curve on which we have combination which are superior to Q. It proves that indifference curve throng Q cannot enter the superior area and inferior area so it must b convex to the origin.

2. Revealed Preference Theory can be used to explain Demand Theory of Fundamental Theorm of consumption theory: According to Marshallôs Law of demand, there is inverse relationship between the price of commodity and quantity demanded. Samuelson in his revealed preference theory establishes inverse relationship between price and quantity demand by assuming positive income elasticity of demand. Samuelson called this demand theorm as a fundamental theorm of consumption theory.

The fundamental theorm or simple demand theorm is based upon the positive income elasticity of demand. There is inverse relationship between price and quantity demanded. This can be proved asunder both when price decreases and price increases.

Inverse price and quantity demanded relationship in the case of price decrease is as shown below :

In Fig. 24, AB price line represents original price income situation. Suppose consumer reveals preference for combination Q over all other

combinations on AB line and within GAB. Its price of x falls, the new price line shifts to AB. Let us take away the increased real income from the consumers so that he is able to buy the old combination Q. Hence a cost difference line CC_1 is drawn parallel to AB, and through point Q. He cannot choose any combination like E on QC portion of the price line. This is because E was rejected when Q was selected. Thus, if he buys a combination such as E_1 it means quantity demanded of X increases when price of X falls. If the money taken away from him is returned, he will definitely buy more of X when price of X falls provided income elasticity of demand for X is positive.

Demand theorm can also be explained in the case of increase in price. It is shown below.

In Fig 25, AB is the original price line under given price income situation. Suppose the consumer prefers combination Q on AB price line. When the price of X increases, the new price line is AB. New Combination Q is not available and suppose we compensate the consumer by giving him extra money equal to decrease in real income so that he can buy the same combination Q. Hence a line CC, is drawn through Q on this new price line. Now the consumer will choose either Q or E, but he cannot choose T as it lies within OAB. T was rejected when O was selected. The choice of E is consistent because it was not available when Q was selected. If E is selected, it means quantity of x would be purchased at higher price. If money income which is given as a compensation when P rises is not paid to the consumer, the consumer will have to buy smaller quantity of X when price of x increases.

Thus Revealed Preference Theory proves if income elasticity of demand is positive price and quantity demanded will be inversely related, i.e., demand curve will be downward sloping.

An Evaluation :

Somuelsonôs Revealed Preference Theory has been regarded as superior to marshallian utility theory and Hicks Ordinal analysis. This is owing to the following :

(1) Revealed Preference Theory is superior because it is behaviouristic whereas utility analysis and indifference curve analysis were psychological and introspective by nature. Somuelsonôs theory is based on actual market behaviour of a consumer.

(2) Revealed Preference Theory avoids the continuity assumption. Which is found in both the utility approach and the indifference curve analysis.

(3) Somuelsonôs theory is superior because it avoids the assumption of maximisation of satisfaction.

(4) Revealed Preference Theory is also useful in the combination of index numbers.

(5) Revealed Preference Theory provides the basis for welfare economics.

Inspite of these merits Revealed Preference Theory has its own weaknesses.

(1) Revealed Preference Theory neglects indifference in consumer behaviour.

(2) Samuelson theory is conditional in the sense that under a given price income situation consumer chooses, something of both the things, but it is very rare that he purchases something, of everything.

(3) The assumption that choice reveals preference is not always valid. Infact choice does not often reveal preference because choice requires rational behaviour and consumer behaviour is not always rational.

(4) Revealed Preference Theory is applicable only where the demand curve is downward sloping. But this approach does not help derive the demand curve.

(5) Revealed Preference Theory is not adequate to explain the demand Theorem.

Lesson - 4

LAW OF DEMAND AND ELASTICITY OF DEMAND

(Author : Bhag Singh Bodla)

4.1 INTRODUCTION

It is essential for the business managers to have a clear understanding of the following aspects of the demand for their products :

i) What are the sources of demand?

ii) What are the determinants of demand?

iii) How do the buyers decide the quantity of a product to be purchased?

iv) How do the buyers respond to the change in a product prices, their income and prices of the related goods?

v) How can the total of market demand for a product for a product be assessed and forecast?

These questions are answered by the **Theory of Demand.** In this and the following lessons we will discuss the theory of individual and market demand.

4.2 MEANING OF DEMAND

The term ódemandô robers to a ódesireô for a commodity backed by ability and willingness to pay for it. Unless a person has an adequate purchasing power or resources and the preparedness to spend his resources, his desire for a commodity would not be considered as his demand. For example, if a man wants to buy a car but he does not have sufficient money to pay for, his want is not his demand for the car. A want with **three attributes - desire to buy, willingness to pay and ability to pay -** becomes effective demand. Only an effective demand figures in economic analysis and business decisions.

The term ódemandô for a commodity (i.e., quantity demanded) has always a reference to óa priceô, óa period of timeô and óa placeô. Any statement

regarding the demand for a commodity without reference to its price, time of purchase and place is meaningless and is of no practical use. For instance, to say ódemand for TV sets is 50,000' carries no meaning for a business decision, nor it any use in any kind of economic analysis. A

4.3 INDIVIDUAL DEMAND FORA COMMODITY

The theory of consumerôs equilibrium provides a convenient basis for the derivation of individual demand curve for a commodity. Marshall was the first economist to explicitly derive the demand curve from consumerôs utility function. Marshall gave the equilibrium condition for the consumption of a commodity, say X, as $MU_x = P_2 (MU_m)$.

The derivation of individual demand for commodity is illustrated in Fig. 4.1 (a) and 3.3 (b). Suppose that the consumer is in equilibrium at point E_1, where given the price of X, $MU_x = P_3 (MU_m)$. Here equilibrium quantity is OQ_1. Now if price of the commodity falls to P_2, the equilibrium condition will be disturbed making $MU_x > P_3 (MU_m)$. Since MU_m is constant, the only way to restore the equilibrium condition is to reduce MU_x, by buying more of commodity X. Thus, by consuming $Q_1 Q_2$ additional units of X he reduces his MU_x to $E_2 Q_2$ and reaches a new equilibrium position at point E_2 where $MU_x = P_2 (MU_m)$. Similarly, if price falls further, he buys and consume more to maximise his satisfaction.

Fig 4.1 (a) reveals that when price is P_3, equilibrium eqantity is OQ_1. When price decreases to P_2 equilibrium point shifts downward to point E_2 where equilibrium quantity is OQ_2 Similarly, when price decreases to P_1 and P

(MU_m) line shifts downward, the equilibrium point shifts to E_1 where equilibrium quantity is OQ. Note that $P_3 > P_2 > P_1$ and the corresponding quantities $OQ_1, < OQ_2 < OQ_3$. It means that as price decreases, the equilibrium

quantity increases. This inverse price-quantity relationship is the basis of the law of demand, explained below.

The inverse price and quantity relationship is shown in part *(b)* of Fig. 3.3. The price quantity combination corresponding to equilibrium point E_3 is shown at point. *J.* Similarly, the price-quantity combinations corresponding to equilibrium points, E_2 and E_1 are shown at points *K* and *L,* respectively. By joining points *J, K* and *L,* we get individualôs *demand curve* for commodity *X* The demand curve *Dx* in the usual downward sloping Marshallian demand curve.

Demand under Variable Mu_m

We have explained above the consumerôs equilibrium and derived his demand curve under the assumption that MU_m remains constant. This analysis holds even if MU_m is assumed to be variable. This can be explained as follows.

Suppose MU_m is variable - it decreases with increase in stock of money and vice varsa. Under this condition, if price of commodity fall and the consumer buys only as many units as he did before the fall in price, he saves some money on this commodity. As a result his stock of money increases and his MU_m decreases, whereas MU_m remains unchanged because his stock of commodity remains unchanged. As a result, his MU_m exceeds his MU_m. When a consumer exchanges money for commodity, his stock of money decreases and stock of commodity increases. As a result, MU_m increases MU_c decreases. The consumer therefore exchans money for commodity until $MU_c = MU_m$. Consequently, demand for a commodity increases when its price fall.

4.4 THE LAW OF DEMAND

The law of demand states that **the demand for a commodity increases when its price decreases and it falls when its price rises, other things remaining constant**. This is an empirical law, i.e., this law is based on observed facts and can be verified with new empirical data. As the law reveals, there is an *inverse relationship* between the price and quantity demanded. The law holds under the condition that ñother things remain constantò. ñOther

thingsò include other determinants of demand, viz., consumersô income, price of the substitutes and complements, taste and preferences of the consumer, etc. These factors remain constant only in the short run. In the long run they tend to change. The law of demand, therefore, hold only in the short run.

Demand Schedule

The law of demand can be presented through a **demand schedule.** *Demand Schedule* is a series of prices placed in descending (or ascending) order and the corresponding quantities which consumers would like to buy per unit of time. Based on the logic of demand curve in Fig. 4.1 (b), a hypothetical demand schedule for a commodity, tea, is given in Table 4.1

Table 4.1 Demand Schedule for Tea

Price **per cup of tea (Rs.) consumer per day**	**No. of cups of tea demand by a combination**	**Points representing Price-quantity**
7	**1**	***i***
6	**2**	***j***
5	**3**	***k***
4	**4**	***I***
3	**5**	***m***
2	**6**	***n***
1	**7**	***0***

4.1 presents seven alternative prices of tea and the corresponding quantities (number of cups of tea) demanded per day. At each price, a unique quantity is demanded. As the table shows, as price of tea per cup decreases, daily demand for tea increases. This relationship between, quantity demanded of a product and its price is the basis of the law of demand.

The Demand Curve

The law of demand can also be presented through a **demand curve. A demand curve is a locus of points showing various alternative price-quantity combinations.** Demand curve shows the quantities of a commodity which a consumer would buy at different prices

Fig. 4.2 Demand Curve for Tea

per unit of time, under the assumptions of the law of demand. By plotting the data given in Table 4.1, we obtain an individual demand curve for tea, as shown in Fig. 4.2. The curve DD is the demand curve. It reads the law of demand. Each point on the demand curve shows one unique price-quantity combination. The combinations read downward along the demand curve show decreasing price of tea and increasing number of cups of tea demanded. Price-quantity combinations read upwards show increasing price of tea per cup and decreasing number of cups of tea per day consumed by an individual. Thus, the demand curve shows a functional relationship between the alternative prices of a commodity and its corresponding quantities which consumer would like to buy during a specific period of time, say day, per week, per month, per season, or per year.

Factors behind the Law of Demand

As Fig. 4.2 shows, demand curve slopes downward to the right. The downward slope of the demand curve depicts the law of demand, *i.e.,* the quantity of a commodity demanded per unit of time increases as its price falls, and *vice verse.* The factors that make the law of demand operate are following.

Substitution Effect

When price of a commodity falls, prices of all other related goods (particularly of substitutes) remaining constant, the goods of latter category become relatively costlier. Or, in other words, the commodity whose price has fallen becomes relatively cheaper. Since utility maximising consumers substitute cheaper goods for costlier ones, demand for the cheaper commodity increases. The increase in demand on account of this factor is known a **substitution effect.**

Income Effect

As a result of fall in the price of a commodity, the real income of the consumer increases. Consequently, his purchasing power increases since he is required to pay less for the same quantity. The increase in real income encourages the consumer to demand more of goods and services. The increase in demand on account of increase in real income is known as income effect. It should however be noted that the income effect is negative in case of inferior goods. In case the price of an inferior goods accounting for a considerable proportion of the total consumption expenditure falls substantially, consumersô real income increases and they become relatively richer: Consequently, they substitute the superior goods for the inferior ones. As a result, the consumption of inferior goods falls. Thus, the income effect on the demand for inferior goods becomes negative.

Utility-Maximising Behavior

The utility-maximising behavior of the consumer under the condition of diminishing marginal utility is also responsible for increase in demand for a commodity when its price falls. As mentioned above, when a person buys a commodity, he exchanges his money income for the commodity in order to maximise his satisfaction. He continues to buy goods and services so long as marginal utility of his money (MU_m) is less than the marginal utility of the commodity (MU_o). Given the price of a commodity, the consumer adjusts his purchases. so that.

$$MU_m = P_o = MU_o$$

When price of the commodity falls, ($MU_m = P_o$) < MU_o, and equilibrium is disturbed. In order to regain his equilibrium, the consumer will have to reduce the MU_o to the level of MU_m. This he can do only by purchasing more of the commodity. Therefore, the consumer purchases the commodity till $MU_m = P_o = MU_o$. This is another reason why demand for a commodity increases when its price decreases.

Exceptions to the Law of Demand

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

(a) ***Expectations regarding further prices.*** When consumers expect a continuous increase in the price of a durable commodity, they buy more of it despite increase in its price with a view to avoiding the pinch of a much higher price in future. For instance, in pre-budget months, prices generally tend to rise. Yet, people buy more of storable goods in anticipation of further rise in prices due to new levies.

(b) ***Status Goods.*** The law does not apply to the commodities which are used as a status symbolô of enhancing social prestige or for displaying wealth and riches, e.g., gold,ô precious stones, rare paintings, antiques, etc. Rich people buy such goods mainly because their prices are high and buy more of them when their prices move up.

(c) ***Giffen Goods.*** Another exception to the law of demand is the classic case of Giffen goods[2]. A Giffen good may be any inferior commodity much cheaper than its superior substitutes, consumed by the poor households as an essential commodity. If price of such goods increases (price of its substitute remaining constant), its demand increases instead of decreasing because, in case of a Giffen good, income effect of a price rise is greater than its, substitution effect. The reason is, when price of, an inferior good increases, income remaining the same, poor people cut the consumption of the superior substitute so that they may buy more of the inferior good in order to meet their basic need.

4.5 SHIFT IN DEMAND CURVE

When demand curve changes its position (retaining its shape though not necessarily), the change is known as shift in demand curve. Consider, for instance, the demand curves viz. D^1, D^2 and D^3 in Fig. 4.3. Let us suppose that demand curve D^2 is the original demand curve for commodity X. As shown in the figure, at price OP_2. Consumer buys OQ_2 units of X, other factors remaining constant. But, if any of the other factor (e.g. consumerôs in come or price of the substitutes) changes, it will change the consumerôs ability and willingness to buy commodity X. For example, if consumerôs disposable income decreases due to increase in income tax, he may be able to buy only OQ_1 units of X instead of OQ_2. The is true for the whole range of prices of X; consumers would be able to buy less at all other prices. This will cause a downward shift in demand curve D_2 to D_1. Similarly, increase in disposable income of the consumer due to, say, reduction in taxes may cause an *upward shift* in D_2 to D_3. Such changes in the location of demand curves are known as *shift in demand curve.*

Reasons for Shift in Demand Curve

Shifts in a price-demand curve may take place owing to the change in one or more determinants of the demand for a commodity. Consider, for example, the decrease in demand for commodity X by Q_1 Q_2 in Fig. 4.3. Given the price OP_2, the demand for x might have fallen from OQ_2 to OQ_1 (i.e., by Q_1Q_2) for any of the following reasons.

(i) Fall in consumerôs income so that he can buy only OQ_1 of X at price OP_2; it is income effect;

(ii) Price of Xôs substitute falls so that the consumers find it worthwhile to substitute Q_1 $Q_2$of X with its substitute; it is substitution effect;

(iii) Advertisement made by the producer of the substitute, changes consumerôs taste or preference against commodity X so much that they replace Q_1 Q_2 of it with its substitute, again a substitution effect ;

(iv) Price of complement of X has increased so much that the consumer can now afford only OQ_1 of X; and

(v) Price remaining the same, demand for X might also decrease for such reasons as X going out it of fashion, deterioration in its quality, change in consumerôs technology and seasonality of the product.

4.6 DETERMINANTS OF MARKET DEMAND

(1) Price of the Product

The price of product is one of the most important determinants of its demand in the long run, and the only determinant in the short run. The price and quantity demand are inversely related. The law of demand states that the quantity demanded of a product which its consumers/users would like to buy per unit of time, increases when its price falls, and decreases when its price increases, *other factors remaining constant.* The assumption other factors remaining constantô implies that income of the consumers, prices of the substitutes and complementary goods, consumerôs taste and preference, and number of consumers, remain unchanged. (The law of demand has already been discussed in detail in the previous chapter).

(2) Price of the Related Goods

The demand for a commodity is also affected by the changes in the price of its related goods. Related gods may be *substitutes or complementary goods.*

Substitutes. Two commodities are deemed to be *substitutes* for each other if change in the price of one affects the demand for the other in the same direction. F or instance, commodities X and Y are considered as substitutes for each other if a rise in the price of X increases demand for Y_2 and vice versa. Tea and coffee, hamburgers and hot-dog, alcohol and drugs are some examples of substitutes in case of consumer goods.

By definition, the relation between demand for a product and price of its substitute is of positive nature. When price of the substitute (say, coffee) of a

product (tea) falls (or increases), the demand for the product falls (or increases). The relationship of this nature is given in Fig. 4.4 (a).

Fig. 4.4 Demand for Substitutes and Complements

Complements. A commodity is deemed to be a *complement* for another when it complements the use of the other or when the use of the two goods goes together so that their demand changes (increases or decreases) simultaneously. For example, petrol is a complement to car and scooters, butter and jam to bread, milk and sugar to tea and coffee, mattress to cot, etc. Two goods are termed as complementary to each other if an increase in the price of one causes a decrease in demand for the other. By definition, there is an inverse relation between the demand for a good and the price of its complement. For instance, an increase (or decrease) in the price of petrol causes a decrease (or an increase) in the demand for car and other petrol-run vehicles, other things remaining the same. The nature of relationship between the demand for a product and the price of its complement is given in Fig. 4.4 (b).

(3) Consumer's Income

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

(a) Essential consumer goods (ECG). The goods and services of this category are called ôbasic needsô and are consumed by all persons of a society, e.g., food grains, salt, vegetable oils, matches, cooking fuel, a minimum clothing and housing. Quantity demanded of this category of goods increases

with increase in consumerôs income but only up to certain limit, even though the total expenditure may increase in accordance with the quality of goods consumed, other factors remaining the same. The relationship between goods of this category and consumerôs income is shown by the curve ECG in Fig. 4.5 As the curve shows, consumerôs demand for essential goods increases only until his income rises to OY_2**.** It tends to saturate beyond this level of income.

(b) Inferior goods. Inferior and superior goods are widely, know to both the consumers and the sellers. For instance, every consumer knows that millet is inferior to wheat and rice; bidi (indigenous cigarette) is inferior to cigarette, coarse textiles are inferior to refiI1ed ones, kerosene is inferior to cooking gas; travelling by bus is inferior to travelling by taxi, so on and so forth. In economic sense, however, a commodity is deemed to be inferior if its demand decreases with the increase in consumerôs income. The relation between income and demand for an inferior goods is shown by the curve IG in Fig. 4.5 under the assumption that other determinants of demand remain the same. Demand for such goods rises only up to a certain level of income (say, OY1) and declines as income increases beyond this level.

Fig. 4.5 Income Demand Curves

(c) Normal goods. Technically, normal are those which are demanded in increasing quantities as consumersô income rises. Clothingôs, household furniture and, automobiles are some of the important examples of this category of goods. The nature of relation between income and demand for the goods of this category is shown by the curve NG in Fig. 4.5. As the curve shown, demand for such good increases with the increases in income of the consumer, but at different rates at different levels of income. Demand for normal goods increases rapidly with the increase in the consumerôs income but slows down with further increase in income.

(d) Prestige and luxury goods. Prestige goods are those which are consumed mostly by rich section of the society, e.g., precious stones, antiques, rare, paintings, luxury cars and such other items of show-off. Though it may look controversial, luxury items include jewellery, costly brands of cosmetics, TV sets, refrigerators, electrical gadgets, cad, etc. Demand for such goods arises beyond a certain level of consumerôs income i.e. consumption enters the area of luxury goods. Producers of such items, while assessing the demand for their product, should consider the income changes in the richer section of the society, not only the per capita income (see curve, LG in Fig. 4.5).

(4) Consumer's taste and preference

Consumerôs taste and preference play an important role in determining demand for a product, Taste and preference depend, generally, on the changing lifestyle, social customs, religious values attached to a commodity, habit of the people, the general levels of living of the society, and age and sex of the consumers. Change in these factors changes consumersô taste and preferences. As a result, consumer reduce or give up the consumption of some goods and add new ones to their consumption pattern. For example, following the change in fashion, people switch their consumption pattern from cheaper, old fashioned goods over to costlier ómodô goods, so long as price differentials are commensurate with their preferences.

(5) Advertisement Expenditure

Advertisement costs are incurred with the objective of promoting sale of the product. Advertisement helps in increasing demand for the product in at least four ways: (a) by informing the potential consumers, about the availability of the product; (b) by showing its superiority to the rival product; (c) by influencing consumersô choice against the rival products; and (d) by setting fashions and changing tastes. The impact of such effects shifts the demand upward to the right. IN other words, other factors remaining the same, as expenditure on advertisement increases, volume of sale increases to an extent. The relation between advertisement outlay and sales is shown in Fig. 4.6.

Fig. 4.6 Advertisement and Sale

(6) Consumers' Expectations

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

(7) Consumer-Credit Facility

Availability of credit to the consumers from the sellers, banks, relations and friends or from any other source encourages the consumers to buy more than what they would buy in the absence of credit availability. That is why, the consumers who can borrow more can consume more than those who cannot borrow Credit facility affects mostly the demand for durable goods, particularly those which require bulk payment at the time of purchase. The car-loan facility may be one reason why Delhi has more cars than Calcutta, Chennai and Mumbai.

(8) Population of the Country

The total domestic demand for a product of mass consumption depends also on the size of the population. Given the price, per capita income, taste and preference etc., the larger the population, the larger the demand for a product with an increase (or decrease) in the size of population, employment percentage remaining the same, demand for the product will increase (or decrease).

(9) Distribution of National Income

The distribution pattern of the national income is also an important determinant of a product. If national income is evenly distributed, market demand for normal goods will be the largest. If national income is evenly distributed, market demand for normal goods will be the largest. If national income is unevenly distributed, i.e., if majority of population belongs to the lower income groups, market demand for essential goods, including inferior ones, will be the largest whereas the demand for other kinds of goods will be relatively less.

4.7 DEMAND FUNCTION

In mathematical language, a function is a symbolic statement of relationship between the dependent and the independent variables. Demand function states the relationship between the demand for a product (the dependent variable) and its determinants (the independent variables). Let us consider a very simple case of demand function. Suppose all the determinants of demand for commodity X, other than its price, remain constant. This is a case of short-run demand function. In case of a short-run demand function quantity demanded of X, (D_x) depends only on its price (P_x). The demand function can then be stated as ódemand for commodity X, (D_x) depends on its price (P_x)ôThe same statement may be symbolically written as

$$D_x = f(P_x) \qquad \text{é} \quad .(4.1)$$

in this function. D_x is a dependent and P_x is an independent variable. The function (4.1) reads ódemand for commodity X (i.e., D_x) is the function of its price (i.e., P_x)ô. It implies that a change in P_x (the independent variable) causes a change in D_x (the dependent variable).

The form of demand function depends on the nature of demand-price relationship. The two most common forms of demand-price relationship are ***linear*** and ***nonlinear.*** Accordingly, the demand function may assume a linear or a nonlinear form.

Linear Demand Function

A demand function is said to be linear when it results in a linear demand curve. Eq. (4.2) represents a linear form of demand function. Assuming that in an estimated demand function a = 100 and b =.5, function (4.2) can be written as

$$D_x = 100 \text{ ï } 5P_x \qquad \text{é } .. (4.3)$$

By substituting numerical values for P_x, a demand schedule may be prepared as follows.

Demand Schedule

P	D_x	=	100	ï	$5P_x$	D_x
0	D	=	100	ï	5×0	100
5	D	=	100	ï	5×5	75
10	D	=	100	ï	5×10	50
15	D	=	100	ï	5×15	25
20	D	=	100	ï	5×20	0

This demand schedule when plotted gives a linear demand curve as shown in Fig. 4.7. Not that the linear demand curve has a constant slope ($\Delta P_x/\Delta D_x$).

From the demand function, one can easily obtain the price function. For example, given the demand function (4.2), the price function may be written as

P

Or P

Assuming $a/b = a_1$ and $1/b = b_1$, the price function may be written as

$$Px = a_1 - b_1 D_x$$

Fig. 4.7 Linear Demand Function.

Nonlinear Demand Function

A demand function is said to be nonlinear or curvilinear when the slope of the demand curve, ($\Delta P/\Delta D$) changes all along the curve. Nonlinear demand function yields a demand curve instead of a demand line, as shown Fig. 4.8. A nonlinear demand function takes the form of a power function, as

$$D = aP^{-h} \qquad a$$

and $$D = \frac{}{} \qquad -b$$

where a > o, b > o and c > o.

It should be noted that the exponent to the price variable in a nonlinear demand function (4.5 a) is the coefficient of price elasticity of demand.

4.8 ELASTICITIES OF DEMAND

4.8.1 IMPORTANCE OF ELASTICITY CONCEPT

We have earlier discussed the **nature of relationship** between demand and its determinants. Form a managerial point of view, however, the knowledge of **nature of relationship alone is not sufficient.** What is more important is the **extent of relationship** or the degree of responsiveness of demand to the changes in its determinants, it, elasticity of demand. The concept of elasticity of demand plays a crucial role in business-decisions regarding maneuvering of prices with a view to making larger profits. For instance, when cost of production is increasing, the firm would want to pass rising cost on to the consumer by raising the price. Firms may decide to change the price even without change in cost of production. But whether this action raising the price following, the, rise in cost or otherwise will prove beneficial depends on (a) the

price elasticity of demand for the products, i.e., how high or low is the proportionate change in its demand in response to a certain percentage change in is price; and (b) price, elasticity of demand for its substitute because when the price of a product increases, the demand for its substitutes increases automatically even if their prices remains unchanged. Raising price will be beneficial only if (i) demand for a product is less elastic; and (ii) demand for its substitute is much less.

In this section, we will discuss various methods of measuring elasticities of demand. The concepts of demand elasticities used in business decisions are: (i) Price-elasticity; (ii) Cross-elasticity; (iii) Income-elasticity; and (iv) Advertisement elasticity, (v) Elasticity ofpri6e expectation.

4.8.2 PRICE ELASTICITY OF DEMAND

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

Percentage change in quantity demanded

e_p == ð ð ð ð ð ð ð ð ð ð ð ð ð ð ð ð ð ð

Percentage change in price

A general formula[2] for calculating coefficient of price-elasticity, derived from this definition of elasticity, is given as follows.

where Q = original quantity demanded, P = original price, ΔQ = change in quantity = demanded, and ΔP = change in price.

It is **important** to note here .that, a minus sign (-) is generally inserted in the formula before the fraction with a view to making elasticity coefficient a nonnegative value.

The elasticity can be measured between two points on a demand curve (called are elasticity) or on a point (called point elasticity).

Arc Elasticity

The measure of elasticity of demand between any two finity points on a demand curve in known as are elasticity. For example, measure of elasticity between points *j* and *k* (Fig. 4.9) is the measure of arc elasticity. The movement from point *j* to *k* on the demand curve (D_x) shows a fall in the price Rs.20 to Rs.10 so that $\Delta P = 20$ ï 10 = 10. The fall in price increases demand from 43 units to 75 units so that $\Delta Q = 43\text{-}75 = -32$. The elasticity between points) and *k* (moving from) to *k)* can be calculated by substituting these values into the elasticity formula as follows :

ep

1.49

It means, a one percent decrease in price of commodity X results into a 1.49 per cent increase in demand for it.

Problem in using are elasticity. The are elasticity should be measured, interpreted and used carefully, otherwise it may lead to wrong decisions. Are elasticity co-efficients differ between the same two finite points on a demand curve if *direction* of change in price is reserved. For instance, as estimated in Eq. (4.7), the elasticity between points j and *k* –

Fig. 4.9 Linear Demand Curve

moving from *j* to *k* - equals 1.49. It may be wrongly interpreted that the elasticity of demand for commodity X between points *j* and *k* equals 1.49 irrespective of direction of price change. But it is not true. A reverse movement in the price, i.e., the movement from point *k* to *j* implies a different elasticity co efficient (0.43). Movement from point *k* to *j* gives P = 10, ΔP = 10 ï 20 = -10, Q = 75, Q= 75, and ΔQ = 75, and ΔQ = 75 ï 43 = 32. By substituting these values into the elasticity formula, we get.

$$e = -\frac{0.43}{0.43} \cdot \frac{10}{p} = \frac{32}{p} \quad \ldots(4.8)$$

Fig. 4.10 Point Elasticity

The measure of elasticity co-efficient in Eq. (4.8) for the reverse movement in price is obviously different from one given by Eq. (4.7). Thus, *the elasticity depends also on the direction of change in price.* Therefore while measuring price elasticity, the direction of price changes should be carefully noted.

Point Elasticity

Point elasticity on linear demand curve. Point elasticity is another way to resolve the problem in measuring the elasticity. The concept of point elasticity is also useful in measuring the elasticity where change in price and quantity combinations is infinitesimally small.

Point elasticity is the elasticity of demand at a finite point on a linear demand curve, e.g., at point P or B on the demand curve MN (Fig. 4.10). This is in contrast to the arc elasticity between point P and B. A movement form point B to wards P implies change in price ΔP becoming smaller and smaller, such

that P is almost reached. Here the change in price is infinitesimally small. Measuring elasticity for an infinitesimally small change in price is the same as measuring elasticity at a point. The formula for measuring point elasticity is given below.

Point elasticity (e) = (4.9)

Note that has -------

been substituted for ~ in the formula for arc elasticity. The derivative ~~ is reciprocal of the slope of the demand curve MN. Point elasticity is thus the product of price-quantity ratio (at a particular point on the demand curve) and reciprocal of the slope of the demand line. The reciprocal of the slope of the straight line MN at point P is geometrically given by o~ so that 8Q QN e = PQ

Note that at point P, price P = PQ and Q = OQ. By Substituting these values in Eq. (4.9), we get

PQ QN ï QN

Ep OQ PQ OQ

Given the numerical value for QN and OQ elasticity at point P can be easily obtained. We may compare here arc elasticity and point elasticity at point *j* in Fig.4.9 At point *j.*

QN 108-43 e p = 1.51 OQ 43

Given that e = 1.51 is different from different measures of are elasticities (i.e., 1.49, 0.43, 0.74, 0.81).

As we will see below, geometrical1y, QN/OQ = PN/PM. Therefore elasticity of demand at point P (Fig. 4.10), may be

PN ep=PM

To conclude, the price elasticity of demand at any point on a linear demand curve is equal to the ratio of lower segment to the upper segments of the line i.e.

Lower segment

Upper segment

By this rule, at mid-point of a linear demand curve, $e_p = 1$, as shown at point P in Fig. 4.10. It follows that at any point to the left of point P, $e_p > 1$, and at any p point to the right of point P, $e_p < 1$. According to the above formula, at the extreme point N, $e_p = 0$, and at extreme point M, e_p is undefined because division by zero is undefined. It must be noted here that these results are relevant between points M and N and that the elasticities at the extreme points M and N are, in effect, undefined.

Fig. 4.1 Point Elasticities of Demand

4.8.3 DETERMINANTS OF PRICE ELASTICITY OF DEMAND

We have noted above that price-elasticity of a product may vary between zero and infinity. The price-elasticity of a product within this range depends on the following factors.

1. Availability of Substitutes. One of the most important determinants of elasticity of demand for a commodity is the availability of its close substitutes. The higher the degree of the closeness of the substitutes, the greater of elasticity

of demand for the commodity. For instance, coffee and tea may be considered as close substitutes for each other. It price of one of these goods increases, the other commodity-become relatively cheaper. Therefore, consumers buy more of relatively cheaper goods, and less of the costlier one, all other things remaining the same. The elasticity of demand for both these goods will be higher. Besides, the wider the range of the substitutes, the greater the elasticity. For instance, soaps, tooth pastes, cigarettes etc., are available in different brands, each brand being a close substitute for the other. Therefore, the price-elasticity of demand for each brand is much greater than the generic commodity. On the other hand, sugar and salt do not have their close substitute and hence their price-elasticity is lower.

2. Nature of Commodity. *The nature of a commodity* also affects the price-elasticity of its demand. Commodities can be grouped as luxuries, comforts and necessities Demand for luxury goods (e.g., high-price refrigerators, TV sets, cars, decoration .items, etc.) is more elastic than the demand for necessities and comforts because consumption of luxury goods can be dispensed with or postponed when their price rise. On the other hand, consumption of necessary goods (e.g. m sugar, clothes, vegetables) cannot be postponed, and hence their demand is inelastic. Comforts have more elastic demand than necessities and less elastic than luxuries. Commodities are also categorised as durable goods and perishable or nondurable goods. Demand for durable goods is more elastic than that of non-durable goods, because when the price of the former increases, people either get the old one repaired instead of replacing it or buy a ósecondhandô.

3. Weightage in the Total Consumption. Another factor that influences the elasticity of demand is the proportion of income which consumers spend on a particular commodity. I proportion of income spent on a commodity is large, its demand will be more elastic, and vice versa. Classic examples of such ócommodities are salt, matches, books, pens, tooth pastes, etc. These goods claim is very small proportion .of income. Demand for these goods is generally inelastic because increase in the price of such goods does not substantially

affect consumerôs budget. Therefore, people continue to purchase almost the same quantity when their prices Increase.

4. Range of Commodity Use. *The range of uses -of a commodity also* influences its demand. The wider the range of uses of a product, the higher the elasticity of demand. As the price of a multi-use commodity decreases, people extend their consumption to its other uses, Therefore, the demand for such a commodity generally increases more than the proportionate increase in its price. For instance, milk can be taken as it is and it may be converted into curd, cheese, ghee and butter-milk. The demand for milk will therefore be highly elastic. Similarly, electricity can be used for lighting, cooking, heating and for industrial purpose. Therefore, demand for electricity has a greater elasticity.

5. Proportion of Market Supplied. The elasticity of market demand depends also on the *proportion of the market supplied at the ruling price* if less than half of the market is supplied at the ruling price, price-elasticity of demand will be higher than one and if more than half of the market is supplied $e < 1$. That is, demand curve is more elastic over the upper half than over the lower half.

4.8.4 CROSS-ELASTICITY OF DEMAND

The cross-elasticity is the measure of responsiveness of demand for a commodity to the changes in the price of its substitutes and complementary goods. For instance, cross-elasticity of demand for tea is the percentage change in its quantity demanded with respect to the change in the price of its substitute, coffee. Formula for measuring cross-elasticity of demand for tea (e.) and the same for coffee ($e_{c.j.}$) is given below :

Percentage change in demand for tea (Q)

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

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

Uses of Cross-Elasticity

An important use of cross-elasticity is that it is used to define substitute goods. If cross-elasticity between tow goods is positive, the two goods may be considered as substitutes of one anthor. Also, the greater the cross-elasticity, the closer the substitute, Similarly, if cross-elasticity of demand for two related goods is negative, óthe two may be considered as complementary of one anthor: the higher the negative cross-elasticity, the higher the degree of complementary.

The concept of cross-elasticity is of vital importance in changing price of products, having substitutes and complementary goods. If cross-elasticity in response to the price of substitutes is greater than one, it would be inadvisable to increase the price; rather, reducing price may prove beneficial. In case of complementary goods also, reducing price may be helpful in maintain the demand in case the price of the complementary goods is rising.

4.8.5 INCOME-ELASTICITY OF DEMAND

A part from the price of a product and its substitutes, consumerôs income is another basic determinant of demand for a product. As noted earlier, the relationship between quantity demanded and income is of positive nature, unlike the negative price-demand relationship. The demand for goods and services increases with increase in consumerôs income and vice-versa. The responsiveness of demand to the change in income is known as income-elasticity of demand.

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

(where X_q = quantity of X demanded; I = disposable income; Δx_q = change in quantity of X demanded; and ΔI = change in income).

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

Unlike price-elasticity of demand, which is always negativeô. Income-elasticity of demand is always positive because of a positive relationship between income and quantity demanded of a product. But there is an exception to this rule income-elasticity of demand for inferior goods is negative, because of inverse substitution effect. The demand for inferior goods decreases with increase in consumerôs income and vise-versa. The reason is when income increased, consumers switch over to he consumption of superior commodities, i.e., they substitute superior goods for inferior ones. For instance, when income rises, people prefer to buy more of rice and wheat and less of inferior food grains; buy more of meat and less of potato, and travel more by plane and less by train.

Nature of commodity and income - elasticity. For all normal goods, income-elasticity is positive though the degree of elasticity varies in accordance with the nature of commodities. Consumer goods of the three categories, viz., necessities, comforts, and luxuries have different elasticities. The general

pattern of income elasticities of different kind of goods of in income and their effect on sales are given in Table 4.2.

Table 4.2 Income-Elasticities

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

1. Except in case of Giffenôs goods.
2. With an exception of inferior goods.

4.9 FURTHER READINGS

BILAS, RICHARD, A. *Microeconomic Theory, McGraw Hll, New York, 1971.*

DOUGLAS, EVANS J. Managerial Economic : Theory and Practice and Problems, Prentice-Hall, inc. New Jersey.

FERGUSON, C. E., *Microeconomics* Theory, Richard D. Irwin, Howe Wood III.

LEFTWICH, RICHARD H., *The price System and Resource Allocation, Dryden Press, Hinsdale, III*

MANSFIELD, EDWIN, *Micro economics: Theory and Application, W. W. Norton, New York.*

NICHOLSON, WALTER, *intermediate Microeconomics and Its Application, Dryden Press, Hinsdale.*

SPENCER, MILTON H., *Managerial Economics, Richard D. Irwin, Homw-wood III.*

WATSON, DONALDs, *Price Theory and Its Uses, Houghton Miffin, Boston Price Theory in Action Houghton Miffin, Boston.*

WEBB, SAMUEL C., *Managerial Economic, Houghton Mifflin, Boston.*

4.10 IMPORTANT QUESTIONS

1. What are the determinants of market demand for a commodity? How do the changes in the following factors affect the demand for a commodity?

   (a) Price, (b) Income, (c) Price of the substitute
   (d) Advertisement and (e) Population.

   Also describe the nature of relationship between demand for a product and these factor (consider one factor at a time assuming other factors to remain constant.)

2. Distinguish between : (i) demand function and demand schedule, (ii) individual demand and market demand, (ii) demand for normal goods and inferior goods.

3. Define and distinguish between :

   (a) Are elasticity and point elasticity.

   (b) Price elasticity and cross-elasticity, and

   (c) Income elasticity and price elasticity.

4. What is meant by demand schedule, demand curve and demand function? Show how market demand is calculated from individual curves.

5. Which of the following commodities has the most inelastic demand and why?

   (a) Soap (b) Salt (c) Penicillin
   (d) Cigarettes, and (e) Ice cream.

6. Explain the following concepts separately :

   (i) Income-elasticity of demand.

(ii) Price-elasticity of supply.

(iii) Elasticity of price expectations.

What useful information do these concepts of elasticity provide to management?

7. Given the demand function

Q_d = 12 ï p

(a) find the demand and marginal revenue schedules,

(b) plot the AR and MR Schedules,

(c) find marginal revenue when P = 10, 6, 2 and

(d) estimate the elasticity co-efficient of the demand curve, when the total is at maximum

8. What is the law of demand? Explain with the help of demand schedule and demand curve what are the exceptions to this law?

9. Why does a demand curve slope downward to the right? Can a demand curve slope upward to the right under any condition?

Lesson : 5

DEMAND FORECASTING

(Author - Dr. B.S. Bodla)

Need of Demand Forecasting : Demand forecasting is predicting future demand for a product. The information regarding future demand is essential for planning and scheduling production, purchase of raw materials, acquisition of finance and advertising. It is much more important where a large-scale production is being planned and production involves a long gestation period. The information regarding future demand is essential also for the existing firms for avoiding under or over-production. Most firms are, in fact, very often confronted with the question as to what would be the future demand for their product. For, they will have to acquire inputs and plan their production accordingly. The firms are hence required to estimate the future demand for their product. Otherwise, their functioning will be shrouded with uncertainty and their objective may be defeated.

An important point of concern in all business activities is to assess the future business trend whether it is going to be favourable or unfavorable. This assessment helps the top management in taking appropriate policy decisions in advance. If sales are expected to rise substantially after, say, 10 years, it will call for measures to build adequate productive capacity well in advance so that future profit potential is not lost to the rival producers. This essentially relates to long-term planning.

On the other hand, if sales of a product are expected to go up in the very near future, it will be prudent on the part of the management to make the needed adjustments in production schedule and take suitable steps immediately to ensure that sufficient stocks are available with given plant capacity as soon as needed. This involves short-term planning.

Irrespective of the length of future time period one is interested in, the planners and policy makers need to know the possible future trends in relation to several variables, which is made possible through forecasting. In this context, forecasting provides knowledge about future trends and deals with the methods of acquiring this knowledge.

Due to dynamic nature of market phenomenon demand forecasting has become a continuous process and requires regular monitoring of the situation.

Demand forecasts are first approximations in production planning. These provide foundations upon which plans may rest and adjustments may be made. *"Demand forecast is an estimate of sales in monetary or physical units for a specified future period under a proposed business plan or program or under an assumed set of economic and other environmental forces, planning premises outside the business organisation for which the forecast or estimate is made".*

Sales forecast is an estimate based on some past information, the prevailing situation and prospects of future. It is based on an effective system and is valid only for some Specific period. The following are the main components of a sales forecasting system :

(i) Market Research Operations to get the relevant and reliable information about the trends in market.

(ii) A data processing and analysing system to estimate and evaluate the sales performance in various markets.

(iii) Proper co-ordination of steps (i) and (ii) and then to place the findings before the top management for making final decision.

In this lesson, we will discuss the important methods of estimating and forecasting demand. The techniques of forecasting are many, but the choice of a suitable method is a matter of experience and expertise. To a large extent, it depends also on the nature of the data available for the purpose. In economic forecasting, classical methods use historical data in a rather rigorous statistical manner for making the future projections. There are also less formal methods where analystôs own judgment plays a greater part in picking, choosing and interpreting the available data than the statistical tools.

TECHNIQUES OF FORECASTING DEMAND

Survey Method : Survey method are generally used where purpose is to make short-run forecast of demand. Under this method, surveys are conducted to collect information about consumerôs intentions and their future purchase-plans. This method includes :

(i) survey of potential consumers to elicit information on their intentions and plan;

(ii) opinion polling of experts, i.e., opinion survey of market experts and sales representative, and through market studies and experiments.

The following techniques are used to conduct the survey of consumers and experts.

Consumer Survey Methods :

The consumer survey method of demand forecasting involves direct interview of the potential consumers. It may be in the form of:

- complete enumeration, or
- sample survey.

These consumer survey methods are used under different conditions and for different purposes. Their advantages and disadvantages are described below.

Direct Interview Method :

The most direct and simple way of assessing future demand for a product is to interview the potential consumers or users and to ask them what quantity of the product they would be willing to buy at different prices over a given period say, one year. This method is known as direct interview method. This method may, cover almost all the potential consumers or only selected groups of consumers from different cities or parts of the area of consumer concentration. When all the consumers are interviewed, the method is known as complete enumeration survey method, and when only a few selected representative consumers are interviewed, it is known as sample survey method. In case of industrial inputs, interview of postal inquiry of only end-users of a conduct may be required. These are described as follows :

Complete Enumeration Method :

In this method, almost all potential users of the product are contacted and are asked about their future plan of purchasing the product in question. The quantities indicated by the consumers are added together to obtain the probable demand for the product. For example, if only n out of m number of households in a city report the quantity (d) they are willing to purchase of a commodity, then total probable demand (D) may be calculated as

$$D_p = d_1 + d_2 + d_3 + \text{é} \ . D_n \quad \text{é é é é} \ . \qquad (1)$$

where d_1, d_2, d_3 etc. denote demand by the individual households 1, 2, 3 etc. This method has certain limitations. It can be used successfully only in case of those products whose consumers are concentrated in a certain region or locality. In case of a widely dispersed market, this method may not be physically possible or may prove very costly in terms of both money and time. Besides, the demand forecast through this method may not be reliable for many reasons : (i) consumers themselves may not be knowing their actual demand in future and hence may be unable or not willing to answer the query; (ii) even if they answer, their answer to hypothetical questions may be only hypothetical, not real; and (ii) their plans may change with the change in factors not included in the questionnaire

Sample Survey Method :

Under this method, only a few potential consumers and users selected from the relevant market through a sampling method are surveyed. Method of survey may be direct interview or mailed questionnaire to the sample consumers. On the basis of the information obtained, the probable demand may be estimated through the following formula :

Dp ------

were D_p = probable demand forecast; H = census number of households from the relevant market; H_s = number of households surveyed or sample households; H_R = number of households reporting demand for the product; A_c = average expected consumption by the reporting households (=total quantity reported to be consumed by the reporting households ÷ number of households).

This method is simpler, less costly, and less time-consuming than the comprehensive survey method. This method is generally used to estimate short-term demand from business firms, government departments and agencies, and also by the households who plan their future purchase.

Sample survey method is widely used to forecast demand. This method, however, has some limitations. The forecaster therefore should not attribute reliability to the forecast more than warranted. Besides, sample survey method can be used to verify the demand forecast made by using quantitative or statistical methods. Although some authors suggest that this method should be used to supplement the quantitative method for forecasting rather than to replace it, this method can be gainfully used where market area is localized.

Expert-Opinion Method :

It is one of the most widely used and influential forecasting technique where the opinions and intuition of management is utilised. The process brings together in an organised manner, personal judgments about the process being analysed Main reliance is on human judgments.

In this method, the executive uses his own anticipation and what he hears from others. Outside experts are also consulted and the other executive heads are also required to give their opinion in the matter. Salesmen are to provide information about customerôs attitude and preferences and the activities of competitors. Thus all possible information from the opinions of various persons is combined together to change the subjective opinions into quantitative forecasts.

No doubt experts and experienced managers can be useful as guides and serve as reliable source of information, but one has to make his own decision from all the opinions. Thus in this method broad guess is made by the executive in charge of a business. There are many advantages and disadvantages of opinion technique of forecasting :

Advantages :

(i) Simple and easy to understand.

(ii) No specialised skill is required.

(iii) Low cost.

(iv) It is based on the information or opinion of the persons who are directly involved in the system.

(v) It can be used in case of new products where satisfactory data is not available.

Disadvantages :

(i) Opinions and intuitions are highly subjective.

(ii) Personal estimates are likely to be biased.

(iii) Time required to take the decision may be more.

(iv) Results can be easily distorted.

(v) This method is not useful for long term planning.

Delphi Method :

Delhpi method of demand forecasting is an extension of the simple expert opinion poll method. This method is used to consolidate the divergent expert opinions and to arrive at a compromise estimate of future demand. The Process is simple.

Under Delphi method, the experts are provided information on estimates of forecasts of other experts along with the underlying assumptions. The experts may revise estimates in the light of forecasts made by other experts. The consensus of experts about the forecasts constitutes the final-forecast. It may be noted that the empirical studies conducted in the USA have shown that unstructured opinions of the experts is most widely uses technique of forecast. This may appear a bit, unusual in as much as this gives the impression that sophisticated techniques, e.g., simultaneous equations model and statistical methods, are not the techniques which are used most often. However, the unstructed opinions of the experts may conceal the fact that information used by experts in expressing their forecasts may be based on sophisticated techniques. The Delphi technique can be used for cross-checking the information on forecasts.

