

Part V

Factor Pricing

31

THE GENERAL THEORY OF FACTOR PRICING

Significance of Factor Pricing

The prices paid by businessmen to the various factors of production in the form of rent, wages, interest, etc., are a