Market Studies and Experiments :

An alternative method of collecting necessary information regarding demand is to carry out market studies and experiments in consumerôs behaviour under actual, though controlled, market conditions. This method is known in common parlance as market experiment method. Under this method, firms first select some areas of the representative markets - three or four cities having similar features, viz., population, income levels, cultural and social background, occupational distribution, choices and preferences of consumers. Then, they carry out market experiments by changing prices, advertisement expenditure, and other controllable variables in the demand function under the assumption that other things remain the same. The controlled variables may be changed over time either simultaneously in all the markets or in the selected markets. After such changes are introduced in the market, the consequent changes in the demand over a period of time (a week, a fortnight, or month) are recorded. On the basis of data collected, elasticity coefficients are computed. These coefficients are then used along, with the variables of demand function to assess the demand for the product.

Alternatively, market experiments can be replaced by consumer clinic or controlled laboratory experiment. Under this method, consumers are given some money to buy in a stipulated store goods with varying prices, packages, displays, etc. The experiment reveals the consumers responsiveness to the changes made in prices, packages and displays, etc. Thus, the laboratory experiments also yield the same information as the field market experiments. But the former has an advantage over the latter because of greater control over extraneous factors and its somewhat lower cost.

Limitations : The market experiment methods have certain serious limitations and disadvantages which reduce the reliability of the method considerably.

(i) The experiment methods are very expensive. It cannot be afforded by small firms.

(ii) Being a costly affair, experiments are usually carried out on a scale too small permit generalization with a high degree of reliability.

(iii) These methods are based on short-term and controlled conditions which may not exist in an uncontrolled market. Hence the results may not be applicable in the uncontrollable long-term conditions of the market.

(iv) The changes in socio-economic conditions taking place during the field experiments, such as local strikes or lay-offs, advertising program by competitors, political changes, natural calamities, may invalidate the results.

(v) ñTinkering with price increases may cause a permanent loss of customers to competitive brands that might have been triedò.

Despite these limitations, however, market experiment method is often used to provide an alternative estimate of demand, and also ñas a check on results obtained from statistical studies.ò Besides, this method generates elasticity coefficients which are necessary for statistical analysis of demand relationships.

Statistical Methods :

Basically all statistical approaches of forecasting, project historical information into the future. These are based on the assumption that future patterns tend to be extensions of past ones and that one can make useful predictions by studying the past behaviour i.e. the factors which were responsible in the past will also be operative to the same extent in future.

Some companies have detailed sales record item wise as well as territory wise. These sales record can be utilised to make useful predictions. The information should be complete with respect to events, policies, quality of the product etc. from period to period. Such information in general is known as Time series data. The time series for any phenomenon is composed of three components (i) Trend (ii) Seasonal variation and (iii) Random fluctuations. Trend exhibits the general tendency of the data and is known as long period or secular trend. This can be either upward or downward, depending on the behaviour.

Mostly trend is used for forecasting in practice. There are many methods to determine trend. Some of the methods are :

(i) Graphical method.

(ii) Least square method.

(iii) Moving average method.

(i) Graphical Method : In this method the period is taken on X-axis and the corresponding sales values on y-axis and the points are plotted for given data on graph paper. Then a free hand curve passing through most of the plotted points is drawn. This curve can be used to forecast the values for future. The method is explained by the following example.

Example 1 : The demand for a product is continually diminishing. Estimate the demand for 2004 with the help of following information:

Year	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**
Demand (in 1000 units)	**75**	**70**	**72**	**69**	**50**	**54**	**37**

Solution : Plot a graph, for the given data to find the demand for 2004 (see fig. 1). From the graph the demand for 1004 comes out to be approximately 20,000 units.

It is an approximate method as the shape of the curve mainly depends on the choice of scale for the graph and the individual who draws the free hand curve.

Fig. 1

(ii) Least Squares Method : This is one of the best method to determine trend. In most cases, we try to fit a straight line to the given data. The line is known as õLine of best fitô as we try to minimise the sum of the squares of deviation between the observed and the fitted values of the data. The basic assumption here is that the relationship between the various factors remains unchanged in future period also.

Let Y denote the demand and X the period for a certain commodity. Then the linear relationship between Y and X is given by

$$Y = a + bX \quad \text{é é é é é é é é é é é é é é é} \quad (3)$$

the nature of the relationship is determined by the values of a and b. The values of a and b can be estimated with the help of the past information about Y and X. If x arid y denote the deviations of X and Y from their respective means, then the least square estimates of a and b are given by

a

b

where n is the number of observations. The calculation of Σy, Σxy and Σx^2 can be done with the help of given data on Y and X. The following example will help you in understanding this method.

Example 2 : The sales of a product is given below :

Years	**1992**	**1993**	**1994**	**1995**
Product Sales (in Rs.)	**1,00,000**	**1,50,000**	**1,25,000**	**1,75,000**

Fit a linear trend and forecast the sales for the year 1996.

Solution : Let years be denoted by X and product sales by Y. Then linear trend of year X is given by

$y = a + bX$

The unknown constant óaô and ôbô can be estimated by least square method. The calculation can be done in the following tabular form.

Table 1

Year X	x-1973.5 x = 5	x^2 Rs. ó000 Y	Sales in XY	
1992	-3	9	100	-300
1993	-1	1	150	-150
1994	1	1	125	125
1995	ó3	9	175	525
n = 4	$\Sigma x^2 = 0$	$\Sigma x^2 = 20$	$\Sigma Y = 550$	$\Sigma XY = 200$

Now a

I:XY 200,000 = 10,000

b = I:X2 = 20

Hence the linear trend is

Y = 1,37,500 + 10,000 (X-l~;3.5)

For X = 1996, forecast of Sales will be

Y = 1.37,500 + 10000 (1996-1993.5)

= 1,37,500 + 10,000 (5)

= 1,87,500

Advantages of least squares method :

(i) There is no need to conduct any sample survey as only past information about sales is required.

(ii) Method is simple and easy to understand.

(iii) Under normal situations the method is likely to give reliable and accurate results.

Disadvantages of least squares method :

(i) The method is based on some mathematical formulate which may not be understood by common man.

(ii) The assumption that other things remaining constant may not hold good in practice.

Exponential trend :

When sales (or any dependent variable) have increased over the past years at an increasing rate or at a constant percentage rate, then the appropriate trend equation to be used is exponential trend equation of the following forms.

(1) Double-log trend of the form

$Y = aT^b$ é ..(4)

or its double logarithmic form

$\lg Y = \log a + b \log T$

This form of trend equation is used when growth rate is increasing.

(2) Polynomial trend of the form

$Y = a + bT + cT^2$ é é é é (5)

In these equations a, band c are constants, Y is sales, T is time and c = 2.718. Once the parameters of the equations are estimated, it becomes quite easy to forecast demand for the years to come.

The trend method is quite popular in business forecasting because of its simplicity. It is simple because only time series data on sales are required. The analyst is supposed to possess only working knowledge of statistics. Since data requirement of

this method is limited, it is also in expensive. Besides, trend method yield fairly reliable estimates of future course of demand.

Limitations :

The trend method has, however, the following limitations.

The **first** limitation of this method arises out of its assumption that, the past rate of change in the dependent variable will persist in future too. Therefore, the forecast based on this method may be considered to be reliable only for the period during which this assumption holds.

Second, this method cannot be used for shot-term estimates. It cannot be used also where trend is cyclical with sharp, turning pints of troughs and peaks.

Thirdly, this method, unlike regression analysis, does not bring out the measure of relationship between dependent variables. Hence, it does not yield the necessary information (e.g., price and income elasticities) which can be used for future policy formulations. The analyst should bear these limitations in mind while making the use of this method.

(c) Box-Jenkins Method

Box-Jenkins method of forecasting is used only for short term predictions. Besides, this method is suitable for forecasting demand with only stationary time-series sales data. Stationary time-series is one which does not reveal a long-term trend. In other words, Box-Jenkins technique can be used only in those cases in which time-series analysis depicts only monthly or seasonal variation or variations that recur with some degree of regularity. When sales data of various commodities are plotted, many commodities will show a seasonal or temporal variation in sales. For examples, sale of woolen clothes will show a hump during months of winter in all the years under reference. The sale of New Year Greeting Cards will be particularly very high in the last week of December every year. Similarly sale of desert coolers is very high during the summers each year. This is called seasonal variation. Box-Jenkins technique is used for predicting demand where time series sales data reveal this kind of seasonal variations.

According to Box-Jenkins approach, any stationary time-series data can be analysed by the following three models :

(i) auto regression model,

(ii) moving average model, and

(iii) auto regressive moving average model.

The three models are, in fact, the three stages of Box-Jenkins method. The auto regressive-moving average model is the final form of the Box-Jenkins model. The purpose of three models is to explain movements in the stationary series with minimised error term, i.e., the unexplained components of stationary series.

The steps and models of Box-Jenkins approach are described briefly here with the purpose of acquainting the reader with this approach rather than providing the entire methodology.

Steps in Box-Jenkins Approach

As mentioned above, Box-Jenkins method can be applied to only stationary time-series. Therefore, the first step in Box-Jenkins approach is to eliminate trend from the time-series data: Trend is eliminated by taking first differences of time-series data, i.e. subtracting observed value of one period from the observed value of the proceeding year. After trend is eliminated, stationary time-series is created.

The **second step** in the Box-Jenkins approach is to check whether there is seasonality in stationary time-series. If a certain pattern is found to repeat over time, there is seasonality in stationary time-series.

The **third step** involves use of models to predict the sales in the intended period. Let us now describe briefly the Box-Jenkins models which are used in the same sequence.

(i) Autoregressive Model

In a general auto regressive model, the behaviour of a variable in a period is linked to the behaviour of the variable in future periods. The general form of the auto regressive model is given below :

$$Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + \text{é} . + a_n Y_{t-n} + e_t \text{é é é é} \quad (6)$$

This model states that the value of Y in period t depends on the values of Y in periods t-1, t-2, ... t-n. The term e_t is the random portion of Y_t that is not explained by

the model. If estimated value of one or some of the coefficients a_1, a_2, é . A_n are different from zero, it reveals seasonality in data. This completes the second step.

The model (6), however, does not specify the relationship between the value of Y and residuals (e_t of previous periods. Box-Jenkins method uses moving average method to specify the relationship between Y_t and e_t values of residuals in previous years. This makes the third step. Let us now look at the moving average model of Box-Jenkins method.

(ii) Moving Average Model

The moving average model estimated Y_t in relation to residuals (e_t) of the previous years. The general form of moving average model is given below :

$$Y_t = m + b_1 e_{t-1} + b_2 e_{t-2} + ... + b_p e_{t-p} + e_t ..... (7)$$

where m is mean of the stationary time series and e_{t-1}, e_{t-2}é é , e_{t-p} are the residuals, the random components of Y in t-1, t-2, é é t-p periods, respectively.

(c) Method of Moving Averages : This method can be used to determine the trend values for given data without going into complex mathematical calculations. The cal1culations are based on some predetermined period in weeks, months, years, etc. The period depends on the nature of characteristics in the time series and can be determined by plotting the observations on graph paper.

A moving average is an average of some fixed or pre-determined number of observations (given by the period) which moves through the series by dropping of top item of the previous averaged group and adding the next item below in each successive average.

The calculation depends upon the period to be odd or even.

In the case of odd order period (3,5,7,é é é .) the average of the observations is calculated for the given period and the calculated value is written in front of central value of the period e.g. for a period of 5 years, the average of the values of five years is calculated and is recorded against the third year. Thus in case of five yearly moving averages, first two years and last two years of the data will not have any average value.

If period of observations is even e.g. four years, then the average of the four yearly observations is written between second and 3rd year values. After this centering

is done by finding the average of the paired values. The method is illustrated by solving example 4.

The even order periods creates the problem of centering between the periods.. Due to this generally odd order periods are preferred.

The calculated values of the moving averages became the basis for determining the expected future sales.

If the underlying demand pattern is stationary i.e. at a constant mean demand level expect, of course, for the superimposed random fluctuations or noise, the moving averages method provided a simple and good estimate. In this method equal weightage is assigned to all the periods chosen for average.

The moving average method for forecasting suffers from the following defects:

(i) records of the demand data have to be retained over a fairly long period.

(ii) if demand series depicts trend as against the stationary level the moving average method would provide forecasts that lags the original series.

Example 3 : The following are the annual sales in thousands of a product during the period 1965-1975. Find the trend of the sales using (i) 3 yearly moving averages and forecast the value for the year 1979.

Year	**Sales in 000 units**	**Year**	**Sales in 000 units**	**Year**	**Sale in 000 units**
1985	**12**	**1989**	**18**	**1993**	**22**
1986	**15**	**1990**	**17**	**1994**	**25**
1987	**14**	**1991**	**19**	**1995**	**24**
1988	**16**	**1992**	**20**		

Solution : The trend values can be calculated in the following tabular form :

Table 12.2

Year	Sale in 000 units	Three yearly moving toia1	3 yearly moving average Trend values
1985	**12**		
1986	**15**	**41**	**41/3 = 13.7**
1987	**14**	**45**	**45/3 = 45**
1988	**16**	**48**	**48/3 = 16**
1989	**18**	**51**	**51/3 = 17**
1990	**17**	**54**	**54/3 = 18**
1991	**19**	**56**	**18.7**
1992	**20**	**61**	**20.2**
1993	**22**	**67**	**22.3**
1994	**25**	**71**	**23.7**
1995	**24**		

e.g. 41 = value of 1985 + value of 1986 + 1987

= 12 + 15 + 14 = 41 written at the central period 1986 of the years 1985, 86 and 87

Business Indicators :

Business indicators refer to the time series data on important business and economic activities in key sectors of the economy. These time series are representative, in one way or the other, of the aggregate business and economic activity in the economy as a whole. It is mote in the sense that the overall behaviour of such aggregate activities has been found to be systematically associated with the pattern of cyclical movements in the indicator series.

An intelligent analysis and understanding of the time duration and the amplitude of cyclical ups and downs in the selected indicators provide useful information regarding the future behaviour of overall cyclical movements. This holds only long as these are specifically related to a particular business activity.

How correctly a business indicator will help predict the immediate future conditions facing a particular business organisation depends to a large extent on the judicious choice of an indicator in terms of its relevance to the type of business in question. The selection of relevant business indicators is so import ant that a large amount of statistical intelligence is required to go into its final choice before any formal statistical technique is applied for purposes of analysis.

The U.S. National Bureau of economic Research, after having carefully studied about 800 time series which could possibly be used as business indicators, have selected around 20 time series; Such series individually follow definite pattern of cyclical movements vis-à-vis those in the general business activity. The cyclical movements in these selected series have been found to be systematically related to successive cycles in the overall business activity in a definite way. While the turning points in the case of a few precede the cyclical turning points in the general business activity, those in some others coincide, and in yet some others follow, the turning points in the latter.

The indicators that were found to precede the general business activity have come to be known as leading series (or leading indicators). The leading indicators are of crucial importance in, providing information about the upward and downward movements, and the consequent peaks and troughs, in the general economic activity at least a few months in advance. This happens because by virtue of their relationship with the general economic activity, the cyclical movements in the leading series tend to occur earlier than the beginning of the turning points to the overall business and economic activity in the economy.

Business indicators that follow the movements in general business activity are termed as lagging series, while those coinciding the movements in general business activity are known as coincident series. The significance of both these series lies in confirming that turning points in the general business activity have actually started occurring. Thus, if the leading indicators have signaled an upward trend in the general business activity, the coincident series will eventually start weakening. Such a development calls for a careful observation of how all the three types of series are likely to behave in the future.

Although the cyclical indicators approach has been found to be quite beneficial in predicting the cyclical turning points. This does not necessarily indicate the existence of any causal relationship between the two series.

Regression Method : Regression analysis is the most popular method of demand estimation. This method combines economic theory and statistical techniques of estimation. Economic theory is employed to specify the determinants of demand and to determine the nature of relationship between the demand for a product and its determinants. Economic theory thus helps in determining the general form of demand function. Statistical techniques are employed to estimate the values of parameters in the equation estimated.

In regression techniques of demand forecasting, the analysis estimate the demand function for a product. In the demand function, quantity to be forecast is a dependent variable and the variables that affect or determine the demand (the dependent variable) are called as óindependentô or óexplanatoryô variables. For example, demand for cold drinks in a city may be said to depend largely on óper capita incomeô of the city and its population. Here demand for cold drinks is a ódependent variableô and óper capit incomeô and ópopulationô are the óexplanatoryô variables.

Simple Regression :

In simple regression technique, a single independent variable is used to estimate a statistical value of the ódependent variableô, that is, the variable to be forecast. The technique is similar to trend fitting. An important difference between the two is that, in trend fitting, independent variable is ótimeô (t) whereas in regression equation, the chosen independent variable is the single most important determinant of demand. Besides, the regression method is less mechanical than trend fitting method of projection.

For an illustration, consider the hypothetical data on quarterly consumption of sugar given in table

Table X : Quarterly Consumption of Sugar

Year	**Population (millions)**	**Sugar Consumed (000) tonnes**
1985-86	10	40
1986-87	12	50
1987 -88	15	60
1988-89	20	70

1989-90	25	80
1990-91	30	90
1991-92	40	100

Suppose we have to forecast demand for sugar for 1994-95 on the basis of 7-year data given in Table. This can be done by estimating a regression equation of the form

$Y = a + bX$ é é é é .. (8)

Where Y is sugar consumed, X is population and a and b are constants

Like trend fitting method, Eq. 8 can be estimated by using the ôleast squareô method. The procedure is the same as shown in Table X That is the parameters a and b can be estimated by solving the following two linear equations:

$\Sigma Y = na + b\Sigma X$ é é é é é é é é é (i)

$\Sigma XY = \Sigma xa + b\Sigma x^2$ é é é é é é é é .. (ii)

The procedure of calculating the terms in equations (i) and (ii) above is presented in Table X.

Table X : Calculation of Terms in Linear Equations

Year	**Population (X)**	**Sugar consumed (Y)**	**X^2**	**XY**
1985-86	10	40	100	400
1986-87	12	50	144	600
1987-88	15	60	225	900
1988-89	20	70	400	1400
1989-90	25	80	625	2000
1990-91	30	90	900	2700
1991-92	49	100	1600	4000
$\Sigma n = 7$	**$\Sigma X_t = 152$**	**$\Sigma Y_t = 490$**	**$\Sigma X^2_t = 3994$**	**$\Sigma X_t Y_t = 12000$**

By substituting the values from Table into equation (i) and (ii), we get

490 = 7a + 152 b é é .. (iii)

12,000 = 152 a + 3994 b (iv)

By solving equations (iii) and (iv), we get

a = 27.42

and b = 1.96

By substituting values for a and bin Eq. (8), we get the estimated regression equation as

Y = 27.44 + 1.96 X

Given the regression equation (8), the demand for sugar for 1994-95 can be easily projected if population for 1994-95 is known. Supposing population for 1994-95 is projected to be 70 million, the demand for sugar in 1994-95 may be estimated as

Y = 27.44 + 1.96 (70) = 164,640 tonnes

The simple regression technique is based on the assumption that (i) independent variable will continue to grow at its past growth rate, and (ii) the relationship between the dependent and independent variables will continue to remain the same in future as in the past.

Multi-variate Regression :

The Multi-variate regression equation is used where demand for a commodity is deemed to be the function of many variables or in cases in which number of explanatory variables is greater than one.

The procedure of multiple regression analysis may be briefly described here. The first step in multiple regression analysis is to specify the variables that are supposed to explain the variations in the demand for the product under reference. The explanatory variables are generally chosen from the determinants of demand, viz., price of the product, price of its substitute, consumersô income, and their taste and preference. For estimating the demand for durable consumer goods, (e.g., TV sets, refrigerators, house, etc.), the other variables which are considered are availability of credit and rate of interest. For estimating demand for capital goods (e.g., machinery

and equipments), the relevant variables are additional corporate investment, rate of depreciation, cost of capital goods, cost of other inputs (e.g., labour and raw materials), market rate of interest, etc. These variables are treated as independent variables.

Once independent variables are specified, the second step is to collect time - series data on the independent variables. After necessary data are collected, the next step is to specify the form of equation which can appropriately describe the nature and extent of relationship between the dependent and independent variables. The final step is to estimate the parameters in the chosen equations with the help of statistical techniques. The multivariate equations cannot be easily estimated manually. They have to be computerised.

Diffusion Index :

Diffusion index as a technique of predicting turning points in the general business activity is an improvement over the business indicator approach in so far as it makes up the deficiency of the latter for lack of uniformity in the duration and amplitude of cyclical fluctuations in the leading series. The computation of diffusion index requires counting of the number of leading series and expressing them as a percentage of the total number of series in the leading group.

For example, if there are 20 leading series in all, and if all of them are expanding cyclically, the diffusion index is 100. If 5 series are declining cyclically, the diffusion index is 75, which means that 15 series are still expanding.

The diffusion index is interpreted as follows :

(i) So long as this index remains above 50 per cent, a decline in the index indicates that the overall business activity is in a state of expansion. Once the index reaches the 50 per cent mark, the overall business activity is considered to have reached the peak of expansion.

(ii) A decline in the index below 50 per cent is indicative of the process of contraction having set in. As long as the index remains below the 50 per cent mark, the overall business activity is in a state of contraction and eventually reaches the trough. Revival starts only when it rises above the 50 per cent mark.

(iii) The 50 per cent mark is also decisive in predicting the turning points in the overall business activity. As the index tends to approach the 50 per cent mark from

above, it is indicative of the beginning of the upward trend in the overall business activity.

Turning points in business cycles predicted in line with the trends in the diffusion index are reliable only so long as all the series behaving in a particular direction move cyclically more or less quite closely with one another, and that all the series have equal importance with respect to the aggregate. In practice these conditions are met fairly well.

However, the use of diffusion Index is not an easy task. It is mainly because the construction of a diffusion index requires determining whether particular series is cyclically expanding or contracting, which is an extremely difficult and laborious task.

References :

Ashton, David and Simister, Leslie (eds.), *The Role of Forecasting in Corporate Planning,* Staples Press, London, 1970, Chs. 1-3, 8-9.

Douglas, Evan, J., *Managerial Economics: Theory Practice and Problems,* Prentice Hall, Inc., N.J. 1970, Ch.5.

Hooda, R.P., Statistics for *Business and Economics,* McMillan India Ltd., 2000.

Wolfe, H.D., *Business Forecasting Methods,* Holt, Rinchart and Winston, Inc., 1966.

Reddy, Mehender, J., *Demand Forecasting,* Light and Life Publications, New Delhi. 1981.

Diwevedi, D.N., *Managerial Economics,* Vikas Publishing House Pvt. Ltd., New Delhi, 1999.

Questions :

1. Discuss meaning and significance of Demand Forecasting.
2. Discuss critically the different methods of demand forecasting.
3. Outline the trend projection method of demand forecasting.
4. What are the possible consequences if a large-scale firm places its project in the market without having estimated the demand for its product?

5. What would be the appropriate variables for estimating demand for (a) steel, (b) sugar, (c) petrol, and (d) toys by the regression method?

6. Plot the following data on a graph and find the trend equation for sales:

Year	1990	1991	1992	1993	1994	1995	1976
Total sales (units)	115	102	305	300	95	306	403

7. The following are the available data of sales for some years :

Years	1990	1991	1993	1994
Sales (in lakhs of Rupees)	50	70	60	80

Assuming the same relationship holds true for future, forecast the sales for the year 2002 by applying least square method.

Hint. To make $\Sigma X = 0$, time deviations from 1992 may be taken.

Ans. 115 in Lakhs of rupees.

8. Explain the regression method of demand forecasting. Compare this method with trend method.

9. You are given the following data :

X	3	6	8	10	13	13	13	14
Y	8	6	10	12	12	14	14	20

Estimate the regression equation $Y = a + bX$

10. What are the different techniques of survey method? Under what conditions are complete enumeration and sample Survey methods are chosen ?

11. What is Delphi method? What is the use of this method in demand forecasting?

12. Explain Business indicator method of demand forecasting.

Lesson - 6

The Organization of Production and the Production Function

(Author : N.K. Bishnoi)

In this section we shall first examine the organization of production and classify inputs into various broad categories and then define the meaning and usefulness of the production function in analyzing the firmôs production activity.

The Organization of Production :

Production refers to the transformation of inputs or resources into outputs of goods and services. For example, if we want to produce wheat, we need land, fertilizer, water, workers and some machinery. These are called inputs or factors of production. The output is wheat. The output can also be service rather than a good. Examples of services are education, medicine, banking, communication, transportation and many others. To be noted is that ñProductionò refers to all of the activities involved in the production of goods and services, from borrowing to setting up of expansion of production facilities, to hiring workers, purchasing raw materials, running quality control, and so on, rather than referring merely to the physical transformation of inputs into outputs of goods and services. In a broader sense, activities adding value to the product are part of the production process.

Inputs are the resources used in the production of goods and services. As a convenient way to analysis, inputs are classified into labour, capital, land and entrepreneur. Each of these broad categories, however, includes a great variety of basic input. For example labour includes farmer, bus driver, assembly line worker, accountants, lawyers, doctors, scientists and govt. officials. Capital consists of all the man made resources helping in the production process. It includes machinery, building, inventory and others. In the same manner land represent the natural resources for which human being has done nothing to bring them about. It includes land, natural resources, minerals, rivers, sunlight and even natural talent in a person. As far as the entrepreneurship is concerned there is a controversy regarding its classification. Some economists call entrepreneurship as a distinct factor of production, which is ultimate risk taker in the production process, while other regard it a distinct type of labour only.

Inputs are also classified as fixed or variable. Fixed inputs are those that cannot be readily changed during the time period under consideration, except perhaps at very great expense. Examples of fixed inputs are the firmôs plant and specialized equipment; it takes several years to build a new thermal power plant. On the other hand, variable inputs are those that can be varied easily and on a very short notice. Examples of variable inputs are most raw materials and unskilled labour.

The time period during which at least one input is fixed is called the short run, while the time period when all inputs are variable is called the long run. The length of the long run (i.e. the time period required for all factors to be variable) depends on the industry. For some, such as the setting up or expansion of dry-cleaning business, the long run may be a few months or weeks. For others, such as construction of integrated iron steel plant, it may be several years. In the short run, the firm can increase output only by using more of the variable inputs (say labour and raw material) together with the fixed inputs (plant and equipment) In the long run, the same increase in. output could very likely be obtained more efficiently by also expanding the firmôs production facilities.

The production Functions :

A production function is an equation, table or graph showing the maximum output of a commodity that a firm can produce per period of time with each set of inputs. Both inputs and outputs are generally measured in physical rather than in monetary units. Technology is assumed to remain constant during the period of the analysis.

The general equation of production function is

$$Q = f(a, b, c, d \text{ é } \text{é } \text{é } .. n, T)$$

Where Q represent the physical quantity of output per unit of time f, denotes functional relationship.

a, b, c, d, represent the quantities of various inputs, per unit of time.

T refers to the prevailing state of technology or know how. The bar (-) is placed on T Just to indicate that technology is assumed to be constant

The equation implies that the output or the quantity (Q) of the product depends on the quantities, of a. b. c. d. n of the various inputs used with the given state of technology in the production process per period of time.

For simplicity, economists assume that a firm produces only one type of output with only two inputs, labour (L) [Entrepreneurship dubbed with labour] and capital (K) [land being passive factor combined with capital]. Thus the simple production function is

Q = f (L; K) é é é é é é 1

Table 1, gives a hypothetical production function, which shows the output (the Qs) that a firm can produce with various combinations of labour (L) and Capital (K). Table shows that by using 1 unit of labour (IL) and I unit of capital (IK) the firm would produce 4 units of output (4Q).

With 2 Land IK, output is 10Q. With 3L and 4K the output is 38Q, and so on. Note that labour and capital can be substituted for each other in production. For example 32Q can be produced using 3L and 2K or with 2L and 4K. Input prices will determine which of these combinations of labour and capital minimizes the firmôs cost.

The production Function with variable input :

In this Section, we present the theory of production when only one input is variable. Thus, we are in the short run. We begin by defining the total, the average and the marginal product of the variable input. We will then examine the law of variable proportion and the meaning and importance of the stages of production.

Total, Average and Marginal Product :

By holding the quantity of an input constant and changing the quantity used of the other input, we can derive the total product (TP) of the variable input. For example by holding capital constant at 1 unit (i.e. with K = 1) and increasing the units of labour used from zero to six units, we generate the total product of labour given in the tales 2 column (2). Note that when no labour is used, total product or output is zero. When one unit of labour (IL) is used, total product (TP) is 4. With 2L, TP= 10 with 3L,

TP = 15 and so on.

Table 2

Total, Marginal, and Average Product of labour; K=1

Labour No. of workers	**output or Total Product TP**	**Marginal Product of Labour (MP)**	**Average Product of Labour(AP)**
(1)	(2)	(3)	(4)
0	0	-	-
1	4	4	4
2	10	6	5
3	15	5	5
4	18	3	4.5
5	18	0	3.6
6	15	-3	2.5

From the total product schedule we can derive the marginal and average product schedules of the variable inputs. The Marginal Product of labour (MP L) is the change in total product per unit of change in 1abour used, while the average product of labour (AP L) equals total product divided by the quantity of labour used. That is

MP_L = + TP / + L ---------- 2

AP_L= TP/L ---------- 3

Column 3 in the table 2 gives the marginal product of labour (MPL). Since labour increases by 1 unit at a time in column, the MPL in column 3 is obtained by subtracting successive quantities of TP in column 2. For example TP increases from 0

to 4 units when the first unit of labour is used. Thus MPL = 4. For an increase in labour from 1L to 2L, TP rises from 4 to 10, So that MPL = 5 and so on.

Column 4 of table 2 gives the APL. This equal TP (Column 2) divided by L (Column 1). Thus with 1 unit of labour 1L, APL = 4, with 2L, APL = 5 and so on.

Plotting the total, marginal and average product of labour of table 2 gives the corresponding product curves shown in figure-(I). Note that TP grows to 18 units with 4L, remains at 18 with 5 L, and then declines to 15 units with 6L.

In figure (2), we see that APL rises to 5 units and than declines. Since the marginal product of labour refers to the change in total product, per unit change in labour used, each value of the MPL is plotted half way between the quantities of labour used.

Fig. 2

The law of variable proportion and stages of production :

In order to show graphically the relationship between the total product, on the one hand and the marginal and average products of labour, on the other hand, we assume that labour time is continuously divisible (i.e. it can be hired for any part of the day). Then the TP, MPL and APL become smooth curves as indicated in figure (3).

The MP_L at a particular point on the TP curve at that point. From the figure3, we see that the slope of the TP rises up to point G (the point of injection on the TP curve), is zero at point J and negative there after. Thus, MPL rises up to point G, is zero at point J, and negative thereafter.

On the other hand, AP_L is given by the slope of a ray from the point of origin to the TP curve. From Figure-3, we see that the slope of the TP curve rises upto point H and falls thereafter but remain positive as long as TP is positive. In same manner, the AP_L rises upto point H and falls after wards.

The point to note here is that at point H the slope of a ray from the origin to the TP curves (or AP_L) is equal to the slope of the TP curve (or MP_L). Thus $AP_L = MP_L$ at point H (The highest point on the AP_L curve). Note AP_L rises as long as MP_L is above it and falls when MP_L is below it.

The relationship between the MP_L and AP_L curves in the bottom panel of figure-3 can be used to define three stages of production for labour (variable input). The range from the origin to the point where AP_L is maximum (Point H at 2.5L) is stage I of production for labour. Stage II of production for labour extends from the point where the AP_L is maximum (MP_L is equal to AP_L) at the point to the point where the MP_L is zero. (i.e. from Point H to Point J at 4.5L). The range over which the MP L is negative (beyond point J or with more than 4.5L) is stage III of production for labour.

The rational producer would not operate in stage III even if labour is available at free of cost, because MP L is negative. This means that using less labour could produce a greater output similarly one Would, not produce in stage I for labour because in this stage Marginal product for labour is negative. As in relation to labour, capital is available much more than required. Obviously by adding more labour output would go up more than proportionately. Thus a rational producer will operate in stage II where the MP of both the factors is positive but declining. The precise point within stage II at which rational producer operates will depend on the prices of inputs and outputs.

Explanation of the Stages :

The operation of the law of variable proportion in three stages is attributed to two fundamental characteristics of factors of production.

i. **Indivisibility of certain fixed factors, and**

ii. **Imperfect substitutability between factors.**

Indivisibility of fixed factors implies that initially when smaller quantities of variable factor inputs are employed alongwith a given set of factors, there is a bit of disproportionality between the two sets of factor components. On technical grounds, thus, fixed factors are not effectively exploited. For instance, a factor like machinery, being lumpy, may remain grossly underutilized when only very few units of variable factor like labour are used. But this is not the whole explanation behind the variable behaviour of the production function. Remaining part of explanation is provided by the notion of substitutability between factors of production. Substitutability means the extent to which one factor can perform the task of other factor. For example food grain production can be increased by using more dosage of fertilizer or more number of workers for better upkeep of the farm. Output would increase in both the cases. But only to a limited extent same is the case with man and machinery. Hence the law of variable proportion.

Assumption's of the Law of variable proportion :

1. Only one factor is varied.
2. The scale of output is unchanged.
3. Technique of production does not change.
4. Units of factor input varied are homogeneous.

Significance of the Law :

The business significance of the law of variable proportion is obvious. A careful producer would not produce in stage I and III. Rationally, the ideal combination off actor proportion (fixed plus variable inputs) will be when the average product of labour is maximum.

Moreover universal occurrence of the low has forced the business to go all out for invention of new technology so as to fend off the operation of the law of variable proportion.

Suggested Questions :

1. What is meant fixed inputs, variable inputs, short run, and long run inputs?

2. How long is the time period of the long run inputs?

3. What is production function? What is its usefulness in the analysis of the firmôs production?

4. What is the relationship between the marginal product and the average product curves of the variable inputs?

5. Explain the law of variable proportion. Under what conditions the stage of diminishing return can be postponed?

6. If the total product curve increase at a decreasing rate from the very beginning (Le. from the point where variable input is zero), what would be the shape of the corresponding marginal and average product curves?

Lesson - 7

Concept of Cost : Short Run and Long Run Cost curves

(Author: N.K. Bishnoi)

This chapter begins by examining the nature of costs of production. These include explicit and implicit costs, opportunity costs and incremental costs. Then the firms short run and long run cost curves total, average and marginal cost curves are derived. Subsequently plant size and economies of scale are examined, that is long run cost curves.

The Nature of Costs :

One crucial distinction in the analysis of costs is between explicit and implicit costs. Explicit Costs refer to the actual expenditures of the firm to hire, rent or purchase the input it requires in production. These include the wages to hire labour, the rental price of capital, equipment and buildings and the purchase prices of raw materials and semi finished products. These are the recorded expenditure during the process of production. They are thus also known as accounting cost or money cost, as these are actual monetary expenditures incurred by the firm.

An economist however is not satisfied with these explicit costs only. In the economic sense there are certain costs which are implicit in nature. This refers to the value of the inputs owned and used by the firm in its own production activity. Even though the firm does not incur any actual expenditure to use these inputs, they are not free since firm can sell them or rent them out to other firms. The amount for which the firm could sell or rent out these owned inputs to other firms represents a cost of production of the firms owning and using them. Implicit costs include the highest salary that the entrepreneur can earn for him, if working for other firms and the-highest return the firm could receive from investing its capital in alternatives uses or renting its land and buildings to the highest bidder rather than using them itself. In general, following are the implicit costs, which should be included in the total cost, but go unrecorded in the account of the firm.

1. Wages of labour rendered by the entrepreneur himself.
2. Interest on capital supplied by the entrepreneurs.

3. Rent of land and premises belonging to the entrepreneurs and used in the production.

4. Normal profit of entrepreneur, compensation for being the ultimate risk taker in the firm.

These items are valued at current market rates for estimating the implicit cost. The distinction between explicit and implicit costs is important in analyzing the concept of profit. In the accounting sense, profit is calculated as the residual of total sales receipts minus explicit costs. In economic sense, however normal profit is included in total cost of production, which consists of explicit and implicit costs taken together.

Economic Cost = Accounting cost (Explicit Costs) + Implicit Cost.

Opportunity Cost : To calculate the market value of implicit cost the concept of opportunity cost is used. Now we elaborate the concept. The opportunity cost of a factor of production is the reward (or value) that factor could have earned in the next best alternative occupation. In fact, a cost is a forgone opportunity; the cost of engaging in an activity is the totality of all the opportunities that the activity requires you to forgo. To avoid double counting only the best alternative is considered as opportunity cost.

Accounting opportunity costs are important for financial reporting by the firm and for tax purposes. For managerial decision making purposes (with which we are primarily interested in economics) opportunity or economic costs is relevant cost concept.

With an example of inventory valuation will clarify the distinction.

Suppose, a firm purchased a raw material for Rs.100/- but its price subsequently rose to Rs.150/-. The accountant would continue to report the cost of the raw material at its original price of Rs.100/-. The economist however, would value the raw material at its current or replacement value. Failure to do so might lead to the wrong managerial decision. This would occur, if the firm decides to continue the production using the raw material, while more beneficial out come would have been to stop output and sell the raw material booking the profit at price Rs.150/-

In the same manner after depreciation accountant could take the value of a machine at zero but economist would have to take its resale value to calculate the true worth.

In discussing production cost, we must also distinguish between marginal cost and incremental cost. Marginal cost refers to the change in total cost for a unit change in output. For example, it total cost is Rs.140/- to produce 10 units of output and Rs.1501- to product II units of output, the marginal cost of 11th. Unit is Rs.10. Incremental cost on the other hand is a broader concept and refers to the change in total cost from implementing a particular management decision, such as the introduction of a new product line, the undertaking of a new advertising campaign or the production of a previously purchased components. The costs that are not effected by the decision are irrelevant and are called sunk cost. In other words, sunk costs are not altered by the change in business activity.

Short Run and Long Run Costs :

Economist usually distinguish between short run and long run costs on the basis of functional or operational time period in production activity.

The short run costs are operating costs associated with the change in output. In the short run, the production function contains a set of fixed factor input and a set of variable inputs. Short run costs vary in relation to the variation in the variable input component only.

The long run costs are the operating costs associated with the changing scale of output and the alteration in the size of plant. In the long run production function all the factor inputs are variable. Their costs are the long run costs.

Behaviour of Costs in the Short-run :

In this section we distinguish between fixed and variable costs and derive the firmôs total and per unit cost functions.

Short Run Total and Per-unit Cost function :

As already defined short-run is the time period during which some of the firmôs inputs are fixed (i.e. cannot be readily changed, except perhaps at very great expense). The total obligations of the firm per time period for all fixed inputs are called total fixed deposits (TFC). These include interest payment, rental expenditures,

property taxes and those salaries (such as for top management) that are fixed by contract and must be paid over the life of the contract whether the firm produces or not.

Total variable costs TYC) : ON the other hand, are the total obligations of the firm per time period for all the variable inputs that the firm use. Variable inputs are those that the firm can change easily and on short notice. Payment for raw materials, depreciation associated with the use of the plant and equipment; most of the labour costs, excise duties are included invariable costs.

Total costs (TC) equal total fixed costs (TFC) plus total variable costs (TVC). That is TC = TFC + TYC.

Within the limits imposed by the given plant and equipment, the firm can vary its out-put in the short run by varying the quantity used of the variable inputs. This gives rise to the TFC, TYC and TC functions of the firm. In defining cost functions, all inputs are valued at their opportunity cost which includes both explicit and implicit cost. Input prices are assumed to remain constant regardless of the quantity demanded of each input by the firm.

From the total fixed, total variable and total cost function, we can derive the corresponding per unit cost function of the firm. Average fixed cost (AFC) equals total fixed costs (TFC) divided by the level of output (Q). Average variable cost (A YC) equals total variable costs (TVC) divided by output. Average total cost (ATC) equal total cost (TC) divided by output. Finally marginal cost (MC) is the change in total costs or change in total variable cost (TVC) per unit change in output.

AFC	TFC/Q
AYC	TVC/Q
TC	TC/Q = AFC +AVC
MC	$\Delta TC / \Delta Q = \Delta TVC / \Delta Q$

Table 1

1	2	3	4	5	6	7	8
Q	TFC	TVC	TC	AFC	AVC	ATC	M
0	100	0	100	-	-	-	-

1	100	25	125	100	125	125	25
2	100	40	140	50	20	70	15
3	100	50	150	33.3	16.6	50	10
4	106	60	160	25	15	40	10
5	100	80	180	20	16	36	20
6	100	110	210	16.7	18.3	35	30
7	100	150	250	14.3	21.4	35.7	40
8	100	300	400	12.5	37.5	50	150
9	100	500	600	11.1	55.6	66.7	200
10	100	900	1000	10.0	90	100	400

Table (1) gives a hypothetical cost function

Behaviour of total Costs :

Examination of the table (1) gives us following observations regarding the total costs.

1. TFC remain constant at all level of output it is unchanged even when the output is nil. Thus TFC is independent of output.

2. TVC varies with the output. it is nil when there is no output. Variable costs are thus direct costs of the output

3 TVC does not change in the same proportion. Initially it is increasing at a decreasing rate, but after a point it increases at an increasing rate. This is due of the law of variable proportion.

4. TC varies in the same proportion as the TVC. In other words, the change

   in total cost is entirely, due to changes in the total variable costs. In fact the distance between TC and TVC is the TFC.

Fig-I

TFG, TVC and TC Curves

Total cost curves are derived by plotting the total cost schedule graphically. A careful observation of fig. 1 reveals the following important characteristics of cost behaviour.

1. The curve TFC is the curve of total fixed costs. Denoting constant characteristics of fixed cost at all level of output, TFC is a straight horizontal line, parallel to the X-axis.

2. The curve TVC represent total variable cost. It reflect the typical behaviour of total variable cost. It initially rises gradually but eventually becomes steeper, denoting a sharp rise in total variable costs.

3. The TC curve represents total cost. It is derived by vertically adding up TVC and TFC curves. Obviously shape of the TVC and TC are identical. The only difference between two is of distance that is total fixed cost.

Short-run per Unit Cost :

From the cost schedule given in table 1, it is clear that costs per unit are derived from the total costs. It is obvious that the firm will have four short period categories of unit costs (I) Average fixed Cost (AFC) (II) Average Variable Cost (A VC) (III) Average Total Cost (A TC) and (IV) Marginal Cost (MC).

Economists, generalize the following relationship with regard to the unit cost data.

1. AFC decreases as output increases. Since AFC = TFC/Q, it is purely a mathematical outcome that with numerator remaining unchanged, the increasing denominator causes a diminishing product:

2. AVC first decreases and them increase as the output increases.

3. Since ATC is the sum of AFC and AVC, it will decrease in the beginning as both component decreases initially. After a point AVC start increasing and pulls up the ATC along with it, out weighing the influence of ever decreasing AFC.

4. Marginal Cost also decreases initially but increases ultimately with the increase in output.

Marginal cost is the rate of change in total costs when output is increased by one unit. In a geometrical sense, marginal cost at any output is the slope of the total cost curve at the corresponding point. In the short run, the marginal cost is independent of fixed cost and is directly related to the variable cost. Hence the MC curve can also be derived from TVC curve. As a matter of fact, AVC curve and MC curve are the reflection and the consequence of the law of variable proportion operating in the short run. As shown in the fig-2 both the curves are U shaped, the explanation of which is as follows. With labour as the only variable input, TVC for any output level (Q) equals the wage rate (W, assumed to be fixed) times the quantity of labour (L) used. Thus

$AVC = TVC/Q = W.L./Q.\ W/Q/L$

$(Q/L = AP_L)\ W/AP_L$

As explained in the previous chapter, Average product of labour usually rises first, reaches the maximum and then falls, it follows that AVC curve first falls, reaches a minimum and then rises. Thus AVC is exactly inverse of AP_L curve whereas MC curve is exactly the reverse of, MP curve. In the last since the C curve is U shaped, the ATC curve is also U shaped. The ATC curve continues to fall after the A VC curve begins to be as long as the decline in AFC exceeds the rise in AVC.

The U shape of MC curve can similarly be explained as follows:

$MC = \Delta TVC / AQ = - \Delta (WL) / AQ$

Since W is constant

$W(\Delta L) / \Delta Q = W/\Delta Q/\Delta L$

As $\Delta Q/\Delta L = MP_L$

$= W/MP_L$

Since the Marginal product of labour (MPL) first rises, reaches a maximum and then falls, it follows that the MC curve first falls reaches a minimum and then rises. Obviously, MC curve is exactly the reverse of MP_L curve.

Fig. 2

Relationship between Marginal Cost and Average Cost :

There is a unique relationship between AC (ATC as well AVC) and MC, that is described as below :

1. When AC is minimum, MC is equal to AC. Thus MC intersect AC at its lowest point.

2. When AC is falling, MC is always below AC. In fact, it is the MC that pulls down AC along with its. The point to note here is that MC may be rising, but will remain below AC.

3. When AC is rising, MC must be above AC.

Long Run Cost Curves :

We now turn to explain the cost curves in the long run. Long run is the period during which all inputs are variable. Thus all costs are variable in the long run (i.e. the firm faces no fixed costs). The length of time of the long run depends on the industry. In some service industries such as photocopying, the period of the long run may be only a few months or weeks. For others, which are very capital intensive, like satellite based commination network, it may take several years. It all depends on the length of time required for the firm to be able to vary all inputs. The Long run cost of production is the least possible cost of Production of producing any given level of output when all the inputs one variable including of course the size of the plant. A long run cost curve depicts the functional relationship between output and the long run cost of production as just defined.

Long run average cost is the long run total cost divided by the level of output. Long run average cost depicts the least possible average cost for producing all possible level of output. In order to understand, how the long run average cost curve is derived, consider the three short run average cost curve as shown in figure-3.

These short run average cost curves are also called plant curves, since in the short run plant is fixed and each of the short run average cost curves corresponds to a particular plant. In the short run the firm can be operating on any short run average cost curve, given the size of the plant? Suppose that only these three are technically possible sizes of the plants. In the long run the firm will examine that which size of the plant or on which short run average cost curve it should operate to produce a given level of output at the minimum possible cost.

It can be seen from fig. 3 that upto OB amount of output, the firm will operate on the short run average cost curve SAC_1, though it could also produce with short-run average cost curve SAC_2 because upto OB amount of output, production on SAC_1 curve entails lower cost than on SAC_2. For instance, if the level of output OA is produced with SAC_1, it will cost AL per unit and if it is produced with SAC_2, it will cost AH per unit. Obviously AL is lower than AH. Similarly all other output levels upto OB can be produced more economically with the smaller plant SAC_1, than with the larger plant SAC_2. It is thus clear that in the long run firm will produce any output upto OB on SAC. If the firm plan to produce any output which is larger than OB but less the OD, then it will not be economical to produce on SAC_1 For the output range between OB to OD, SAC_2 provides cheaper option. Thus the output OC if produced on SAC_2 costs CK per unit which is lower than CJ which is the cost incurred if produced on SAC_1. Therefore, if the firm plans to produce between OB and OD, it will employ the plant corresponding to short run average cost curve SAC_2. In the same manner, for output larger than OD, economically, SAC3 provides the best possible alternative. Given that only three sizes of plants as shown in figure-3, then the long

run average cost curve are having scallops on it. This heavily scalloped long run average cost curve consists of some segments of all the short run average cost curves as explained above.

Suppose now that the size .of the plant can be varied by infinitely small gradations so that there are infinite number of plants corresponding to which there will be numerous short run average cost curves. In that case, the long run average cost curve will be smooth and continuous line without any scallops. Such smooth long run average cost curve has been shown in fig. 4 and has been labeled as LAC. There will be infinite number of short-run cost curves though only, eight SACS are shown in the fig. 4. In fact the long run average cost curve is locus of all tangency points with some short run average cost curves. If a firm decides to produce particular output in the long run it will pick a point on the long run average cost curve corresponding to that output and it will than build relevant plant and operate on the corresponding short-run average cost curve.

Fig. 4

In fact the long run average cost curve is locus of all tangency points with some short run average cost curves. If a firm desides to produce particular output int he long run it will pick a point on the long run average cost curve corresponding to that output and it will than build relevant plant and operate ont he coff esponding short-run average cost curve.

It can be seen from the fig. 4 that the long run average cost curve first falls and then beyond a point it rises, that is, the long run average cost curve is U shaped, though U shape of long run average cost curve is less pronounced. In other words long run average cost is flatter in comparison to short run average cost curve.

Long-Run Marginal-Cost Curve (LMQ) :

Like the short run marginal cost curve, the long run marginal cost curve is also derived from the slope of total cost curve at the various points relating to the given

output each time. The shape of LMC curve has also a flatter U shape indicating that initially as output expands in the long run, LMC tend to decline. At a certain stage however, LMC, tends to increase. The behaviour of LMC is shows in fig. 5.

Fig. 5

From the fig. 5, the relationship between LAC and LMC may be traced as follows:-

1. When LAC is decreasing, LMC is below LAC.
2. LMC is equal to LAC, when LAC is at its minimum point.
3. LMC is above LAC, when LAC is rising.

Explanation of the U shape of the Long-run average Cost curve :

The economists generally believe that the LAC is U shaped. Now what is the proper explanation for such behaviour of the LRAC?

We have seen that U shape of SAC curve is explained with the law of variable proportions. But the LAC depends upon the returns to scale. And return to scale in turn depends on the internal economies of scale. In other words our problem is what are the reasons that the firms first enjoy internal economies of scale and then beyond a point it has to suffer internal diseconomies of scale.

We will first, discuss in detail the nature of the internal economies of scale, that is, economies, which arise from the firm increasing its plant size. Economies of scale are distinguished into real economies and strictly pecuniary economies of scale.

Pecuniary economies are economies realized from paying lower prices for the factors used in the production and distribution of the product due to bulk-buying by the firm as its size increases. Such strictly monetal-N, economies do not imply an actual decrease in the quantity of inputs used but accrue to the firm from lower prices paid for raw material, low interest rates, lower wages and salaries due to firmôs better

bargaining power. These are called pecuniary because they accrue to the firm at the cost of counter party without increasing the economic efficiency.

Real economies are those associated with a reduction in the physical quantity of inputs, raw material, various types of labour and various types of capital, We may distinguish the following main types of real economies (i) Production economies (ii) Selling or marketing economies (iii) Managerial economies (iv) Transport and Storage economies.

Production Economies of Scale :

Production economies may arise from labour, capital (technical) and inventory requirement of the firm.

Labour economies are achieved as the scale of output increases for several reasons:

(a) Better specialization become possible with higher level of output.

(b) Higher level of output allows the use of more efficient automated machines.

(c) Division of labour, those increases with the increase in output, results in saving of time usually lost in going from one work to another.

Technical economies are associated with theô fixed capitalô, which includes all types of machinery and other equipment. These main technical economies arises from (a) more specialized and efficient machines are available generally for larger output level.

(b) Set up costs are normally a fixed amount, obviously larger the size of machine lower the set up cost in proportion to total cost of capital.

(c) Generally, as size is increased, machine cost does not go up proportionally. In fact -in engineering there, is a rule of thumb of 0.6. It means if size is increased by 100 percent, cost will go up by 60% only.

Inventory economies occur when with the increase in the level of output; requirement for reserve inventory does not increase proportionately.

Selling or Marketing economies :

Selling economies are associated with the distribution of the product of a firm. The main types of such economies are (a) advertising economies (b) economies from special arrangement with exclusive dealers (c) model change economies.

Advertising Economies : It is generally agreed that advertising space (in newspapers or magazines) and time (on television or Radio) increase less than proportionately with scale, so that advertising cost per unit of output fall with scale. The advertising budget is usually decided on the basis of available-funds, profits and similar activities of competitors rather than on the basis of output. Obviously advertising budget is almost like a fixed cost, hence the larger the output the smaller the advertising cost per unit.

Special arrangement Economies : Large firms can enter into exclusive agreements with distributors to provide after sales services for the, products of the firm, reducing the need for the firm to have massive arrangement for the purpose.

Model Change Economies : In modem industry, firms need to change the style of their product quite frequently in order to meet the demand of their customers and the competition of the rival firms. A change in the model or style of the product often involves considerable expenses in research and development and possibly on new material and equipment the spreading of such overheads is lower per unit if the scale of output is large.

Manaiterial Economies : Managerial economies arise for various reasons, the most important being (i) specialization of management and (ii) mechanization of managerial functions.

Specialized managerial economies occur when large scale operation make it feasible for the firm to employ production manager, sales manager, personnel manager, finance manager and so on. This division of managerial work increases the experience of managers in their own areas of Specialization and leads to the, more efficient working of the firm.

Mechanization economies : Large firms apply techniques of management involving a high decree of mechanization such as computerized managerial information system reducing the cost of information flow substantially.

Transport and Storage economies :

Transport costs are incurred partly on the production side (transportation of raw materials or intermediate products) and partly on the selling side of the firm (transportation of final product to its market). The same holds for storage costs.

Storage economies : Storage costs will clearly fall with size Geometry tells us that volume increases more than proportionately with the increase of surface area. Similarly maintenance, supervision cost of storage will not increase proportionately with the increase in output.

Analysis of transport cost is more complicated still higher output provides more flexible transport planning of goods giving rise to economies of scale.

Thus, economies of scale are the reason behind the falling portion of the LAC. But what causes it to go up beyond a point? It is the diseconomies of scale.

Diseconomies of scale arises primarily because as the scale of operation increases, it becomes more difficult to manage the firm effectively and coordinate the various operations and divisions of the firm. The number of meeting, the paper work and telephone bills increases more than proportionately to the increase in the scale of operation and it becomes increasingly difficult for top management to ensure that their subordinates properly carry out their directives and guidelines. Thus, efficiency decreases and cost per unit tend to rise.

In the real world, the forces for increasing and decreasing economies of scale operate side by side. In the beginning economies outweigh diseconomies, while beyond a point diseconomies becomes more powerful minimum point LAC is achieved when economies and diseconomies balance each other completely.

Model Ouestioris

1. Distinguish between -

   i) Marginal and Incremental Cost

   ii) Accounting Cost and Economic Cost.

   iii) Explicit cost and implicit cost.

   iv) Outlay Cost and opportunity cost.

2. (i) What is meant by opportunity-cost?

(ii) What is its significance in managerial decision making?

3. Why is short run cost curves U shaped?

4. Why long run average cost curve is flatter than the S.R.A.C. ?

5. Explain the various economies of scale.

Lesson - 8

Concept of Revenue; and Break Even Analysis

Section A – Concept of Revenue

(Author : N.K. Bishnoi)

Introduction :

The revenue of a film together with its cost determines the profits. We therefore, turn to the study of the concept of revenue. Revenue means sales receipts. It is the receipts obtained by action from selling various quantities of its products Revenue depends on the price at which the quantities of output are sold by firm.

A firmôs revenue may be classified as: (i) Total Revenue (ii) Average Revenue (iii) Marginal revenue.

Total Revenue (TR) :

Total revenue is the total sales receipt of the output sold over a given period of time. Total revenue depends on two- factor (i) Price of the product and (ii) the quantity of the product. It is obtained by multiplying the quantity sold (Q) by its selling price (P) per unit. In symbolic terms TR = P × Q.

For example, if the selling price of a pen is Rs.20 per pen and 80 pens are sold during the week, total revenue would be TR = 20 × 80 = Rs.1600/-

Average Revenue (AR) :

Revenue obtained per unit of output sold is termed óaverage revenueô. It is simply the total revenue divided by the number of units of output sold. Thus

AR = TR/Q

In our example, AR = 1600/80 = Rs.20/-

Thus, the revenue earned per unit is Rs.20/-. That is equal to the price in the example. Is this average revenue always equal to the price? If seller charges different price for different units (like bulk discount) or charges different price from different customers. (Doctors charging different price from different patients), then price will not be equal to the average revenue.

Marginal Revenue : Marginal revenue is the addition made to the total revenue by selling one more unit of the item, or simply, it is the revenue or sales receipt of the marginal (latest addition) unit of the firmôs sale.

$MR_n = TR_n ï TR_n$

In our example TR_{80} = Rs.1600/- when 80 units are sold, and price is given at Rs.20 per unit TR_{79} would be 1580.

Hence MR_{80th} unit =1600 - 1580 = 20

Otherwise, it is the rate of increase in total revenue when the increment in the sale of output is assumed unit wise i.e. MR = DTR/ DQ

Relationship between AR and MR curves :

The relationship between AR and MR depends on the market form, within which the firm under consideration is operating. For the purpose of revenue analysis market form can be classified in to perfect competition and imperfect competition. The reason behind this classification is that in perfect competition the firm is a price taker hence AR = MR at all levels of sale. While in imperfect competition AR and MR are different to the firm under study.

1) Under Perfect Competition :

Under perfect competition a very large number of firms are producing identical product. Hence the market forces of supply and demand determine the price and that price prevails for all the firms in the industry. It is as shown in fig. 1 (A). Each firm can sell as much as it wishes at the ruling market price OP. Thus the demand for the product is infinitely elastic, (Fig .IB). Since the demand curve is the average revenue curve for the firms and AR is unchanged at all levels hence MR is equal to AR at all levels of demand.

Fig. 1

Table-I

Q	AR	TR	MR
1	20	20	20
2	20	40	20
3	20	60	20
4	20	80	20
5	20	100	20
6	20	120	20
7	20	140	20

2) Under Imperfect Competition :

When competition is not perfect, the firm will face downward sloping demand curve, whether market, form is monopolistic competition, oligopoly or monopoly. Downward sloping demand curve means firmsô can sell larger quantity of output only lowering the price of the product. In other words in imperfect competition AR curve would be downward sloping for the firm. And when average revenue curve is downward sloping marginal revenue curve would be below AR

Imperfect Competition :

Q	**AR (P)**	**1R**	**MR**
1	20	20	20
2	18	36	16
3	16	48	12
4	14	56	8
5	12	60	4
6	10	60	0
7	08	56	(-) 4

The relation between MR and AR is explained in the table 2. To increase the demand of the product the producer reduces the price by Rs.2/- in each case and MR

incoming down by Rs.4/- in each case because producer gets lower price on previous units also.

This relationship is shown in fig. 2

Geometrical Relationship between AR and MR Curve

A typical Geometrical relationship is observed between the linear AR and MR curves. That is in the case of linear data the MR falls twice of the fall in price at each level of output. Thus when demand curve (AR) is straight line, the MR is also straight line and lies in Mid-way between price axis (Y-axis) and average revenue curves.

Proof : To prove the statement a point P is taken on the price axis (Y -axis) in fig. 3

Fig- 3

At price OP, quantity OQ is demanded. The point B is thus obtained on the AR curve. Line PB is drawn. The MR curve cuts the line PB at point T. PB is the distance between AR curve and the Y-axis. In order to prove that the MR curve lies exactly at half the distance, we have to prove PT = BT. for this a perpendicular BQ is drawn. The MR curve cuts BQ at point N.

Since total Revenue. (TR) = Price × Output

TR = OP × OQ = Area OPBQ ---------- 1

Again since TR = ∑MR

TR = Area ODBQ ---------- 2

It follows thus OPNQ = ODNQ ---------- 3

Geometrically, it is clear that

OPBQ = OPTNQ + BIN ---------- 4

And ODNW = OPTNQ-+ PTD ---------- 5

Form eq (3) I follows that

OPTNQ + BTN= OPTNQ +PTD ---------- 6

OPTNQ being common it follows

Therefore, BTN = PTD ---------- 7

This means D triangle BTN and are equal in area

Again in these triangles

$\angle$ DPT = $\angle$ TBN (being right angle)

$\angle$ PTD = $\angle$ BTN (being vertically opposite angles)

$\angle$ PDT = $\angle$ TNB (being alternate angles)

Hence, both triangles are equiangular.

ΔPTD and ΔBTN are similar.

Since both these triangles are equal in area and also similar, it follows that both are congruent. Hence their corresponding sides are equal.

PT = BT

BN = PD

TN = DT

Hence point p lies exactly in the middle of line PB. This means, when the MR curve passes through point T, it has exactly at half the distance between AR curve and price axis (y-axis).

If the demand curve is non - linear, then also the MR curve will be below to AR but it will not be at mid way.

Fig. 4

The Relationship between AR, MR and elasticity.

The marginal revenue (MR) is related to the price elasticity of demand with the formula:

MR = P ï 1 ï i/e

Proof :

Where MR ï Marginal Revenue

P - Price

£ - Price elasticity of Demand

P = f(Q)

The total revenue is TR = PQ = [f(Q)]Q.

The MR is

MR =

On the basis of this formula the relationship between AR and MR is explained.

e = -

rearranging we obtain -

Substituting dp/dQ in the expression of MR we find

MR = P + Q

MR = P + Q P /-eQ = P - Q P/eQ = P ï

MR = P (1 ð)

And P = AR

MR = AR (1 ð)

On the basis of this formula the relationship between AR and MR is explained. MR is positive only if the price elasticity of demand is greater the 1.

Example i.e. = 1

MR = AR (1- 1/e) MR = AR (1-1/1) = AR (0) =0

MR = AR (1-1/2) = AR (1/2)

(MR = 1/2) AR and so on

Questions :

1. Define (i) Total Revenue (ii) Average Revenue (iii) Marginal Revenue.
2. Trace the relationship between price, total average and marginal revenues of a firm under perfect competition.
3. Trace the relationship between price and total revenue under imperfect competition.
4. Explain the geometrical relationship between the linear AR and MR curve.
5. Write short notes on :
   - (i) AR, MR and elasticity of Demand.
   - (ii) Firm demand under perfect competition.
   - (iii) Firm demand under imperfect competition.

Section : B

Break Even Analysis

The break - even analysis (BEA) has considerable significance for economic research, business decision making company management, investment analysis and public policy.

Break even analysis is an important technique to trace the relationship between cost, revenue and profits at the varying levels of output or sales. In BEA, the break even point is located at that level of output or sales at which the net income or profit is zero. At this point total cost is equal to total revenue. Hence the break -even point is the no profit no loss point. However the object or the BEA is not just to determine the break - even point (BEP), but to

understand the financial relationship among cost, revenue and the rate of output .It is also called cost -volume - profit analysis.

Fig - 5

In the fig. 5 total revenue and total costs are plotted on vertical axis, where sales or output per period are plotted on the horizontal axis.

The slope of the TR curve refers to the constant price of Rs.10 per unit at which the firm can sell-its output. The TC curve indicates total fixed costs (TFC) of Rs.200 (the vertical intercept) and constant average variable cost (AVC) of Rs.5 (the slope of TC curve). In the figure it is clear that the firm break even (with TR = TC = Rs.400) at Q = 40 per time period. (Point B in the figure). The firms incurred loss at smaller output and earn profit at higher output levels.

The cost volume profit or break - even chart is a flexible tool to quickly analyse and plan accordingly the effect of changing conditions on the firm. For example an increase in the price of commodity can be shown by increasing the slope of the TR curve on shown an increase in the total fixed costs of the firm can be shown an increase in the vertical intercept of the TC curve and a decrease in the average of a variable cost by decrease in the TC curve.

Break even Analysis (BEA) can also be performed algebraically, as follow. Total revenue is equal to the selling price (P) per unit times the quantity of output or sales (Q). That is :

$$TR = (P)\,(Q) \quad \text{.....} \quad (1)$$

Total costs (TC) equal total fixed cost (TFC) plus total variable cost (TVC) Since TVC is equal to the average (per unit) variable cost. (AVC) times the quantity of output or sales we have

TC = TFC +AVC (Q) ---------- (2)

Setting total revenue equal to total costs and substituting Qb (the break - even output) for Q, we get

TR = TC ..... (3)

(P) (QB) = TFC + AVC (QB) ---------- (4)

Solving the equation (4) for QB, we have

(P)(QB) - AVC (QB) = TFC

QB (P-AVC) = TFC

QB = TFC/

P-AVC ---------- (5)

For example, with TFC = Rs.200, P = B 10 and AVC = Rs.5

QB = 200 / 10-5 = 40

This is the same break -even output shown on the cost - volume profit fig . (6). The denominator in equation (5) (P-A VC) is called contribution margin per unit because it represent the portion of the selling price that can be applied to cover the fixed costs of the firm and to provide for profits.

Profit Planning :

More generally, suppose, the firm wishes to earn a specific profit and want to estimate the quantity that they must sell to earn that profit. Cost volume profit or break even analysis can be used in determining the target output (QT) at which a target profit (πT) can be achieved. To do so, we simply add p T to the numerator of eq . 5 and we have

QT = TFC + πT/P - AVC

For example, if the firm wanted to earn a target profit of Rs.200 in our Previous example, the target output would be

QT = 200 +200 /10 - 5 = 400 / 5 = 80

To see that the output of Q = 80 does indeed lead to the target profit (πT) of Rs.200, note that

TR = (P)(Q) = (10) (80) = 800

TC = TFC + AVC (Q) = 200 + 5 (80) = 200 + 400 = 600

= TR - TC = 800 - 600 = 200

While linear cost volume profit analysis can be very useful and are frequently used by business executives government agencies and other organizations, care must be taken to apply them only in uses where the assumption of constant price and average variable costs hold.

It prices and average variable costs are not constant, a non-linear Break-even-analysis can be applied, that is an advance technique to be covered under the under graduate course level.

QUESTIONS :

1) (a) What is break even analysis.

(b) What is the assumption under lying the linear BEA

1) What are the limitations of BE A

2) Explain and illustrate a break - even chart. Point out the usefulness of the break-even analysis.

Lesson - 9

Price Determination Under Perfect Competition

(Author : Anil Kumar)

õPerfect Competitionô is a phrase used often in everyday discussion and any people have an institute and vague understanding of what it means. The concept of perfect competition is very old and was discussed in a casual way by Adam Smith in his õWealth of Nationsô. Edgeworth was the first to attempt (in his book õMathematical Psychicsô; 1881) a systematic and vigorous definition of perfect competition. The concept received its complete formulation in Frank Knights book õRisk, Uncertainty and Profitô (1921).

The concept of perfect competition is based on large number of assumptions, but following are the most important.

i) Every firm in the market is so small that it cannot exert any perceptible reference on price. Thus the firm is a price taker and not price maker.

ii) The product is homogenous. In the eyes of the consumer, the product of one seller is identical to that of another seller. This ensures that buyers are indifferent as concerned to the firm from which they purchase.

iii) The industry is characterized by freedom of entry and exit. Any new firm is free to setup production if it so wishes, and any existing firm can stop production and leave the industry according to its will.

iv) There is free mobility of factors of production. All resources are perfectly mobile. For instance, labour is mobile geographically and among jobs.

v) The participants in the market have perfect knowledge. Consumers know prices; producers know costs; workers know wage rate; and So on In addition every one has complete knowledge of the market.

vi) There is no government interference in the market. Tariffs, subsidics and so on are ruled out.

vii) There is absence of transport cost as all firms are closer to the market, and all firms are supposed to be equally far away from the market.

Pure and Perfect Competition

A distinction is often made between pure competition and perfect competition. But this distinction is more a matter of degree than of find. For a market to be purely competitive, three fundamental conditions must prevail.

(i) A large number of buyers and sellers.

(ii) A homogeneity of product and

(iii) The free entry or exit of firms.

For the market to be perfectly competitive, following additional conditions must be fulfilled,

i) Perfect knowledge of market.

ii) Perfect mobility of factors of production.

iii) Absolutely no government interference and

iv) No transport cost difference incidently, the term perfect competition is traditionally used by British economists while discussing the price theory. American economists, however, prefer to construct a pure competition market model realistically assuming that additional conditions for perfect competition, such as perfect mobility of labour, perfect knowledge etc. may not be attainable.

Equilibrium of Firm

A firm is said to be in equilibrium when it has no tendency either to increase or to contract its output. A firm is in equilibrium when it is earning maximum profit.

Conditions of Equilibrium

A firm would be in equilibrium when the following two conditions are fulfilled :

1. MC = MR

2. MC curve cuts MR curve from below.

Under perfect competition, an individual firm has to accept, price determined by industry. The firm under perfect competition is a price taker and not price maker. Demand curve or average revenue curve of the firm is a horizontal straight line (i.e. parallel to X-axis). Since perfectly competitive firms sell additional units of output at the same price, marginal revenue curve coincides with average revenue curve.

to decide about its equilibrium output, the firm will compare marginal cost with marginal revenue; It will be in equilibrium at the level of output at which marginal cost is equal to marginal revenue and marginal cost curve cuts marginal revenue curve from below.

Consider the Fig. I in which price OP is prevailing in the market. Marginal cost curve cuts MR curve at two different points E_o and E_1 and marginal cost and marginal revenue are equal at these two points. E_o can not be the position of equilibrium since at E_o second order condition of the firms equilibrium is not satisfied.

The firm can increase its profits by increasing production beyond E_o because marginal revenue is greater than marginal cost. The firm will be in equilibrium at point E_1 or output OQ_1 since at E_1 marginal cost equals to marginal revenue as well as marginal cost curve cuts marginal revenue curve from below.

Equilibrium of the firm in the short period

Short run means period of time within which the firms can alter their level of output only by increasing or decreasing the amount of variable factors such as labour and raw material, while fixed factors like capital equipment

remain unchanged. Moreover, in the short run, new firms can neither enter the industry nor the existing firms can leave it.

For the sake of simplicity of study, let us suppose that in an industry all factors of production, are homogenous. All the firms are equally efficient such as they have identical cost curve. Under the circumstances each firm of a given industry, in equilibrium may get either.

i) Super normal profit.

ii) Normal profit.

iii) Suffer losses

All the three situations depend upon the price determined by the industry.

All the three situations faced by the firms in equilibrium in short run are explained diagrammatically.

i) Equilibrium with Super Normal Profits

A firm is in equilibrium when its marginal cost is equal to marginal revenue and marginal cost curve cuts marginal revenue curve from below A firm in equilibrium earns super normal profits, when average revenue (Price) determined by industry is more than its average cost. In the Fig. 2 SAC and SMC are short run average and marginal cost curves of the firm. PP, is the average and marginal revenue curves, which are parallel to X-axis. The reason being, under perfect competition, firm is a price taker not price maker. The firmôs equilibrium will be at point E. A perpendicular parallel to the Y-axis is drawn at point E connecting the

Fig. 2

X-axis at Q. EQ is the equilibrium price because point E lies on the demand curve, and price is determined by demand curve. Average cost is equal to CQ. Since average revenue is greater than average cost. Thus, firms per unit excess profit is EC which is the difference between price (EQ) and the corresponding average cost (CQ). Total supernormal profit of a firm is PECD.

Equilibrium with Normal Profit

In the short period, it is possible that firm earns only normal profit. This happens only when the average cost curve of the firm is tangent to its average revenue curve. Equilibrium of the firm has been explained in the Fig. 3.

E is the equilibrium point because at this point MC = MR. MC curve cuts MR curve from below. OQ is the equilibrium output. At OQ level of output the firms AC curve is tangent to AR curve. Thus the firm will earn only normal profit because average revenue (EQ) being equal to average cost (EQ).

Equilibrium with Losses

A firm in equilibrium may incur losses when at the equilibrium level of output firmôs average cost is greater than average revenue. The equilibrium of the firm can be explained with the help of Fig. 4.

In the Fig. 4 marginal cost is equal to marginal revenue at point E. MC curve cuts MR curve from below. OQ is the equilibrium level of output.

Average revenue and average cost of the firm are equal to EQ and FQ respectively. At OQ level of output, firms average cost is greater than average revenue. Firmô s per unit loss is equal to EF and total loss is equal to area EFPG.

Now the question arises why the-firm continues production even at losses. The reason being, in the short period fixed factors like machinery and plants cannot be changed. Therefore, if the firm stops production due to loss, then it will have to bear losses equivalent to fixed cost.

If the firm in the short period earns revenue which covers not only its average variable cost but also some part of fixed cost, the firm will continue its production. In such circumstances, firm will incur more losses if it stops production. Therefore, it is better for the firm to continue to produce so long as it earns revenue more than or equivalent to minimum average variable cost, then firm will incur minimum losses. But when the firmôs price or average revenue falls below minimum average variable cost the firm will prefer to discontinue its production. The firm can avoid cost of variable factors of production.

The above argument has been elaborated by the Fig 4. When price is OP then firmôs equilibrium is at point E and it will produce OQ level of output. The firm experiences loss equivalent to area FEPG. The firm will continue its production in such situation, because price is greater than minimum average variable cost.

If the price of the commodity is OP_1 then the equilibrium of the firm will be at point E and price is equivalent to minimum of average variable cost.

At point E_o, the firm is covering its minimum average variable cost. But at this point no part of fixed cost is being covered. Therefore, the loss of firm is equivalent to total fixed cost, at OQ_o level of output Point E_o is known as óshut down pointô. If price falls below OP_1 then production will be stopped because firmôs loss is more than total fixed cost.

Long-Run Equilibrium of the Firm

The long run is a period of time which is sufficiently long to allow the firm to make changes in all factors of production. The firms in the long run, can increase their output by changing their capital equipment, they may expand their old plant or replace the old lower capacity plants by the new higher capacity plant. Besides, in the long run new firm can enter the industry to complete with existing firm.

The long run equilibrium refers to the situation where free and full adjustment in the capital equipment as well as in the number of firms has been allowed to take place.

A firm is in equilibrium under perfect competition when MC = MR and MC curve must cut MR curve from below. But for the firm to be in long run equilibrium, besides the equality of MC and MR, there must be equality of AR and AC. In other words, the firm will get only normal profits. If the price is greater than the average cost, the firms will earn super normal profits. The supernormal profits will attract other firms into the industry. The price of the product will go down as a result of increase in supply of output and the cost will go up as a result of more intensive competition for factors of production. The firms will continue entering into the industry until the price is equal to average cost so that all firms are earning only normal profits.

On the contrary, if the price is lower than the average cost, the firm would make losses. These losses will induce some of the existing firms to quit the industry. Supply of output will decrease and price will increase because of increase in the average cost. Thus, the firms will get only normal profit, in the long run.

From this analysis we conclude that for the firm to be in equilibrium in the long run following two condition should be fulfilled.

(i) MC = MR and MC curve must cut MR curve from below.

(ii) Average Revenue must be equal to average Cost (AR = AC).

Because in the perfect competition, AR = MR, the above the condition can also be written as :

Price = AR = MR = LMC = LAC.

Price = LMC = LAC

The relationship MC and AC also reveals that MC curve cuts AC curve at its minimum point.

These, conditions for long run equilibrium of the firm can also be written as:

Price = MC = Minimum Average Cost

The Fig. 5 represents long run equilibrium of firm under perfect competition.

LAC and LMC are the long run average and marginal Cost curves, respectively. The firm will be in equilibrium at point E, at which marginal Cost is equal to marginal revenue and marginal Cost curve is rising.

Fig. 5

The firm will get only normal profits because at point E, LAC curve is tangent to AR curve.

If price is increases from OP to OP_1, where the firm is earning abnormal profits. There will be tendency for new firms to enter and compete away these abnormal profits. The firms cannot be in long run equilibrium at any price higher than OP.

On the contrary, if price declines from OP to OP_2 then price will be less than marginal cost, and consequently the firms will incur losses. Some of existing firms will quit the industry due to which supply of the commodity will decline. The price will increase due to decrease in supply. In the long run the

equilibrium of the firm will be at OP price because firm will get only normal profits at the price.

Equilibrium of Industry under Perfect Competition

The industry will be in equilibrium when industry has no tendency to either increase or decrease its level of output. An industry is said to be in equilibrium when t1kre is no tendency for it to expand or contract. It means demand for the product of industry and supply of it are in equilibrium. The industry has no tendency to vary its output. If at a prevailing price, demand for the commodity is more than supply, the industry will try to expand its output. On the other hand, if at prevailing price, quantity demanded of a product falls short of quantity supplied, the price and output of the industry will tend to fall.

When demand for the commodity is equal to supply of commodity, then industry will have no tendency to vary its output. Thus we conclude that industry will be in equilibrium at that level of price end output, where demand curve and supply curve intersect each other.

Conditions of Equilibrium of the Industry

For the industry to be in equilibrium following three conditions should be fulfilled :

i) Demand for and supply of product of the Industry must be equal.

ii) All the firms in the industry should be in equilibrium

iii) There should be no tendency to change the number of firms in the industry i.e. the firms are earning only normal profits.

Short Run Equilibrium of the Industry

In the short run, new firms can neither enter in the industry nor the old firms exit from the industry. Therefore, industry will be in equilibrium when above given first two conditions are fulfilled. The short run equilibrium of industry has been shown in the Fig. 6.

Fig. 6

In part A of the diagram, the equilibrium of the industry has been shown. Demand curve and supply curve of the industry intersect each other at point E. OP is the equilibrium price and OQ is the equilibrium output.

The firm will take OP price as given and adjust its output in such a way that it may earn maximum profit. In part B of the diagram equilibrium of the firm has been shown. EO is the firmôs equilibrium. OM is the equilibrium output. Average revenue and average cost are equal to EOM and CM respectively. Since average revenue is greater than average cost, the firm is earning super normal profit equal to area EOCGP. Suppose; cost of all the firms are identical, all the firms are earning normal profit. If the demand for the product declines, the price of the product will also decline and the equilibrium will be at lower level of output. The industry will be in equilibrium, although firms might be incuring losses.

In this case too the industry will be in short run equilibrium.

Long-Run Equilibrium of the Industry

Long run is that period of time under which new firms can enter and old firms can leave the industry. If firms in the industry are earning super normal profits, new firms will enter in the industry. On the other hand if the firms in the industry are incurring losses, then some existing firms will eave the industry. Therefore, the industry will be in equilibrium, when above given three conditions are fulfilled.

In part A of Fig. 7, industry equilibrium is shown. E is the equilibrium point. OP and OQ are the equilibrium level of price and output. The firms will adjust their output in such a way that it may earn maximum profits. In part B of diagram, equilibrium of the firm has been shown

Fig. 7

OM is the equilibrium level of output. The firm will get only normal profits because LAC curve is tangent to AR curve at equilibrium level of output OM. If cost curve of all the firms are identical all the firms in the industry will earn only normal profits.

Under these circumstances, there will be no tendency for the firms to enter or leave the industry.

Price Determination under Perfect Competition

In the previous section we have already discuses that under perfect competition price of the commodity is determined by industry and firm has to accept the price prevailing in the market. In other words, under the perfect competition firm is a price taker and not price maker. Individual firm can not influence the price of the commodity. The question arises, how the price is determined under perfect competition. We will give the answer to this question in this part. Analysis of price determination is the effort of many economists. Classical and non-classical economists developed the ideas. Modem economists improved and further developed it. There were some differences among classical economists, regarding the price determination. Before Marshall, there were two schools of thoughts in this regard. According to Adam Smith, Ricardo etc., the believers of one school of thought, price of the commodity is determined by its cost of production or supply of the commodity.

According of Walras, Jevons etc. the believers of other school of thought, price of the commodity is determinded by its marginal utility or demand for the commodity. ñBut each school of thought took one sided view of the pricing problem.

At the end of nineteenth century, the credit of finding the true answer to pricing problem goes to Marshall, who held the view that price of the commodity is influenced by the forces of both demand and supply. In other words, the price of the commodity depends on cost of production as well as on the-marginal utility of the commodity. He linked the price determination to the cutting of a piece of paper by both blades of scissor. It means to cut a piece of paper, coordination of both blades of scissor is essential.

Equilibrium Price

The price at which quantity demanded equals quantity supplied is called equilibrium price. The quantity of goods which is bought and sold at this equilibrium price is called equilibrium amount. The intersection of demand and supply curves determines the price-quantity equilibrium. At the equilibrium price both the buyer and seller could be satisfied. If price of the commodity is more than equilibrium price, then the seller will offer more quantity for sale as compared to the demand. It means seller will not be able to sell all the commodity at that price. The seller will reduce the price of the commodity. This tendency will continue till the demand for the commodity becomes equal to supply of the commodity.

If price of the commodity is less than equilibrium price, the quantity demanded will be more than quantity supplied. It means some of the consumers would not be able to purchase the commodity. There will be the tendency of price to increase till quantity demanded becomes equal to quantity supplied. It means, the price which is determined in the markets will not be more or less than equilibrium price. The determination of equilibrium price can be explained with the help of Fig. 8.

In this Fig. on X-axis quantity and on Y-axis price have been taken. DD is the demand curve and SS is the supply curve.

E is the equilibrium price, where demand and supply curve inter-sect each other. OP is the equilibrium price and OQ is the equilibrium quantity.

If price is more than equilibrium price, suppose it is OP_1, at this price quantity demanded is P_1A and quantity supplied is P_1B. It means there is excess supply over demand equal to AB. To sell this excess supply the sellers will compete with each other and this process will bring down the price. Thus, there will be tendency for the price to fall to the level of equilibrium price OP.

Now suppose, if price of the commodity. is less than equilibrium price, at this price the quantity demanded is equal to P_2D and quantity supplied is P_2C. It means there is excess demand over supply equal to CD. The buyers will demand more commodity and there will the tendency of price to increase due to increase in demand till it becomes equal to equilibrium price.

From above it follows that if the price of the commodity is more or less than equilibrium price, certain forces in the system will operate to bring the price equal to the equilibrium price.

Effect of Changes in Demand and Supply on Price

Equilibrium price will change if either the demand or the supply curve changes due to changes in demand or supply conditions. The change in demand is due to changes in income, taste, preferences and prices of related commodities. Similarly, the change in supply is due to change in the cost of production, change in the method of production etc. The change in demand and supply will shift the position of demand and supply curves. Consequently, the equilibrium price is determined at the new position. If the supply curve remains constant, an increase in demand will shift the demand curve to right and equilibrium price will increase. On the contrary; if demand decreases the demand curve will shift to left and equilibrium price will decrease. This process can be explained with the help of the Fig. 9.

In the fig., on the X-axis quantity and on Y-axis price has been taken. DD and 88 are the demand and supply curves respectively. E is the equilibrium price where demand and supply curves intersect each other. OP and OQ are the equilibrium price and quantity respectively. If supply remains constant, increase in demand will cause the shi ft of demand curve to right of the original demand curve (DD). The new demand curve (D_1D_1) will intersect original supply curve (88) at E_1. Thus E_1 will be new equilibrium point. OP_1 and OQ_1 are the new equilibrium price and quantity respectively. Thus, the equilibrium price will increase due to increase in demand. On the contrary, if the demand decreases, it will cause the shift of demand curve to left of original demand curve (DD).

Fig. 9

The new demand curve (D_2D_2) intersects original supply curve 88 at thus, E_2 will be the equilibrium point OP_2 and OQ_2 are the new equilibrium price and quantity, respectively. It means decrease in demand will reduce the price and quantity demanded.

On the other hand, if the demand for the commodity remains constant, the increase in supply, will cause the shift of supply curve to the right of the original supply curve, the equilibrium price will fall and vice-versa. This has been explained with the help of the Fig. 10.

On the X-axis quantity and on Y-axis price have been taken. DD and SS are original demand and supply curves. Both curves intersect each

Fig. 10

other at point E. OP and OQ are the equilibrium price and quantity respectively. If now supply increases, supply curve will experience shift from SS to. S_1S_1, the equilibrium price will decrease from OP to OP_1 and equilibrium quantity will increase from OQ to OQ_1. If supply decreases, the supply curve will undergo shift from SS to S_2S_2 the equilibrium price will increase from OP and OP_2. Thus, the price will increase due to increase in supply and vice-versa.

Influence of Time Element on Price Determination

In previous part, we have analysed that under perfect competition price of the commodity is determined by the forces of demand and supply. Marshall, who propounded the theory says the price is determined by the forces of demand as will as supply. He also laid emphasis on the time element in the determination of price. According to him, time plays a vital role in the determination of the price of the commodity, because when the demand for the commodity changes the supply can not be changed in the same proportion. It takes time to bring changes in the supply of commodity.

Marshall has divided the time into .three categories from the view point of supply

1. Market Period
2. Short Period
3. Long Period
4. Secular period

It is worth mentioning that Marshall has not classified time on the basis of clock time, rather it has been done on the basis of operational time. Operational time means the time during which supply adjusts itself according to the change in demand.

Technical conditions of production do not allow the supply to adjust according to changes in demand conditions. It takes time to change size, scale and Organisation of firms as well as industry.

1. Market Period :

Market period is that period during which supply of commodity cannot be changed. It means supply can not be increased beyond the stock of the commodity. In case demand increases the supply can not be increased beyond the stock available. In market period, supply of the commodity remains constant, It is the demand that plays a Vital role in determining the price of the commodity. The price will increase due to increase in demand and vice-versa.

2. Short Period

Short period is that period under which the supply can be adjusted to a limited extent. During this time, the firms cannot bring change in the size of the plant. Production can be increased only by changing the variable factors of productions. It means production can be increased only by using the existing factors, of production intensively. . In short period also, neither the new firms can enter in industry nor the existing firms can leave the industry.

In the short period demand will influence price more as compared to supply. The reason being, supply, can be increased upto a limited extent. Supply can not be adjusted fully according to change in demand. The supply of commodity will be more in the short period as compared to that in market period.

If demand for the commodity increases in the short period, supply of commodity can increase upto a limited extent.

3. Long Period :

Long period is that period of time under which factors of production can be adjusted fully according to the change in demand. In long period, the firms can change the size of the existing plants. New firms can enter in the industry and old firms can leave the industry. Thus, in long period supply can be adjusted according to change in demand.

Secular Period :

It is also called very long period in which habits, population and technology etc. also undergo a change.

It is clear from the above analysis that time plays a vital role in determining the price of commodity. The shorter the time, the more will be the influence of demand as compared to the supply.

Determination of Market Price

Market price is the price of commodity, which prevails at any given point. Market price is determined by the equilibrium between demand and supply in a market period. Supply of commodity in the market period is limited by existing stock of the commodity. The market Period is so short that supply can not be increased in response to increase in demand supply can not be more than stock of commodity. It is not essential that whatever, is available in stock, is offered for sale. Commodity offered for sale, out of the stock, depends upon the nature of the commodity, i.e. whether it is perishable or durable one. Perishable commodities like vegetables, milk, etc can not be stored for a longer period of time, due to nature of commodity. The whole of stock is to be offered for sale, whatever may be the price of the commodity. Thus, the supply of the commodity will be perfectly inelastic in the short period. It means the supply of the commodity will remain constant. The price of the commodity is influenced by demand of commodity alone. Price of perishable commodities in the market period has been explained with the help of Fig. 11.

Fig. 11

In the fig. 11 SM is supply curve of the perishable commodity. It is parallel to Y-axis, Suppose DD is the original demand curve. It intersects supply curve at point E. OP will be the market price. If demand increases, the demand curve will shift to the right of the original demand curve (DD). E_1 will

be new equilibrium point. The price will increase from OP to OP_1. On the other hand, if demand decreases, the demand will shift to the left of the original demand curve. The new demand curve D_2D_2 intersects supply curve at point E_2. The market price will reduce from OP to OP_2.

As we have already explained, supply of the perishable commodities are limited by the existing stock. But in case of durable commodities, it is not so. The reason being, such commodities can be stored for a longer period of time. Therefore, the seller will not sell all the stock of commodity at a given price. He will wait for some time in anticipation of earning more profit. Seller will sell less quantity at low price and vice-versa. At a particular price level he will be ready to sell entire stock of commodity and at a certain minimum price, he will prefer to keep all commodity as a stock. In the former case, supply is equal to stock and in latter case, supply is equal to zero. Supply of the commodity will be elastic in these two extremes. The price at which a seller will refuse to sell his commodity is called Reserve Price.

There are several factors which govern the reserve price. These are as follows.

1. It depends upon the sellerôs expectation regarding future price of the commodity. If he expects higher future price, reserve price will be higher and vice-versa.

2. The sellerôs liquidity preference is another factor. The reserve price will be lower in case of higher liquidity preference and vice-versa.

3. The reserve price also depends upon the durability of the commodity. Higher the durability of the commodity, higher will be the reserve price.

The above analysis usually explains that market price is influenced more by demand factor. Supply of commodity remains constant during market period. The price of the commodity changes due to change is demand.

The price determination of durable commodity has been explained with the help of Fig. 12.

The supply curve RS is elastic from R to E and inelastic beyond E, Total stock of the commodity is OM_1. Suppose DD is the original demand curve. It intersects supply curve at point E. OP price is equilibrium price. At this price OM quantity is offered for sale, which is less

Fig. 12

than total stock. The seller will keep MM, quantity as a stock if the demand for commodity increases. The demand curve will shift from DD to D_1D_1. The new equilibrium price will be OP_1. At this price woe stock of the commodity will be offered for sale. On the other hand, if demand decreases, the demand curve will shift from DD to D_3D_3. The equilibrium price will be OP_3 and OM_O quantity is offered for sale. M_OM_1 amount of quantity will be kept in stock. In the fig., at OR reserve price, the supply of the commodity is zero. When the price increases beyond OR, the supply of the commodity also increases. At OP_1 price, supply of the commodity becomes perfectly inelastic.

If demand for commodity increases further, the new demand curve will be D_2D_2 and the new price will be OP_2. Because supply of the commodity can not be increased.

Determination of Short-Period Price

Short period price is determined by forces of demand and supply. Under perfect competition, supply curve of the industry in the short period is the summation of short run cost curves of the firms. Supply curve of industry is positively sloped in the short period. The supply curve of industry lies above minimum of average variable cost.

The process of price determination has been explained with the help of the Fig. 13.

In the fig. MS and SRS are market period and short run supply curves respectively. D_1D_1 is the original demand curve. E_1 is the equilibrium point, where demand curve (DD) and short run supply curve (SRS)

Fig. 13

intersect each other. OP_1 is the equilibrium price. This is also market price because market supply curve (MS) also intersects demand curve (D_1D_1) at point E_1. If demand increases, the new demand curve (D_2D_2) intersects market supply curve (MS) at point E_3 and short run supply curve (SRS) at point E_2**.** The price in the short period increases from OP_1 to OP_2 and supply of the commodity increases from OQ to OQ_1. In comparison to short run price, market price is fixed at OP_3**.** which is higher than short run period price due to inelastic supply of commodity. But the short period price OP, is higher than original price OP. The reason being, when the production is increased, the marginal cost of commodity increases.

On the contrary, when the demand decreases, it will cause shift of the demand curve to the left of original demand curve (D_1D_1). The new demand curve will be D_3D_3. The new demand curve (D_3D_3) intersects market supply (MS) at A and short period supply curve (SRS) at E_O**.**

OP_O and OB will be price in the short period and market period respectively. The price in the short period is higher than the market period. The qua1J.tity supply in the short period will decline from OQ to OQ_O, where it is fixed at OQ in the market period. The reason being in the short period production can be changed by changing the variable factors of production, whereas, it is not feasible in the market period.

Determination of Long Period or Normal Price

Long period price is also known as normal price. Normal price is determined by the long run forces of demand and supply. Firms in the industry can vary the size of plant. New firms can enter in the industry and existing firms can leave the industry. Supply can be adjusted fully according to the change in demand. Normal price never remains constant. Normal price undergoes change with change in demand and supply forces. The process of normal price determination has been explained in the Fig. 14.

In the fig., LRS and MPS are the long run supply curve and market period supply curve respectively. D_OD_O is the original demand curve. E_O is the equilibrium price and OM_O is the equilibrium output. With increase in demand, the new demand curve will be D_1D_1 and this demand curve cuts LRS curve at point E_2. OP_2 will be new equilibrium normal price, the new demand curve (D_1D_1) also intersects MPS curve at E_1 and

Fig. 14

market price is fixed at OP_1. The fig. vividly reveals that increase in normal price is less as compared to market price. The reason being, in market period, supply of commodity remains constant whereas it is not so in the long run. It means in long run supply can be changed according to change in demand. Supply of commodity also increases from OM to OM_2 at OP_2 price.

On the contrary, if demand decreases, the new demand curve will be D_2D_2 It intersects LRS and MPS curves at points E_4 and E_3 respectively. New equilibrium prices will be OP_4 and OP_3 respectively. Normal price declines with decrease in demand.

Normal Price and Returns to Scale

Long run normal price is determined by the long run equilibrium between demand and supply. In the long run the supply curve does not have any definite slope. The reason being, cost of production is influenced by returns to scale. Thus, the slope of supply curve will be different accordingly. In the long run, under perfect competition supply curve can have three possible slopes.

i) When the production in the industry is according to increasing-returns or diminishing cost the slope of the industry supply curve will be negative.

ii) When the production in the industry is according to diminishing returns or increasing cost, the slope of the industry supply curve will be negative.

iii) When the production in industry is according to constant returns or constant cost, the supply curve will be parallel to X-axis.

Thus returns to scale influence the normal price to a consider able extent.

Determination of Normal Price in Decreasing Returns or Increasing Cost Industry

Increasing cost industry means when the size of the industry expands, the cost of production of firms in the industry enhances considerably. The reason being, it experiences certain external economies and diseconomies. But diseconomies in case of increasing cost industry overweight the external economies. It will cause increase in the cost of production. Consequently, supply curve of the industry rises from left to right. When industry expands then average minimum cost of production of the firms enhances. Normal price under increasing cost has been explained in the Fig. 15.

Fig. 15

In the fig. LRS supply curve has a positive slope. DD is the original demand curve. OP is the equilibrium curve. If demand increases, it will cause the shift of demand curve to right of original demand curve (DD). The new demand curve (D_1D_1) intersects (LRS) supply curve at point E. OP_1 is the equilibrium price. Price increases due to increase in cost of production. On the contrary, if demand decreases demand curve will shift to left of the original demand curve (DD). The new demand curve (D_2D_2) intersects LRS curve at point E_2. The equilibrium price reduces from OP to OP_2. The industry supplies OM_2 quantity at this equilibrium price.

Determination of Normal Price in Increasing Returns or Diminishing Cost Industry

Diminishing cost industry means when industry expands, the cost of production of firms in the industry declines. The reason being, firms experience more external economies as compared to external .diseconomies. In other words external economies overweigh the external diseconomies in case of diminishing cost industry. Thus, in case of a decreasing cost industry, the additional supplies of the product will be forthcoming at reduced cost. The supply curve of industry will have a negative slope.

The normal price in case of diminishing cost industry has been explained with the help of Fig. 16. In the fig LRS supply curve slopes

Fig. 16

downwards from left to right. DD is the original demand curve. E is the equilibrium point and OP is the equilibrium price.

If demand increases, the demand curve will shift from DD to D_1D_1. The new demand curve (D_1D_1) intersects LRS supply curve at point E_1. OP_1 is the new equilibrium price, which is lower than the original equilibrium price (OP). The reason being; law of increasing returns operates in the industry. On the contrary, if demand decreases, the demand curve will shift to the left of original demand curve (DD). The new demand curve (D_2D_2) intersects (LRS) supply curve at point E_2. The new equilibrium price is OP_2, which is higher than .the original equilibrium price (OP).

Supply will decrease, due to decrease in demand. Therefore, in case diminishing cost industry, with increase in demand industry offers more quantity at reduced price.

Determination of Normal Price in Constant Returns or Constant Cost Industry

Constant cost industry is that industry in which external commodities as well as diseconomies cancel each other. The cost of firms in the industry remains constant, with the change in industry.

The long run supply curve of the constant cost industry is a horizontal straight line at the level of long run minimum average cost.

The process of price determination under constant cost industry has been explained in the Fig. 17.

Fig. 17

In the fig. LRS supply curve is the horizontal straight line parallel to X-axis. DD is the original demand curve, OP is the equilibrium price. If demand increases the demand curve will shift from DD to D_1D_1**.** The new demand curve

(D_1D_1) intersects LRS curve at point E_1. The above diagram reveals that the price has not undergone any change despite the increase in demand. But the supply of the commodity has increased from OM to OM_1. The reason being there is no change in cost of production. On the contrary, if demand decreases, the demand curve will shift to left of original demand curve (DD). E_2 is the equilibrium point. In this situation, there is no change in price of the commodity, but supply of the commodity decreases. Thus, under the constant cost industry, with the change in demand the price of the commodity remains constant. The reason being, cost of product of industry remains constant.

From the above discussion, it is clear that as demand increases, the long run normal price increases, remains the same or decreases depending upon whether the industry in question is an increasing cost, constant cost or decreasing cost industry.

Questions :

1. Explain equilibrium of the firm under perfect competition in the short run and the long run.

2. What is the meaning of equilibrium of Industry? Describe the equilibrium of industry under perfect competition.

   (i) in the short run and

   (ii) in the long run

3. Describe the importance of time element in the determination of price.

4. Give the meaning of normal price. How is the normal price is determined?

Selected References

1. Alfred Marshall : Principles of Economics.

2. Stonier and Hague : A Text Book of Economic Theory.

3 C.E. Ferguson : Microeconomic Theory.

4. Richard H. Leftwich : The Price System and Resource Allocation.

Lesson - 10

PRICE DETERMINATION UNDER DUOPOLY AND OLIGOPOLY

(Author : N.K. Bishnoi)

(A) Meaning of Duopoly

Duopoly is a special case of the theory of oligopoly in which there are only two sellers. Both the sellers are completely independent and no agreement exists between them. Even thought they are independent, a change in he price and output of one will affect the other, and may set a chain of reactions. A seller may, however, assume that his rival is unaffected by what he does, in that case he takes only his own direct influence on the price. Thus the duopoly problem can be considered as either ignoring mutual dependence or recognizing it. The Cournot solution refers to the former where mutual dependence is ignored while the Chamberlin solution relates to the latter problem where mutual dependence is recognized.

The Cournot Model

The oldest *determinate* solution to the duopoly problem is by the French economist, A.A. Cournot in 1838, who took the case of two mineral water springs situated side by side and owned by two firms A and B.

Its Assumptions

The Cournot model is based on the following assumptions :

(i) There are two independent sellers. In other words, interdependence of the duopolists is ignored.

(ii) They produce and sell a homogeneous product-mineral water.

(iii) The total output must be sold out, being perishable and non storable.

(iv) The number of buyers is large.

(v) The cost of production is assumed to be zero.

(vi) Both have identical costs and identical demands.

(vii) Each seller decides about the quantity he wants to produce and sell in each period.

(viii) But each is ignorant about his rivalôs plan about output.

(ix) At the same time, each seller takes the supply/output of its rival as constant.

(x) Neither of them fixes the price for its product, but each accepts the market demand price at which the product can be sold.

(xi) The entry of other firms is blocked.

(xii) Each seller aims at obtaining the maximum net revenue or profit.

Given these assumptions, suppose there are two mineral water springs exploited by two firms, A and B. The market demand curve is DD. and its marginal revenue curve is MR. as shown in Figure (10.1). The marginal costs of both A and B are assumed to be zero, so that they coincide with the horizontal axis. Suppose firm A is the only producer in which case it produces and sells OA (=½OD_1) quantity when its MR. equals its marginal cost curve (horizontal axis) at point A. It charges the monopoly price AS (=OP) and earns OASP as monopoly profits. Now firm B enters the market and expects that A will not change its output level OA. It, therefore regards SD_1 segment of the market demand cure as its demand curve. Its corresponding marginal revenue curve is MR_2 which interests the horizontal axis (its marginal cost curve) at point B. Thus it produces and sells AB (=½AD_1=BD_1) quantity at BO (=OP_1) price and it expects to earn BGTA profits.

Figure 10.1

Firm A finds that with the entry of B, price has fallen to OP_1 from OP. As a result, its expected profits decline to OP_1TA. In this situation, it tries to adjust its price and output. Accordingly, assuming that B will continue to sell the same quantity AB (=BD_1), it regards the remaining portion of the market OB available to it. It thus sell ½OB. The reduction in its output from OA (=½OD_1) to ½OB causes the price to rise (not shown in the figure to simplify the analysis). As a result, to Aôs reduction in output, B reacts by increasing its output to ½ (OD_1 = ½OB) which causes the price to rise and Bôs reaction in increasing its output and causing the price to fall, will ultimately lead to an equilibrium price OP_2**.** At this price, the total output of mineral water is OF, which is equally divided between the two firms. Each duopolist sells 1/3 of the market i.e. A sells OC and B sells CF. At this price, Aôs profits $OCLP_2$ equal that of Bôs profits CFRL.

It is evident that both the producers sell 2/3 of the total output, OPI. If there are *n* producers, the equilibrium output would be *n/n*+1 times of the total output. The total output of both the producers A and B is 2/(2+1) = 2/3.

Let us compare the Cournot duopoly solution with the perfectly competitive solution. The duopoly firms A and B, in equilibrium charge OP_2 price and sell OF output. Under perfect competition, the total output will be OD_1 at zero price. The price is zero because the marginal cost is zero. When the MRI curve intersects the horizontal axis, which is the MC curve, the price is zero at point A in the figure. The total output OD_1 will be divided between A and B equally as OA and AD_1. Notice that in the Cournot solution, the price OP_2 exceeds the zero marginal cost and the price under perfect competition and the output OF is less than OD_1 under perfect competition. However, in the Cournot solution the output (OF) is greater than it would under monopoly (OA). But the price under monopoly (OP) would be higher than under the Cournot solution (OP_2). Algebraically, in the Cournot solution, output will be 4/3 of the monopoly output and 2/3 of the perfectly competitive output.

Conclusion

The Cournot model can be extended even to more, than two firms. As more and more firms enter the oligopoly industry, the equilibrium output and price of the industry will approach the perfectly competitive output OD_1 and the zero price.

Its Criticisms

The main defect in Cournotôs solution is that each seller assumes his rivalôs supply fixed despite repeatedly observing changes in it. Joseph Bertrand, a French mathematician, criticizing Cournot in 1883 pointed out the seller A in order to regain all the customers lost to B, will fix a price slightly below that fixed by B and price cutting may continue until the price becomes zero. Thus, Bertrand argued that there would not be any limit to the fall in price since each seller could by doubling his produce, underbid his rival. This would tend to drive down the price to the competitive level in the long-run.

Second, the model is silent about the period within which one firm reacts and adjusts its output to the moves of the other. Thus, it is a static model.

Third, it is closed model because is does not allow entry of firms.

Fourth, the assumption that each duopolist can act without any output reaction from the other is unrealistic. It is, in fact, a no-learning-ly-doing model.

The Chamberlin Model

Prof. Chamberlin proposed a stable duopoly solution recognizing mutual dependence between the two sellers. He criticized and rejected the Cournot model on the ground that it does not conform perfectly to the hypothesis that each seller acts so as to render his profit a maximum. In order to do this, he will take account of his total influence upon the price, indirect as well as direct. When a seller remains passive to changes in price or output of his rival, it is a direct influence. On the other hand, when a seller reacts to the price or output changes of his rival and changes his own price or output, the influence is indirect. According to Chamberlin, when interpendence is recognized between sellers both direct and indirect influences of a change in the price or output of a seller lead to a stable industry equilibrium with monopoly price and output.

The Chamberlin solution can be explained both in terms of output adjustment and price adjustment. Let us take Figure (10.2) where seller A enters the market as a monopolist first as in the Cournot model and maximizes his profit by selling OA output at OP_1 price, thereby earning $OASP_1$ monopoly profit. Seller B enters the market after him and considers

SD_1 segment of the market demand curve (DD_1) as his demand curve. Under the Cournot assumption that his rival A will not change his output, he will sell AB output at OP_2 (=BG) price. The difference arises in Chamberlinôs solution from this point. In the Cournot model, each rival acts independently. But Chamberlin assumes their interdependence. So seller A does not react to Bôs move and compromises with the existence of B. Accordingly, he decides to reduce his output from OA to OE equal to Bôs output AB. Seller B also recognizes interdependence and realizes that by selling E output at a higher price OP_1, he will share the monopoly profit. Thus by recognizing their interdependence, each seller shares equally the monopoly industry output OA, seller A selling OE and seller B selling EA. They also share the total monopoly profit $OASP_1$ equally between them, A earning $OEKP_1$ and B earning EASK at the monopoly price OP_1**.** Thus OP_1 (=AS) is a perfectly stable price, for either seller can bring disaster upon himself as well as upon his rival by behaving differently.

Figure 10.2

Chamberlin also shows that the result would be identical if sellers adjust their prices rather than their supplies. Suppose the price is anywhere between OP_1 and OP_2 In order to earn maximum profit, if. A increases his price to OP_1. B will at once follow suit. Thus B also raises his price to OP_1 to get the maximum possible profit. Once the price OP_1 is set no one will cut it, for each seller realizes that by so doing, he will reduce his profit. The equilibrium is again, stable and determinate.

The Chamberlin solution involves a kind of agreement between the two sellers. They do not sign it, but each seller is intelligent enough to realize the importance of mutual dependence. Each acts rationally, looks beyond his nose and understands that sharing monopoly profit is to the best of his advantage. Thus, in Chamberlinôs model the sellers are independent, yet they are in a kind of collusion which leads to stable equilibrium, a sort of monopoly equilibrium.

Its Criticisms

The Chamberlin model is also not free from certain weaknesses.

1. Like the Cournot model, it ignores entry of firms and is thus a closed model.

2. This model involves joint-profit sharing with zero enforcement costs by rivals. But problems might arise in sharing profit by the two sellers.

3. Fellner does not agree with Chamberlin that monopoly (Joint Profit Maximization) solution is possible under duopoly interdependence. A (202) firm often underestimates the elasticity of the market demand curve. The underestimation of the market demand curve leads to the wrong estimation of the market MR curve. This may lead the collusive firms to charge a price higher than the monopoly price. A high price yielding very high profits may lead to the entry of firms into the industry, thereby make the Chamberlin solution an impossibility.

(B) Meaning of Oligopoly

Oligopoly is a market situation in which there are a few firms selling homogeneous or differentiated products. It is difficult to pinpoint the number of firms in the oligopolist market. There may be threeò four of five firms. It is also known as competition among the few. With only a few firms in the market, the action of one firm is likely to affect the others. An oligopoly industry produces either a homogeneous product or heterogeneous products. The former is called pure or perfect oligopoly and the latter is called imperfect or differentiated oligopoly. Pure oligopoly is found primarily among producers of aluminium, cement, copper, steel, electricity, etc. Differentiated oligopoly is found among

producers of such consumer goods as automobiles, cigarettes, soaps & detergents, TVs, rubber tyres, refrigerators, etc.

Characteristics of Oligopoly

In addition to fewness of sellers, most oligopolistic industries have several common characteristics which are explained below.

(1) Interdependence : There is recognized interdependence among the sellers in the oligopolistic market. Each oligopolist rum knows that changes in its price, advertising, products characteristics, etc. may lead to countermoves by rivals. When the sellers are few, each produces a considerable fraction of the total output of the industry and can have a noticeable effect on market conditions. He can reduce or increase the price for the whole oligopolist market by selling more quantity or less and affect the profits of the other sellers. Each seller has direct and ascertainable influence upon every other seller in the industry. Thus, every move by one seller leads to countermoves by the others.

(2) Advertisement : The main reason for this mutual interdependence in decision making is that one producerôs fortunes are dependent on the policies and fortunes of the other producers in the industry. It is for this reason that oligopolist firms spend much on advertisement and customer services. For example, if all oligopolists continue to spend a lot on advertising their products and one seller does not match up with them, he will find his customers gradually going in for his rivalôs product. If, on the other hand, one o.1igopolist advertises his product, others have to follow him to keep up their sales.

(3) Competition : This leads to another feature of the oligopolistic market, the presence of competition. Since under oligopoly, there are a few sellers, a move by one seller immediately affects the rivals. So each seller is always on the alert and keeps a close watch over the moves of its rivals in order to have a counter move. This leads to intense competition on the basis of advertisement quality improvement, cost reduction, better service / delivery etc.

(4) Barriers to Entry of Firms : As there is keen competition in an oligopolistic industry, there are no barriers to entry into or exit from it in legal

sense. However, in the long-run, there are some other types of barriers to entry which tend to restrain new firms from entering the industry. They may be : (a) economies of scale enjoyed by a few large firms; (b) control over essential and specialized inputs; (c:) high capital requirements due to plant costs, advertising costs, etc. (d) exclusive patents; and licenses; (e) government policy i.e. licence, permit and the other control measures. When entry is restricted or blocked by such natural and/or artificial barriers the oligopolistic industry can earn long-run supernormal profits.

(5) Indeterminate Demand Curve : It is not easy to trace the demand curve for the product of an oligopolist. Since under oligopoly the exact behaviour pattern of a producer cannot be ascertained with certainty, his demand curve cannot be drawn accurately and with definiteness. How does an individual sellerôs demand curve look like in oligopoly is most uncertain because a sellerôs price or output moves led to unpredictable reactions on price output policies of his rivals, which may have further repercussions on his price and output. The chain of action reaction as a result of an initial change in price or output is all a guess-work.

Price Determination Under Oligopoly

With these characteristics of oligopoly in the background, we study the determination of prices and outputs by oligopolistic firms. Prof. Machlup has given a detailed classification of oligopolies. But we shall confine our study to the non-collusive oligopoly model of Sweezy (the kinked demand curve) and to the collusive oligopoly models relating to cartels and price leadership.

I. Non-collusive Oligopoly

1. The Sweezy Model of Kinked Demand Curve

In his article published in 1939, Prof. P.A. Sweezy presented the kinked demand curve analysis to explain price rigidities often observed in oligopolistic markets.

Sweezy assumes that if the oligopolistic firm lowers its price, its rivals will react by matching that price cut in order to avoid losing their customers.

Thus the firm lowering the price will not be able to increase its demand much. This portion of its demand curve is relatively inelastic. On the other hand, if the oligopolistic firm increases its price, its rivals will not follow it and raise their prices. Thus the quantity demanded of this firm will fall considerably. This portion of the demand curve is relatively elastic. In these two situations, the demand curve of the oligopolistic firm has a kink at the prevailing market price which explains price rigidity.

Its Assumptions

(i) There are few firms in the oligopolistic industry.

(ii) The product is of the same quality.

(iii) There is an established or prevailing market price for the product at which all the sellers are satisfied.

(iv) Any attempt on the part of a seller to push up his sales by reducing the price of his product will be counteracted by other sellers who will follow his move. If he raises the price others will not follow him, rather they will stick to the prevailing price and cater to the customers, leaving the price-raising seller alone.

(v) The marginal cost curve passes through the dotted portion of the marginal revenue curve so that changes in marginal cost do not affect output and price.

The Model

Given these assumptions, the price-output relationship in the oligopolist market is explained in Figure (10.3) where KPD is the kinked demand curve and OP_0 is the prevailing price in the oligopoly market for the OR product of one seller. Starting from point P_1 corresponding to the current price OP_0, any increase in price above it will considerably reduce his sales, as his rivals are not expected to follow his price increase. This is so because the KP portion of the kinked demand curve is elastic, and the corresponding portion KA of the MR

curve is positive. Therefore, any price increase will not only reduce his total sales but also his total revenue and profit.

Figure 10.3

On the other hand, if the seller reduces the price of the product below OP_0 (or P), his rivals will also reduce their prices. Though he will increase his sales, his profit would be less than before. The reason is that the PD portion of the kinked demand curve below P is less elastic and the corresponding part of marginal revenue curve below R is negative. Thus in both the price raising and price-reducing situations the seller will be a loser. He would stick to the prevailing market price OP_0 which remains rigid.

In order to study the working of the kinked demand curve, let us analyze the effect of changes in cost and demand conditions on price stability in the oligopolistic market.

Changes in Costs : In oligopoly under the kinked demand curve analysis changes in costs within a certain range do not affect the prevailing price.

Suppose the cost of production falls so that the new MC curve is MC. to the right, as in Figure (10.4). It cuts the MR curve in the gap AB so that the profit maximizing output is OR which can be sold at OP_0 price. It should be noted that with any cost reduction the new MC curve will always cut the MR curve in the gap because as costs fall the gap AB continues to widen due to two reasons: (i) As costs fall, the upper portion KP of the demand curve becomes more elastic because of the greater certainty that a price rise by one seller will not be followed by rivals and his sales would be considerably reduced. (ii) With the reduction in costs the lower portion PD of the kinked curve becomes more inelastic, because of the greater certainty that a price reduction by one seller will be followed by the other rivals.

Figure

Thus the angle KPD tends to be a right angle at P and the gap AB widens so that any MC curve below point A will cut the marginal revenue curve inside the gap. The net result is the same output OR at the same price OP_0 and large profits for the oligopolistic sellers.

In case the cost of production rises the marginal cost curve will shift to the left of the old curve MC as MC_2. So long as the higher MC curve intersects the MR curve within the gap upto point A, the price situation will be rigid., However, with the rise in costs the price is not likely to remain stable indefinitely and if the MC curve rise above point A, it will intersect the MC curve in the portion KA so that a lesser quantity is sold at a higher price. We may conclude that there may be price stability under oligopoly even when costs change so long as the MC curve cuts the MR curve in its discontinuous portion. However, chances of the existence of pricer rigidity are greater where there is a reduction in costs than there is a rise in costs.

The analysis of the kinked demand curve points out that price rigidity in oligopolistic markets is likely to prevail if there is a price reduction move on the part of all sellers. Changes in costs and demand also lead to price stability under normal conditions so long as the MC curve intersects the MR curve in its discontinuous portion. But price increase rather than price rigidity may be found in response to rising cost or increased demand.

Reasons for Price Stability

There are a number of reasons for price rigidity in certain oligopoly markets. First, individual sellers in an oligopolistic industry might have learnt through experience the futility of price wars and thus prefer price stability. Second, they may be content with the current prices, outputs and profits and avoid any involvement in unnecessary insecurity and uncertainty. Third, they may also prefer to stick to the present price level to prevent new firms from

entering the industry. Fourth, the sellers may intensify their sales promotion efforts at the current price instead of reducing it. They may view non-price competition better than price rivalry. Fifth, after spending a lot of money on advertising his product, a seller may not like to raise its price to deprive himself of the fruits of his hard labour. Naturally, he would stick to the going price of the product and in the last, if a stable price has been set through agreement or collusion, no seller would like to disturb it, for fear of unleashing a price war and thus engulfing himself into an era of uncertainty and insecurity.

Its Shortcomings

But the theory of kinked demand curve in oligopoly pricing is not without shortcomings.

(i) Even if we accept all its assumptions it is not likely that the gap in the marginal revenue curve will be wide enough for the marginal cost curve to pass through it. It may be shortened even under conditions of fall in demand or costs thereby making price unstable.

(ii) Price stability may be illusory because it is not based on the actual market behaviour. Sales do not always occur at list prices. There are often deviations from posted prices because of trade-ins, allowances and secret price concessions. The oligopolistic seller may outwardly keep the price stable but he may reduce the quality or quantity of the product. Thus price stability becomes illusory.

(iii) Critics point out that the kinked demand curve analysis holds during the short-run, when the knowledge about the reactions of rivals are low. But it is difficult to guess correctly the rivalsô reactions in the long-run. Thus the theory is not applicable in the long-run.

(iv) Prof. Stigler points out that case in oligopoly industries where the number of sellers is either very small or somewhat large, the kinked demand curve is not likely to be there. He concludes that ñthe empirical evidence reveals neither price experiences that would lead oligopolists to believe in the existence of a kink nor the pattern of changes of price quotations that the theory

leads us to expect.ò Thus the empirical evidence does not support the existence of a kink.

However, the analysis does show how the oligopolistic firmôs view of competitive reaction patterns can affect the changeability of whatever price it happens to be charging.

II. Collusive Oligopoly

Collusive oligopoly is a situation in which firms in a particular industry decide to join together as a single unit for the purpose of maximizing their joint profits and to negotiate among themselves so as to share the market. The former is known as the joint profit maximization cartel and the latter as the market-sharing cartel. There is another type of collusion, known as leadership, which is based on tacit agreements. Under it, one firm acts as the price leader and fixes the price for the product while other firms follow it. Price leadership is of three types : low-cost firm, dominant firm, and barometric firm.

(i) Cartels

A cartel is an association of independent firms within the same industry. The cartel follows common policies relating to prices, outputs, sales and profit maximization and distribution of products. Cartels may be voluntary or compulsory, open or secret depending upon the policy of the government with regard to their formation. Thus cartels have many forms and use many devices in order to follow varied common policies depending upon the type of the cartel. We discuss below the two most common types of cartels: (i) joint profit maximization or perfect cartel; and (ii) market-sharing cartel.

1. Joint Profit Maximization Cartel

The uncertainty to be found in an oligopolistic market provides an incentive to rival firms to form a perfect cartel. Perfect cartel is an extreme form of perfect collusion. In this, firms producing a homogeneous product form a centralized cartel board in the industry. The individual firms surrender their price-output decisions to this central board. The board determines output quotas for its members, the price to be charged and the distribution of industry profits.

Since the central board manipulates prices, outputs, sales and distribution of profits, it acts like a single monopoly whose main aim is to maximize the joint profits of the oligopolistic industry.

Its Assumptions

The analysis of joint profit maximization cartel is based on the following assumptions :

(i) Only two firms A and B are assumed in the oligopolistic industry that form the cartel.

(ii) Each firm produces and sells a homogeneous product that is a perfect substitute for each other.

(iii) The number of buyers is large.

(iv) The cartel aims at joint profit maximization.

Joint Profit Maximization Solution

Given these assumptions, and given the market demand curve and its corresponding MR curve, joint profits will be maximized when the industry MR equals the industry MC. Figure (10.5) illustrates this situation where D is the market (or cartel) demand curve and MR is its corresponding marginal revenue curve; The aggregate marginal cost curve of the industry SMC is drawn by the lateral summation of the MC curves of firms A and B, so the SMC = MC_a + MC_b. The cartel solution that maximizes joint profit is determined at point E where the SMC curve intersects the industry MR curve. Consequently, the total output is OQ which will be sold at OP = (QF) price. As under monopoly, the cartel board will allocate the industry output by equating the industry MR to the marginal cost of each firm. The share of each firm in the industry output is obtained by drawing a straight line from E to the vertical axis which passes through the curves MC_b and MC a of firms B and A at points E_b and E_a respectively. Thus the share of firm A is OQ_a and that of firm B is OQ_b which equal the total output OQ (=OQ_a + OQ_b). The price OP and the output OQ distributed between A and B firms in the ratio of OQ_a **:** OQ_b is the monopoly

solution. Firm A with the lower costs sells a larger output OQ_a the firm B with hirer costs so that $OQ_a > OQ_b$. But this does not man that A will be getting more profit than B. The Joint maximum profit is the sum of RSTP and ABCP earned by A and B respectively. It will be pooled into a fund and distributed by the cartel board according to the agreement arrived at by the two firms at the time of the formation of the cartel. A pooling agreement of this type will make it possible for both firms to maximize their joint profit provided the total profits earned by them independently do not exceed the former.

Figure 10.5

Thus perfect collusion by oligopolistic firms in the form of a cartel has certain advantages. It avoids price wars among rivals. The firms forming a cartel gain at the expense of customers who are charged a high price for the product. The cartel operates like a monopoly organization which maximizes the joint profit of firms. Joint profits are generally more than the total profits earned by them if they were to act independently.

Difficulties of a Cartel

The above analysis is based on perfect collusion in which all firms relinquish their individual price-output decisions to a central board of the cartel which acts like a multi-plant monopolist. But this is only a theoretical possibility in the short-run because in practice the joint profit maximization objective cannot be achieved by a cartel. In the long-run, there are a number of difficulties faced by a cartel which tend to break it down. They are as under :

(i) It is difficult to make an accurate estimate of the market demand curve. Each firm thinks that its own demand curve is more elastic than the market demand curve because its product is a perfect substitute for the product of its rivals. Thus if the market demand curve is underestimated so will be its

corresponding MR curve which will make the estimation of the market price inaccurate by the cartel.

(ii) Similarly, the estimation of the market MC curve may be inaccurate because of the supply of wrong data about their MC by individual firms to the cartel. There is every possibility that the individual firms may supply low-cost data to the central cartel board in order to have a larger share of output and profits. This may ultimately lead to the break down of the cartel.

(iii) The formation of a cartel is a slow process which takes a long time for the agreement to arrive at by firms especially if their number is very large. In the meantime, there may be changes in the cost structure and market demand for the product. This renders the cartel agreement useless and it breaks down soon.

(iv) If a firmôs product is preferred more by consumers than that of the other members of the cartel, the market demand for it may be higher than the quota fixed by the cartel. It may, therefore, secretly sell more than its quota and if followed by other firms, the cartel will break down.

(v) The larger the number of firms in a cartel, the less are its chances of survival for long because of the distrust, threatening and bargaining resorted to by them. The cartel will, therefore, break down.

(vi) When a cartel raises the price of the product and increases the profits of its members, it creates an incentive for new firms to enter the industry. Even if the entry of new firms is blocked, it is only a short-run phenomenon because the success of the cartel will lead to the entry of firms in the long-run. This will force the cartel to break down. It the new firms are allowed to enter the cartel, it will become unmanageable, increase the defectors and brings its end.

Thus the chances are greater for individual firms to leave the cartel on account of personal bickerings and antagonism of member firms over allotment of qU0tas and division of profits which are likely to affect adversely joint profit maximization and end the cartel agreement.

Besides these problems in the working of a cartel, it is more difficult to form and run a cartel for long in the case of a differentiated product than in the case of a homogeneous product. For, it is not possible to rationalize and sort out the differences in the qualities of the product.

(ii) Market-Sharing Cartel

Another type of perfect collusion in an oligopolistic market is found in practice which relates to market-sharing by the member firms of a cartel. The firms enter into a market-sharing agreement to form a cartel but keep a considerable degree of freedom concerning the style of their output, their selling activities and other decisions. There are two main methods of market-sharing : (a) non-price competition; and (b) quota system. They are discussed as under :

(a) Non-Price Competition Cartel : The non-price competition agreement among oligopolistic firms is a loose form of cartel. Under this type of cartel, the low cost firms press for a low price and the high cost firms for a high price. But ultimately, they agree upon a common price below which they will not sell. Such a price must allow them some profits. The firms can compete with one another on a non-price basis by varying the colour, design, shape, packing, etc. of their product and having their own different advertising and other selling activities. Thus each firm shares the market on a non-price basis while selling the product at the agreed common price.

This type of cartel is inherently unstable because if one low-cost firm cheats the other firms by charging a lower price than the common price, it will attract the customers of other member firms and earn larger profits. When other firms come to know of this, they will leave the cartel. A price war will start and ultimately the lowest-cost firm will remain in the industry.

In case the cost curves of the firms forming a cartel differ, the low-cost firms may not stick to the common price. They may try to increase their share of the market by means of secret price concessions. They may also resort to better sales promotion methods. Such policies tend to change their demand-cost conditions further. Consequently, price variations among firms become more

common. Ultimately, the cartel agreement becomes a farce and a price war starts. This leads to be breaking up of the cartel agreement.

(b) Market Sharing by Quota Agreement : The second method of market sharing is the quota agreement among firms. All firms in an oligopolistic industry enter into a collusion for charging an agreed uniform price. But the main agreement relates to the sharing of the market equally among member firms so that each firm gets profits on its sales.

Its Assumptions

This analysis is based on the following assumptions :

(i) There are only two firms that enter into market-sharing agreement on the basis of the quota system.

(ii) Each firm produces and sells a homogeneous product which is a perfect substitute for each other.

(iii) The number of buyers is large.

(iv) The cost curves of the two firms arc identical.

(v) Both firms share the market equally.

(vi) Each sells the product at the agreed uniform price.

(vii) There is not threat of entry by new firms.

Market-Sharing Solution

Given these assumptions, he equal market sharing between the two firms is explained in terms of Figure (10.6) where D is the market demand curve and d/MR is its corresponding MR curve ΣMC is the aggregate MC curve of the industry. The ΣMC curve intersects the d/MR curve at point E which determines QA (=OP) price and total output OQ for the industry. This is the monopoly solution in the market-sharing cartel.

Figure 1 0.6

How will the industry output be shared equally between the two firms? Now assume that the d/MR is the demand curve of each firm and mr is its corresponding MR curve. AC and MC are their identical cost curves. The MC curve intersects the mr curve at point e so that the profit maximization output of each firm is Oq. Since the total output of the industry is OQ which is equal to 2 × Oq = (OQ = 20q), it is equally shared by the two firms as per the quota agreement between them. Thus each sells Oq output at the same price qB (OP) and earns RP per unit profit. The total profit earned by each firm is RP × Oq and by both is RP × 20q or RP × OQ.

However, in actuality, there are more than two firms in an oligopolistic industry which do not share the market equally. Moreover, their cost curves are also not identical. In case their cost curves differ, their market shares will also differ. Each firm will charge an independent price in accordance with its own MC and MR curves. They may not sell the same quantity at the agreed common price. They may be charging a price slightly above or below the profit maximization price depending upon its cost conditions. But each will try to be nearest the profit maximization price. This will ultimately lead to the breaking up of the market sharing agreement.

We may conclude that collusive oligopoly pricing has not any set pattern of price behaviour. The resultant price and output will depend upon the reaction of the collusive oligopolists towards the profit maximization price and their attitude towards the existing and potential rivals.

(iii) Price Leadership

Price leadership is imperfect collusion among the oligopolistic firms in an industry when all firms follow the lead of one big firm.

There is a tacit agreement among the firms to sell the product at a price set by the leader of the industry. Sometimes there is a formal meeting and a definite agreement with the leader-firm. If the products are homogeneous, a

uniform price is established. In case of a differentiated product also prices can be uniform. Whatever price changes take place the leader announces from time to time, and the other firms follow him.

Price leadership is of various types. But there are three most common price leadership models which we discuss now.

(a) The Low-Cost Price Leadership Model

In the low-cost price leadership model, an oligopolistic firm having lower costs than the other firms sets a lower price which the other firms have to follow. Thus the low-cost firm becomes the price leader. Maruti Car, Bajaj Scooter are prominent example of this type in India.

Its Assumptions

The low-cost firm model is based on the following assumptions :

(i) There are two firms A and B.

(ii) Their costs differ. A is the low-cost firm and B is the high-cost firm.

(iii) They have identical demand and MR curves. The demand curve faced by them is 1/2 of the market demand curve.

(iv) The number of buyers is large.

(v) The market industry demand curve for the product is known to both the firms.

The Model

Given these assumptions, both firms enter into a tacit agreement whereby the high-cost firm B will follow the price set by the price leader firm A and to share the market equally. The price policy to be followed by both is illustrated in. Figure (10.7). D is the industry demand curve and d/MR is its corresponding marginal revenue curve which is the demand curve for both the firms and mr is their marginal revenue curve. The cost curves of the low-cost firm A are AC and MC and of the high cost firm B are AB_a and MC_b.

Figure 10.7

If the two firms were to act independently, the high cost firm B would charge OP price per unit and sell OQ_b quantity, as determined by point B where its MC_b curve cuts the mr curve. Similarly, the low-cost firm A would charge OP_1 price per unit and sell OQ_a quantity, as determined by point A where its MC a curve cuts the mr curve. As there is a tacit agreement between the two firms, the high-cost firm B has no choice but to follow the price leader firm A. It will, therefore, sell OQ quantity at a lower price OP_1 even though it will not be earning maximum profits. Oil the other hand, the price leader A will earn much higher profits at OP_1 price by selling OQ_a quantity. Since both A and B sell the same quantity OQ_a the total market demand OQ is equally divided between the two, $OQ = 20Q_a$**.** But if firm B sticks to OP price, its sales will be zero because the product being homogeneous, all its customers will shift to firm A.

The price-leader firm A can, however, drive firm Bout of the market by setting a lower price than OP_1, lower than the average cost AC_b of firm B. Firm A :would become a monopoly firm. But in such a situation it will have to face legal problems. Therefore, it will be in its interest to fix OP_1 price and tolerate firm B in order to share the market equally and maximize its profits.

(b) The Dominant Firm Price Leadership Model

This is a typical case of price leadership where there is one large dominant firm and a number of small firms in the industry. The dominant firm fixes the price for the entire industry and the small firms sell as much product as they like and the remaining market is filled by the dominate firm itself.

It will, therefore, select that price which brings more profits to itself. BSNL, SAIL can be taken as dominant price leader.

Its Assumptions

This model is based on the following assumptions :

(i) The oligopolistic industry consists of a large dominant firm and a number of small firms.

(ii) The dominant firm sets the market price.

(iii) The dominant firm sets in a position to predict the supplies of other firms at each price set by it.

The Model

Given these assumptions, when each firm sells its product at the price set by the dominant firm, its demand curve is perfectly elastic at that price. Thus its marginal revenue curve coincides with the horizontal demand curve. The firm will produce that output at which its marginal cost equals marginal revenue. The MC curves of all the small firms combined laterally establish their aggregate supply curve. All these firms behave competitively while the dominant firm behaves passively. It fixes the price and allows the small firms to sell all they wish at that price.

The case of price leadership by the dominant firm is explained in terms of Figure (10.8) where DD_1 is the market demand curve. ΣMC_s is the aggregate supply curve of all the small firms. By subtracting ΣMC_s from DD_1 at each price, we get the demand curve faced by the dominant firm, $PNMBD_1$ which can be drawn as follows. Suppose the dominant firm sets the price OP. At this price, it allows the small firms to meet the entire market demand by supplying PS quantity. But the dominant firm would supply nothing at the price OP. Point P is, therefore, the starting point of its demand curve. Now take a price OP_1 less than OP. The small firms would supply P_1C ($=OQ_1$) output at this price OP_1 when their ΣMC_s curve cuts their horizontal demand curve P_1R at point C. Since the total quantity demanded at OP_1 price is P_1R ($=OQ$) and the small firms supply PIC quantity, CR ($=Q_1Q$) quantity would be supplied by the dominant firm. By taking $P_1N = CR$ on the horizontal line P_1R, the dominant firmôs supply becomes P_1N ($=OQ_d$). Thus we derive point N on the dominant

firmôs demand curve by subtracting the horizontal distance from point P_1 to N from the demand curve DD_1. Since the small firms supply nothing at prices below OP_2 because their ΣMC_s curve exceeds this price, the dominant firmôs demand curve coincides with the horizontal line P_2B over the range MB and then with the market demand curve over the segment BD_1. Thus the dominant firmôs demand curve is $PNMBD_1$.

Figure 10.8

The dominant firm will maximize its profits at that output where its marginal cost curve MC_d cuts its MR_d, the marginal revenue curve. It establishes the equilibrium point E at which the dominant firm sells OQ_d output at OP_1 price. The small firms will sell OQ_s output at this price for ΣMC_s, the marginal cost curve of the small firms equals the horizontal price line PIR at C. The total output of the industry will be $OQ = OQ_d + OQ_s$. If OP_2 price is set by the dominant firm, the small firms would sell P_2A and the dominant firm AB. In case a price below OP_2 is set the dominant firm would meet the entire industry demand and the sales of the small firms would be zero. The above analysis shows that the price-quantity solution is stable because the small firms behave passively as price-takers.

But this does not mean that the dominant firm charges the same price that is charged by a monopolist operating in the same market. As observed by Professor Markham : ñThe rationale of price making by the dominant or partial monopolist differs but little from that employed by the pure monopolist. They both, presumably, have complete control over prices, but the partial monopolist, unlike the pure monopolist, must take account of the quantity that the competitive sector of the industry will offer at any price he may setò.

However, the real test of a dominant firmôs price leadership is the extent to which the other firms follow its lead. The moment the firms cease to follow the price leader, the model breaks down. Besides, if the other firms have

different cost curves the same price may not maximize short-run profits for all the firms.

The dominant-firm model of price leadership can have a number of variations. There may be two or more large firms among a number of small firms which may enter into a collusion for sharing the market at various prices. There may be product differentiation. Nevertheless, the conclusions arrived at help to explain price-output policies in all such situations.

(c) The Barometric Price Leadership Model

The barometric price leadership is that in which there is no leader firm as such but one firm among the oligoplistic firms with the reputed management which announces a price change first which is followed by other firms in the industry. The barometric price leader may not be the dominant firm with the lowest cost or even the largest firm in the industry. It is a firm which acts like a barometer in forecasting changes in cost and demand conditions in the industry and economic conditions in the economy as a whole. On the basis of a formal or informal tacit agreement, the other firms in the industry accept such a firm as the leader and follow it in making price changes for the product. Newspaper and hotels industry in India can be taken as examples under this category.

The barometric price leadership develops due to the following reasons :

(i) As a reaction to the earlier experience of violent price change and cut-throat competition among oligopolistic firms, they accept one firm as the price leader.

(ii) Most firms do not possess the expertise to calculate cost and demand conditions of the industry. So they leave their estimation to one leader firm which has the ability to do so.

(iii) Oligopolistic firms accept one among them as the barometric leader firm which possesses better knowledge and predictive power about changes in direct costs or style and quality changes and changes in the economic conditions as a whole.

Exercises

1. Explain with diagrams the main characteristics of an oligopolistic market and equilibrium of a firm facing kinked demand curve.

2. Some economists have argued that the demand curve facing an oligopolist must have a kink in it. Discuss their reasons. Show in a diagram the kinky demand curve and, the equilibrium of a firm facing such a curve.

3. Explain price and output determination under price leadership by a dominant firm.

4. Explain price determination under conditions of price leadership in an oligopolistic market.

5. Define a cartel. How cartel maximizes joint profits? Explain the factors which lead to the breakup of cartel.

6. Explain the characteristic features of the Cqurnot model of duopoly. What are its limitations?

7. Write notes on : Market-sharing Cartel, Barometric Price Leadership, Chamberlinǿs Duopoly Model.

Lesson - 11

Price Determination Under Monopoly, Monopolistic Competition and Discriminating Monopoly

(Author :N.K. Bishnoi)

Analysis of the working of a competitive system was the main task done by the classisia1economists such as Adam smith, David Ricardo and J.S. Mill. Considering the earlier views, later economists of the 19th century developed the óidealô system of perfect competition. Many economists, since the time of Adam smith, where more interested in theoretical perfections than in the actual development of the capitalist system. They tried to explain the meaning of an economic system based on the model of perfect competition. According to them perfect competition would mean

(i) Production at minimum possible cost

(ii) Consumer satisfaction at its maximum

But in real word we hardly come across such a system of perfect competition. The exception to perfect competition which attracted serious attention during the 19th century was the concept of monopoly. This is in fact, the antithesis of perfect competition.

Monopoly market is one in which there is only one seller of the product having no close substitutes. The cross elasticity of demand of a monopolised product is either zero or negative. There being only one firm, producing that product, there is no difference between the firm and industry in case of monopoly. Monopoly is a price maker not the price taker.

In the words of koutsoyiannis, ñMonopoly is a market situation in which there is a single seller, there are no close substitutes for commodity it produced there are barriers to entry of other firmsò.

Features of Monopoly

Following are the features of monopoly :

(i) One seller of the product: In case of monopoly there is only one seller of product. He may be sole proprietor or a partnership firm or a joint stock company or a state enterprise. There is no difference between firm and industry. The firm is a price maker and not price taker.

(ii) No close substitute : The commodity which the monopolist produces has no close substitutes. Lack of substitutes means no other firm in the market is producing same type of commodity.

(iii) Restriction no the entry of the new firm : There are powerful restrictions to the entry of new firms in the industry, under the Monopoly.

Revenue and cost curves under Monopoly

A monopoly firm face a downward sloping demand curve, unlike a competitive firm, a monopolist can reduce the price and sell more. In a monopoly situation, there is no difference between firm and industry. Accordingly, under monopoly firmôs demand or average revenue curve (AR) and marginal revenue (MR) curves are separate from each other Both are down word sloping from left to right.

Fig. 1 Shows AR and MR curves under monopoly.

Under monopoly, shape of different cost curves is exactly like under perfect competition. Both AC and MC curves, will be U-shaped on account of law of variable proportions.

Price and Output Determination

In case of monopoly, one can know about price determination or equilibrium position with the help of marginal revenue and marginal cost

analysis. According to this analysis, a monopolist will be in equilibrium when two conditions are fulfilled.

(i) MC = MR and

(ii) MC curve cuts MR curve from below.

Study of price and equilibrium determination under monopoly is conducted in two time periods.

(i) Short Period and

(ii) Long Period

Price Determination under Short Period or Short Run Equilibrium :

In the short run, a monopolist has to work with a given existing plant. He can expand or contract output by varying the amount of variable factors. He cannot adjust the size of plant in the short run.

A monopolist in equilibrium may face three situations in the short run

(1) Excess Profit

(2) Normal Profit

(3) Minimum Losses

The process of price determination under monopoly has been explained as follows : (i) Super Normal Profit

If the price (AR) fixed by monopolist in equilibrium is more than

Fig.2

the average cost (AC) than he will earn excess profits.

The revenue and cost conditions faced by monopolist firm are presented in the Fig. 2.

AR and MR are the average and marginal revenue curves of the firm respectively. SAC and SMC are the short run average cost and marginal cost curves of the firm, respectively. To maximise profits, the monopolist firm chooses a price and output combination for which SMC = MR, and SMC curve cuts MR from below. As shown in the fig.2, E is the equilibrium where monopolist SMC curve cuts MR curve from below. A perpendicular parallel to y-axis is drawn at point E connecting the x-axis at Q and the demand curve at A. OQ is the equilibrium output. AQ is the equilibrium price, because the price is determined by demand curve or average revenue curve. The average cost is BQ, because line AQ cuts SAC curve at point B. Thus the monopolistôs per unit excess profit is AB, which is the difference between the price (AQ) and the corresponding average cost of production (BQ) The ABPPI represent total monopolistôs profit. The total profit of the monopolist will be maximum only at OQ level of output.

Normal Profit

In the short period it is possible that monopolist may earn normal profit. This happens only when the average cost curve of the monopolist is tangent to its average revenue curve.

Fig.3

In fig.3 the monopolist is the equilibrium at OQ level of output, because at this level of output his marginal cost curve (SMC) cuts MR curve at point E. Also at same level of output (OQ) the monopolist SAC curve touches his AR curve at point A. Thus AQ or OP is the monopolist price (which is determined by AR curves) is also equal to the cost per unit (AQ). The monopolist will earn

only normal profit and the normal profit is included in the average cost of production.

Loss Minimization in the Short Period

In the short run, the monopolist may incur losses also. The monopolist may continue his production so long as price of his product is high enough to cover his average variable cost. If the price falls below average variable cost, the monopolist prefers to stop production. Accordingly, a monopolist in equilibrium, in the short run, may bear minimum losses equivalent to fixed costs. The situation of minimum losses has been illustrated in the fig. 4.

Fig. 4

The monopolist is in equilibrium at point E, where SMC = MR and SMC curve cuts MR curve below. OQ is the equilibrium level of output.

The price of equilibrium output OQ is fixed at BQ or OP.

At this price, average variable cost (AVC) curve AR curve at point B. It means firm will set at only average variable cost from the prevailing price. The firm will bear the loss of fixed cost equivalent to AB per unit. The firm will bear total loss equivalent to $ABPP_1$. If its price falls below (BQ) the monopolist will prefer to stop the production. The point B is also known as ósHut-down pointô.

From the above analysis of short run price and output equilibrium it may be content that profit maximisation or loss minimization or attainment of normal profit will be accomplished only at that level of output at which marginal cost is equal to marginal revenue and marginal cost cuts MR curve from below.

Price Determination under long - Run

In the long run the monopolist has the time to expand his size of the plant, or to use his existing plant at any level which will maximise his profit. With entry blocked, however, it is not necessary for the monopolist to reach an optimal scale, what is certain is that the monopolist will not stay in business, if he makes losses in the long run. However, the size of his plant and the degree of utilisaion of any given plant size depends entirely on the market demand. After these adjustments are completed, the monopoly rum will have a long run

equilibrium determined by the equality of long run marginal cost and marginal revenue as shown in fig. 5.

E is the equilibrium point of the monopolist firm. Corresponding to this equilibrium point, OQ is the equilibrium level of output. The monopoly will fix price AQ in the long run. Average cost is BQ. Profit per unit is AB. Total profit is equal to $ABPP_1$.

It may be noted that there is always a tendency for the monopolist firm to secure excess profits, even in the long run, since entry into the industry is prohibited.

Price Determination under Monopolistic Competition

Monopoly and perfect competition are really two extremes, and many industries fall in between. There are very few pure monopolies, since there are very few commodities for which close substitute does not exist. Similarly, there are very few commodities that are entirely homogenous to make the assumption of perfect competition realistic. There is, thus, a large grey area between these two extremes. Although the French Economist Cournot pointed this out in

1838, it was in the early 1930s that economist began turning their attention to the middle ground between monopoly and perfect competition. In 1933, Edwin H. Chamberlin of Harvard University published the book, ñThe Theory of monopolistic Competition : A Re-orientation of the Theory of Valueò. It was received very enthusiastically and many economist talked of the ñChamberlin Revolutionò. In the same year, but six months later, Joan Robinson of Cambridge University in England published a similar book titled ñThe Economics of imperfect Competition.ò Although there are similarities in the books, there are major differences as well. For instance, Chamberlin treated at length product differentiation and advertising, which were neglected by Joan Robinson. Joan Robinson discussed problems such as price discrimination, monopolistic and monopsonistic exploitation not covered by Chamberlin. We shall start our discussion by defining monopolistic competition. Monopolistic competition is said to exist when there are many firms, as in perfect competition, but each firm produces a product that is slightly differentiated from that of others. Examples of these are numerous such as retail clothing stores, restaurants, barber shops etc. There are several distinguishing characteristics of monopolistic competition.

(i) Product differentiation : The products are heterogeneous rather than homogenous. However, products are only differentiated. The output of one firm is close (but not perfect) substitute of the output of other firms. Differentiation grants each firm some monopoly power. Whereas the presence of close substitutes provides competition. There are many sources of differentiation. Some of these are : Chemical composition, advertising, packaging, brand names, location and design.

(ii) Non price competition : Since the products are slightly differentiated, the different firms try to play up the difference in their products in order to increase their demand. They do this in a variety of ways such as advertising the differences or adding same frills. Take for instance, firms producing washing powder, óSurf and Detô. With one pack of surf: the company gives a free gift of one glass-tumbler.

Short Run Equilibrium of the Firm lU!der ~onopolistic Competition

Short run equilibrium of a monopolistic competitive firm is very similar to that of monopoly find. Only difference is that find under monopolistic competition produces differentiated products, have some degree of monopoly power. The demand curve facing the monopolistic firm is more elastic. The firm is in equilibrium when

i) MC = MR and

ii) MC curve cuts MR from below

The amount of profit earned by the firm in equilibrium, in Short run, depends on demand of the product and the efficiency of the find. The firm may face any of the three situations in this period i.e. excess profit, normal Profit and minimum losses.

The short run equilibrium of a firm under the monopolistic competition is explained with the help of following diagrams.

Excess Profit

When the price is more than average cost, the firm under the monopolistic competition will earn excess profit. In the fig. 6 equilibrium is at point E.

Where MC = MR and OQ is the equilibrium level of output, at which price is Q or OP, because A is a point on AR, and average cost is BQ (B is on SAC curve). Therefore, AB is the excess profit per unit of output. The excess profit will be measured by the area of rectangle $ABPP_1$ i.e. output multiplied by excess profit per unit of output.

Fig. 6

Long Run Equilibrium of the Firm

Long run refers to that time period in which production capacity of each firm can be changed as required. Firm can change the size of the plant. New firms can enter in the group and old firms can leave it. In the long run the firm in the group earns only normal profit. In long run no firm can bear losses. In short run if firms are earning profits, then in long run new firm will enter in the group. With the entry of new firms, supply will increase and share of each firm in the total output will decline. AR and MR curves of each firm will shift to left of original AR and MR curves. Each firm uses various devices including those of publicity and advertisement to change the shape of demand curves.

Equilibrium under condition of the entry of new firms

In the long run new firm can enter in the group. The equilibrium output and price of firm can be explained with the help of fig. 7. LAC

Fig. 7

and LMC are long run average and marginal cost curves of the firm. Ar_2 and MR_2 are the original average and marginal revenue curves of the firms. In the short run, E_2 is the equilibrium point. Firmôs equilibrium output is OQ_2 and average revenue or price is OP_2 (= EQ_2). Average cost is BQ_2. Profit per unit is EB.

New firms will enter in the group due to supernormal profit. With the entry of new firm in the group, production will increase, but contribution of each firm in the total supply decreases. Average revenue curves become AR_1 and marginal revenue MR_1. Firms will be in equilibrium at point E. Long run average cost curve (LAC) is tangent to AR_1 curve at OQ_1 level of output. Firm will get only normal profits. Because firms average revenue AQ (=OP_1) becomes equal to average cost AQ_1 (=OP_1). Thus, in long run, the firms in the group will be in equilibrium when all the firms in group earn only normal

profit. No new firm will enter in the group. Thus the long run equilibrium indicates the blend of competition and monopoly.

Selling Cost

Unlike a perfectly competitive firm, in order to increase the sale of the product the monopolist has to undertake huge advertising campaigns. In perfect competitions, product manufactured by different firms are homogenous, there must be uniformity of price. But in case of monopolistic competition, product differentiation is the common feature and for this reason selling cost would be crucial.

According to Chamberlin, ñSelling costs are those costs which are to be incured in order 10 alter the position or shape of the demand curve for a product.ò In words of Cairn cross, ñSelling costsinc1ude all expenditure designed to create, increase or maintain me demand for a firms output.

Difference Between Selling Cost and Production Cost

Production costs are those costs which are incured to produce the particular quantity of output. These costs include expenditure purchase of raw material, energy, packaging etc. Thus, the production costs are incured in order to create utility, whereas, selling costs are outlay made in order to secure a demand for the product.

In words of Cairn cross, ñProduction costs have no influence on demand, whereas, selling costs are incured in order to influence demand. According to Chamberlin, ñThose costs which are incured to adapt the product to the demand are production cost, while those which adapt the demand to production are selling cost.

Effect of Selling Costs on Price Determination under Monopolistic Competition

Under monopolistic competition, selling costs increase demand for the product as well as the cost of production. The main objectives the firm is to maximise the profit. When the firm incurs an expenditure on selling costs, there

will be shift of its demand and cost curves. The firm will continue to incur expenditure on advertisement, so long as, revenue earned is more than or equal to cost of advertisement. The equilibrium of the firm under the selling cost has been explained with the help of fig. 8.

Fig. 8

Suppose the firm incurs Rs.5000/-, Rs.6000/- and Rs.7000/-in the form of selling cost. In such situation, the firm will have three average production costs (APC) and demand curves. With increase in selling costs, demand curve of the firm will shift to right. The firm price and output equilibrium have been explained in the Fig. 8.

Suppose, the firm incurs selling cost Rs.5000/- initially. Then, APC_1 and D_1D_1 are the firmôs average production cost curve and demand curve respectively. OM_1 and OP_1 is the equilibrium output and price of the firm respectively.

The firm earns supernormal profit equivalent to area $A_1B_1P_1L_1$. If the firm incurs Rs.6000/- as selling costs the demand curve of the firm shift to right of average demand curve (D_1D_1). The new demand curve of the firm will be D_2D_2 and average production cost curve of the firm will shift upwards to APC_2. OP_2 and OQ_2 will be the firmôs equilibrium output and price respectively. The firm will earn supernormal profit equivalent to area $A_2B_2P_2L_2$. If firm incurs more incurs more expenditure on selling costs i.e. Rs.7000/-, then D_3D_3 and APC_3 will be firms new demand and average production cost curves respectively. OP_3 and OM_3 will be equilibrium price and output of the firm respectively. The firm will earn supernormal profits equivalent to $A_3B_3P_3L_3$.

This process will continue till revenue earned by the firm is greater than or equivalent to expenditure incured on advertisement.

Discriminating Monopoly

Price discrimination exists when the same product is sold at different prices to different buyers. The cost of production is either the same or it differs but not as much as the difference in the charged price. The product is basically same, it may have slight difference. (For example, different binding of same book; different location of seats in a theatre; different seats in an aircraft or a train, etc.). Here, we will concentrate on the typical case of an identical product, produced at the same cost, which is sold at different prices, depending on the preference of the buyers, income, location and the ease of availability of substitutes. These factors give rise to demand curves with different elasticities in the various sectors of the market of a firm. It is also common to charge different prices for the same product at different time periods. The necessary conditions, which must be fulfilled for the implementation of price discrimination are the following :

i) The market must be divided into submarkets with different price elasticities.

ii) There must be effective separations of the sub-markets, so that no reselling can take place from a low-price market to a high price market.

These conditions show why price discrimination is easier to apply with commodities like electricity, and services (Like service of a doctor etc.), which are consumed by buyer and cannot be resold.

Degrees of Price Discrimination

i) Discrimination of first degree : In first degree price discrimination the monopolist charge different prices for every unit of commodity. It means he charges the price accordingly, to extract entire amount of consumers surplus. Mrs. Joan Robinson refers to this kind of discrimination as Perfect Discrimination.

ii) Second degree price discrimination : In Second degree price discrimination the monopolist charges different prices for a specific quantity or block of output. It means monopolist will sell one block of product at one price and another block at lower price. Second degree price discrimination is more common than first degree price discrimination.

iii) Third degree price discrimination : In third degree price discrimination the price charged by monopolist is different in different market of same commodity. The division of whole market into the two or more than two sub-markets is essential for third degree price discrimination. The third degree price discrimination is most common in practice.

Technique of Price Discrimination

The reason for a monopolist to apply price discrimination is to obtain an increase in the total revenue and his profits.

We will start from the simplest case of a monopolist who sells his product at two different markets. It is assumed that the monopolist will sell this product in two segregated markets. Each of them having a demand curve with different elasticities. Let us assume there are two markets A and B. In the market A demand for the product is less elastic while in market B, it is more elastic. This process has been explained in the fig. 9. AR_1 and MR_1 are average and marginal revenue curve of market A. AR_2 and MR_2 are the average and margined revenue curves of market B. CMR is the combined marginal revenue (CMR) of market A and B. It is derived by lateral summation of MR curves of market A and B. It has a kink ôKô due to differences in elasticities of demand of both markets. MC is the marginal cost curve of monopolist. The condition of equilibrium of discriminating monopolist is $MC = CMR = MR_1 = MR_2$.

E is the equilibrium point where MC curve cuts combined marginal revenue (CMR) curve. At that Point (i.e. E) a horizontal straight line parallel to the X-axis is drawn to figs. 9(a) and 9(b). At points B_1 and A_1, MR_1 and MR_2 becomes equal to MC. By dropping perpendicular from point B_1 and A_1 connecting the X-axis and the respective demand curves, the prices in the two markets are found out :

	Market A	Market B	Aggregate
Price	**OP**	**OP_1**	**-**
Quantity	**OB**	**OA**	**OQ = (OB + OA)**

The price in the first market A is higher than the price in the market B. We do not calculate any price in Fig. C. as there is no aggregate price. Thus, in the technique of price discrimination the prices should be determined ill such a way that MR in the different markets is equal to MC.

Dumping - A Special Case of Discrimination

Dumping means charging a higher price in the domestic market

Fig. 9

and lower price in the foreign market for the same product. For dumping, the total market for the product is divided between domestic and foreign markets. The necessary condition for dumping is that in the domestic market the demand for the product is less elastic and more elastic in the foreign market. The process has been explained in the fig. 10.

Fig. 10.

In the fig. AR_d and MR_d are the average and marginal revenue curves for the domestic market. AR_f and MR_f are the average and marginal revenue curve for the foreign market. AR_d and MR_d are downward sloping average revenue and marginal revenue curves in the domestic market. AR_f and MR_f are average revenue and marginal revenue curves in the foreign market and are parallel to X-axis. Reason being, in the former case producer has a monopoly and in the latter case he faces perfect competition like situation in the foreign market.

Aggregate marginal revenue curve (ARED) is obtained by adding the marginal revenue curves of domestic and foreign market. MC curve represents marginal cost curve of monopolist total output. The marginal cost curve cuts combine marginal revenue curve (ARED) at point E. OQ is the equilibrium output. Now the total output is distributed between domestic and foreign market in such a way that marginal revenue of both markets should not only be equal to each other but equal to its marginal cost also. It is vivid from the fig., when amount OM is sold in the domestic market, marginal revenue MR is equal to marginal cost QE. Thus, out of total output OQ, OM will be sold in the domestic market. In the domestic market, OP_1 price will be charged. MQ output will be sold in the foreign market at OP price. Thus area TERA is the total profit earned by the monopolist. Thus, monopolist will maximize his profit by charging higher price in the domestic market and lower price in the foreign market.

Comparison Between Monopoly and Perfect Competition

There are certain similarities and dissimilarities between monopoly and perfect competition which are as follows :

1. Similarities :

(i) Main objective of the firm in both these market is to earn maximum profit.

(ii) Condition of equilibrium of the firm in both the markets are similar i.e. MC = MR, MC curve cuts MR curve from below.

2. Dissimilarities

(i) Nature of average and marginal revenue curves : Under perfect competition there are large number of firms selling homogenous products. The price of the product is determined by the industry. The firm is a price taker and not the price maker. Demand or average revenue curve is a horizonta1line parallel to X-axis. As the price of the product is constant, marginal revenue coincides with average revenue. On the contrary, in case of monopoly there is only single firm of the product. There is no difference between firm and industry under monopoly. The firm is a price maker. In order to sell more he

will have to reduced the price of the commodity. That is why average revenue curve is negatively sloped. When average revenue curve is negatively sloped the marginal revenue curve will also be negative.

(ii) Entry of firms : Under perfect competition the new firms can enter the industry and existing firms can leave the industry. But in case of monopoly it is not possible due to certain obstacles for the firm to enter in the market.

(iii) Price discrimination : Under monopoly price discrimination is feasible but it is not so under perfect competition.

(iv) Long run output and price : Output under perfect competition is more as compared to that under monopoly. The reason being, under perfect competition production is being done at the minimuin of long run average cost curve. But in case of monopoly production is being on the downward sloping part of long run average cost curve. In the long run, firm under perfect competition earns only normal profit, where as firm under monopoly always earns supernormal profit.

(v) Consumer welfare : Production under perfect competition causes an increase in consumer welfare. The reason being, under perfect competition more quantity is produced at less cost and sold at less price. In contrast, during monopoly the price of the commodity is higher due to low production and high cost.

Comparison Between Monopoly and Monopolistic Competition

There are certain similarities and dissimilarities between monopoly and monopolistic competition which are as follows :

1. Similarities

(i) Condition of equilibrium : in both the situations, firm will be in equilibrium when following two conditions are fulfilled i.e. MC = MR and MC curve cuts MR curve from below.

(ii) **Under utilization of capacity :** In both the situations production capacity is not being fully utilized. The reason being, the demand curve is not tangent to long run average cost curve at its minimum points.

2. Dissimilarities

(i) **Number of firms under monopoly :** There is only one firm, whereas there are a large number of firms under monopolistic conditions.

(ii) **Nature of commodity :** All the units of commodity produced by firm under monopoly is homogenous, whereas, it is not so in case of firm working under monopolistic competition.

(iii) **Elasticity of demand :** The demand curve is more elastic under monopolistic competition, whereas, it is not so under monopoly. The reason being, substitutes of the commodity are available in the former case and in the latter case substitutes are not available.

(iv) **Nature of profit :** In the long run, firm working under monopoly will always get that under monopoly. The reason being, under perfect competition production is being done at the minimum of long run average cost curve. But in case of monopoly production is being on the downward sloping part of long run average cost curve. In the long run, firm under perfect competition earns only normal profit, where as firm under monopoly always earns supernormal profit.

(v) **Consumer Welfare :** Production under perfect competition causes an increase in consumer welfare. The reason being, under perfect competition more quantity is produced at less cost and sold at less price. In contrast, during monopoly the price of the commodity is higher due to low production and high cost.

Comparison Between Monopoly and Monopolistic Competition

There are certain similarities and dissimilarities between monopoly and monopolistic competition which are as follows :

1. Similarities

(i) **Condition of equilibrium :** In both the situations, firm will be in equilibrium when following two conditions are fulfilled i.e. MC = MR and MC curve cuts MR curve from below.

(ii) **Under utilization of capacity :** In both the situations production capacity is not being fully utilized. The reason being, the demand curve is not tangent to long run average cost curve at its minimum points.

2. Dissimilarities

(i) **Number of firms under monopoly :** There is only one firm, whereas there are a large number of firms under monopolistic conditions.

(ii) **Nature of commodity :** All the units of commodity produced by firm under monopoly is homogenous, whereas, it is not so in case of firm working under monopolistic competition.

(iii) **Elasticity of demand :** The demand curve is more elastic under monopolistic competition, whereas, it is not so under monopoly. The reason being, substitutes of the commodity.

(iv) **Nature of Profit :** In the long run, firm working under monopoly will always get supernormal profits and this supernormal profits disappear under monopolistic competition. The reason being, in the latter case large number of firms are producing the commodity having close substitutes and there is free entry and exit of the firms.

(v) **Control overprice :** Firm under monopolistic competition has to keep in mind prices being charged by their rivals. If difference in prices are substantial, the firm will not be able to retain their customers. However, the monopolist can follow an independent price policy. The reason being, entry of the firm is restricted under monopoly ,

(vi) **Price discrimination :** The monopolist firm can follow the policy of price discrimination due to inherent advantages associated with the firm. But the firm under monopolistic competition can not do so due to large number of sellers selling close substitute.

(vii) Selling Cost : The monopolist has to incur huge expenditure on selling cost in order to retain and attract new customers. But this type of expenditure has no relevance under monopoly.

Important Questions

1. Define monopoly. How is price and equilibrium output determined under monopoly?

2. Define price discrimination. How is the price and output determined under discriminating monopoly?

3. Write notes on :

(a) Nature of AR and MR curves under monopoly.

(b) Price determination under conditions of dumping.

4. What is monopolistic competition? How is price determined under monopolistic competition?

5. What is difference between production costs and selling costs? How do the selling costs affect he equilibrium output determination under monopolistic competition.

6. Write notes on :

(a) Make a comparison between monopoly and monopolistic competition.

(b) Make a comparison between monopoly and perfect competition.

Selected References

1. Alfred Marshall : Princip1es of Economics
2. John Robinson : The Economics of Imerfect Competition.
3. Edward H. Chamberlin : The Theory of Monopolistic Competition.
4. G.J. Stigler : The Theory of Price.

Lesson - 12

WELFARE ECONOMICS

(Author : Ved Paul)

In previous chapters, the theoretical framework we have discussed involves the efforts of individuals and firms to make themselves as well-off as possible Individuals seek to attain the highest possible indifference curve and firms seek the maximum possible profit. Their efforts constitute the reason why trade take place. When a consumer purchases a good from a producer, each is better off or at least not worse off, otherwise no trade would have occurred in first place. The gain from trade is affect-ed by various changes in market conditions like changes in taxes, price controls, subsidies, quotas, rationing etc. Economics has different types of criteria to see who gains and who loses from such policies and to evaluate the size of gains and losses.

1. Measuring Gains from Trade

When a consumer buys a good from a producer, each one gains from the trade. We measure this gain with the help of consumer surplus and producer surplus.

1.1 Consumer Surplus : Suppose we have a consumer who wants to purchase apples. Let this consumer is willing to pay Rs.20 for acquiring an apple. The value of one apple to this particular consumer would be Rs.20. If the consumer is willing to have two apples. For these two apples this consumer wants to pay less than Rs.40. Because when he has only one apple, he uses it in the way in which he would most like to use an apple, and uses the second apple where he considers it to be of second most important use. In other words, the marginal value of the second apple is somethings less than Rs.20, and so the total value of the first two apples is something less than Rs.40. As the consumer acquires more apples, their marginal value continues to decrease. Let us assume that these marginal values are as given in table 1.

Table 1

Quantity	Total value	Marginal value
1	20	20
2	35	15
3	44	9
4	49	5
5	52	3
6	54	2

If, plot the marginal values of apples, then we well get a downward sloping curve as shown in fig. (1).

Fig. 1

This figure may also be considered as demand curve of the apples for this consumer. Now question arises how many apples a consumer will buy at a particular market price. Suppose the market price of an apple is Rs.5/-. He will certainly buy a first apple because he values it at Rs.20 and can get it for Rs.5. He will also buy a second apple, which he values at Rs.15/- and can also get for Rs.5/-. Similarly he will buy third and fourth. But he will not buy apples more than four. The fifth apple provides only Rs.3 worth of additional value and costs Rs.5/- to acquire, Likewise, sixth and more than it. Thus the number of apples purchased by this particular consumer is 4. The consumer buys, apples as long as the marginal value of an apple exceeds its price, and stops when the two become equal.

Suppose this consumer purchases 4 apples. Thetota1 values of these apples for this consumer is equal to sum of the marginal values of first four apples i.e. Rs.49. It is shown in the following figure no. 2.

Fig. 2

Assuming that apples are fully divisible then the total value of the consumerôs purchase is equal to the shaded area under the demand curve. But in acquiring these four apples the consumer has paid only Rs.20 (Rs.5×4). The difference between the total value of the apples he buys, minus what he actually pays for them is called consumer surplus. It can be shown in fig. 3.

Fig. 3

Here total value of the consumerôs apples is equal to the entire shaded area, at the cost to him of acquiring those apples is the area labeled A. The gain to the consumer is the remainder, namely area B. Thus, consumer surplus is it difference between what a consumer is willing to pay and what actually he pay for a certain amount of a commodity.

1.2 The Producer Surplus : The producer supplies the goods to the consumer. In this trade producer also gain. For calculating the gain of the producer from trade, suppose there is a producer of apple. The marginal cost of producing the apple for this producer is given in the table-11.

Table 11

Quantity	Marginal Cost	Total Cost
1	0.5	0.5
2	2.0	2.5
3	3.5	6.0
4	5.0	11.0
5	7.0	18.0
6	8.5	26.5

The Total cost of supplying, these six apples in the market in the sum of the marginal cost of supplying first apple (Rs.5), the marginal cost of second (Rs.2.5) and so on.

It can be shown with the help of following figure No. 4. Here it is

Fig. 4

assumed that there is prefect competition in the market. In this market structure, marginal cost curve will work as market curve. In case our good is fully divisible then the total cost of supplying the four apples in the market is area D in the lower part of curve in the figure. If the market prices of one apple is Rs.5, then the producer will supply only 4 apples in the market because the cost of porduction of each apple next to four apples is more than Rs.5. The total revenue for supplying 4 apples in the market for the producer is Rs.20. It is equal to the area C + D in the figure no. 5.

Fig. 5

Here total revenue of the producer is C+D and his cost of production are D. The difference, area C, is called the producerôs surplus and represents the gain to the producer as a result of his participation in the market place. Thus, the producer surplus is the difference between what the producer willing to charge and what he actually charges for supplying the good in the market.

2. Social Welfare

The sum of consumerôs surplus and producerôs surplus is called social gain or welfare gain. The social welfare depends upon this social gain. In economics there are different criteria for measuring the social welfare. These criteria are related to the evaluation of different situations from the point of view of societyôs well being.

Some of these criteria are discussed below :

2.1 Growth of GNP (Gross National Product) as a criterion on welfare:

Adam Smith implicitly accepted that growth of gross national product as a criterion for growth in welfare of the society. He believed that invisible hand took an automatic care of accumulation of wealth and output on which social welfare depended. It is held that economic growth resulted in the increase of social welfare because growth increased employment and good available for consumption to the community. As all individuals were capable of looking after their respective individual interests and maximizing their satisfactions, and as the society was just collection of individuals, the social welfare was automatically taken care of and would be automatically maximized.

2.2 Betham's Criterion : Jermy Betham, as English economist, argued that welfare is improved when the greatest good secured for greatest number. It has assured that welfare is the sum of the utilities of the individuals of the society. To illustrate let us assume that the society consists of three individuals A, Band C, that

$$\Delta W = U_A + U_B + U_C$$

Where Us are the utilities of respective individuals and W is the total welfare of the society. According to Bentham $\Delta W > O$ if $(\Delta U_A + \Delta U_B + \Delta U_C) > 0$, Hence, Δ stands for change.

2.3 Pigou's Criterion : According to A.C. Pigou ñEconomic Welfare is that part of general (social) welfare that can be brought directly or indirectly into relation with measure of money.ò Pigovian welfare economics is related to the satisfaction derived from the use of exchangeable goods and services, Pigovian welfare economics is based on the following assumptions:

(a) Every individual attempts to maximise his satisfaction from the use of limited monetary resources.

(b) Satisfaction derived from the consumption of goods and services can be compared interpersonally and interpersonally.

(c) Marginal utility of money income decreases with every increase in it which implies that marginal of a unit of money to the poor is greater than that to the rich.

(d) Different individual have equal capacity for satisfaction, that is, different people derive the same satisfaction from the same real income.

Pigou formulated dual criterion to maximize welfare as : (i) An increase in national income brought about either by increasing some goods without reducing the others or by transferring factors from less productive to more productive activities, increases economic welfare (ii) Any re-organisation of the economy, which increases purchasing power of the poor without reducing the national income, increases social welfare

2.4 The Kaldor-Hicks Compensation Critetion : Nicholas Kaldor and John Hicks gave the following criterion for the welfare of the society.

Assume that a change in the economy is being considered, which will benefit some (gainers) and hurt others (losers). One can ask the gainers how much money they would be prepared to pay in order to have the change and the losers how much money they would prepared to pay in order to prevent the

change. If the amount of money of the gainers is greater than the amount of the losers, the change constitutes an improvement in social welfare, because the gainers could compensate the losers and still have some net gain. Thus, the Kaldor-Hicks compensation criterion states that a change constitutes an improvement in social welfare if those who benefit from it could compensate those who are hurt, and still be left with some net gain.

2.5 The Pareto-Optimality Criterion : This criterion refers to economic efficiency, which can be objectively measured. It is called Pareto criterion after the famous Italian economist Vilfreds Pareto (1848-1923). According to this criterion any change that makes at least one individual better-off and no one worse off is an improvement in social welfare. Conversely, a change that makes no one better off and at least one worse-off is a decrease in social welfare. Pareto optimal is a situation in which it is impossible to make anyone better off without making someone worse-off. It is a situation in a system where the social welfare will be maximized. Three basic conditions must be satisfied if Pareto efficiency is to be attained. The economy must achieve :

i) Efficiency of distribution of commodities among consumers.

ii) Efficiency of allocation of factors of productions.

iii) Efficiency in matching production to consumption.

Let us examine how these three conditions are satisfied to attain the Pareto-Optimality.

(a) Efficiency of distribution of commodities among consumers :

It is also called efficiency in consumption, which requires that it, is impossible to redistribute a given set of goods among consumers in a manner that would improve one personôs welfare at nobodyôs expense. In economic terms it means that if X any Y and are two goods, then the marginal rate of substitution of X for Y (MRS X for Y) should be same for all individuals consuming both the goods.

Suppose, in an economy there are two consumers A and B, and two goods X and Y. Let X = apple and Y = orange. Suppose for individual A, MRS X for Y = 2 and for individual B, MRS X for Y = 1. This means that individual A is willing to exchange 2 oranges for 1 apple. Individual B is willing to exchange 1 orange for 1 apple. We can now re-allocate apples and oranges between them to make at least one of them better off, without making the other one worse-off. What we do is to take away 1 apple from B and give it to A. He will give us 2 oranges. Now we give one of these to B. He is no worse-off because he is willing to exchange 1 apple for 1 orange. But we have 1 orange left. We can give it to either A (or B) and, thus, make A (or B) better off without making the other person worse-off. Thus, the initial allocation was not efficient.

We cannot make any such re-distribution if the MRS X for Y is same for all consumers. In that case, we could make one person better off only by making another worse-off. In other words if the MRS X for Y is the same for everyone, and there is no redistribution of goods that would constitute Pareto improvement.

(b) Efficiency of allocation of factors among firm producers :

It is also called production efficiency which requires that it is impossible to redistribute inputs to produce more of one product without reducing the output of another product. An increase in one product could make some better off at nobodyôs expense means more efficiency in utilisation of re sources. If this is impossible, then the old allocation of inputs was efficient. Suppose, there are two factor of production namely labour and capital in the economy. With the help of these two factors of production two commodities are produced namely apple and orange. Then, production efficiency requires that the marginal rate of technical substitution of L for K (MRTS L for K) must be same for (1) all the products that a single firm produces using these two inputs and (2) all producers producing the same output.

The first condition is sometimes referred to as the requirement for managerial efficiency, because it deals with input allocation within a single

firm. If this condition is not satisfied and two products have different MRTS L for K, then we can redistribute the inputs so that this firm can produce more of one good without reducing the production of other good. Suppose that the two products are apple and orange, and suppose the MRTS of L for K is 2 for apple and 1 for orange. This means that we can substitute 1 unit of labour for 2 units of capital and keep apple output constant. Similarly, we can substitute 1 unit of labour for 1 unit of capital and keep orange output constant.

So, all we do is take, one unit of labour out of apple production and switch to orange production. This release of 2 units of capital from apple production, 1 unit of which is transferred to the production of oranges. Now the output of apple and oranges is unaltered, but we left with an extra unit of capital. We can allocate this to apple (or orange) and get more apples (or oranges). Thus, one output is increased without reducing the other output. The second condition deal with the efficient allocation of inputs between firm. If the condition is not satisfied then a redistribution of inputs between firms will produce an increase in at least one output with no reduction in the other.

(c) Efficiency in Matching Production and Consumption :

This efficiency requires that we produce the correct mix of outputs. The condition for efficiency in the matching of production and consumption is that it be impossible to rearrange outputs in manner that would constitute a Pareto improvement. This type of efficiency necessitates that for two goods x and y, the marginal rate of transformation (in production) of x for y (MRT x for y) is the same as the marginal rate of substitution (in consumption) of x for y (MRS x for y).

That is,

MRT x for y for all producers = MRS x for y for all consumers.

Suppose, this condition is not satisfied and for a producer A we have MRT x for y = 2 and for a consumer B we have MRS x for y = 1

X = Kilograms of Apples

Y = Kilograms of Oranges

Then, since MRT x for y = 2, the producer can decrease production of apples by 1 kg. and increase production of oranges by 2 kg. with the same total in puts. Now the producer can give consumer 1 kg less of apples and 1 kg. more of oranges. Since MRS x for y = 1 for consumers, they are neither better off nor worse-off. But the producer is better-off as he has 1 kg of orange left. If he gives it away to the consumer, the consumer is better-off and the producer is not worse-off. Thus, at least one of two can be made better off without the other being made worse off, by the change. Thus, the original situation is not Pareto Optimal.

Thus, there are two basic principles for Pareto Optimality (1) any MR_1 must equal any corresponding MRS, (2) Any MR_1 must equal any body elseôs MRT, and any MRS must equal anybody elseôs MRS.

When all the conditions of economic efficiency are fulfilled simultaneously, a society is said to have achieved a Pareto optimum. As long as these conditions are not fulfilled and inequalities persist, a reallocation of resources or goods can be made, that will increase total economic welfare.

Pareto conditions for efficiency would be satisfied if profit maximizing firms and utility maximizing households were to determine the optimum quantities of goods and services that they wish to trade with the help of equilibrium prices established in perfectly competitive markets. In this case.

	MRS x for y	=	Px Py is the same for all consumers
	MRISL for k	=	PL/Pk is same for all producers
	MRT x for y	=	Px/Py for all producers
		=	Px/Py for all consumers
		=	MRS x for y
Where	Px	=	price of commodity x
	Py	=	price of commodity y

PL = price of labour

PK = price of capital

3. GENERAL EQUILIBRIUM ANALYSIS

In previous lessons to study the economic problem partial equilibrium approach is used. This approach is related to the decisions in a particular segment of the economy in isolation of what was happening in other segments, under the ceteris paribus assumption.

In consumer behaviour lesson, the utility maximising behaviour of household was examined under the assumption that its income was given, although income depends upon the amount of labour and other factors of production that the consumer owns and on their prices. The ceteris paribus assumption was useful in that it enabled us to study the individuals demand for different commodities in isolation from influences arising from other parts of the economy.

In producer behaviourôs lesson, the production decision of a firm was examined under the assumption that factor prices, the state of technology and prices of commodities were given. The ceteris paribus assumption allowed us to study the cost minimisation behaviour of a firm in isolation from such factors such as demand for the products, which in turn are influenced by the level of employment, income and tastes of consumers.

Thus, the basic characteristics of partial equilibrium analysis in the determination of the price and quantity in each market by demand and supply curves drawn by assuming that other things are remaining constant. Each market in this approach is regarded independently of others.

General equilibrium analysis, by contrast, is concerned with the interdependence of all economic units and all markets in the economy. The markets of all commodities and all productive factors are inter-related, and the prices in all markets are simultaneously determined. For example, consumersô demand for various goods and services depend upon their tastes and incomes. Consumersô incomes in turn depend on the amounts of resources they own and

factor prices. Factor prices depend on the demand and supply of various inputs. The demand for factors by firm depends not only on the state of technology but also on the demand for final goods they produce. The demands for these goods depend on consumers incomes, which as we saw depend on the demand for factors of production. This circular interdependence of the activity within an economic system can be illustrated with a simple economy composed of two sectors, a consumer sector, which includes households and a business sector, which includes firms. It is assumed that (i) all production take place in the business sector, (ii) all factor of production are owned by the households (iii) all factors are fully employed, (iv) all income is spent.

The economic activity in the system takes the form of two flows between consumer sector and the business sector: a real flow and a monetary flow. These flows are shown in the following figure no.5.

Fig. 5

The real flow is the exchange of goods for services of factors of production firms produce and offer final goods to the household sector, and consumer offer to the firm the services of factors which they own.

The monetory flow is the real flow expressed in monetory terms. The consumers receive income payments from the firms for offering their factor services. These incomes are spent by consumers for the acquisition of the finished goods produced by business sector. The expenditures of firms become the money incomes of the households. Similarly, the expenditures of households become the receipts of the firms, which they once again pay the households for the factor services which they supply.

The real flow and monetary flow, which represent the transactions and the interdependence of the two sectors, more in opposite direction. They are

linked by the prices of goods and factor services. The economic system is in equilibrium when a set of prices is attained at which the magnitude of the income flow from firms to households is equal to the magnitude of the money expenditure flow nom households to firms.

An economic system consists of millions of economic decision making units, who are motivated by self-interest. Each one pursues his own goal and strives for his own equilibrium independently of others. The problem is to determine whether the independent, self-interest mo1ivated behaviour of economic decision makers is consistent with each individual agentôs attaining equilibrium. All economic units, whether consumers, producers or suppliers of factors, are independent. General equilibrium theory deals with the problem of whether the independent action by each decision maker leads to a position in which equilibrium is reached by all. A general equilibrium is defined as a state in which all markets and all decision making units are in simultaneous equilibrium. A general equilibrium exists if each market is cleared at a positive price, with each consumer maximizing satisfaction and each firm maximising profit. The scope of general equilibrium analysis is the examination, whether this state can, if ever, be reached that is how prices are determined simultaneously in all markets, so that there is neither excess demand nor excess supply, while at the same time the individual economic units attain their own goals.

EXERCISES

Q. No. 1 How we measure the consumer and producer surplus?

Q. No. 2 What is Welfare Economic? Explain different methods for measuring the social Welfare.

Q. No. 3 What is social welfare? How we measure it with the help of paretocriteria.

Q. No. 4 Explain the Pareto-Optimality criteria for measuring the social welfare.

Q. No. 5 What is difference between partial equilibrium and general equilibrium analysis. Explain the general equilibrium analysis in simple economic system.

Suggested Readings :

1. Koutsoyiannis, A Modem Microeconomics, 2nd Edition, Macmillan Publication, Landon.
2. Landsberg, Steven E. Price Theory and Application. The Dryden Press, Landon.
3. Salvatore, S., Managerial Economics, Mc Graw Hill Publications.

The Theory of Distribution
Francis Y. Edgeworth

Quarterly Journal of Economics, February, 1904.

Distribution is the species of Exchange by which produce is divided between the parties who have contributed to its production.[1] Exchange being divided according as both, or one only, or neither of the parties have competitors, Distribution is similarly divided. The case in which both parties have competitors will here be first and principally considered.

The simplest type of this distributive exchange would be of a kind which is effected once for all, without reference to a series of future productions and exchanges. For example, to adapt an illustration used by Mr. Henry George,[2] let it be supposed that on a particular occasion each out of a number of white men hires one or more black men to assist in catching seals, on the agreement that each white man shall give his black assistants a certain proportion of the take, the terms having been settled in an open market in which any one white is free to bid against any other white and any one black against any other black. A conception more appropriate to existing industry is that each white agrees to pay in exchange for a certain amount of service a definite quantity of produce, not in general limited to the result of a particular operation. On a particular day less seal may be taken than the employer has agreed to give the employee for the day. In this case, even if payment is not made till the end of the day, the employer must pay for help on a particular day in part with seal caught on a previous day. He must pay altogether out of past accumulations when payment is made before the work is done. When the employer agrees to pay a definite amount, he cannot expect to gain on each day's transaction, but on an average of days.

This example is suited to illustrate some general properties of Exchange which attach to Distribution as a species of Exchange. Such are the laws which connect a change in the supply or demand upon one side of the market with a change in the advantage resulting from the transaction to the parties on either side. Thus, competition on both sides being presupposed, a decrease of supply in a technical sense of the term on the one side is, *ceteris paribus*, universally attended with detriment to the other side, but is not universally attended with detriment to the side on which the supply is decreased. Accordingly, a limitation of supply on one side may be advantageous to that side, though not to both sides. The case of Distribution compared with Exchange in general in respect to such limitation of supply has only this peculiarity,—that the danger of this policy defeating itself is in the case of Distribution specially visible and threatening. There is an evident limit to what the black man dealing with the white man can get in exchange for a certain amount of his service; namely, the total product which that service utilised by the white man will on an average produce. To be sure, there is here but a case of the general principle that no one will give more for a thing, whether article of consumption or factor of production, than the equivalent of its total utility to him, which total diminishes as the quantity of the Commodity is reduced. But this limit is less liable to escape attention when it is fixed by the material conditions of production rather than by the

---

1. This definition, if not made more specific, includes some kinds of International Trade, *just* as the generic definition of International Trade includes some kinds of Distribution. See II. 5, 19.

2. *Progress and Poverty*, Book I. chap. iii.

desires of consumers. Conspicuous warning is given to parties in the position of our black men not to attempt to benefit themselves by a considerable reduction in their supply of service; for, though they might possibly obtain a larger proportion, they would probably obtain a smaller portion, of the average product. The laws which have been stated and other general laws of Exchange are equally true in more complicated cases of Distribution.

So far, we have supposed only a single factor—the service of the black man, or, more generally, the factor ß—offered by the competitors, say, B_1, B_2, etc., in exchange for some of the produce a offered by the competitors, say, A_1, A_2, etc. Let us now introduce other kinds of factors, γ, δ, etc. And let us no longer suppose payment to be made by parties of the type A, in the kind of commodity which is produced, namely, *a*. A more concrete conception is that, besides the group A, B, C, D, there is another and another group,—A', B', C', D'; A", B", C', D';— where each capital letter typifies a set of competing individuals. It may be supposed that each A purchases out of the finished product that he turns out—namely, *a*—portions of the products a', a', etc., which he distributes according to the law of supply and demand among parties of the type B, C, D. In fine, each A may pay for the factors of production altogether in some one product, a"',— "*numéraire*," as happily conceived by M. Walras, or, less generally, money,—which the purveyors of the factors can exchange for the articles which they want. These articles need not be all commodities ready for consumption: some of the parties may care to purchase factors of production wherewith to play the role which has been assigned to A.

Having now obtained a general idea of the machinery by which distribution in a regime of competition is effected, let us go on to consider in more detail the parts of the mechanism. And, first, of the party that takes factors of production in exchange for products or the means of purchasing the same, the party above represented by the white man and labelled A. The functions of this party may be investigated by an ancient method which Sidgwick has proposed to rehabilitate[3] for the purposes of modern economics,—the search for a definition. What is an entrepreneur? Amid the diversified combinations of attributes which the industrial world presents—innumerable as the varieties in which vegetable nature riots—we ought to fix certain characters agreeably to the rule laid down by Mill under the head of Definition by Type. "Our conception of the class" should be "the image in our minds which is that of a specimen complete in all the characteristics."[4] *Four* such type-specimens may be distinguished, ranged in a descending order according to the extent of functions ascribed to the entrepreneur. There is, *first*, the party whom the classical writers designate as the Capitalist, "who from funds in his possession pays the wages of the labourers, or supports them during the work; who supplies the requisite buildings, materials, and tools, or machinery; and to whom, by the usual terms of the contract, the produce belongs to be disposed of at his pleasure."[5] This party will here be considered as devoting his care and savings to a single business. There is, *second*, the entrepreneur as portrayed by the late President Walker, "not an employer because he is a capitalist, or in proportion as he is a capitalist."[6] There is, *third*, the party to whom Mr. Hawley

---

3. *Political Economy*, Book I, chap. ii, § 1.

4. *Logic*, Book III, chap. vii.

5. Mill, *Political Economy*, Book II, chap. xv, §1.

6. *The Wage Question*, p. 228.

would wish to restrict the term "entrepreneur,"[7] the man who undertakes risks, of which class the most prominent, though not the only, species is the investor in joint stock companies.[8] *Fourth*, at the extreme degree of tenuity, is the entrepreneur who makes no profit. It might seem, indeed, as if this class did not call for special treatment, as differing only in the amount, not in the kind of remuneration. A fig tree which bears no fruit is not therefore a tree of a distinct species. The horse which the Scotchman its owner had just trained to live upon a minimum, when the animal unfortunately died, was not therefore a new variety of the equine genus, requiring mention in a treatise on Natural History. However, as imposing theories have been connected with this last category, it comes within the scope of the present inquiry.

As our aim in comparing definitions should be, as Sidgwick says, "far less to decide which we ought to adopt than to apprehend the grounds on which each has commended itself to reflective minds,"—the hunt for a definition being followed not so much for the sake of the quarry as of the views which are incidentally presented,—let us go on to consider the principal propositions which the several conceptions are adapted to bring under our notice. In this inquiry much assistance will be obtained from a series of articles on cognate subjects in the *Quarterly Journal of Economics*,[9] which forms a sort of economic symposium.

The *first* definition is particularly suited to inquiries in which the parties who are in the habit of saving are contrasted as to their actions and interests with the parties who do not save, approximately, the working classes. Specimens of such inquiry may be found in the fifth chapter of Mill's first book, and in Professor Taussig's important article on "The Employer's Place in Distribution."[10] It sounds paradoxical to add that the classical conception is not particularly adapted to illustrate the Ricardian theory of rent. But the definition of the capitalist above given is not easily reconciled with the received representation, that the capitalist's remuneration is equal to the number of doses which he lays out, multiplied by the remuneration of the last dose, the ordinary rate of profit. For, as Sidgwick argues, there is no adequate reason for expecting that "remuneration for management" as well as interest should tend to be at the same rate for capitals of different sizes.[11] Doubtless, the proposition is accurate enough to support the practical consequences which have been deduced from it. But, while fully admitting this, one may still agree with Sidgwick that "even Mill's

---

7. *Quarterly Journal of Economics*, Vol. VI (1892) p. 283; VII, p. 459 et seq.; XV, p. 77 et seq.

8. Compare Mangoldt, *Unternehmergewinn*, pp. 41–43. A person who does not work, "wie der stille Gesellachafter, hört darum nicht auf, wahrer Unternebmer zu sein." This type is the limiting case, short of which the trouble of management in various degrees is combined with what Mr. Hawley calls "the irksomeness of risk." As Professor Taussig says, "The corporation of modern times presents all possible varieties of the relation between active manager and idle investor. Nominally, the stockholders are a group of associated active capitalists. Practically, they range from shrewd managers to the most helpless of inactive investors." *Quarterly Journal of Economics*, Vol. X (1895) p. 83. Cp. Marshall, *Principles of Economics*, Book IV, chap. xii, §§8 and 9.

9. References to the series up to November, 1900, are given in the *Quarterly Journal of Economic.*, Vol. XV, p. 75.

10. *Quarterly Journal of Economics*, Vol. X, p. 72.

11. *Political Economy*, 3d edition, Book II, chap. ix, §3. Cp. chap. ii, §8.

expositions" is "highly puzzling." For the idea of an economic person laying out doses up to the margin and obtaining the remuneration equal to the number of doses multiplied by the marginal productivity of each dose is only proper to the case in which the doses are for sale. But it is only in the conditions proper to our third definition that doses of capital are put on a market in exchange for profit. Perhaps the classical writers, having an eye to practice and not restricted by a sharp definition, often tacitly introduce the supposition that it is open to the "capitalist" to take part in some other business besides his own.[12]

The classical formula for surplus may be employed along with our *second* definition if we use the phrase "amount of outlay multiplied by average rate of return" to designate the amount which the entrepreneur of the Walker type pays in the way of interest from year to year to those who have lent him the means of carrying on his business. The surplus, according to this conception, will include not only the landlord's rent, but also the entrepreneur's net income. The portion of this surplus which accrues to the entrepreneur is not given by any simple formula. The conditions by which it is determined may be considered under two heads, corresponding to Cairnes's categories,—commercial and industrial competition. This distinction becomes clearest when, in conformity with the division of employments, we conceive different occupations to be separated by great gulfs, so that they who would pass from one to the other must make a complete, or at least a considerable, change in their business arrangements.[13] In virtue of the first kind of competition the entrepreneur endeavours to make the best possible arrangements within the occupation which he has chosen. In virtue of the second kind of competition he endeavours to choose the occupation which will afford to him the greatest net advantage.

His motive under the first head may be understood by likening him to a monopolist who does not control the prices of the factors of production, nor yet the price of the product, the latter being fixed by a maximum law, or, rather, the case being that in which the monopoly is just becoming extinct, as Cournot would say, by the introduction of competitors, so that this entrepreneur can no longer sensibly alter at will the price of the product. Under such circumstances each entrepreneur will vary all the variables under his control up to the margin at which his own advantage becomes greatest. If he or we be content with a rough estimate of this advantage, it may be measured by the

---

12. Cp. Mill on various employments of capital, *Political Economy,* Book II. chap. xv. §1, par. 4.

13. See note to the present writer's Address to the British Association, Section F, 1889 (a, vol. ii.), which, written before the publication of Marshall's *Principles of Economics,* does not sufficiently emphasise the "principle of continuity." It may he observed that the two kinds of competition involve respectively two mathematical operations, the determination of a maximum, and of the greatest among maxima. There is the distinction between finding the top of a hill and finding the highest hill-top. The demarcations between the two species of competition and between the two mathematical operations are not coincident, so far as an entrepreneur, without leaving his business, may introduce considerable and, so to speak, integral changes in its organisation, in accordance with the "principle of substitution" (Marshall). This principle seems to cover both the species of competition and both the mathematical operations. Doubtless, it is convenient to have a term applicable to every method by which maximum advantage is sought. Among such methods ought, perhaps, to be placed the *calculus of variations,* where the "margin of profitableness" is considered as "a sort of boundary line, cutting one after another every possible line of business organisation." *Principles* of *Economics,* Book VI. chap. vii. §7, 4th edition.

difference between his incomings and outgoings. His incomings may be regarded as the product multiplied by the price thereof, the amount of the product depending in some definite manner on the amounts of the factors of production which are employed.[14] The outgoings may be regarded as a sum of terms, each of which is the amount of a factor of production multiplied by its price.[15] It follows[16] that in a state of equilibrium the increment of value produced by the last increment of a factor is just equal to its price. **"**The *marginal* shepherd... adds to the total produce a net value just equal to his own wages."[17]

---

14. Some *function of* the amounts.

15. Or, rather, the *accumulated* price, in the sense explained by Professor Marshall *(Principles of Economics Book* V. chap. iv, §2, p. 432, 4th edition): "Looking backwards, we should sum up the net outlays, and add in accumulated compound interest on each element of outlay." Compare note xiv. of his mathematical Appendix Abstraction was made of this sort of correction in the British Association Address to which reference has been made. For instance, it Was tacitly assumed that the entrepreneur might have as much labour as he could pay for (at a prevailing rate of wages) at the time when the value of the finished product was realised. Professor Barone has pointed out the need of greater accuracy and a means of obtaining it by employing his remarkable conception of "capital of anticipation." *Giornale degli Economisti,* February, 1896.

16. Marshall, *Principles of Economics,* Book VI. chap. i, §8, 4th edition. Mr. J. A. Hobson's criticism of this doctrine exemplifies the difficulty of treating the more abstract parts of Political Economy without the appropriate mathematical conceptions An elementary discipline in the differential calculus would have corrected the following passage and its context: "In order to measure the productivity of the last dose of labour, let us remove it. The diminution of the total product may be 5 per cent. This 8 per cent, according to Marshall's method, we ascribe to the last dose of labour. If now, restoring this dose of labour, we withdrew the last dose of capital, the reduction of the product might be 10 per cent. This 10 per cent, is regarded as the product of the last dose of capital. Similarly, the withdrawal of the last dose of land might seem to reduce the product by 10 per cent. What would be the effect of a simultaneous withdrawal of the last dose of each factor? According to Marshall's method, clearly 28 per cent. But is this correct?" *The Economics of Distribution*, p. 146. Quite correct, if in the spirit of the differential calculus we understand by dose an increment as small as possible, not as large as the objector pleases. He goes on: "Put the same experiment upon its broadest footing, and the overlapping fallacy becomes obvious. Take the labour, capital, and land as consisting of a single dose each; now withdraw the dose of labour, and the whole service of capital and land disappears. Is the destruction of the whole product a right measure of the productivity of the labour dose alone? "(loc. *cit.,* p. 147). Imagine an analogous application of the differential calculus in physics, "put upon its broadest footing," an objector substituting x wherever a mathematician had used dx or Δx!

17. It being assumed that the function expressing the product in terms of the factors of production is such that for the values of the variables with which we are concerned the net income of the entrepreneur may be a maximum, let P be the amount of the product, π its price, a, *b,* c, amounts of factors of production, p_1, p_2, p_3, etc., their respective prices—their actual prices—for a first approximation, their *accumulated* prices for a more accurate statement. The net income of the

So far supposing the entrepreneur's work to be a constant quantity. In a more exact estimate the quantity which the entrepreneur seeks to maximise is the utility to be derived from his net income *minus* the disutility incident to its production. From this consideration it follows that the increment of utility due to the increment of product which is produced by the last increment of entrepreneur's work is just balanced by the increment of disutility due to that work.

To this condition is superadded the tendency towards equal net advantages in different occupations, resulting, as Professor Marshall has shown, not so much in the equal advantageousness as in the equal attractiveness of different occupations. The remuneration of the entrepreneur thus corresponding to his services may be classed along with the remuneration of the workman as "earnings," from a certain point of view, which is doubtless proper to the publicist and philosopher. As Mangoldt points out, "the circumstance that certain services do or do not attain a market price" does not **"**essentially alter the measure of their compensation." But there is another point of view which is proper to those who study the mechanism of distribution. As Professor Taussig well observes, "The cobbler who works alone in his petty shop gets in the main a return for labour as much as the workman in the shoe factory"; but "with regard to the machinery by which distribution is accomplished he [the cobbler] belongs in a different class from the hired labourer."[18]

The tendency to equality of net advantages of course only exists with respect to positions between which there is industrial competition. Accordingly, if the union in one person of natural abilities and money constitutes him a member of a "non-competing group," there is no presumption that the remuneration of such an entrepreneur will be exactly equal to the interest which he might have obtained by lending his money plus the salary which a person of his ability could command as a hired manager. There exists an excess above that sum, corresponding to what Mangoldt calls *Unternehmergewinn.* There may be excesses somewhat similarly caused by different degrees of ability and resources; the various rents" enumerated by Mangoldt, which, as he observes, tend to diminish with the progress of society, so far as education becomes more diffused and it becomes easier for persons properly qualified to obtain the use of capital.

Some additional light on the functions of the entrepreneur may be obtained by comparing the profits in businesses of a different size. Suppose (for the sake of the argument) that the work and

---

entrepreneur may then be written (abstraction being made of the entrepreneur's own effort). $P = \pi f(a,b,c) - p_1 a - p_2 b - p_3 c$. In order that this expression may be a maximum, the law of decreasing returns must hold in the *first* of the two senses elsewhere distinguished. The condition must still be postulated when account is taken of the entrepreneur's subjective feelings,—effort and sacrifice in the way of production balanced by satisfaction immediate or prospective in the way of consumption. Nor is the case essentially altered when account is taken of the possibility (noticed by Professor Pareto, *Cours,* Art. 718) that the factors are not independent. Suppose that the amount of labour must always be in proportion to, or on any definite function of, the amount of land. Then, eliminating one of these quantities, we may treat the other as independent.

18. *Quarterly Journal of Economics,* Vol. X (1895) p. 88. Professor Taussig goes on, "For an understanding of the machinery by which distribution is accomplished in modern times, the classification of sources of income should thus be different from that to be adopted for an explanation of the fundamental causes."

worry of the "boss" do not increase[19] with the scale of operations, how is the equality of net advantages which theory leads us to expect brought about? *Ceteris paribus,* might we not expect the entrepreneur's residue to be larger in the large industries?[20] The answer seems to be that, as equilibrium is approached under the joint influence of Commercial and Industrial Competition, the amounts of the factors[21] are so varied as to fulfil the condition that equal efforts and sacrifices on the part of the entrepreneur are attended with equal remuneration.[22] This equality is irrespective of identity in the relation between factors and product.[23] It may exist whether that identity is supposed to be present between industries of different sizes or, as in general to be supposed, there is no identity in the relation between factors and product for different individuals and industries.

The sort of adjustment thus postulated may be illustrated by a more familiar kind of surplus, that which accrues to the landlord according to the received theory of rent. Let there be a homogeneous tract of land equally adapted to the cultivation of wheat and barley, owned by a set of competing landlords, who accordingly obtain an equal rent per acre whether wheat or barley is to be grown thereon.[24] Now let a tax be imposed on the rent of land used for growing barley. There must result a new equilibrium, in which it remains true that owners of homogeneous land obtain equal rent per acre for whichever purpose used, and that cultivators of wheat and barley obtain, *ceteris paribus,* equal profits. These conditions can be fulfilled if the extent of the land applied to the cultivation of wheat is increased while the intensity of cultivation is diminished, and contrariwise for barley the extent is diminished and the intensity increased. This proposition holds good whether or not the relation between outlay and product[25]—corresponding to the shape of the curve in the illustration which Professor Marshall has made familiar[26]—is supposed identical for wheat and barley, and even if the cultivator seeking the greatest possible profits is able to vary that relation in accordance with the **"** law of substitution." It is here assumed that the case of manufacture is not so different from agriculture, but that an analogous adjustment of "margins" must be considered to take place between large and small businesses under the conditions specified, and generally between different industries where industrial competition acts.

---

19. That the trouble does not increase proportionately would be a more concrete supposition. As Sidgwick says, "Though it is more troublesome to manage a large factory than one half the size, it can hardly be twice as troublesome." *Political* Economy, Book II, chap. ix, §3.

20. Cp. Marshall, *Economics of Industry,* Book II. chap. xii. §4, 1st edition.

21. The factors generally, and sometimes also the form of the function expressing the quantity of the product in terms of the quantities of the factors used, the function designated *f* in note 17.

22. The equality is that of an ordinary equation, not an identity.

23. The function which expresses the amount of the product in terms of the factors (including entrepreneur's work).

24. Compare II. 78.

25. The function expressing the product in terms of the outlay.

26. *Economics of Industry,* 1st edition, p. 83. *Principles of Economics,* 4th edition**,** p. 232.

A similar adjustment must be postulated when we entertain the *third* definition of entrepreneur, and consider competing investors in the stock of companies which may at first be supposed equal in respect of risk, though not in size. The competitors being free to invest units consisting, say of £100 or less in any kind of business (of the given riskiness), large or small, it follows that a return to a dose anywhere invested tends, *ceteris paribus*, to be the same.[27] This result, which is by no means a deduction from the general formula considered under our second head, may be supposed to be brought about by an adjustment of margins of the sort which has been explained.

Now at length the Ricardian theory of rent as ordinarily stated becomes exact,—the payment for land rented by a joint stock company ought to be just the difference between the returns (after capital has been replaced and labour paid) and the amount of capital laid out, multiplied by an average rate of profit.

Though the class of shareholder is the principal, it is not the only species, of the third kind of entrepreneur, if defined so as to include all risk-takers. As Mr. Hawley observes,[28] workmen take some risk, entrepreneurs who have no capital of their own run the risk of not being paid for their trouble. Enterprise may be taken as the essential attribute of a wide class entitled to a share in the national dividend along with the purveyors of land, labour, and capital. It does not seem to be a fatal objection that enterprise is hardly to be found in the concrete, separate from other factors of production. As Mr. Hawley replies,[29] labour and waiting, the attributes of familiar classes, are not to be found in abstract purity.

To some there may seem a more serious scruple: whether the undertaking of risk does even in thought constitute a fourth factor, whether the distinction between interest and the reward for risk is radical. It is all very well for Jevons to distinguish by different coefficients, p and q, the depreciation of future goods due to uncertainty and to remoteness. But, since the distant pleasure is always uncertain, can we really disentangle the two causes of depreciation?

Fortunately, these questions of logical definition and psychological analysis do not affect the important lessons respecting the participation of risk which have been taught by Professor J. B. Clark,—"that a corporation can run risks which the individual could not with prudence," that by forming corporations "we reduce the initial terrors of business enterprises."[30] It is an exemplification of the old maxim not to put all one's eggs in one basket. If a hundred persons are carrying each a hundred eggs, each independently running the risk of tripping and by the loss of all or many of his eggs being exposed to great privation, this great danger will be averted, this chance of great disaster will be commuted for a somewhat higher probability of a much more easily borne loss, if each person carries only one of his own eggs and one belonging to each of the rest, the total to be redistributed at the end of the journey to market or after sale.

It is noticeable that in Professor Clark's nomenclature this risk is borne by the capitalist. "The

---

27. Accordingly, in order that equilibrium should be stable in this regime, investment in each industry ought to be pushed up to a point at which the law of decreasing returns is fulfilled in its *second* sense,—that the rate of total cost to total product increases with the increase of product.

28. *Quarterly Journal of Economics,* Vol. VII, (1893) p. 470.

29. Ibid., Vol. XV, (1900) p. 78.

30. "Insurance and Business Power," Ibid**.,** Vol. VII, (1892) p. 40, *et seq.*

hazard of business falls on the capitalist." "Business repays men not only for their labours, but their fears." But this repayment is "not a part of mercantile profit": it is realised by the capitalist "as such." Admitting a real remuneration for risk, while giving a different name to the recipient from that which others have preferred, Professor Clark is perhaps not committed to the paradox which Mr. Hawley would affix upon the conception of the entrepreneur with vanishing profits,—our *fourth* species.[31]

"To eliminate profit, wholly static conditions must be more absolute. ... There must be a cessation of all variations due to the changeableness of the environment due to fire, lightning, hail. We must imagine industrial society in the static condition as an automatic machine,... working without friction in an absolutely unchangeable environment."[32]

This idea of perfect tranquillity is certainly inappropriate to the troubled world in which we live. "Things are always finding their level," like a fluctuating and, in nautical phrase," confused" sea. The oscillating character of the waves is quite consistent with a gradual change of level, as when the tide is flowing. It is a legitimate conception, familiar in statistics, to regard a phenomenon as hovering about an average, even though that average is known to be changing. Let the great tidologist calculate the dynamics of the flow, but let him not convey the impression that but for the action of this flow there would be the level of the proverbial mill-pond. Very probably, however, Professor Clark would recognise the continuance of risk not involving secular progress,—due to unpredictable weather or credit cycles, for example,—but would regard the remuneration for undergoing such risk as accruing to the "capitalist as such" rather than, with Mangoldt and others, as a part of the entrepreneur's gain. With regard to other elements of remuneration it is more doubtful whether Professor Clark would accept Mangoldt's statements as to the permanence of the entrepreneur's gain,—statements which read with their context, and attention being paid to Mangoldt's terminology, deserve much consideration.

We must suppose the existence of undertaker's gain [*Unternehmergewinn*],—otherwise what object has the entrepreneur to increase his business? (substance of p. 50).

The undertaker's gain *(Unternehmergewinn)* is "not simply something transitory," but a "permanent species of income" (p. 51).

"The undertaker's remuneration [*Unternehmerlohn*] preserves its position, though in a limited form" (p. 105. *Cf.* p. 169).

Perhaps Professor Clark would be satisfied with the "limited form" of the remuneration and the disappearance of certain other elements.

It is always pleasant to believe that one's differences with high authorities are only verbal. This satisfaction may now be enjoyed with respect to M. Walras's doctrine that the entrepreneur makes neither gain nor loss. Professor Pareto[33] has made it clear that, as the object of the entrepreneur is to procure the greatest amount of satisfaction, so his income is not to be considered as *nil*, in the ordinary sense of the term. Rightly interpreted, the doctrine that "the entrepreneur makes neither gain nor loss," taken in connection with the "coefficients of production," appears to cover all the conditions of equilibrium, both those which are involved in what Cairnes called

---

31. See the appended note, referring to the observations on Risk in Mr. Keynes' *Probability.*

32. *Quarterly Journal of Economics.* Vol. XV. (1900) p. 91.

33. *Cour. d'Économie Politique,* passages referring to "entrepreneur."

"industrial competition" and those which would be satisfied even if we made abstraction of the tendency to equal advantages in different occupations. But, while we accept the ideas, we are not bound to adhere to the words of a master; and the expression in question may be objected to on several grounds which will repay examination. It is violently contrary to usage; it lends itself to a dangerous equivoque; and it has led distinguished economists to paradoxical conclusions.

No amount of authority and explanation can make it other than a strange use of language to describe a man who is making a large income, and striving to make it larger, as "making neither gain nor loss." There is an oddity about the phrase which recalls the use of "gratis" by Sir Murtagh's lady in *Castle Rackrent*: "My lady was very charitable in her own way. She had a charity school for poor children where they were taught to read and write gratis, and where they were kept well to spinning gratis for my lady in return."

A more serious objection is that the term "making neither gain nor loss" has to be used in two different senses almost in the same breath. It is a sufficiently difficult lesson for the plain man to learn that the maximum of income which the entrepreneur aims at realising is zero. But the difficulty is doubled when he comes to learn—as he must in dealing with a maximum problem—that the increment to that income due to the last increment of any factor of production is also zero. There is apt to arise a confusion between conditions belonging to the total and to the marginal quantity,—an ambiguity of a kind which has before now proved detrimental in economics.[34] A hasty reader of Professor Walras might suppose that it was intended to affirm that the entrepreneur made neither gain nor loss at the *margin*: whereas the meaning is, rather, that nothing remains to be distributed—on an average and apart from oscillations—after that the entrepreneur has paid a normal salary to himself.[35]

The implication that the remuneration of entrepreneur labour may be treated like that of any other labour presents some difficulty. It is the one obscure topic in Professor Barone's brilliant studies on Distribution.[36] His observations deserve to be quoted at some length. He first (in a note on p. 132) announces as true in a particular case, what is here regarded as true in general, that " there must be left to the entrepreneur's profit (*profitto dell' impresa*) the differentiating character of 'residual claimant'; and nothing else can be said but that profit is formed by the difference between the entire product and the remunerations of the various factors corresponding to (*ragguagliate*) their respective marginal productivities." But Professor Barone regards this enunciation as only provisional. He promises to show in a later section that "with the increase in the number of the competing entrepreneurs the profit of the undertaking tends to lose more and more the character of residual claimant, and tends to conform to that of the law of marginal productivity."

In the later section he says:—

---

34. Mill's hesitation between equal sacrifice and least sacrifice as the criteria of taxation may seem due to a confusion of this kind, is pointed out by the present writer in the *Economic Journal*, 1897. (Cp. *Mathematical Psychics*, p. 118.) Mill's ambiguity had already been noticed by Professor Carver in his article on "The Ethical Basis of Distribution" in the *Annals of the American Academy* for 1895, p. 95.

35. Cp. Pareto, *Cours*, Art. 87, "his salary as director of the enterprise being comprised in the expenses of production"; and the similar expressions of Professor Barone, quoted below.

36. *Giornale degli Economisti*, February, 1896.

"If on the market there is only one entrepreneur, Titius, and if he does not monopolize the product, that is, if he in the management of his business arranges [*fa in modo di*] to obtain not indeed the greatest monopoly profit, but the greatest profit Obtainable in a regime of free competition,... his profit will be [a surplus indicated by a figure which is not here reproduced]. But, if there is an entrepreneur Caius capable of entering into competition with the preceding,... the profit of Titius will be reduced below what he had when he was alone on the market. And, if there is a third employer also capable of entering into competition with the first two, the profit of Titius will be reduced still more. The more the number of employers increases, the more there is a necessary tendency to a limiting state in which all the employers who continue to produce have a remuneration which, like that of any other labour, satisfies the condition that the marginal disutility [*penosita*] of the same labour [*medesimo*] shall be equal to the marginal utility of the returns which that labour procures, *and not more than this*. And, since it is this equality which characterizes the return to labour, it follows (*ne viene*) as a legitimate consequence that in this limiting state the remuneration of the entrepreneur may be treated like the remuneration of any other species of labour."

The fact that wages are usually paid in advance is not to the point, as Professor Barone very properly observes. He proceeds:—

"These considerations seem to me to prove to demonstration how profound and correct is Walras's conception of an entrepreneur who under the conditions postulated makes neither gain nor loss after having paid himself (or others, it is indifferent which) the remuneration of the labour of direction and conduct of production. And, if it is no wonder that this conception should not be comprehended by economists who have really very vague ideas of quantity, it is absolutely astounding that the conception should have been also made the subject of criticism by other economists to whom the notions of quantity are quite familiar.... I frankly must confess myself absolutely incapable of understanding how any difficulty whatever can arise as to the validity [literally, the affirmation] of this conception, which is indeed most simple."

Having called once more attention to the abstract character of the conditions, Professor Barone reiterates.—

"In such conditions the law of marginal productivity extends to the remuneration of the entrepreneur; and, after having remunerated all the factors (the work of the entrepreneur included) in proportion to their marginal productivity [with a discount corresponding to the time elapsing between the service and the product], there remains no undistributed residue."

If there could be any about the meaning of this thesis, it would be removed by the unequivocal language of symbols employed in the Appendix,[37] where, by way of illustration, the labour of the entrepreneur is expressed by the total number of hours of work that he devotes to the business.

Upon this it may be remarked that the last state of Titius, After Caius and the rest have entered as competitors, seems identical with the case of "extinct" monopoly which was above adduced, in order to exhibit the motives of the entrepreneur. As there appears, both before and after the competitors have entered the remuneration of the entrepreneurs, in Professor Barone's phrase, "satisfies the condition that the marginal disutility of the labour shall be equal to the marginal utility of the return which that labour procures." But neither before nor after the competitors have entered is there any reason for regarding the remuneration of the entrepreneur as the product of the number of doses (e. g., hours worked) and the marginal productivity of a dose (multiplied by a coefficient

37. Loc. cit.

depending on the length of the productive process[38]). It is only with respect to factors of production which are articles of exchange that the proposed law of remuneration, the "law of marginal productivity," is fulfilled in a regime of competition. Thus, in our typical example of black men assisting white men to catch seals, what the black man gets in a perfect market is an amount of seal equal to the number of units of service which he supplies, multiplied by the quantity of seal for the sake of which he is just induced to offer an additional unit of service, the unit employed being a small quantity. Likewise, what the white man gets in exchange is an amount of service equal to the amount of seal which he distributes to the black man, multiplied by the quantity of service for the sake of which he is just induced to offer an additional unit of produce. If the amount of service rendered may be taken as the measure of the black man's labour (or of some other factor of production supplied by him), the proposed law holds good for his share of the distributed produce. But, as the amount of produce given by the white man in exchange for services cannot be taken as the measure of his work, the proposed law does not hold for his share of the distributed produce.

This discussion will appear otiose to the economists who are not conversant with the science of quantity. The proposition that the remuneration of the entrepreneur is equal to the amount of his work multiplied by its marginal productivity will be interpreted by them as signifying simply that he will get more, *ceteris paribas*, the more work he does and the greater the addition to the produce which he would effect by doing a little more work. For them a *product* will do duty for a function of two variables which increases with the increase of either variable. But this easy interpretation is not open to mathematical economists. They must be aware that the formulae in question affirm some thing more than the simple truth just stated. If nothing more than that simple truth can be deduced from the theory of Exchange, it ought not to be a matter of surprise that the "law of marginal productivity" applied to the entrepreneur should be challenged by those who affect mathematical precision.

The law of marginal productivity, then, is not fulfilled in the sense that the portion of the national dividend accruing to entrepreneurs is a sum of terms each of which is the product of an entrepreneur's work reckoned in hours, or similar doses, and the marginal productivity of a dose (multiplied by a certain coefficient). Let us see whether the law is fulfilled when we take a larger dose, the total work of an entrepreneur. The law will then be fulfilled if the net gains of any entrepreneur tend to be equal to what society would lose if he were removed. Can this be generally affirmed? Let us look at the typical case of distribution between whites and blacks above instanced. It may be granted that the white entrepreneur does not normally obtain more than he adds to the common stock. For otherwise the society would gain through his removal, his black assistants eat er hunting by themselves or being taken on by other entrepreneurs. And neither of these suppositions is possible in a state of equilibrium; for, if either were possible, it would have been already brought about by the free play of self-interest, in a regime of competition. The gain of a white man, then, cannot be greater, but where is the proof that it cannot be less, than the loss which would be occasioned to the society by his removal?

Such a proof might be forthcoming if the white men were not, as hitherto supposed, genuine entrepreneurs, but managers acting under entrepreneurs of our third species, the stockholder. The income of the managers will fulfil the marginal law of productivity if the new entrepreneurs are conceived as competing against each other in such wise as to bring about the result that no manager

---

38. Remark that the correction proposed by Professor Barone for the effect of time is not identical with Professor Marshall's accumulation of price.

earns more or less than what he adds to the profits of his employers. The income of the new entrepreneurs also fulfils the law; for the remuneration of this species of entrepreneur—unlike that of entrepreneurs in general—is proportional to the amount of the factor which they contribute,—namely, capital invested.

The affinity between entrepreneurs and salaried managers in modern industry supplies the missing link for the general proof of the new law. For, normally, it may be presumed that an independent entrepreneur (of our second species) does not make less (in addition to the profits that he makes or might have made by investing in some other business money of his own) than a manager of like abilities. And perhaps he does not make much more. The difference is possibly small,[39] probably diminishing, certainly difficult to verify statistically, perhaps hardly worth fighting about. Interpreted cautiously, the law holds good approximately. If the remuneration of the manager, like that of the "marginal shepherd," is just equal to the amount that he produces, then the remuneration of the entrepreneur is not very different from the amount that he produces. But, if the law of marginal productivity is fulfilled for the manager only while we consider doses less than his total work, say hours of work, then the law is fulfilled for the entrepreneur only so far as it is presumed from the similarity in nature and habits between the manager and entrepreneur that, when the total remuneration of each is nearly the same, the amount of work and its marginal productivity are not very different.

According to the interpretation which has been suggested, the new law of distribution would be fulfilled by an adjustment of the quantities involved, the amount of each factor, not simply in virtue of the relation which subsists between the product and the factors of production.[40] The sense in which the law is fulfilled is otherwise conceived by a distinguished mathematical economist, Mr. Wicksteed, who regards the law as following from "the modern investigations into the theory of value,"[41] and seems to treat it as a clue whereby to investigate the nature of the relation between the product and the factors of production, including the work of the entrepreneur.[42] In fact, he finds that the product depends upon the factors by a relation which mathematicians designate a "homogeneous

---

39. Mainly and apart from "rents" of the order of quantity called by Mangoldt *Unternehmerlohn*.

40. The form of a function such as that represented by *f* in a preceding note, or rather what that function becomes when the work of the entrepreneur enters as a variable.

41. *Essay on the Co-ordination of the Laws of Distribution* (1894), §2, and prefatory note.

42. The product being a function of the factors of production, we have $P = f(a, b, c, \ldots)$; and the form of the function is invariably such that if we have $\pi = f(\alpha, \beta, \gamma, \ldots)$, we shall also have $v\pi = f(v\alpha, v\beta, v\gamma, \ldots)$ (*loc. cit.*, p. 4).

"Let the special product to be distributed (P) be regarded as a function (F) of the various factors of production ($A, B, C, \ldots$)" (*loc. cit.*, p. 8).

$$\frac{dP}{dA}A + \frac{dP}{dB}B + \frac{dP}{dC}C + \ldots = P$$

"under ordinary conditions of competitive industry" (*loc. cit.*, pp. 33–38).

function of the first degree."[43] This is certainly a remarkable discovery; for the relation between product and factors is to be considered to hold good irrespectively of the play of the market: "an analytical and synthetical law of composition and resolution of industrial factors and products which would hold equally in Robinson Crusoe's island, in an American religious commune, in an Indian village ruled by custom, and in the competitive centres of the typical modern industries."[44] There is a magnificence in this generalization which recaps the youth of philosophy. Justice is a perfect cube, said the ancient sage; and rational conduct is a homogeneous function, adds the modern *savant*. A theory which points to conclusions so paradoxical ought surely to be enunciated with caution.

To sum up this criticism, as Distribution is a species of Exchange, it seems undesirable to employ a phrase so foreign to the general theory of Exchange as the dictum that one of the parties to an exchange normally gains nothing. Innocently used at first, such paradoxes are calculated to lead to confusion and misrepresentation.

A similar remark applies to another form of the gainless entrepreneur, involved in Walker's analogy between profits and agricultural rent.[45] Even on the simpler and provisional view which is confined to short periods and commercial competition, this form of expression has no advantage over the terminology proper to the general theory of Exchange. When we consider long periods and industrial competition, Walker's theory has the graver disadvantage of not distinguishing between rent and quasi-rent. It seems to be generally admitted that Walker's masterly portrait of the industrial captain was not improved by his representation of profits as rent.[46]

Having now considered the party that takes factors of production in return for products, or the proceeds thereof, let us look at the other side of the counter,—the triangular counter across which we may imagine the three factors of land and labour and capital to be exchanged, if we place in the interior of the triangle an entrepreneur of Walker's type, our second species, dealing with three parties in quick succession, and in some sense simultaneously.[47]

At the height of abstraction from which it is here attempted to survey the economic world, what appears the most salient feature in the transactions respecting *land* is the circumstance that the

---

43. As pointed out by Professor in his review of Mr. Wicksteed's essay *Economic Journal*, Vol. IV. p. 311. In Mr. Wicksteed's notation the function f must be of the general form $A\psi\left(\frac{B}{A}+\frac{C}{A}+\ \right)$, where ψ is an arbitrary function.

See Forsyth, *Differential Equations*, Art. 189, or Boole, *Differential Equations*, chap. xiv, Art. 6.

44. *Loc. cit.*, p. 42.

45. As argued by the present writer in his Address to the British Association for the Advancement of Science, 1889, written before the publication of Professor Marshall's weightier judgment in the *Principal of Economics*.

46. Compare Mr. J. H. Curran's temperate criticism in his study on Walker (in Conrad's *Abhandlungen*).

47. In the sense in which equations are called simultaneous.

quantity of ground, or at least space,[48] is limited, not capable of being increased by human effort. From this property flow most of the general theories relating to the landlord's share in distribution,—that a tax on rent (proper) falls wholly on the land, that the remission of agricultural rent by landlords would not benefit the consumer,[49] and other propositions often connected with the formula that " rent does not enter into the cost of production." Some remarks on that time-honoured formula seem called for here. It would not be consistent to have complained of the expression that "the entrepreneur makes no gain" as perplexing and apt to mislead, however innocently used by high authorities, and to pass over in silence this dictum about rent, against which and in favour of which much the same is to be said. Certainly, it is supported by very high authority,— the authority not only of Ricardo and Professor Marshall, but also of Hume, who in the letter which he wrote to Adam Smith on the publication of *The Wealth of Nations* (the letter which, written a few months before Hume's death, may be considered his economic testament) says, "I cannot think that the rent of farms makes any part of the price of the produce, but that the price is determined altogether by the quantity and the demand."[50] On the other hand, it can hardly be denied that the dictum in question is calculated to obscure the truth that "land is but a particular form of capital from the point of view of the individual manufacturer or cultivator";[51] that, as he doses land with capital and labour, so he doses capital and labour with land,[52] up to a margin of profitableness. And, in fact, the similarity of the factors of production from the entrepreneur's point of view does not seem to have been apprehended in all its generality by the classical writers. Thus Fawcett, who may be taken as a type, when explaining rent seems to posit the size of the farm as something fixed and constant.[53] J. S. Mill

---

48. Cp. Marshall on "extension" as the "fundamental attribute of land." *Principles of Economics*, Book IV, chap. ii. p. 221 et seq., 4th edition. Not even the enterprise of Boston, which converted marshes into the site of noble streets, can form an exception to the law so stated. But the more familiar statement is accurate enough. For, as Professor Bullock has said (at the banquet of the Massachusetts Single Tax League, 1902), "it may be safely contended that the additions which man can make to the land surface of the globe are so small as to be a negligible quantity when we compare land with the things that human labour places upon it."

49. The received proposition is of the nature of a first approximation, as pointed out in II 76. When the writer there observed that "there might be now required a higher rate of remuneration to evoke the same exertion from the cultivator," et seq., he was not aware that he had been anticipated by the very first writer who stated the true theory of rent, James Anderson, who says that the only consequence of remitting rents "would be the enriching one class of farmers at the expense of their proprietors, without producing the smallest benefit to the consumers of grain,—perhaps the reverse, as the industry of the farmer might be slackened." *Enquiry into the nature of the Corn-laws* (1777), p. 48, note.

50. Burton's *Life of Hume,* Vol. II, p. 486.

51. Marshall, *Principles of Economics,* Book V, chap. ii, §5.

52. The propriety of reversing the classical formula so as to make dose and patient change places is well expressed by Mr. Wicksteed, *Laws of Distribution,* p. 20.

53. *Manual of Political Economy,* Book III, chap. iii.

argues that "there is always some agricultural capital which pays no rent,"[54] not noticing the counter-argument that there is a portion of land which pays no interest.[55] These imperfections belong now, it may be hoped, to past history. And yet that the description of rent as not entering into price is apt to prove misleading may be inferred from the many protests which eminent critics eve raised against Professor Marshall's use of the time-honoured phrase.[56] Their criticisms attest the correctness of their own views rather than their capacity of appreciating the views of others. What should we say of critics who should think fit to read Mill a lecture on the errors of the Mercantile system, because Mill had employed the terms "favourable and unfavourable" exchanges! To have attributed to Professor Marshall the very error which he by his doctrine of the "Margin-of-building" has done more than any other economist to obviate would be unpardonable if it were not excused by the misleading associations of an unfortunate phrase.

To return to the real, from the seeming, import of the phrase, we see that, as the offer of land is in general attended with no real cost, a tax upon the payment for land does not disturb production.[57] On grounds of distribution, too, a sort of income which increases without any effort on the part of the recipient is *prima facie* a suitable object for a specially heavy impost. On these grounds Mill's proposal to tax away the future unearned increment of rent is defensible, if accompanied with Mill's proviso, that existing interests should not be disturbed. For, as argued elsewhere, a special tax on existing incomes from land would violate the two principal conditions of a good tax: it would both tend to diminish the amount of production, and also to impair the equality in the distribution of burdens between the owners of incomes derived from land and from other kinds of property.

The practical importance of Mill's proposal is greatly reduced by the proviso with which it is accompanied. For, in order that the State may make a good bargain by giving the market price for a certain class of future goods, the State must be able to look further ahead—must exercise the telescopic faculty of prospectiveness in a higher degree—than the ordinary capitalist. And it may well be doubted whether this condition is fulfilled by the politicians who act on behalf of the State. We hear much of instances, like that of Chicago, where the value of sites is said to have multiplied

---

54. *Political Economy,* Book II. chap. xvi. §4.

[55] As noticed by Professor J. B. Clark and others mentioned by Professor Fetter in the *Quarterly Journal of Economics,* Vol. XV, not to p. 436.

56. See in particular Hobson's *Economics of Distribution,* chap. iv, Fetter, "The Passing of the Old Rent Concept," v and vii, (3), *Quarterly Journal of Economics,* Vol. XV, (1901); J.B. Clark, *Political Science Quarterly,* March, 1891; Wicksteed, *Laws of Distribution,* p. 47 (the last critic not referring *nominatim* to Professor Marshall). For a more sympathetic criticism of Professor Marshall's doctrine see *Economic Journal,* Vol. V. p. 589.

57. As Professor Carver said lately (at the banquet of the Massachusetts Single Tax League, 1902), a person who thinks that the repressive effect of a tax on land is at all comparable with the repressive effect of a tax on the products of industry must have an eye for exceptions like "a certain senator of whom it was said that he could see a fly on a barn-door without being able to see the barn or the door either." The incident in question may be elucidated by representing the "supply curve" of land as a perpendicular line. Cp. II. 69.

some eighty-fold in half a century; but we hear little of proposals to buy up at their present market value the site of some future Chicago, unless, indeed, as part of a scheme for Land Nationalisation, which does not include compensation to vested interests. Unlike the husbandman, who plants trees the fruit of which he will not himself see, the advocates of a single tax and other socialist agitators grasp at the standing crop which has been sown by others, heedless whether cultivation in the future is thereby discouraged.

But, even if their outlook were as distant as it is bounded, there would remain the possibility that, though looking far ahead, they might not discern distant objects clearly. Mill cannot be accused of the shortsightedness which sacrifices the future to the present. He looked very far ahead. But he did not see what was coming, the fall of English rents. Actuated by the highest motives, he proposed an arrangement which was perfectly just to the landlords, and would have proved perfectly disastrous to the State.

Passing in the traditional order from *Land* to *Labour*, we may begin by considering a very abstract labour market, in which the difficulty caused by the "advance" of wages is kept out of sight.[58] The following example of such a labour market may be worth reproducing, although it is not a genuine case of Distribution:—

Let us suppose several rich men about to ascend some an easy mountain, some a difficult one, each ascent occupying a day. And let these rich travellers enter into negotiations with a set of porters who may be supposed many times more numerous than the employers. An arrangement according to which the remuneration for ascending the easy and the difficult mountains was the same could not stand: it would not be renewed from time to time. For some of the porters employed on the difficult mountains, seeking to minimise the disutility of their task, would offer their services to travellers on the easy mountains at a rate somewhat less than the temporarily prevailing one. Nor would equilibrium be reached until each porter employed on a difficult mountain received an excess above the fee for the ascent of an easy one sufficient to compensate him for the extra toil. At the same time—simultaneously, in a mathematical sense—the increment of satisfaction due to the last porter taken on by each traveller would just compensate the purchaser of that labour for his outlay on it.[59]

In this example the great number of the employees as compared with the employers is not an accidental circumstance. Suppose that the arrangement which is common in the Tyrol—that each amateur ascensionist should be accompanied by only one guide—were for technical reasons universal. Then the bargain between travellers, on the one hand, and guides, on the other, would not in general be perfectly determinate. It would still indeed be true that "an arrangement according to which the remuneration for ascending the easy and the difficult mountains was the same could not stand." But it would no longer be true that the remuneration for the easy mountain—or, rather, for the average mountain, from which the fares both of the easier and the more difficult ascents might be measured—would be in general determinate.[60] There would in general exist no force of

---

58. There is an abstract point of view from which, as Professor Barone well observes (*Giornale degli Economisti*, loc. cit.), the circumstance that wages are paid in advance is of secondary importance.

59. *Economic Journal*, Vol. IV, p. 225.

60. As argued in *Mathematical Psychics*, p. 42.

competition by which any particular arrangement (as to the average mountain) initiated by custom and accident could be disturbed. That is, still supposing the service of a guide or porter to be sold as a whole. For, if the labour of the assistants can be sold by the hour, or other sort of differential dose, the phenomenon of determinate equilibrium will reappear. There seems no reason to think that the case of indeterminate equilibrium which has been illustrated is other than exceptional in the actual labour market, even where the bargain appears to be made for totals as distinguished from doses of labour,— situations rather than tasks. For there is, in fact, such a variety of situations attended with different amounts of work[61] as probably in practice to realise that divisibility of the thing supplied—here labour—which, together with the divisibility of the thing demanded,—here money,—constitutes a condition of a perfect market[62] with determinate equilibrium.[63] Still, the point of theory is worth notice. Perhaps the friction in the labour market would be less if labour were sold freely by the hour (or other small "dose").

It ought to be mentioned that a different view of Exchange has been taken by a high authority on Distribution. Professor Böhm-Bawerk presents as the general type of a market that very case which is here regarded as exceptional. On one side of the markets are put dealers each with a horse—or it may be a batch of several horses[64]—which he will not sell under a certain price, on the

---

61. Cp. Marshall, *Principles of Economics*, Book VI, chap. ii, §2, note, p. 599, 4th edition. Consider the case of managers.

62. Though one condition of a perfect market is thus secured, it does not follow that the labour-market will be perfect. Let us start with any system of bargains between entrepreneurs and work-people (presumed not to be capable of serving two masters at the same time). Then, there being supposed a variety of situations and tasks, let the round men in square berths change places with the square men in round berths with advantage to all (entrepreneurs included). There will thus be reached a settlement such that it cannot be disturbed with advantage to each and all; except by the employers competing with each other for workmen. Suppose the settlement to be such and so favourable to the work-people that it cannot be disturbed by the competition of the employers; then, the market win be indeterminate, just as if the work-people were all equally efficient. Accordingly, "There is no determinate and very generally unique arrangement towards which the system tends under the operation of, may we say, a law of Nature, and which would be predictable if we knew beforehand the real requirements of each, or of the average, dealer; but there are an indefinite number *a priori* possible settlements."

63. *Mathematical Psychics*, p. 18.

64. In the criticism of the *Positive Theory of Capital*, at p. 333 of the *Economic Journal*, Vol. II, repeated from the Address to the British Association, Section F. 1889 (reprinted in the *Journal of the Statistical Society*, December, 1889), it was too leniently suggested that the author, in a subsequent note (p. 214, Smart's translation of *Positive Theory*), brought in the essential circumstance which his main illustration omits; namely, doses with varying marginal utility. It would rather seem, however, that the stud of horses permitted in the said note does not differ essentially from the single horse of the main illustration. It seems to be treated as a mass of commodity which the seller offers, the buyer takes or leaves, as a whole At any rate, the writer has failed to see the significance of divisibility in the commodity For, otherwise, he would not have attributed so much

other side buyers each of which will not go beyond a certain price. The following scheme is given as an example of such data:[65]—

Buyers.		**Sellers**.	
A_1 values a horse at (and win buy at any price under)	£30	B_1 values a horse at (and will sell at any price over).	£10
A_2	28	B_2	11
A_3	26	B_3	15
A_4	24	B_4	17
A_5	22	B_5	20
A_6	21	B_6	21 10*s*.
A_7	20	B_7	26
A_8	18	B_8	26
A_9	17		
A_{10}	15		

From these data it is deduced that the price of a horse must be between £21 and 21 10*s*. But, if the data had been different, the price might not have been thus determinate. "If there are, for instance, ten buyers who each value the commodity at £10, and ten sellers who each value it subjectively at £1, obviously all the ten pair can come to terms, and the zone which lies between the valuation of the last buyer and the last seller represents the wide latitude between £1 and £10." Of this character, according to the writer, are the circumstances of the labour market.[66] In such a case some further datum is required to determine price. "That this latitude should be narrowed down, the further circumstance must be present that the desire of the buyers is directed to an unlimited number of goods, while at the same time the total amount of means of purchase must be strictly limited, and the buyers must be determined to spend the whole of this sum in purchase of the commodities in question."[67] This condition is fulfilled, according to Professor Böhm-Bawerk, by the "general subsistence market."

This example will hardly be accepted as typical of a market by the mathematical economists who walk in the way of Gossen. Agreeing with the Austrian leader that value rests at bottom on subjective estimates, they will accept his scheme, just as they would accept the description of a common auction, as illustrative of that attribute. But they may complain that the illustration does not illustrate another attribute which they regard as essential to the determination of value in a market,—the circumstance that each party on the one side is free, in concert with some party or parties on the other side, to vary the amounts of those quantities on which depends his

---

"latitude" (*loc. cit*. quoted in the text) to the case in which the severs (and likewise the buyers) do not differ from each other in their subjective valuation of a horse.

65. *Positive Theory of Capital* (translated), Book IV. chap. iv.

66. Op. cit., Book IV. chap. v. p. 217; Book VI, chap. v. ("On the General Subsistence Market").

67. Loc. cit.

advantage—the quid and the *pro quo*—up to a limiting point, or *margin* at which he estimates his advantage to be a maximum. The " marginal pair " of the Austrian scheme hardly exemplifies the *law of marginal utility*. We require to know, not so much the least price which each horse dealer will take for his horse or stud, but how much horseflesh each individual, or at least all collectively, will offer at each of several prices, with similarly graduated data for the would-be buyers. Granted data of this sort, the mathematical economist need not trouble himself much about a matter which is vital according to the Austrian scheme,—whether the "subjective valuation" of a horse is the same (or very similar) for all the sellers, while the dispositions of the buyers are likewise identical. The case of like dispositions does not constitute a special variety of the problem, one which is insoluble without additional data. Far from being anomalous, that case may be normally assumed as a harmless and convenient simplification, very proper to an introductory statement of the general theory.[68]

> "Neo Deus intersit, nisi dignus vindioe nodus
> Inciderit "—

The case of like dispositions does not present any peculiar difficulty calling for so very mechanical a *Deus ex machina* as the hypothesis that "the total amount of means of purchase must be strictly limited and the buyers must be determined to spend the whole of this sum in purchase of the commodities in question." It is riding a one-horse illustration to death to put the accidents of an exceptional sort of auction as representative of the actual transactions by which the great mass of national income is distributed.

This criticism, it must be freely admitted, involves an issue about which legitimate differences of opinion may exist,—what is the most appropriate conception of the process by which value is determined through the higgling of the market? Any simple conception must involve a considerable element of hypothesis, not admitting of decisive proof. The hypothetical character of the inquiry will appear if we look back to that model labour market in which guides or porters were supposed to be hired by amateur mountaineers. It was tacitly assumed that each party has certain dispositions as to the amount of money that he is willing to give or take in exchange for a certain amount of work,—a scale of subjective estimates[69] which is supposed to be formed before the parties come into communication, and not to be modified by the chaffering of the market. The constancy of these dispositions being assumed, it is presumed that somehow a state of equilibrium will be brought about, such that the party on one side cannot improve his position by entering into new contracts with some party or parties on the other side. The better opinion is that only the position of equilibrium is knowable, not the path by which equilibrium is reached. As Jevons says, "It is a far more easy task to lay down the conditions under which trade is completed and interchange ceases than to attempt to ascertain at what rate trade will go on when equilibrium is not attained."[70]

---

68. It is so assumed in *Mathematical Psychics*.

69. Whether expressed by a demand-curve (or schedule, cf. Marshall, *Principles* Book III) or by way of *indifference curves*, as Professor Pareto has suggested (*Giornale degli Economisti*, 1900).

70. *Theory*, 2nd edition, pp. 101–2. The context seems to impose an unnecessary limitation: "Holders of commodities will be regarded not as continuously passing on these commodities in

Particular paths may be indicated by way of illustration, "to fix the ideas," as mathematicians say.[71]

In this spirit two kinds of higgling may be distinguished as appropriate respectively to short and long periods. First, we may suppose the intending buyers and sellers to remain in communication without actually making exchanges, each trying to get at the dispositions of the others, and estimating his chances of making a better bargain than one that has been provisionally contemplated. By this preliminary tentative process a system of bargains complying with the condition of equilibrium is, as it were, rehearsed before it is actually performed. Or, second, one may suppose a performance to take place before such rehearsal is completed. On the first day in our example a set of hirings are made which prove not to be in accordance with the dispositions of the parties. These contracts terminating with the day, the parties encounter each other the following day,[72] with dispositions the same as on the first day,—like combatants *armis animisque refecti*,[73]—in all respects as they were at the beginning of the first encounter, except that they have obtained by experience the knowledge that the system of bargains entered into on the first occasion does not fit the real dispositions of the parties. The second plan of higgling was supposed in the example,—the plan which is more appropriate to "normal" value.

Contemplating the theory of exchange in the abstract, we may exclaim with Burke, " Nobody, I believe, has observed with any reflection what market is without being astonished at the truth, the correctness, the celerity, the general equity, with which the balance of wants is settled."[74] But, when we come to the labour market, or any particular market, we must carefully inquire with what degree of approximateness the above-stated fundamental postulate holds good. When the bargaining extends over a considerable time, changes are apt to occur in the dispositions of the parties, whether independently of each other and sporadically, or in a manner even more fatal to the theory, by way of imitation.[75] Also, where there occurs a series of encounters between buyers and sellers, the results of the earlier encounter may affect the dispositions with which the later ones are entered on. The terms which the labourer is ready to offer and accept are altered by the alteration in his habits and efficiency which is the consequence of previous bad bargains.[76]

---

streams of trade, but as possessing certain fixed amounts which they exchange until they come to equilibrium." The "fixed amount" may be considered as renewed from time to time for each of the individuals placed along a "stream of trade."

71. This view of the subject is presented at greater length in an article in the *Revue d'Économie Politique*, January, 1891.

72. They *recontract*, in the phraseology of *Mathematical Psychics*.

73. *Aeneid*. xii. 788.

74. *Thought and Details on Scarcity*. He is speaking with special reference to the labour market.

75. See Pigou on "Utility" in the *Economic Journal* for March, 1901. Compare, as to the absence of predeterminateness in the dispositions of parties to the labour market, Walker, *Political Economy*, Art. 320.

76. Cp. Marshall, *Principles of Economics*, Book VI, chap. iv, and Walker *Political Economy*, Art. 308 et seq.

The peculiarities of the labour market pointed out by Professor Marshall go far to modify the general presumption in favour of *laisser faire*. But less careful writers are less successful in supporting the burden of proof which lies on those who profess to add to or take away from that outlined theory of Exchange which seems to express all that is known *in general* about the working of a market. A warning example of such modification not warranted by specific experience is the doctrine of the wagefund, which is now universally discredited, and ought always to have excited suspicion and challenged proof because, as already intimated in another connection, it is a supposition repugnant to the general theory of Exchange that "the total amount of means of purchase must be strictly limited, and the buyers must be determined to spend the whole of this sum in purchase of the commodities in question."[77] Perhaps, as Sir Leslie Stephen says with reference to the classical writers, "the assumption slipped into their reasoning unawares."[78] Sometimes it may have been intended only to convey that early lesson which is contained in our opening paragraphs,—that no party to production can expect to earn more than the total produce. Sometimes there was contemplated a more definite statement true of short periods,—a truth which has been well stated by Professor Taussig in his article on "The Employer's Place in Distribution," and at greater length in his book on *Wages and Capital*—

"The whole of the real income available for the community is not in any substantial sense at the disposal of the capitalists.... A large part of the commodities now on hand would not serve their turn. The supply of bread and flour and grain at any moment is adjusted to the expected needs of the whole mass of consumers.... The effective choice which the capitalists would have . . . would be thus confined, for the time being at least, within limits not very elastic."[79]

Let us suppose that the working classes live on bread only, while the capitalist classes consume buns also. On a day, after a conference between employers and employed, the partition of the national dividend is altered in favour of the capitalists. Yet they will be unable to benefit immediately by the change. On that day more buns will not be forthcoming, all the bakers' ovens being preoccupied with bread.

For the purpose of illustration there has been chosen a specially simple case in which the articles consumed by the two classes are formed out of the same material, and by a process which is identical up to the penultimate stage. The stream of production does not bifurcate till it debauches into the mouths of the two parties to Distribution.

When we consider longer tracts of that stream, there comes into view a circumstance to be discussed under the head of *Capital*, the influence of time on value. To illustrate the distribution of produce between those who have contributed at different times to its production, let us at first make abstraction of other differences, and imagine economic men uniting the functions of workman and capitalist-entrepreneur, differing only in the amount of capitalization, the length of time during which their labour is invested. One labours at proximate means, another at remote means, tending

---

77. Quoted from Böhm-Bawerk, who himself compares his theory with that of the wage-fund (*Positive Theory*, p. 419). Both theories seem true of short periods. The context accords with the view here taken of the theory, as true of short periods, inadequate to long periods.

78. *The English Utilitarians*, Vol. III, p. 216.

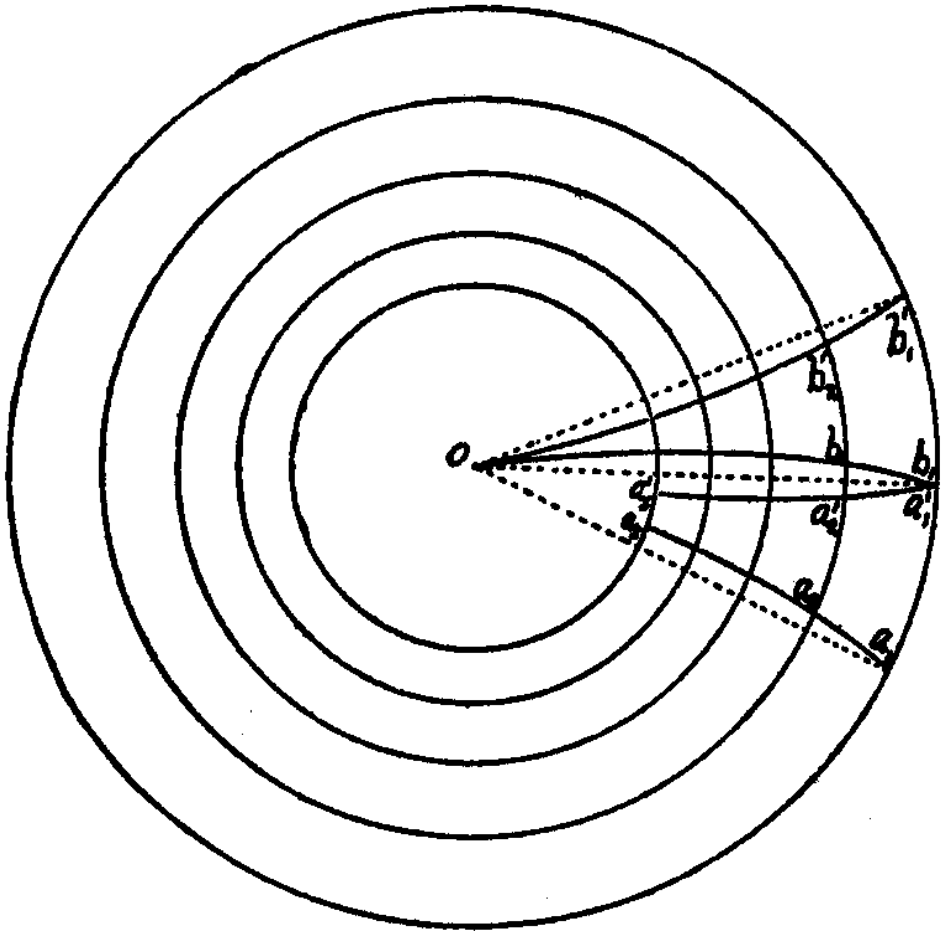
79. *Quarterly Journal of Economics*, Vol. X, p. 74.

to the ultimate product out of which all the producers are remunerated. An idea of a train of production formed by successive operations directed to an ultimate product may be obtained by watching any factory. Here you have the raw cotton-wool put in, there you see a "sliver" of carded cotton flowing from one machine en route to another, until at the last stage there comes out the finished article. To illustrate the process of distribution, we must now conceive a backward flow of the ultimate product to the several producers. We might imagine each one's share to be conveyed to him by some contrivance like those wondrous little vehicles in the Boston Public Library, which, as if gifted with human intelligence, find their way about the building to the particular place where each book belongs. To illustrate the effect of distance in time on distribution, we must further modify the model presented by an ordinary factory. We must suppose the interval of time between the processes to be greatly magnified, months being substituted for minutes. Then there will come into view the circumstance to which attention is particularly directed,—that a larger share will be conveyed to each producer (other things being equal), the greater his distance from the final stage. There will thus be a continual flow of materials in process of manufacture onwards and of products ready for consumption backwards, if the work at each stage is steadily maintained,— provided that there is a continual stream of raw material, and that the machines are continually renewed. Considering the continuous round of production and consumption, we realise the important truth which Mill has thus expressed:—

"The miller, the reaper, the ploughman, the plough-maker, the wagoner and wagon-maker, and the sailor and ship-builder, when employed, derive their remuneration from the ultimate product,—the bread made from the corn on which they have severally operated or supplied the instruments for operating."[80]

To represent the continual expansion of value as the present ripens into the future, a series

[80] *Political Economy*, Book I. chap. ii. §§1, 2.



of concentric circles has been happily employed by Professor Böhm-Bawerk.[81] Varying his illustration, let us suppose the circles to be drawn on ground which rises uniformly from the outmost circle towards the centre O in the accompanying diagram at which the apex tapers to a needle-point.[82] The circles are drawn at equal distances as measured on the surface, and therefore, in a bird's-eye view which the diagram is intended to represent, become huddled together in the neighbourhood of the central height. Across the circles, down the hill, flow streams with uniform velocity, so as to pass from circle to circle in a unit of time. The breadth of a stream increases with its length,—not in direct proportion to the length, but according to the law of *accumulated* price.[83] The volume of the stream is proportioned to its breadth and to its depth (not shown on the figure).

---

81. *Positive Theory*, Book II, chap v.

82. The series of highering circles is not shown in the diagram after the fifth circle.

83. Marshall, as cited above.

The stream takes its rise at some position on the channel (e.g., at $a_5a'_5$), the flow per unit of time at that point being proportioned to the energy put forth in pumping from a certain source. As the volume thus originated rolls down the channel, it continually increases by infiltration from the neighbouring soil without any additional pumping, so that, the depth being preserved constant, the volume is proportioned to the increasing breadth.[84] Besides this increase due to its defluxion, the volume may also in the course of its downward flow be increased by additional pumping from a second source (e.g., $a_2a'_2$). This second increase corresponds to an increase in depth (not shown in the figure); and this second contribution is augmented, like the first, by the infiltration which attends defluxion. There may be as many sources as there are circles out by the descending stream. But there need not be a source at each interval. The equidistant circles correspond to successive lines, not always coincident with successive stages of production at each of which additional labour is applied.[85] The train of production thus represented terminates in a product ready for consumption—it may be loaves or ribbons, wine or shoes—on the shore of a circumfluent sea of commodities. As in the natural world rivers are replenished by the melting of the snow, which is formed on mountains by the congelation of vapour, which is wafted up from the ocean, into which the rivers flow down, so in the *mundus economicus*, by a compensation carried into more just detail, labour is restored and recreated by a refreshing rain of commodities derived from that sea into which all finished commodities are discharged. Volatile shoes and wine, and other commodities in due admixture up to a certain value, find their way to each point upon the heights from which a source has been tapped, the volume of this return corresponding to the volume of the original contribution,—not indeed the same, but the same increased by a factor of accumulation, the ratio which the breadth of the stream at the littoral bears to its breadth at the point of origin (e.g., $a_1a'_1$: $a_5a'_5$). The flight of the commodities from the littoral to the heights need not be supposed to occupy an appreciable time.

The idea of a Flow which has been illustrated is primarily applicable to the case in which materials and consumable commodities are used up once for all within a unit of time. But the case of labour invested for longer periods is easily assimilated. Suppose that a plough lasts five years, and that in each year of its existence it makes an equal addition to the consumable crop the year being taken as the unit of time. Then, although the plough may have been made in a week or month, the labour of its production is to be considered as invested in five unequal portions at unequal distances in time from the epoch at which the invested labour meets with its return. The total labour of making

---

84. The broadening of the stream corresponds to the two consilient facts, that future pleasures are discounted and that production is increased by "roundabout" methods. As to the first of these facts, see in Marshall's *Principles of Economics* the passages which relate to discounting future pleasures, and the remarks on those passages in the review of the second edition of the *Principles* in the *Economic Journal*, Vol. I. (1891) p. 613. See also the admirably clear explanation and illustration given by Professor Carver in his article on "Abstinence and the Theory of Interest," *Quarterly Journal of Economics*, Vol. VIII (1893) p. 48. As to both the first and second facts, see Böhm-Bawerk's well-known expositions. But as to the congilience of the two facts see, rather, Professor Marshall on the "fundamental symmetry" between the action of Supply and Demand (noticed in the review referred to). See also Professor Carver s explanation of the double statement that interest is payment for the sacrifice of abstinence, and that interest is paid because capital is productive (loc. cit. p. 43).

85. Corresponding to the machines in the illustration given in the preceding paragraphs.

the plough may be considered as applied at several positions (a'_1a_1, $a_2a'_2$,... $a_5a'_5$) in several contributions, respectively proportioned to the breadth of the stream at these points. If labour is invested in the production of a machine, imagined by economists, which lasts for ever,[86] or, what comes to the same, an improvement, such as the draining of land or opening a mine, or cutting an isthmus, which is calculated to yield a constant income for an Indefinitely long series of years, then the series of positions along the stream at which the labour is supposed to be invested must be carried back indefinitely (see the channel of which the mouth is $b_1b'_1$) up to that needle-point whose tapering dimensions correspond to the perspective of an indefinitely distant future.

Eternal machines are not very common; but the conception may serve to illustrate a species of tool or implement of which the race remains immortal, though the individual is worn out and perishes. Of this kind are implements which are directed not only to produce goods immediately ready for consumption or implements of a kind different from their own, but also to reproduce their own kind. Hammers and axes are presumably of this kind in a primitive society; in an advanced state of industry, some more complicated engines.[87] Such machines may be compared to horses, if used not only as beasts of burden, but also as stallions. The demand for such creatures is presumably influenced by the expected series of future generations, so fat as commercial prospectiveness may extend. In the stationary state of steady motion, here provisionally contemplated, reproductive machines would be illustrated by beasts of burden of which the breed does not sensibly improve in successive generations.

Two channels only have been represented in the diagram, one of finite, the other of infinite length, with breadth exaggerated for the sake of clearness. Properly, there should be as many channels as there are categories of articles ready for immediate consumption,—"goods of the first order," as the Austrians say; and the breadth should be such as to allow of the corresponding number of sectors being fitted into the circle. Another circumstance which must be left to the imagination is the introduction of one and the same article into several streams of production at different distances from the final stage. Coal, for instance, so far as it is used for warming dwelling-houses, is a good of the first order; so far as it is used to drive machines,—themselves perhaps used only to produce other machines,—coal is to be placed among the higher orders.

The distinction which has been drawn between work which is applied in the neighbourhood of and at a distance from the final stage of production is not coincident with the distinction between the saving and the non-saving classes. The shower of commodities apportioned to each spot according to its height above the littoral as well as to the volume of value which there took its rise, is not "like the gentle rain from heaven." It does not drop impartially on all who have been concerned with the work of eliciting the stream. Those who have done the common labour of pumping—the drawers of water—fare no better than if that work had been done at the littoral. In fact, it is proper to conceive that it was done at the littoral. As the energy generated at the Falls of Niagara is transmitted for use to a point higher up on the river, so on the stream of production the work of pumping is mostly done at the littoral, though it is applied at the heights. For instance, on the first stream an amount of work proportioned to $a_5a'_5$ might be done at the littoral, and be paid for in commodities at the rate current on the littoral; that is, without the augmentation of value which

---

86. Mill, *Political Economy*, Book I, chap. vi §2.

87. Or rather a certain system of machinery. Cp. Marx on machines produced by machinery. *Capital*, ch. xv.

is due to defluxion. The remainder of the volume of value which is discharged per unit of time flies off to those who occupy the height represented by $a_5a'_5$.

If now it is asked where rent comes into this representation of distribution, the answer is to be found in the theory that from the point of view of the entrepreneur the use of land appears in the same light as the use of labourers,—as a factor of production. The idea of a steady cyclic flow which we are striving to win becomes not much more complicated when we magma that those who, placed on the heights, preside over the origination of productive streams, obtain the material that is to form the current, the precious fluid which it is their office to start upon its downward flow, not solely from a pumping proletariat, but also from the fortunate owners of springs which gush spontaneously. There is, indeed, this difference between the labourer and the land-owner: that, whereas the former (even in the present age and still more when the classical economists flourished) has to spend a great proportion of his daily wage upon his daily necessaries, and therefore in respect of the bulk of his income must be placed at the littoral line, the latter may save a great part of his income, when it is greatly in excess of his daily necessaries, and in particular, with respect to that great portion, may defer fruition until the stream shall have flowed down from the point at which his contribution is applied to the point at which production becomes merged in consummation. Another difference between land and labour in their relation to capital and enterprise arises from the circumstance that, unlike the labourer (in a free country), land itself, as well as its use, is sold. Whence arises a well-known correspondence between rent and interest in their relation to the capital value of land. This similarity will not be mistaken for identity[88] by those who find the essential attribute of rent in the limitation of the objects for which rent is paid.[89]

To complete the analysis of the parties to Distribution, it may next be required to distinguish the capitalist from the entrepreneur. They are both easily distinguished from the salaried manager in that he is at the littoral, in that respect like the common workman, while they are both above that line. But to draw a line in the series of shades which intervene between the employer of Walker's type and the mere shareholder, to determine at what point the capitalist ends and the entrepreneur begins, appears to defy analysis. As Thought and Emotion are inseparably blended, though one may so far preponderate as to give its name to the state of consciousness at any time, such is the inseparable connection, such the intelligible but not exactly definable distinction, between Enterprise and Saving. The indefiniteness of the relation is illustrated by the shifting use in economic literature of the term Profit.[90]

That profit other than remuneration for managerial work should be transmitted to those who occupy a position on the heights—Often the easy position of a dormant shareholder—is certainly invidious and difficult to justify to those who toil below. Yet it may be reflected that the condition of those below would have been worse if those above, or those from whom they purchased or inherited their position, had not been content to wait for future goods instead of grasping at

---

88. "The attempt of certain writers to refine away this traditional distinction between land and capital, rent and interest, impresses me as a subtle obscuration of plain facts," well remarked one of the speakers at the recent banquet of the Massachusetts Single Tax League (1902).

89. Marshall, *Principles*, *sub voce* "Rent."

90. As instructively pointed out by Mr. L. L. Price in his article on "Profit-sharing" published in the *Economic Journal*, Vol. II (1892), and in his *Economic Science and Practice*, p. 75 and *ante*.

immediate pleasure. The Flow so beneficial to all classes would never have been set up without abstinence.[91] It could not continue in its present magnitude but for the continued abstinence of each one who has a right to dispose of wealth which is in course of production,— make a bonfire of it, if he can get a momentary pleasure from that extravagance, or by some less simple, though more familiar increase of unproductive consumption "eat up his capital."

The consequences of an increase in unproductive consumption may be contemplated by reversing the consequences of an increase in parsimony. The latter increase forms part of a larger subject, economic progress. The progressive change in the volume of value and channels of production cannot be understood until there has been attained what was the object of the preceding paragraphs,—the clear idea of a steady flow in channels for a time unchanged.[92] The study of this stationary state is perhaps the part of economic science which principally deserves to be described as theory of Distribution. In these pages it is not attempted to go far beyond the comparatively narrow round of steady motion in fixed cycles of production and consumption. It must suffice to indicate three species of progressive alteration in the economic mechanism. There is, *first*, a uniform increase in the number of both capitalists and labourers, or, more generally, capital and labour, other things being the same. This change presents no difficulty: it may be represented by an increase in the depth of all the channels. *Second*, the rate at which the breadth of the channels diminishes as one ascends from the littoral—in other words, the rate of interest—might be diminished. A limiting case of this species is put by Mill when he supposes unproductive expenditure of capitalists to be " reduced to its lowest limit." Conceivably, this change might have no other effect than to reduce the portions accruing to the capitalists—such as $a_1a'_1$— $a_2a'_2$—to a minimum. The capitalists with new eagerness bid against each other for the service of the labourers; but, if the latter do not give more work for higher pay, the consequences might be a new equilibrium in which the same volume of value is steadily rolled down the same channels of trade, though the portion which flies back to the heights is a minimum. But, even if the quantity of value continued constant, it is hardly to be supposed that the quality[93] of the commodities which make up the amount would remain unchanged. And, in fact, an increase of wages would probably be followed by an increase in the number and efficiency of the wage-earning classes.[94] And these results would favour the occurrence of a *third* kind of progress which may, however, be considered as arising independently of the others; namely,

---

91. Compare Adam Smith. "By what a frugal man annually saves he not only affords maintenance for an additional number of productive hands for that or the ensuing year, but, like the founder of a workhouse, he establishes, as it were, a perpetual fund for the maintenance of an equal number in an times to come." *Wealth of Nations*, Book II, chap. iii. In our metaphor, talking up a new position on the heights corresponds to this establishment of a perpetual fund.

92. On the nature of the steady flow with which we are concerned see Marshall *Economic Journal*, Vol. VIII, p. 40, and *Principles of Economy*, *sub voce* "Stationary State."

93. Cp. Mill, loc. cit.,—"there would no longer be any demand for luxuries on the part of capitalists."

94. Cp, Marshall, *Principles*, Book IV, ch. xiii.

the lengthening of the trains of production.[95] It may be doubted whether any great lengthening of the trains is possible without a concomitant improvement in the arts of production; yet, as Sidgwick observes,[96] invention is not necessarily followed by increase of capitalisation.[97]

The third head of progress even more surely than the second will be attended with changes in the channels of production. As already observed with reference to the portion of truth contained in the wage-fund theory, time will in general be required for the carrying out of such changes. The means of production which are rolling down the channels at the instant when the change begins must all or in great part be suffered to run out: otherwise there will probably be a considerable waste of labour, and interruption to consumption. One delicate adjustment which would be deranged can only be alluded to here—the monetary circulation, especially that form of it which consists of debts that are continually "cleared," or cancelled. We might imagine the flow of factors in the channels of production and the flight of finished products backward on the way to consumption to be attended each with a displacement of air in a direction opposite to the main movement,—light counter-currents which have their use in facilitating the movements of solid wealth, and in the fulfilment of their useful function continually meet and neutralize each other. But, evidently, we have reached the degree of complexity at which the illustration becomes more difficult to understand than the thing which is to be illustrated. For a more concrete embodiment of a more complete theory the student is referred to the *Principles of Economics*,—a reference of which the value is, if possible, enhanced by the solid work which Mr. N. G. Pierson has published under the same title.[98]

The preceding hints and metaphors and warnings may assist the student to obtain a general idea of the process by which distribution of the national income is effected. An outline of theory so abstract is not to be despised as useless. It satisfies a legitimate curiosity. It is part of a liberal education. It is comparable in these respects with an elementary knowledge of astronomy. Such knowledge will not be of much use in navigation. And yet it has a certain bearing on real life. The diffusion of just notions about astronomy has rendered it impossible for astrologers any longer to practice on the credulity of mankind. A knowledge of first principles affords a test by which the authority of those who offer themselves as guides may be estimated. A little science has a further use: it is of assistance in obtaining more.

As the astronomer will proceed from a first approximation to a second, so economists should soften the hard outline of abstract theory by a regard to particular circumstances. As he in dealing

---

95. It is possible, as Mill shows, *Political Economy*, Book I, chap. vi, §2 (cp. Ricardo on machinery and Mr. Pierson, *Principles of Economics* p. 311), that lengthening the period of investment, and also invention, while it increases the amount of goods accruing to the capitalist, may diminish the amount accruing to the workers. What Mill says in this connection of the "fresh creation" of capital and "additional saving consequent on improvements" is made more intelligible by the use of the illustration here offered.

96. *Political Economy,* Book I, chap. iv, §8.

97. Loc. cit. Mill treats capital and arts of production as independent variables. *Political Economy*, Book IV, chap. iii.

98. Translated into English from the Dutch by Wotzel.

with a new object will make certain of his first approximation,— will consider, for example, whether an ellipse or a parabola fits better to the orbit of a new comet,—so it behaves us to consider whether the classical hypothesis presupposed in the preceding pages—two-sided competition—is appropriate to the conditions of modern industry. The hypothesis of two-sided monopoly[99] is strongly suggested by what we see before us,—consolidated capital confronted by consolidated trade unions. But it is alleged that beneath that appearance the forces of competition are effectively at work; that the settlement which is apt to be, and ought to be, agreed to between a combination of Capital and a combination of Labour is no other than that which would have been determined by competition if the individuals now combined had been free to act competitively. No one has expressed this view with more authority and decision than Walker:—

"Competition, perfect competition, affords the ideal condition for the distribution of wealth."[100]

"Competition affords the only absolute security possible for the equitable and beneficial distribution of the products of industry."[101]

To the same effect, Professor Clark, when he teaches that—

"The question whether the labourer is exploited and robbed depends on the question whether he gets his product."[102]

What is meant by getting his product appears from the following passages:—

"What we are able to produce by means of labour is determined by what a final unit of mere labour can add to the product that can be created without its aid."[103]

"If each productive function is paid for according to the amount of its product [thus reckoned], then each man gets what he himself produces."

The ideal of just arbitration is that—

"Men should get something approximating the part of that joint product which they may fairly regard as solely the fruit of their own labour.[104] The basis of the claim that a workman makes is that his presence in a mill causes a certain increase in the output of it."[105]

If these views are generally accepted, the analysis of bargains in a regime of competition will its importance. But it may well be doubted whether these views will be generally accepted, even by the thoughtful few, much less by the more numerous of the concerned parties. First, it may be objected that the same principle will give very different results according to the relative numbers of the parties. Put a case which has actually existed, or at least may be well supposed to have existed, in order to test the general application of the principle,—the case in which the number of the

---

99. The useful phrases of Dr. Böhm-Bawerk.

100. *Political Economy*, par. 466.

101. Ibid., par. 467. Cp. par. 343 et seq.

102. *The Distribution of Wealth*, chap. i.

103. Ibid., p. 180.

104. "Authoritative Arbitration," *Political Science Quarterly*, December, 1892, p. 559.

105. Ibid., p. 559.

employees is not much greater than, say not more than twice as great as, the number of the employers. In such a case, if labour is sold by the hour,—openly, or virtually in a fashion that probably prevails at present,[106]—there would be a determinate equilibrium of the labour market such that each labourer would earn an amount equal to the number of hours worked, multiplied by the final productivity of each hour. That arrangement might appear just, on a certain interpretation of the dictum that one's product " is determined by what a final unit of mere labour can add to the product." But the arrangement would not be just if "the basis of the claim that a workman makes is that his presence in the mill will cause a certain increase in the output of it." All turns on the unit employed. If it is allowable to take the hour as the unit, and find the wage of the individual man by multiplying the number of hours worked by the final productivity of the unit, why should it not be allowable to take a gang of men as the unit, and find the wage of the individual man by dividing the number of men in a gang into the final productivity of a gang? Not to rest the argument on supposed cases, take the case of the "capitalist" as he existed in Ricardo's time, or even the modern entrepreneur who is not a salaried manager. If such a one is to be paid on the basis that " his presence in a mill causes a certain increase in the output of it," it is quite possible that he would be justified in claiming a much larger share of the joint product than he now obtains.[107] The assertion that the entrepreneur receives just as much as he adds to product is at best an empirical law, not possessing the sort of universality proper to a general canon of distributive justice. Thus the coincidence of perfect competition with ideal justice is by no means evident to the impartial spectator: much less is it likely to be accepted by the majority of those concerned, whose views must be taken into account by those who would form a theory that has some relation to the facts. One who has closely observed popular movements in America testifies to "the growing belief that mechanical science and invention applied to industry are too closely held by private interests."[108] "An enormous private ownership of industrial mechanism, especially if coupled with lands and mines," forms the gravamen of the complaints. To advert for a moment to the accessory grievance with the view of understanding the main one, can we suppose that in a case such as Ireland was supposed to constitute before the Gladstonian land legislation, the land leaguers would have been content if they had obtained a perfect market in land, an equation of supply and demand undisturbed by hustling or delay, intimidation or cornering?[109] This perfection of the market might have served only to bring out the disadvantage at which the many were placed by the vesting of the complete ownership of land in the hands of a few. The prevailing sentiment about the "enormous private ownership of industrial mechanism " may well be similar. It is true that the expediencies governing " judicial rents " are very

---

106. See Marshall, *Principles of Economics*, Book VI, chap. ii, §2, note to p. 499, 4th edition.

107. The attribution of a portion of the product to a unit of productive factor is only significant when the unit can be treated as a final increment. Cp. Marshall, *Principles of Economics*, note to p. 465, 4th edition. When this condition is not fulfilled,—e.g., Professor Clark's *Distribution of Wealth*, p. 326, where "the amount that is attributable to one-half of the capital" ("the capital that is used in the industry") is specified,—this doctrine of attribution becomes perilously like the Austrian doctrine of "imputation."

108. Graham Brooks, *The Social Unrest*, p. 122.

109. Such a market as is analysed in *Mathematical Psychics*, p. 141.

different from those which are opposed to the legal regulation of wages. But we are now considering how the matter appears to the many, what regime they can be got to accept. It seems not to be competition pure and simple.[110]

Are we, then, to abandon the guidance of competition, and follow a higher, an ethical, standard? Does the theory of distribution require a definition of distributive justice? What is justice? The result of Plato's prolonged inquiry would not be satisfactory to the modern asserter of the rights of labour. If a new Socrates were to go about inquiring, what is the ideally just distribution between the employing and employed classes, he would probably find the wisest to be those who confessed their ignorance. As Jevons says, nothing at first sight can seem more reasonable and just than the "favourite saying that a man should have a fair day's wages for a fair day's work.... But, when you examine its meaning, you soon find that there is no real meaning at all. There is no way of deciding what is a fair day's wages."[111] It has been well observed that an intuition as to the just rate of wages, the labourer's share of the total product, involves an intuition as to the capitalist's share,—a share which depends on the rate of interest.[112] Can any one seriously pretend that the dictates of a moral sense are clear and decisive in such a matter?

Let it be remembered also that the path of justice is not only dark, but dangerous. Striving to secure the rights of labour, you are very likely to hurt the interests of labour. The action of trade unions by lowering interest and harassing employers may result, as pointed out by Professor Marshall,[113] in checking the accumulation of capital and the supply of business power. The increase in personal capital may indeed compensate for this check, but also it may not. Greater efficiency does not follow higher wages as the night the day.[114]

In view of these considerations it is doubtful whether in the near future an influential majority will aim at setting aside competition. Moreover, even if this consummation were aimed at, it is not likely to be attained. So invincible in human nature is the "propensity to truck,"[115] so true is it that, "when one person is willing to sell a thing at a price which another is willing to pay for it, the two manage to come together in spite of prohibitions of King or Parliament, or of the officials of a Trust or Trade Union."[116] Competition is like the air we breathe, which it is not only dangerous, but difficult to exclude.

Between two guides, of which neither can be followed implicitly, let us walk warily. On the one hand, let us not aim at impossible ideals. But, on the other hand, let us not deserve the criticism

---

110. It is possible that competition purified in the manner suggested below might be accepted by moderate trade unionists of the type of Applegarth and Dunning, as to whom see *History of Trade Unionism*, S. and B. Webb.

111. *Scientific Primer*, chapter on "Wages."

112. Margaret Benson, *Capital, Labour, and Trade*, chap. xvi.

113. *Elements of Economics of Industry* (1892), Book VI. chap. xiii.

114. See the careful statement of the relations by Mr. Pierson in his *Principles of Economics*.

115. Adam Smith, *Wealth of Nations*, Book I, chap. ii.

116. Marshall, *Quarterly Journal of Economics*, Vol. XI, (1897) p. 129.

which the advocates of trade unionism have with too much truth directed against "the verdict of the economists" respecting trade unions.[117] Let us not be as trenchant in act as we have been in thought. Let us be cautious in applying our abstract theory to flesh and blood.

To one seeking a representation at once clear and appropriate, the actual conditions of industry present the appearance of a viscous and deliquescent body,[118] not so easy to be treated by simple formulae as a perfect liquid or a perfect solid. An adequate theory of Distribution must in these days take some account of the action proper to combinations, effecting collective treaties between employers and employed: competition pure and simple no longer constitutes an adequate hypothesis. Exactly how these twos principles are to be conceived as coexistent it is premature to state dogmatically: the economist whose aim is to "teach, not preach," to show what is or will be rather than what ought to be, may well hesitate to pronounce on this question. He can at best invent hypotheses which may facilitate the conception of a compromise between the opposed principles of competition and combination. For example, the required compromise might be attained if it were arranged that the agreement between employers and employed under some heads might be settled by collective treaty between combinations, but under other heads by competitive bargaining between individuals,—as the German students in their duels expose only certain parts, not all parts, of the body to the brunt of the combat.[119] To determine what matters should be the subject of treaty would indeed itself require some sort of treaty.[120] But it would be a kind of treaty for which there is good precedent in laws and institutions. For instance, there might grow up, or be enacted by law, the practice that the hours of labour in a trade should be a matter for collective treaty between a trade union and a combination of employers, the particular number of hours to be settled by such treaty, while other terms, such as the rate of wages, should be settled by the play of competition.

So far as competition has free play, the received theory of supply and demand, even in its severest mathematical form, would be applicable. Indeed, severer forms would be peculiarly appropriate in that they do not lend themselves to the contemplation of cornering and other dodges of the market, but assume the "true price"[121] to be worked out honestly. Presumably, the competition which all parties agreed to retain would have to be conducted in a similar spirit. The conditions of the duel, already prescribed, would be further limited by forbidding certain strokes.

A similar regulation may be suggested for the working of an imaginary sort of competition which seems to be contemplated by some who are conversant with the practical problems of

---

117. Sidney and Beatrice Webb, *Industrial Democracy*, Part III. chap. i.

118. Cp. J. B. Clark, *Philosophy of Wealth*: "The present state of industrial society is transitional and chaotic.... The consolidation of labour is incomplete," that of capital also (p. 148 and context).

119. Cp. J. B. Clark, op. cit., p. 208: "A spirit of Justice is ever standing over the contestants, and bidding them compete only thus and thus."

120. "No individual competitor can lay down the rules of combat." Sidney Webb, *Contemporary Review* (1889), p. 869.

121. Condillac's phrase, appropriate to the ideal market above described.

industry. Their view appears to be[122] that two combinations might, without resorting to actual competition, agree to accept those terms which would probably result from the play of free competition. In playing this sort of *Kriegspiel*, it might be laid down as a rule of civilised industrial warfare that the workman should not be treated as living from hand to mouth. Suppose him freed from the imminence of starvation for a time at least, and then consider what sort of arrangement of the terms to be settled would constitute a steady flow of the type above described, in which each individual's final sacrifice is normally equivalent to the final utility which he procures thereby.[123] Other rules might be suggested for the working of such imaginary competition.[124] But it may be questioned whether the method admits of precision, for a reason urged by Fir. L. L. Price with reference to a proposed principle of arbitration, "that the arbitrator should endeavour to award such wages as would be attained if combination on either side were absent." "Where is the arbitrator to discover this ideal standard?" pertinently asks Mr. Price.[125]

The terms forming the subject of a collective treaty would be settled by a method essentially different from competition. For instance, in the case above proposed, the length of a working day, let there be a law removing this article from the category of terms which are to be settled by the play of competition between individuals. Those who hold that such a law is based on the utilitarian first principle, the greatest happiness of those concerned,—here the citizens who have enacted the law,—will be prepared for the further suggestion that the particular number of hours to be settled will also be regulated by the utilitarian first principle, only that those concerned, whose maximum advantage constitutes the criterion, are not now the citizens,—if the citizens generally have no interest in the particular number of hours in the trade,— but only the parties to the distribution, the members of the contracting combination. That this undergrowth of utilitarianism may, like the parent tree, prove fruitful, has been argued elsewhere.[126] Here it need only be repeated that, when the

---

122 It is difficult to attach any other interpretation to Walker's dicta referred to above. He is presumably supposing that all the terms of contract are settled by ideal competition, a limiting case of the regime here suggested that some of the terms should be settled by competition, actual or imaginary.

123. The "method of mutual insurance" practiced by trade unions, according to Mr. and Mrs. Webb (*Industrial Democracy*), seems to confer this sort of advantage on its members.

124. E. g. in order to estimate that result, it might be thought consonant to the amount of industrial solidarity actually existing not to treat each individual workman as an economic atom, but rather to suppose comparatively few independent bodies, each formed by the solidification of many individual atoms. Compare T. J. Dunning, *Trade Unions and Strikes* (a work mentioned by J. S. Mill with approval), p. 21, where reply is made to the question, "Why cannot a man sell his labour for what he likes, as a shopkeeper tickets his goods under the price of those of his neighbour?" "The shopkeepers," replies Dunning, "are not obliged to be always together." "But the matter assumes a very different aspect" in the case of wage-earners who work together. Though, as will presently appear, a preliminary use of the sort of potential competition which has just been described may be required.

125. *Economic Science and Practice*, p. 198 and context.

126. *Mathematical Psychics*, p. 53.

utilitarian arrangement is defined as the basis of conciliation between self-interested parties to a contract, it is presupposed that both parties gain by the contract:[127] that it does not seem to either party to be their interest, rather than accept such an arrangement, to give up dealing at all with the other party—seek, it may be, some third party, some other employment of their capital and labour,[128] or at least to defer agreement with the other party, in view of the probability that they will reduce their terms.[129]

The rationale of conciliation thus presented will doubtless not commend itself to many who accept substantially identical principles invested in a different form. Uniformity is not to be expected in the enunciation of first principles. The vital tenet is that each party must take account of and enter into the wants and motives of the other party. When competition is no longer umpire, the economist must abandon—if he ever maintained—the position of extreme *solipsisms* which Jevons in a solitary but remarkable passage has propounded:—

Every mind is thus inscrutable to every other mind, and so no common denomination of feeling seems to be possible.... The motive in one mind is weighed only against other motives in the same mind, never against the motives in other minds. Each person is to other persons a portion of the outward world.... Hence the weighing of motives must always be confined to the bosom of the individual.[130]

Jevons himself has not remained consistently on this pinnacle of solitude. It is abandoned by economists in general in the received theory of taxation, founded, as Mill says, on " human wants

---

127. Consider the weighty passage referring to the principles on which courts of arbitration and boards of conciliation should act, in Marshall's *Economics of Industry* (1879), Book III, chap. viii, §2: "They must not set up by artificial means arrangements widely different from those which would have been naturally brought about," et seq. Compare Marshall's Preface to (L. L. Price's) *Industrial Peace*, p. xxiii. "The arbitrator is compelled to take some account of the fighting forces of the two sides; the necessity to be practical may compel him to go further than he would otherwise have done away from an absolute standard of fairness."

128. In the technical terms of *Mathematical Psychics* the utilitarian point in the contract-curve must not be outside the points at which that curve is cut by the indifference curve." It is significant that this abstract representation is adapted to the first rather than the second of the two cases, in which the utilitarian arrangement would not be accepted,—the case, for example, in which the capitalist combination refuses the arrangement, because, considering it as permanently at work, they would be worse off than if they were to transfer their capital to some other held of enterprise; not the ease In which they defer making an agreement for strategic reasons, because, being better supplied for siege, so to speak, than the other party, they hope to reduce them in ease of a stoke to submission. Compare what was said above as to the advisability of not admitting this kind of strategy into industrial combat waged under ideal conditions.

129. Compare Marshall, *Economics of Industry*, loc. cit.: "Mischief almost always results in the long run from an award which gives to one side terms much worse than those which it knows it could obtain by a strike or a lockout."

130. *Theory of Political Economy*, edition 3, p. 14.

and feelings."[131] Self-regarding self-interest, the gospel of Adam Smith, is not alone sufficient for industrial salvation: a leaf must be taken from his older and less familiar testament, of which the cardinal doctrine was sympathy. Sympathy does not necessarily imply sentimental attachment: sympathy, according to Adam Smith, is the basis of a not very sociable emotion,—ambition. A distinguished psychologist has not hesitated to pronounce "sympathy compatible with dislike."[132] It is, then, no counsel of perfection to cultivate sympathy, in the sense of mutual understanding, between the parties to distribution. No Utopian eradication of self-love is contemplated. It may be hoped, indeed, that through the practice of conciliation, in the course of generations, the dispositions of which the gratification constitutes self-interest may become more social, so that, for instance, an advantage founded on the extreme privation of others would not appear desirable to the capitalist employer of the future. But such "moralization" of the saving classes, though it may be expected, need not be postulated for the working of conciliation. Intellectual sympathy alone might effect much. The arts[133] by which the sympathetic imagination may be cultivated form a supremely important topic, but one which hardly falls under the *theory of Distribution*.

Note

[On the remuneration for risk some additional light is derivable from Mr. Keynes' great treatise on Probability; where he shows that mathematical expectation—the product of advantage and the probability of obtaining it—is not the measure of expediency (ch. xxvi. p. 311 et seq.; discussed by the present writer in *Mind*, 1922, vol. xxxi p. 276 et seq.). The motives of the entrepreneur may be illustrated by the position of Paul in the classical problem which Mr. Keynes thus restates: "Peter engages to pay Paul one shilling if a head appears at the first toss of a coin, two shillings if it does not appear until the second, and in general 2^{r-1} shillings if no head appears until the *r*th toss. What is the value of Paul's expectation?" If the number of tosses is limited to a finite number *n*, the mathematical expectation is ½ *n*. But, if *n* is large, no sensible person would give anything like that sum for the chance. Now Paul may be taken as typical of the entrepreneur. Peter in this case may fix what Paul must pay for a trial—corresponding, say, to the outlay on factors of production required

---

131. *Political Economy*, Book V. chap. ii. § 4.

132. Bain, *Emotion and Will* (Table of Contents).

133. For example, co operation, as many economists have pointed out, would have among its good effects that of enabling workmen to realize the position of employers. Again, the training of future business men in economies at the universities as Professor Marshall has lately urged, would tend to develop the sympathetic use of the imagination. "It has been found," he says, "by experience in England n in America that the young man who has studied both sides of labour questions in the frank and impartial atmosphere of a great university is often able to throw himself into the point of view of the working-men and to act as interpreter between them and persons of his own class with larger experience than his own." See his address on "Economic Teaching at the Universities," published in the Review of the Charity Organisation Society, January, 1903, noticed in the *Economic Journal*, Vol. XIII. p. 155, and his *Plea* for the creation of a curriculum in economies (addressed to the Cambridge Senate), noticed in the Economic Journal, Vol. XII. p. 289.

Compare the expressions in the *Report of the Anthracite Coal Commission*, USA (1908), on the Importance of "a more conciliatory disposition in the operators and their employees."

for a unit of product. But Paul will have a say as to the amount which he stands to win by that outlay. Say the payment is ½ *n* shillings or pounds, *n* not now indefinitely large; Paul will demand a higher prize than the bare actuarial 2^{r-1}; unless he is a fatuous gambler (cp. Marshall, *Principles*, Bk. V. ch. vii. §4, and p. 613, note, 5th edition; and Pigou on uncertainty-bearing). At what terms above the actuarial limit Paul will touch the point of indifference, what is his demand-schedule in respect of such transactions, depends upon his mentality, his "dispositions," in the phrase of Walras relative to supply and demand in general. Thus the share of the entrepreneur in the product equally with the share of the workman depends on the play of demand and supply. It is no more predetermined than the wage-fund.]

Alternative Theories of Distribution

According to the Preface of Ricardo's Principles, the discovery of the laws which regulate distributive shares is the " principal problem in Political Economy ". The purpose of this paper is to present a bird's eye view of the various theoretical attempts, since Ricardo, at solving this " principal problem ". Though all attempts at classification in such a vast field are necessarily to some extent arbitrary, and subjective to the writer, in terms of broad classification, one should, I think, distinguish between four main strands of thought, some of which contain important sub-groups. The first of these is the Ricardian, or Classical Theory, the second the Marxian, the third the Neo-Classical or Marginalist Theory and the fourth the Keynesian. The inclusion of a separate " Keynesian " theory in this context may cause surprise. An attempt will be made to show however that the specifically Keynesian apparatus of thought could be applied to the problem of distribution, rather than to the problem of the general level of production ; that there is evidence that in its early stages, Keynes' own thinking tended to develop in this direction—only to be diverted from it with the discovery (made some time between the publication of the *Treatise on Money* and the *General Theory*) that inflationary and deflationary tendencies could best be analysed in terms of the resulting changes in output and employment, rather than in their effects on prices.

The compression of a whole army of distinguished writers, and schools of thought, between Ricardo and Keynes (Marx aside) under the term of Neo-Classical or Marginalist Theory is harder to justify. For apart from the marginalists proper, the group would have to include such " non-marginalists " or quasi-marginalists (from the point of view of distribution theory) as the Walrasians and the neo-Walrasians,[1] as well as the imperfect competitionists, who though marginalist, do not necessarily hold with the principle of Marginal Productivity. But as I shall hope to show, there are important aspects which all these theories have in common,[2] and which justifies bringing them under one broad umbrella.

Ricardo prefaced his statement by a reference to the historical fact that " in different stages of society the proportions of the whole produce of the earth which will be allotted to each of these (three) classes under the names of rent, profit and wages will be essentially *different*."[3] To-day, a writer on the problem of distribution, would almost be inclined to say the opposite—that " in different stages of (capitalist) society the proportions of the national income allotted to wages, profits, etc., are *essentially similar* ". The famous " historical constancy " of the share of wages in the national income—and the similarity of these shares in different capitalist economies, such as the U.S. and the U.K.—was of course an unsuspected feature of capitalism in Ricardo's day. But to the extent that recent empirical research tends to contradict Ricardo's assumption about the variability of relative shares, it makes the question of what determines these shares, more, rather than less, intriguing. In fact no hypothesis as regards the forces determining distributive

[1] By the term " neo-Walrasians " I mean the American " linear programming " and " Activity analysis " schools, as well as the general equilibrium model of von Neumann (*Review of Economic Studies*, 1945-46, Vol. XIII (1)) whose technique shows certain affinities with Walras even though their basic assumptions (in particular that of the " circularity " of the production process) are quite different. From the point of view of distribution theory however, the approach only yields a solution (in the shape of an equilibrium interest rate) on the assumption of constant real wages (due to an infinitely elastic supply curve of labour) ; it shows therefore more affinity with the classical models than with the neo-classical theories.

[2] With the possible exception of the " neo-Walrasian " group referred to above.

[3] Preface (my italics).

shares could be intellectually satisfying unless it succeeds in accounting for the relative stability of these shares in the advanced capitalist economies over the last 100 years or so, despite the phenomenal changes in the techniques of production, in the accumulation of capital relative to labour and in real income per head.

Ricardo's concern in the problem of distribution was not due, or not only due, to the interest in the question of distributive shares *per se*, but to the belief that the theory of distribution holds the key to an understanding of the whole mechanism of the economic system—of the forces governing the rate of progress, of the ultimate incidence of taxation, of the effects of protection, and so on. It was through ". the laws which regulate distributive shares " that he was hoping to build what in present-day parlance we would call " a simple macro-economic model ".[1] In this respect, if no other, the Ricardian and the " Keynesian " theories are analogous.[2] With the neo-Classical or Marginalist theories, on the other hand, the problem of distribution is merely one aspect of the general pricing process ; it has no particular theoretical significance apart from the importance of the question *per se*. Nor do these theories yield a " macro-economic model " of the kind that exhibits the reaction-mechanism of the system through the choice of a strictly limited number of dependent and independent variables.

I. THE RICARDIAN THEORY

Ricardo's theory was based on two separate principles which we may term the " marginal principle " and the " surplus principle " respectively. The " marginal principle " serves to explain the share of rent, and the " surplus principle " the division of the residue between wages and profits. To explain the Ricardian model, we must first divide the economy into two broad branches, agriculture and industry and then show how, on Ricardo's assumptions, the forces operating in agriculture serve to determine distribution in industry.

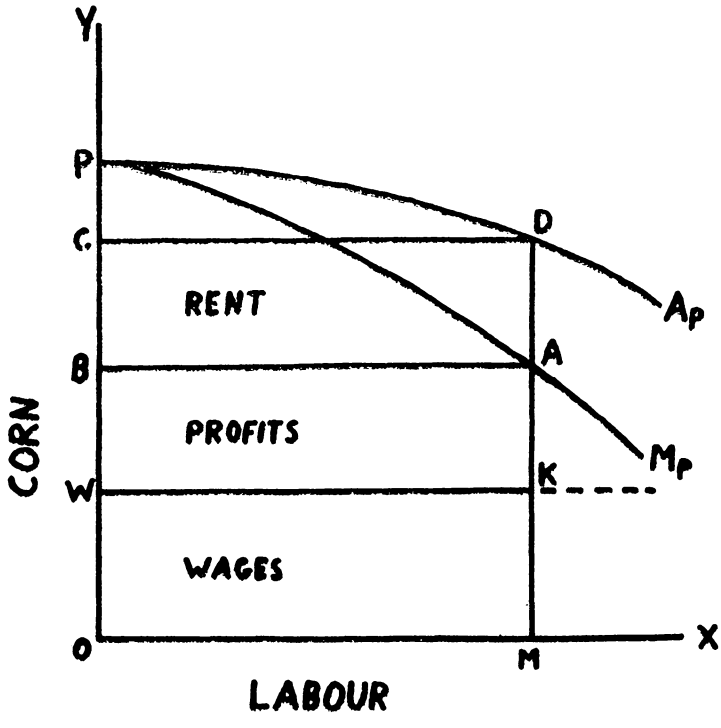
The agricultural side of the picture can be exhibited in terms of a simple diagram (FIG. 1), where Oy measures quantities of " corn " (standing for all agricultural produce) and Ox the amount of labour employed in agriculture. At a given state of knowledge and in a given natural environment the curve $p—Ap$ represents the product per unit of labour and the curve $p—Mp$ the marginal product of labour. The existence of these two *separate* curves, is a consequence of a declining tendency in the average product curve—i.e., of the assumption of diminishing returns. Corn-output is thus uniquely determined when the quantity of labour is given :[3] for any given working force, OM, total output is represented by the rectangle $OCDM$. Rent is the difference between the product of labour on " marginal " land and the product on average land, or (allowing for the intensive, as well as the extensive, margin) the difference between average and marginal labour productivity which depends on the elasticity of the $p—Ap$ curve, i.e., the extent to which diminishing returns operate.

The marginal product of labour (or, in classical parlance, the " produce-minus-rent ") is not however equal to the wage, but to the sum of wages and profits. The rate of wages is determined quite independently of marginal productivity by the supply price of labour

[1] " Political Economy " he told Malthus " you think is an enquiry into the nature and causes of wealth —I think it should rather be called an enquiry into the laws which determine the division of the produce of industry amongst the classes who concur in its formation. No law can be laid down respecting quantity, but a tolerably correct one can be laid down respecting proportions. Every day I am more satisfied that the former enquiry is vain and delusive, and the latter only the true objects of the science." (Letter dated 9 Oct., 1820, Works (Sraffa edition) vol. VIII, pp. 278-9.)

[2] And so of course is the Marxian : but then the Marxian theory is really only a simplified version of Ricardo, clothed in a different garb.

[3] This abstracts from variations in output per head due to the use of more or less fixed capital relative to labour—otherwise the curves could not be uniquely drawn, relative to a given state of technical knowledge. As between fixed capital and labour therefore the model assumes " fixed coefficients ' ; as between labour and land, variable coefficients.



which Ricardo assumed to be constant in terms of corn. In modern parlance, the Ricardian hypothesis implies an infinitely elastic supply curve of labour at the given supply price, *OW*.[1] The demand for labour is not determined however by the *p—Mp* curve, but by the accumulation of capital which determines how many labourers can find employment at the wage rate *OW*. Hence the equilibrium position is not indicated by the point of intersection between the *p—Mp* curve and the supply curve of labour, but by the aggregate demand for labour in terms of corn—the " wages fund ".[2] As capital accumulates, the

[1] The basis of this assumption is the Malthusian theory of population, according to which numbers will increase (indefinitely) when wages are above, and decrease (indefinitely) when they are below, the " subsistence level ". In Ricardo's hands this doctrine had lost its sharp focus on a biologically determined quantum of subsistence to which the supply price of labour must be tied ; he emphasized that habits of restraint engendered in a civilized environment can permanently secure for labour higher standards of living than the bare minimum for survival. Yet he retained the important operative principle that in any given social and cultural environment there is a " *natural* rate of wages " at which alone population could remain stationary and from which wages can only deviate temporarily. The hypothesis of an infinitely elastic supply curve of labour thus did not necessarily imply that this supply price must be equal to the bare minimum of subsistence. Yet this assumption was inconsistent with another (implied) feature of his model discussed below, that wages are not only *fixed* in terms of " corn " but are entirely (or almost entirely) *spent* on corn.

[2] Total wages depend on—and are " paid out of "—capital simply because production takes time, and the labourers (unlike the landlords) not being in the position to afford to wait, have their wages " advanced " to them by the capitalists. This is true of fixed as well as circulating capital, but since with the former, the turnover period is relatively long, only a small part of annual wages is paid out of fixed capital ; the amount of circulating capital was therefore treated as the proper " wages fund ". Despite his analysis of the effect of changes in wages on the amount of fixed capital used relative to labour, i.e., on the proportions of fixed and circulating capital employed in production (Professor Hayek's celebrated " Ricardo effect ") for the purpose of his distribution theory this ratio should be taken as given, irrespective of the rate of profit.

labour force will grow, so that any addition to the total wage fund, through capital accumulation—the *agricultural* wages fund is indicated by the area *OWKM*—will tend to be a horizontal addition (pushing the vertical line *KM* to the right) and not a vertical one (pushing the horizontal line *WK* upwards).[1]

For any given *M*, profits are thus a residue, arising from the difference between the marginal product of labour and the rate of wages. The resulting ratio, $\frac{\text{Profits}}{\text{Wages}}$, determines the rate of profit per cent on the capital employed ; it is moreover *equal* to that ratio, on the assumption that the capital is turned over once a year, so that the capital employed is equal to the annual wages-bill. (This latter proposition however is merely a simplification, and not an essential part of the story).

In a state of equilibrium, the money rate of profit *per cent* earned on capital must be the same in industry and in agriculture, otherwise capital would move from one form of employment to the other. But it is the peculiarity of agriculture that the money rate of profit in that industry cannot diverge from the rate of profit measured in terms of that industry's own product, *i.e.*, the corn-rate of profit. This is because in agriculture both the input (the wage outlay) and the output consist of the same commodity, " corn ". In manufacturing industry on the other hand, input and output consist of heterogeneous commodities—the cost per man is fixed in corn, while the product per man, in a given state of technical knowledge, is fixed in terms of manufactured goods. Hence the only way equality in the rate of profit in money terms can be attained as between the two branches is through the prices of industrial goods becoming dearer or cheaper in terms of agricultural products. The money rate of profit in manufacturing industry therefore depends on the corn-rate of profit in agriculture,[2] the latter on the other hand, is entirely a matter of the margin of cultivation, which in turn is a reflection (in a closed economy and in a given state of technical knowledge) of the extent of capital accumulation. Thus " diminishing fertility of the soil," as James Mill put it, " is the great and ultimately only necessary cause of a fall in profit ".

To make the whole structure logically consistent it is necessary to suppose, not only that wages are fixed in terms of " corn " but that they are entirely spent on " corn ", for otherwise any change in the relation between industrial and agricultural prices will alter real wages (in terms of commodities in general) so that the size of the " surplus ", and the rate of profit on capital generally, is no longer derivable from the " corn rate of profit "—the relationship between the product of labour and the cost of labour working on marginal land. Assuming that (" corn ") agricultural products are wage-goods and manufactured products are non-wage goods (i.e., ignoring that *some* agricultural products are consumed by capitalists, and *some* non-agricultural products by wage-earners), the whole corn-output (the area *OCDM* in the diagram) can be taken as the annual wages fund, of which *OWKM* is employed in agriculture and *WCDK* in the rest of the economy. Any increase in *OWKM* (caused, *e.g.*, by protection to agriculture) must necessarily lower the rate of profit (which is the source

[1] The feature which the modern mind may find most difficult to swallow is not that capital accumulation should lead to a rise in population but that the reaction should be taken as something so swift as to ignore the intervening stage, where the increase in the wages fund should raise the rate of wages rather than the numbers employed. The adjustment of population to changes in the demand for labour would normally be treated as a slow long-run effect whereas changes in the demand for labour (caused by capital accumulation) may be swift or sudden. Ricardo however conceived the economy as one which proceeds at a more or less steady rate of growth in time, with the accumulation of capital going on at a (more or less constant) rate ; while he conceded that *changes* in the rate of capital accumulation will temporarily raise or lower wages, he assumed that the rate of population growth itself is adapted to a certain rate of capital accumulation which had been going on for some time.

[2] The analytical basis for this conclusion, given above, was never, as Sraffa remarks, stated by Ricardo in any of his extant letters and papers though there is evidence from Malthus's remarks that he must have formulated it either in a lost paper on the Profits of Capital or in conversation (cf. *Works*, Vol I., Introduction, p. xxxi.).

of all accumulation) and thus slow down the rate of growth.[1] Similarly all taxes, other than those levied on land, must ultimately fall on, and be paid out of, profits, and thus slow down the rate of accumulation. Taxation and agricultural protection thus tend to accelerate the tendency (which is in any case inevitable—unless *continued* technical progress manages to shift the *p—Ap* and *p—Mp* curves to the right sufficiently to suspend altogether the operation of the Law of Diminishing Returns) to that ultimate state of gloom, the Stationary State, where accumulation ceases simply because " profits are so low as not to afford (the capitalists more than) an adequate compensation for their trouble and the risk which they must necessarily encounter in employing their capital productively ".[2]

II THE MARXIAN THEORY

The Marxian theory is essentially an adaptation of Ricardo's " surplus theory ". The main analytical differences are :— (1) that Marx paid no attention to (and did not believe in) the Law of Diminishing Returns, and hence made no analytical distinction between rent and profits ; (2) that Marx regarded the supply price of labour (the " cost of reproduction " of labour) as being fixed, not in terms of " corn ", but of commodities in general. Hence he regarded the share of profits (including rent) in output as determined simply by the surplus of the product per unit of labour over the supply price (or cost) of labour—or the surplus of production to the consumption necessary for production.[3]

There are important differences also as between Marx and Ricardo in two other respects. The first of these concerns the reasons for wages being tied to the subsistence level. In Marx's theory this is ensured through the fact that at any one time the supply of labour—the number of workers seeking wage-employment—tends to exceed the demand for labour. The existence of an unemployed fringe—the " reserve army " of labour—prevents wages from rising above the minimum that must be paid to enable the labourers to perform the work. Marx assumed that as capitalist enterprise progresses at the expenses of pre-capitalistic enterprise more labourers are released through the disappearance of the non-capitalist or handi-craft units than are absorbed in the capitalist sector, owing to the difference in productivity per head between the two sectors. As long as the growth of capitalist enterprise is at the cost of a shrinkage of pre-capitalist enterprise the increase in the supply of wage labour will thus tend to run ahead of the increase in the demand for wage labour.

Sooner or later, however, the demand for labour resulting from accumulation by capitalist enterprise will run ahead of the increase in supply ; at that stage labour becomes scarce, wages rise, profits are wiped out and capitalism is faced with a " crisis ". (The crisis in itself slows down the rate of accumulation and reduces the demand for labour at any given state of accumulation by increasing the " organic composition of capital," so that the " reserve army " will sooner or later be recreated.)

The second important difference relates to the motives behind capital accumulation. For Ricardo this was simply to be explained by the lure of a high rate of profit. Capitalists accumulate voluntarily so long as the rate of profit exceeds the minimum " necessary compensation " for the risks and trouble encountered in the productive employment of capital. For Marx however, accumulation by capitalist enterprise is not a matter of choice

[1] The evil of agricultural protection is thus not only that real income is reduced through the transfer of labour to less productive employments, but that owing to the reduction in the rate of profit, industrial prices fall in terms of agricultural prices ; income is thus transferred from the classes which use their wealth productively to classes which use it unproductively.

[2] Ricardo, *Principles*, p. 122 (Sraffa Edition).

[3] Ricardo himself abandoned in the *Principles* the idea that wages *consist* of corn (to the exclusion of manufactures) but whether he also abandoned the idea that the agricultural surplus is critical to the whole distribution process through the fixity of wages in terms of *corn only* is not clear. (Cf. Sraffa, *op. cit.*, pp. xxxii-xxxiii.)

but a necessity, due to competition among the capitalists themselves. This in turn was explained by the existence of economies of large scale production (together with the implicit assumption that the amount of capital employed by any particular capitalist is governed by his own accumulation). Given the fact that the larger the scale of operations the more efficient the business, each capitalist is forced to increase the size of his business through the re-investment of his profits if he is not to fall behind in the competitive struggle.

It is only at a later stage, when the increasing concentration of production in the hands of the more successful enterprises removed the competitive necessity for accumulation—the stage of " monopoly capitalism "—that in the Marxian scheme there is room for economic crises, not on account of an excessive increase in the demand for labour following on accumulation but on account of an insufficiency of effective demand—the failure of markets resulting from the inability of the capitalists either to spend or to invest the full amount of profits (which Marx called the problem of " realising surplus value ").

Marx has also taken over from Ricardo, and the classical economists generally, the idea of a falling rate of profit with the progressive accumulation of capital. But whereas with the classicists this was firmly grounded on the Law of Diminishing Returns, Marx, having discarded that law, had no firm base for it. His own explanation is based on the assumed increase in the ratio of fixed to circulating capital (in Marxian terminology, " constant " to " variable " capital) with the progress of capitalism ; but as several authors have pointed out,[1] the law of the falling rate of profit cannot really be derived from the law of the " increasing organic composition " of capital. Since Marx assumes that the supply price of labour remains unchanged in terms of commodities when the organic composition of capital, and hence output per head, rises, there is no more reason to assume that an increase in " organic composition " will yield a lower rate of profit than a higher rate. For even if output per man were assumed to increase more slowly than (" constant " plus " variable ") capital per man, the " surplus value " per man (the excess of output per man over the costs of reproduction of labour) will necessarily increase faster than output per man, and may thus secure a rising rate of profit even if there is diminishing productivity to successive additions to fixed capital per unit of labour.

While some of Marx's predictions—such as the increasing concentration of production in the hands of large enterprises—proved accurate, his most important thesis, the steady worsening of the living conditions of the working classes—" the immiseration of the proletariat "[2]—has been contradicted by experience, in both the " competitive " and " monopoly " stages of capitalism. On the Marxian model the share of wages in output must necessarily fall with every increase in output per head. The theory can only allow for a rise of wages in terms of commodities as a result of the collective organisation of the working classes which forces the capitalists to reduce the degree of exploitation and to surrender to the workers some of the " surplus value ".[3] This hypothesis however will only yield a constant share of wages on the extremely far-fetched assumption that the rate of increase in the bargaining strength of labour, due to the growth of collective organisation, precisely keeps pace with the rate of increase in output per head.

[1] Cf. in particular, Joan Robinson, *An Essay in Marxian Economics*, pp. 75-82.

[2] It is not clear, in terms of Marx's own theoretical model, why such a progressive immiseration should take place—since the costs of reproduction of labour appear to set an *absolute* limit to the extent to which labour can be exploited. Some parts of *Das Kapital* could however be construed as suggesting that wages can be driven below the (long run) reproduction cost of labour, at the cost of a (long run) shrinkage in the labour force : and with the increasing organic composition of capital, and the rise of monopolies, the demand for labour may show an equally declining tendency.

[3] Marx himself would have conceived a reduction in the " degree of exploitation " in terms of a reduction in the length of the working day rather than a rise in real wages per day. In fact both have occurred side by side.

III THE NEO-CLASSICAL THEORIES

(A) MARGINAL PRODUCTIVITY

While Marx's theory thus derives from Ricardo's surplus principle, neo-classical value and distribution theory derives from another part of the Ricardian model : the " marginal principle " introduced for the explanation of rent (which explains why both Marx and Marshall are able to claim Ricardo as their precursor). The difference between Ricardo and the neo-classics is (1) that whereas Ricardo employed the " principle of substitution " (or rather, the principle of " limited substitutability "—which is the basic assumption underlying all " marginal " analysis) only as regards the use of labour relative to land, in neo-classical theory this doctrine was formalized and generalized, and assumed to hold true of any factor, in relation to any other ;[1] (2) whereas Ricardo employed the principle for showing that a " fixed " factor will earn a surplus, determined by the gap between the average and marginal product of the variable factor, neo-classical theory concentrated on the reverse aspect—i.e., that any factor variable in supply will obtain a remuneration which, under competitive conditions, must correspond to its marginal product. Thus if the total supply of *all* factors (and not only land) is being taken as given, independently of price, and all are assumed to be limited substitutes to one another, the share-out of the whole produce can be regarded as being determined by the marginal rates of substitution between them. Thus in terms of our diagram, if we assumed that along Ox we measure the quantity of any particular factor of production, x, the quantities of all the others being taken as fixed, p—Mp will exhibit the marginal productivity function of the variable factor. If the actual employment of that factor is taken to be M, AM will represent its demand price per unit, and the rectangle $OBAM$ its share in the total produce. Since this principle could be applied to any factor, it must be true of all (including, as Walras and Wicksell have shown, the factors owned by the entrepreneur himself) hence the rectangle $BCDA$ must be sufficient, and only just sufficient, for remunerating all other factors but x on the basis of their respective marginal productivities. This, as Wicksteed has shown[2] requires the assumption that the production function will be homogeneous of the first degree for all variables taken together—an assumption which he himself regarded as little more than a tautology, if " factors of production " are appropriately defined.[3] From the point of view of the theory, however, the *appropriate* definition of factors involves the elimination of intermediate products and their conversion into " ultimate " or "original" factors, since only on this definition can one assume the properties of divisibility and variability of coefficients. When factors are thus defined, the assumption of constant

[1] As well as of any particular commodity in the sphere of consumption. The utility theory of value is really Ricardian rent-theory applied to consumption demand. In fact, as Walras has shown, limited substitutability in consumption might in itself be sufficient to determine distributive shares, provided that the proportions in which the different factors are used are different in different industries. His solution of the problem of distribution, based on " fixed coefficients " of production (intended only as a first approximation) is subject however to various snags since the solution of his equations may yield negative prices for the factors as well as positive ones and it cannot be determined beforehand whether this will be the case or not. If the solution of the equations yields negative prices the factors in question have to be excluded as " free goods "; and the operation (if necessary) successive repeated until only factors with positive prices are left. Also, it is necessary to suppose that the number of different " factors " is no greater than the number of different " products " otherwise the solution is indeterminate.

[2] *The Co-ordlination of the Laws of Distribution* (1894).

[3] *Ibid.*, p. 53 " We must regard every kind and quality of labour that can be distinguished from other kinds and qualities as a separate factor ; and in the same way, every kind of land will be taken as a separate factor. Still more important is it to insist that instead of speaking of so many £ worth of capital we shall speak of so many ploughs, so many tons of manure, and so many horses or footpounds of power. Each of these may be scheduled in its own unit." Under these conditions it is true to say that " doubling all factors will double the product ", but since these " factors " are indivisible in varying degrees, it does not mean that the production function is a linear and homogeneous one in relation to incremental variations of output. Also a change in output may be associated with the introduction of *new* factors of production.

returns to scale is by no means a tautology ; it is a restrictive assumption, which may be regarded, however, as being co-extensive with other restrictive assumptions implied by the theory—i.e., the universal rule of perfect competition, and the absence of external economies and diseconomies.

The basic difficulty with the whole approach does not lie, however, in this so-called " adding-up problem " but in the very meaning of " capital " as a factor of production.[1] Whilst land can be measured in acres-per-year and labour in man-hours, capital (as distinct from " capital goods ") cannot be measured in terms of physical units.[2] To evaluate the marginal product of labour it is necessary to isolate two situations containing identical " capital " but two different quantities of labour, or identical amounts of labour and two differing quantities of " capital ", in precise numerical relationship.[3]

Marshall, without going into the matter in any detail, had shown in several passages that he was dimly aware of this ; and in carefully re-defining marginal productivity so as to mean " marginal *net* productivity " (*net* after deduction of all associated expenses on other " factors ") he shied away from the task of putting forward a general theory of distribution altogether.[4]

In fact, in so far as we can speak of a " Marshallian " theory of distribution at all, it is in the sense of a " short period " theory, which regards profits as the " quasi-rents " earned on the use of capital goods of various kinds, the supply of which can be treated as given for the time being, as a heritage of the past. The doctrine of the " quasi-rent " assimilates capital as a factor of production to Ricardian land : the separate *kinds* of capital goods being treated as so many different kinds of " land ". Here the problem of the measurement of capital as a factor of production does not arise : since, strictly speaking, no kind of change or reorganization in the stock of intermediate products is permitted in connection with a change in the level or composition of production. It was this aspect of Marshall which, consciously or sub-consciously, provided the " model " for most of the post-Marshallian Cambridge theorizing. Prices are equal to, or determined by, marginal prime costs ; profits are determined by the difference between marginal and average prime costs ; prime costs, for the system as a whole, are labour costs (since raw-material costs, for a closed economy at any rate, disappear if all branches of industry are taken together) ; ultimately therefore the division of output between profits and wages is a matter depending on the existence of diminishing returns to labour, as more labour is used in conjunction with a *given* capital equipment ; and is determined by the elasticity of labour's average productivity curve which fixes the share of quasi-rents.

Marshall himself would have disagreed with the use of the quasi-rent doctrine as a distribution theory, holding that distributive shares in the short period are determined by long-period forces.[5] Clearly even if one were to hold strictly to the assumption that " profit margins " are the outcome of short-period profit-maximisation, this " short-

[1] For a general equilibrium system, capital goods cannot be regarded as factors of production *per se* (in the manner suggested by Wicksteed) otherwise the same things are simultaneously treated as the parameters and the unknowns of the system.

[2] Measurement in terms of value (as so many £'s of " capital ") already assumes a certain rate of interest, on the basis of which services accruing in differing periods in the future, or costs incurred at differing dates in the past, are brought to a measure of equivalence.

[3] The product of the " marginal shepherd " is the difference, in terms of numbers of sheep, between 10 shepherds using 10 crooks and 11 shepherds using 11 slightly inferior crooks, the term " slightly inferior " being taken to mean that the 11 crooks in the one case represent precisely the same amount of " capital " as the 10 crooks in the other case. (Cf. also, Robertson, " Wage Grumbles," in *Economic Fragments*, 1931.)

[4] " The doctrine that the earnings of a worker tend to be equal to the net product of his work, has by itself no real meaning ; since in order to estimate the net product, we have to take for granted all the expenses of production of the commodity on which he works, other than his own wages ". Similarly, the doctrine that the marginal efficiency of capital will tend to equal the rate of interest " cannot be made into a theory of interest, any more than a theory of wages, without reasoning in a circle ". (Cf. *Principles*, 8th edition, Book VI, ch. I, paras 7-8.)

[5] Cf., in particular, *Principles*, 8th edition, Book V, ch. V, and 6, and Book VI, ch. VIII, paras. 4.

period " approach does not really get us anywhere : for the extent to which diminishing returns operate for labour in conjunction with the capital equipment available to-day is itself a function of the price-relationships which have ruled in the past because these have determined the quantities of each of the kinds of equipment available. The theory does not therefore really amount to more than saying that the prices of to-day are derived from the prices of yesterday—a proposition which is the more true and the more trivial the shorter the " day " is conceived to be, in terms of chronological time.

For the true neo-classical attempt to solve the general problem of distribution we must go to Wicksell who thought that by integrating the Austrian approach to capital with Walrasian equilibrium theory he could provide a general solution, treating capital as a two-dimensional quantity, the product of time and labour. The " time " in this case is the investment period or waiting period separating the application of " original " factors from the emergence of the final product, and the marginal productivity of capital the added product resulting from an extension of " time ". This attempt, again, came to grief (as Wicksell himself came near to acknowledging late in life[1]) (i) owing to the impossibility of measuring that period in terms of an " average " of some kind ;[2] (ii) owing to the impossibility of combining the investment periods of different " original " factors in a single measure.[3]

In fact the whole approach which regards the share of wages and of profits in output as being determined by the marginal rate of substitution between Capital and Labour—with its corollary, that the constancy of relative shares is evidence of a unity-Elasticity of Substitution between Capital and Labour[4]—is hardly acceptable to present-day economists. Its inadequacy becomes evident as soon as it is realized that the "marginal rate of substitution " between Capital and Labour—as distinct from the marginal rate of substitution between labour and land—can only be determined once the rate of profit and the rate of wages are already known. The same technical alternatives might yield very different " marginal rates of substitution " according as the ratio of profits to wages is one thing or another. The theory asserts in effect, that the rate of interest in the capital market, (and the associated wage rate in the labour market) is determined by the condition that at any lower interest rate (and higher wage rate) capital would be invested in such " labour-saving " forms as would provide insufficient employment to the available labour ; whilst at any higher rate, capital would be invested in forms that offered more places of employment than could be filled with the available labour.

Quite apart from all conceptual difficulties, the theory focuses attention on a relatively unimportant feature of a growing economy. For accumulation does not take the form of " deepening " the structure of capital (at a given state of knowledge) but rather in keeping pace with technical progress and the growth in the labour force. It is difficult to swallow a theory which says, in effect that wages and profits are what they are for otherwise there would be too much deepening or too little deepening (the capital/output ratios would be either too large or too small) to be consistent with simultaneous equilibrium in the savings-investment market and in the labour market.

[1] Cf. the concluding passage of his posthumous contribution to the Wieser Festschrift. *Die Wirtschaftstheorie der Gegenwart* (1928) Vol. III, pp. 208-9 ; also his Analysis of Akerman's Problem, reprinted in *Lectures*, Vol. I, p. 270.

[2] Since owing to compound interest, the weights to be used in the calculation of the average will themselves be dependent on the rate of interest.

[3] For a more extended treatment cf. my articles on capital theory in *Econometrica*, April 1937 and May 1938 ; also Joan Robinson, The Production Function in the Theory of Capital, *Review of Economic Studies*, Vol. XXI (1953-54) p. 81, and *Comment* by D. G. Champernowne, *ibid* page 112.

[4] Cf. Hicks, *The Theory of Wages* (1932) ch. VI, passim.

(B) The "Degree of Monopoly" Theories of Distribution

Monopoly profit was always regarded as a distinct form of revenue in neo-classical theory, though not one of any great quantitative importance since the mass of commodities was thought of as being produced under competitive conditions. But the modern theories of imperfect competition emphasised that monopoly profit is not an isolated feature. Profits in general contain an *element* of monopoly revenue—an element that is best defined as the excess of the actual profit margin in output over what the profit margin would have been under perfectly competitive conditions. Under Marshallian "short-period" assumptions the perfectly-competitive profit margin is given by the excess of marginal cost over average prime costs. The additional monopoly element is indicated by the excess of price over marginal cost. The former, as we have seen, is a derivative of the elasticity of labour's productivity curve where capital equipment of all kinds is treated as given. The latter is a derivative of the elasticity of demand facing the individual firm. The novel feature of imperfect competition theories is to have shown that the increase of profit margins due to this element of monopoly need not imply a corresponding excess in the rates of profit on capital over the competitive rate ; through the generation of excess capacity (i.e., the tendency of demand curves to become "tangential" to the cost curves) the latter may approach a "competitive" or "normal" rate (as a result of the consequential rise in the capital/output ratio) even if the former is above the competitive level.

Kalecki[1] built on this a simplified theory of distribution, where the share of profits in output is shown to be determined by the elasticity of demand alone. This was based on the hypothesis that in the short period, labour and capital equipment are largely "limitational" and not "substitutional" factors, with the result that the short-period prime cost-curve is a reverse —L shaped one (prime costs being constant up to full capacity output). In that case marginal costs are equal to average prime costs ; the ratio of price to prime costs (and hence, in a closed economy, the ratio of gross profits to wages) is thus entirely accounted for by the elasticity of the firm's demand curve.

On closer inspection, however, the elasticity of the demand curve facing the individual firm turned out to be no less of a broken reed than its counterpart, the elasticity of substitution between factors. There is no evidence that firms in imperfect markets set their prices by reference to the elasticity of their sales-function, or that short-period pricing is the outcome of any deliberate attempt to maximize profits by reference to an independent revenue and a cost function. Indeed the very notion of a demand curve for the products of a single firm is illegitimate if the prices charged by different firms cannot be assumed to be independent of each other.[2]

In the later versions of his theory Kalecki abandoned the link between the "degree of monopoly" and the elasticity of demand, and was content with a purely tautological approach according to which the ratio of price to prime costs is *defined* simply as the "degree of monopoly". Propositions based on implicit definitions of this kind make of course no assertion about reality and possess no explanatory value. Unless the "degree of monopoly" can be defined in terms of market relationships of some kind (as, for example, in terms of the "cross-elasticities" of demand for the products of the different firms)[3] and an attempt is made to demonstrate how these market relationships determine

[1] The original version appeared in *Econometrica*, April 1938. Subsequent versions appeared in *Essays in the Theory of Economic Fluctuations* (1938) ch. I, *Studies in Economic Dynamics* (1943) ch. 1, and *Theory of Dynamic Economics* (1954) Part 1.

[2] The theory of the "kinked" demand curve is in fact no more than a recognition of the fact that the demand curve of the firm (in the sense required for the purpose of deriving price from the postulate of profit maximisation) is non-existent. Since the position of the "kink" *depends* on the price, it cannot *determine* the price ; it thus leaves the profit margin completely undetermined.

[3] The "cross-elasticities" of demand indicate the degree of interdependence of the markets of different firms and are thus inversely related to monopoly power in the usual sense of the word.

the relation between prices and costs, the theory does not provide a hypothesis which could be affirmed or refuted.

There is no need, of course, to follow Kalecki in the attempt to lend spurious precision to the doctrine through implicit theorizing—a vice which afflicts all theories which we grouped together as " neo-classical " in varying degrees. Fundamentally, the proposition that the distribution of income between wages and profits depends on market structures, on the strength or weakness of the forces of competition, is not a tautological one ; it asserts *something* about reality (which may in principle be proved false) even if that " something " cannot be given a logically precise formulation. Just as the positive content of the marginal productivity theory can be summed up by the statement that the rate of profit on capital (and the margin of profit in output) is governed by the need to prevent the capital/output ratio from being either too large or too small, the positive content of the " degree of monopoly " theory can be summed up in the sentence that " profit margins are what they are because the forces of competition prevent them from being higher than they are and are not powerful enough to make them lower than they are ". Unfortunately neither of these statements gets us very far.

Dissatisfaction with the tautological character and the formalism of the " marginal revenue-equals-marginal cost " type of price theory led to the formulation of the " full cost " theories of pricing,[1] according to which producers in imperfect markets set their prices independently of the character of demand, and solely on the basis of their long run costs of production (including the " normal " rate of profit on their own capital). If these theories asserted no more than that prices in manufacturing industry are *not* determined by the criterion of short-run profit-maximization, and that profit margins can be fairly insensitive to short-period variations in demand,[2] (the impact effect of changes in demand being on the rate of production, rather than on prices) they would provide a healthy antidote to a great deal of facile theorising. When, however, they go beyond this and assert that prices are determined quite independently of demand, they in effect destroy existing price theory without putting anything else in its place. Quite apart from the fact that a " full cost " theory is quite unable to explain why some firms should be more successful in earning profits than others, the level of the " normal profit " on which the full cost calculations are supposed to be based is left quite undetermined. The very fact that these full cost theories should have received such widespread and serious consideration as an alternative explanation of the pricing process is an indication of the sad state of vagueness and confusion into which the neo-classical value theory had fallen.

[1] Cf. Hall and Hitch, *Oxford Economic Papers*, 1939 ; P. M. S. Andrews, *Manufacturing Business* (1949).

[2] This, I believe, was the intention of the original Hall-Hitch article. Cf. Marshall, *Principles*, Book VI, ch. VIII, paragraph 4 : " We see then that there is no general tendency of profits on the turnover to equality ; but there may be, and as a matter of fact there is, in each trade and in every branch of each trade, a more or less definite rate of profits on the turnover which is regarded as a " fair " or normal rate. Of course these rates are always changing in consequence of changes in the methods of trade ; which are generally begun by individuals who desire to do a larger trade at a lower rate of profit on the turnover than has been customary, but at a larger rate of profit per annum on their capital. If however there happens to be no great change of this kind going on, the traditions of the trade that a certain rate of profit on the turnover should be charged for a particular class of work are of great practical service to those in the trade. Such traditions are the outcome of much experience tending to show that, if that rate is charged, a proper allowance will be made for all the costs (supplementary as well as prime) incurred for that particular purpose, and in addition the normal rate of profits per annum in that class of business will be afforded. If they charge a price which gives much less than this rate of profit on the turnover they can hardly prosper ; and if they charge much more they are in danger of losing their custom, since others can afford to undersell them. This is the " fair " rate of profit on the turnover, which an honest man is expected to charge for making goods to order, when no price has been agreed on beforehand ; and it is the rate which a court of law will allow in case a dispute should arise between buyer and seller." Cf. also Kahn, *Economic Journal*, 1952, p. 119.

IV THE KEYNESIAN THEORY

Keynes, as far as I know, was never interested in the problem of distribution as such. One may nevertheless christen a particular theory of distribution as " Keynesian " if it can be shown to be an application of the specifically Keynesian apparatus of thought and if evidence can be adduced that at some stage in the development of his ideas, Keynes came near to formulating such a theory.[1] The principle of the Multiplier (which in some way was anticipated in the *Treatise* but without a clear view of its implications) could be alternatively applied to a determination of the relation between prices and wages, if the level of output and employment is taken as given, or the determination of the level of employment, if distribution (i.e., the relation between prices and wages) is taken as given. The reason why the multiplier-analysis has not been developed as a distribution theory is precisely because it was invented for the purpose of an employment theory—to explain why an economic system can remain in equilibrium in a state of under-employment (or of a general under-utilization of resources), where the classical properties of scarcity-economics are inapplicable. And its use for the one appears to exclude its use for the other.[2] If we assume that the balance of savings and investment is brought about through variations in the relationship of prices and costs, we are not only bereft of a principle for explaining variations in output and employment, but the whole idea of separate " aggregate" demand and supply functions—the principle of "effective demand "—falls to the ground ; we are back to Say's Law, where output as a whole is limited by available resources, and a fall in effective demand for one kind of commodity (in real terms) generates compensating increases in effective demand (again in real terms) for others. Yet these two uses of the Multiplier principle are not as incompatible as would appear at first sight : the Keynesian technique, as I hope to show, can be used for both purposes, provided the one is conceived as a short-run theory and the other as a long-run theory—or rather, the one is used in the framework of a static model, and the other in the framework of a dynamic growth model.[3]

[1] I am referring to the well-known passage on profits being likened to a " widow's cruse " in the *Treatise on Money*, Vol. I, p. 139. " If entrepreneurs choose to spend a portion of their profits on consumption (and there is, of course, nothing to prevent them from doing this) the effect is to *increase* the profit on the sale of liquid consumption goods by an amount exactly equal to the amount of profits which have been thus expended . . . Thus however much of their profits entrepreneurs spend on consumption, the increment of wealth belonging to entrepreneurs remain the same as before. Thus profits, as a source of capital increment for entrepreneurs, are a widow's cruse which remains undepleted however much of them may be devoted to riotous living. When on the other hand, entrepreneurs are making losses, and seek to recoup these losses by curtailing their normal expenditure on consumption, i.e., by saving more, the cruse becomes a Danaid jar which can never be filled up ; for the effect of this reduced expenditure is to inflict on the producers of consumption-goods a loss of an equal amount. Thus the diminution of their wealth, as a class is as great, in spite of their savings, as it was before." This passage, I think, contains the true seed of the ideas developed in the *General Theory*—as well as showing the length of the road that had to be traversed before arriving at the conceptual framework presented in the latter work. The fact that " profits ", "savings" etc. were all defined here in a special sense that was later discarded, and that the argument specifically refers to expenditure on consumption goods, rather than entrepreneurial expenditure in general, should not blind us to the fact that here Keynes regards entrepreneurial incomes as being the resultant of their expenditure decisions, rather than the other way round—which is perhaps the most important difference between " Keynesian " and " pre-Keynesian " habits of thought.

[2] Although this application of Keynesian theory has been implicit in several discussions of the problem of inflation. (Cf. *e.g.* A. J. Brown, *The Great Inflation*, Macmillan, 1955.)

[3] I first thought of using the Multiplier technique for purposes of a distribution theory when I attempted the ultimate incidence of profits taxation under full employment conditions in a paper prepared for the Royal Commission on Taxation in 1951. The further development of these ideas, and particularly their relationship to a dynamic theory of growth, owes a great deal to discussions with Mrs. Robinson, whose forthcoming book, *The Accumulation of Capital*, contains a systematic exploration of this field. I should also like to mention here that I owe a great deal of stimulus to a paper by Kalecki, " A Theory of Profits " (*Economic Journal*, June-Sept. 1942) whose approach is in some ways reminiscent of the " widows' cruse " of Keynes' *Treatise* even though Kalecki uses the technique, not for an explanation of the share of profits in output, but for showing why the *level* of output and its fluctuations is peculiarly dependent on entrepreneurial behaviour. (In doing so, he uses the restrictive assumption that savings are entirely supplied out of profits.) I have also been helped by Mr. Harry Johnson and Mr. Robin Marris, both in the working out of the formulae and in general discussion.

We shall assume, to begin with, a state of full employment (we shall show later the conditions under which a state of full employment will *result* from our model) so that total output or income (Y) is given. Income may be divided into two broad categories, Wages and Profits (W and P), where the wage-category comprises not only manual labour but salaries as well, and Profits the income of property owners generally, and not only of entrepreneurs ; the important difference between them being in the marginal propensities to consume (or save), wage-earners' marginal savings being small in relation to those of capitalists.[1]

Writing S_w and S_p for aggregate savings out of Wages and Profits, we have the following income identities :

$$Y \equiv W + P$$
$$I \equiv S$$
$$S \equiv S_w + S_p.$$

Taking investment as given, and assuming simple proportional saving functions $S_w = s_w W$ and $S_p = s_p P$, we obtain :

$$I = s_p P + s_w W = s_p P + s_w(Y - P) = (s_p - s_w)P + s_w Y$$

Whence
$$\frac{I}{Y} = (s_p - s_w)\frac{P}{Y} + s_w \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

and
$$\frac{P}{Y} = \frac{1}{s_p - s_w}\frac{I}{Y} - \frac{s_w}{s_p - s_w} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

Thus, given the wage-earners' and the capitalists' propensities to save, the share of profits in income depends simply on the ratio of investment to output.

The interpretative value of the model (as distinct from the formal validity of the equations, or identities) depends on the " Keynesian " hypothesis that investment, or rather, the ratio of investment to output, can be treated as an independent variable, invariant with respect to changes in the two savings propensities s_p and s_w. (We shall see later that this assumption can only be true within certain limits, and outside those limits the theory ceases to hold). This, together with the assumption of " full employment ", also implies that the level of prices in relation to the level of money wages is determined by demand : a rise in investment, and thus in total demand, will raise prices and profit margins, and thus reduce real consumption, whilst a fall in investment, and thus in total demand, causes a fall in prices (relatively to the wage level) and thereby generates a compensating rise in real consumption. Assuming flexible prices (or rather flexible profit margins) the system is thus stable at full employment.

The model operates only if the two savings propensities differ and the marginal propensity to save from profits exceeds that from wages, i.e. if :

and
$$s_p \neq s_w$$
$$s_p > s_w$$

The latter is the stability condition. For if $s_p < s_w$, a fall in prices would cause a fall in demand and thus generate a further fall in prices, and equally, a rise in prices would be cumulative. The degree of stability of the system depends on the *difference* of the marginal propensities, *i.e.*, on $\frac{1}{s_p - s_w}$ which may be defined as the " coefficient of sensitivity of income distribution ", since it indicates the change in the share of profits in income which follows upon a change in the share of investment in output.

[1] This may be assumed independently of any skewness in the distribution of property, simply as a consequence of the fact that the bulk of profits accrues in the form of company profits and a high proportion of companies' marginal profits is put to reserve.

If the difference between the marginal propensities is small, the coefficient will be large, and small changes in $\frac{I}{Y}$ (the investment/output relationship) will cause relatively large changes in income distribution $\frac{P}{Y}$; and *vice versa.*

In the limiting case where $s_w = 0$, the amount of profits is equal to the sum of investment and capitalist consumption, *i,e,* :

$$P = \frac{1}{s_p} \dot{\mathrm{I}}.$$

This is the assumption implicit in Keynes' parable about the widow's cruse—where a rise in entrepreneurial consumption raises their total profit by an *identical* amount—and of Mr. Kalecki's theory of profits which can be paraphrased by saying that " capitalists earn what they spend, and workers spend what they earn."

This model (*i.e.*, the " special case " where $s_w = 0$) in a sense is the precise opposite of the Ricardian (or Marxian) one—here wages (not profits) are a residue, profits being governed by the propensity to invest and the capitalists' propensity to consume, which represent a kind of " prior charge " on the national output. Whereas in the Ricardian model the ultimate incidence of all taxes (other than taxes on rent) fall on profits, here the incidence of all taxes, taxes on income and profits as well as on commodities, falls on wages.[1] Assuming however that $\frac{I}{Y}$ and s_p remain constant over time, the share of wages will also remain constant—*i.e.*, real wages will increase automatically, year by year, with the increase in output per man.

If s_w is positive the picture is more complicated. Total profits will be reduced by the amount of workers' savings, S_w ; on the other hand, the sensitivity of profits to changes in the level of investment will be greater, total profits rising (or falling) by a greater amount than the change in investment, owing to the consequential reduction (or increase) in workers' savings.[2]

The critical assumption is that the investment/output ratio is an independent variable. Following Harrod, we can describe the determinants of the investment/output ratio in terms of the rate of growth of output capacity (G) and the capital/output ratio, v :

$$\frac{I}{Y} = Gv \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

[1] The ultimate incidence of taxes can only fall on profits (on this model) in so far as they increase s_p, the propensity to save out of *net* income after tax. Income and profits taxes, through the " double taxation " of savings, have of course the opposite effect : they reduce s_p, and thereby make the share of *net* profits in income larger than it would be in the absence of taxation. On the other hand, discriminatory taxes on dividend distribution, or dividend limitation, by keeping down both dividends and capital gains, have the effect of raising s_p. (All this applies, of course, on the assumption that the Government *spends* the proceeds of the tax—*i.e.*, that it aims at a balanced budget. Taxes which go to augment the budget surplus will lower the share of profits in much the same way as an increase in workers' savings.)

[2] Thus if $s_p = 50\%$, $s_w = 10\%$, $\frac{I}{Y} = 20\%$, $\frac{P}{Y}$ will be 15% ; but a rise in $\frac{I}{Y}$ to 21% would raise $\frac{P}{Y}$ to 17·5%. If on the other hand $s_w = 0$, with $s_p = 50\%$, $\frac{P}{Y}$ would become 40%, but an increase in $\frac{I}{Y}$ to 21% would only increase $\frac{P}{Y}$ to 42%. The above formulae assume that average and marginal propensities are identical. Introducing constant terms in the consumption functions alters the relationship between $\frac{P}{Y}$ and $\frac{I}{Y}$ and would reduce the *elasticity* of $\frac{P}{Y}$ with respect to changes in $\frac{I}{Y}$.

In a state of continuous full employment G must be equal to the rate of growth of the " full employment ceiling ", *i.e.*, the sum of the rate of technical progress and the growth in working population (Harrod's " natural rate of growth "). For Harrod's second equation :

$$\frac{I}{Y} = s$$

we can now substitute equation (1) above :

$$\frac{I}{Y} = (s_p - s_w)\frac{P}{Y} + s_w.$$

Hence the " warranted " and the " natural " rates of growth are not independent of one another ; if profit margins are flexible, the former will adjust itself to the latter through a consequential change in $\frac{P}{Y}$.

This does not mean that there will be an *inherent* tendency to a smooth rate of growth in a capitalist economy, only that the causes of cyclical movements lie elsewhere—not in the lack of an adjustment mechanism between s and Gv. As I have attempted to demonstrate elsewhere[1] the causes of cyclical movements should be sought in a disharmony between the entrepreneurs' *desired* growth rate (as influenced by the degree of optimism and the volatility of expectations) which governs the rate of increase of output capacity (let us call it G') and the natural growth rate (dependent on technical progress and the growth of the working population) which governs the rate of growth in output. It is the excess of G' over G—not the excess of s over Gv—which causes periodic breakdowns in the investment process through the growth in output capacity outrunning the growth in production.[2]

Problems of the trade cycle however lie outside the scope of this paper ; and having described a model which shows the distribution of income to be determined by the Keynesian investment-savings mechanism, we must now examine its limitations. The model, as I emphasized earlier, shows the share of profits $\frac{P}{Y}$, the rate of profit on investment $\frac{P}{vY}$, and the real wage rate $\frac{W'}{L}$, as functions of $\frac{I}{Y}$ which in turn is determined independently of $\frac{P}{Y}$ or $\frac{W}{L}$. There are four different reasons why this may not be true, or be true only within a certain range.

(1) The first is that the real wage cannot fall below a certain subsistence minimum. Hence $\frac{P}{Y}$ can only attain its indicated value, if the resulting real wage exceeds this minimum rate, w'. Hence the model is subject to the restriction $\frac{W}{L} \geqslant w'$, which we may write in the form :

$$\frac{P}{Y} \leqslant \frac{Y - w'L}{Y} \quad \ldots\ldots\ldots\ldots \quad (4)$$

[1] *Economic Journal*, March 1954, pp. 53-71.

[2] $\frac{I}{Y}$ will therefore tend to equal $G'v$, not Gv. It may be assumed that taking very long periods G' is largely governed by G but over shorter periods the two are quite distinct, moreover G itself is not independent of G', since technical progress and population growth are both stimulated by the degree of pressure on the " full employment ceiling ", which depends on G'. The elasticity of response of G to G' is not infinite however : hence the greater G', the greater will be G (the *actual* trend-rate of growth of the economy over successive cycles) but the greater also the ratio $\frac{G'}{G}$ which measures the strength of cyclical forces.

(2) The second is that the indicated share of profits cannot be below the level which yields the minimum rate of profit necessary to induce capitalists to invest their capital, and which we may call the " risk premium rate ", r. Hence the restriction :

$$\frac{P}{vY} \geqslant r \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

(3) The third is that apart from a minimum rate of profit on capital there may be a certain minimum rate of profit on turnover—due to imperfections of competition, collusive aggreements between traders, etc., and which we may call m, the " degree of monopoly " rate. Hence the restriction :

$$\frac{P}{Y} \geqslant m \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

It is clear that equations (5) and (6) describe *alternative* restrictions, of which the higher will apply.

(4) The fourth is that the capital/output ratio, v, should not in itself be influenced by the rate of profit, for if it is, the investment/output ratio Gv will itself be dependent on the rate of profit. A certain degree of dependence follows inevitably from the consideration, mentioned earlier, that the value of particular capital goods in terms of final consumption goods will vary with the rate of profit,[2] so that, even with a *given technique* v will not be independent of $\frac{P}{Y}$. (We shall ignore this point). There is the further complication that the relation $\frac{P}{Y}$ may affect v through making more or less " labour-saving " techniques profitable. In other words, at any given wage-price relationship, the producers will adopt the technique which maximizes the rate of profit on capital, $\frac{P}{vY}$; this will affect (at a given G) $\frac{I}{Y}$, and hence $\frac{P}{Y}$. Hence any rise in $\frac{P}{Y}$ will reduce v, and thus $\frac{I}{Y}$, and conversely, any rise in $\frac{I}{Y}$ will raise $\frac{P}{Y}$. If the sensitiveness of v to $\frac{P}{Y}$ is great, $\frac{P}{Y}$ can no longer be regarded as being determined by the equations of the model ; the *technical* relation between v and $\frac{P}{Y}$ will then govern $\frac{P}{Y}$ whereas the savings equation (equation (2) above) will determine $\frac{I}{Y}$ and thus (given G) the value of v.[3] To exclude this we have to assume that v is invariant to $\frac{P}{Y}$,[4] *i.e.* :

$$v = \bar{v} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

[1] Where L = labour force.

[2] Cf. p. 90 above. In fact the whole of the Keynesian and post-Keynesian analysis dodges the problem of the measurement of capital.

[3] This is where the " marginal productivity " principle would come in but it should be emphasized that under the conditions of our model where savings are treated, not as a constant, but as a function of income distribution, $\frac{P}{Y}$, the sensitiveness of v to changes in $\frac{P}{Y}$ would have to be very large to overshadow the influence of G and of s_p and of s_w on $\frac{P}{Y}$. Assuming that it is large, it is further necessary to suppose that the value of $\frac{P}{Y}$ as dictated by this technical relationship falls within the maximum and minimum values indicated by equations (4)-(6).

[4] This assumption does not necessarily mean that there are " fixed coefficients " as between capital equipment and labour—only that technical innovations (which are also assumed to be " neutral " in their effects) are far more influential on the chosen v than price relationships.

If equation (4) is unsatisfied, we are back at the Ricardian (or Marxian) model. $\frac{I}{Y}$ will suffer a shrinkage, and will no longer correspond to Gv, but to, say, αv where $\alpha < G$. Hence the system will not produce full employment ; output will be limited by the available capital, and not by labour ; at the same time the classical, and not the Keynesian, reaction-mechanism will be in operation : the size of the " surplus " available for investment determining investment, not investment savings. It is possible however that owing to technical inventions, etc., and starting from a position of excess labour and underemployment (*i.e.*, an elastic total supply of labour) the size of the surplus will grow ; hence $\frac{I}{Y}$ and α will grow ; and hence α might rise above G (the rate of growth of the " full employment ceiling ", given the technical progress and the growth of population) so that in time the excess labour becomes absorbed and full employment is reached. When this happens (which we may call the stage of *developed* capitalism) wages will rise above the subsistence level, and the properties of the system will then follow our model.

If equations (5) and (6) are unsatisfied, the full employment assumption breaks down, and so will the process of growth ; the economy will relapse into a state of stagnation. The interesting conclusion which emerges from these equations is that this may be the result of several distinct causes. " Investment opportunities " may be low because G' is low relatively to G, *i.e.*, the entrepreneurs' expectations are involatile, and/or they are pessimistic ; hence they expect a lower level of demand for the future than corresponds to potential demand, governed by G. On the other hand, " liquidity preference " may be too high, or the risks associated with investment too great, leading to an excessive r. (This is perhaps the factor on which Keynes himself set greatest store as a cause of unemployment and stagnation.) Finally, lack of competition may cause " over-saving " through excessive profit margins ; this again will cause stagnation, unless there is sufficient compensating increase in v (through the generation of " excess capacity " under conditions of rigid profit margins but relatively free entry) to push up Gv, and hence $\frac{I}{Y}$.

If however equations (2)-(6) are all satisfied there will be an inherent tendency to growth and an inherent tendency to full employment. Indeed the two are closely linked to each other. Apart from the case of a developing economy in the immature stage of capitalism (where equation (4) does not hold, but where $\gamma < G$), a tendency to continued economic growth will only exist when the system is only stable at full employment equilibrium—*i.e.* when $G' \geqslant G$.

This is a possible interpretation of the long-term situation in the " successful " capitalist economies of Western Europe and North America. If G' exceeds G, the investment/output ratio $\frac{I}{Y}$ will not be steady in time, even if the *trend* level of this ratio is constant. There will be periodic breakdowns in the investment process, due to the growth in output capacity outrunning the possible growth in output ; when that happens, not only investment, but total output will fall, and output will be (temporarily) limited by effective demand, and not by the scarcity of resources. This is contrary to the mechanics of our model, but several reasons can be adduced to show why the system will not be flexible enough to ensure full employment in the short period.

(1) First, even if " profit margins " are assumed to be fully flexible, in a downward, as well as an upward, direction the very fact that investment goods and consumer goods are produced by different industries, with limited mobility between them, will mean that profit margins in the consumption goods industries will not fall below the level that ensures full utilization of resources in the consumption goods industries. A *compensating* increase

in consumption goods production (following upon a fall in the production of investment goods) can only occur as a result of a transfer of resources from the other industries, lured by the profit opportunities there.

(2) Second, and more important, profit-margins are likely to be inflexible in a downward direction in the short period (Marshall's " fear of spoiling the market ") even if they are flexible in the long period, or even if they possess short period flexibility in an upward direction.[1]

This applies of course not only to profit margins but to real wages as well, which in the short period may be equally inflexible in a downward direction at the *attained* level, thus compressing $\frac{I}{Y}$, or rather preventing an *increase* in $\frac{I}{Y}$ following upon a rise in the entrepreneurs' desired rate of expansion G'. Hence in the short period the shares of profits and wages tend to be inflexible for two different reasons—the downward inflexibility of $\frac{P}{Y}$ and the downward inflexibility of $\frac{W}{L}$—which thus tend to reinforce the long-period stability of these shares, due to constancy of $\frac{I}{Y}$, resulting from the long period constancy of Gv and $G'v$.

We have seen how the various " models " of distribution, the Ricardo-Marxian, the Keynesian and the Kaleckian are related to each other. I am not sure where " marginal productivity " comes in in all this—except that in so far as it has any importance it does through an extreme sensitivity of v to changes in $\frac{P}{Y}$.

Cambridge. NICHOLAS KALDOR.

[1] Cf. the quotation from Marshall, note 2, page 93 above.

[2] This operates through the wage-price spiral that would follow on a reduction in real wages ; the prevention of such a wage-price spiral by means of investment rationing of some kind, or a " credit squeeze ", is thus a manifestation of downward inflexibility of $\frac{W}{Y}$.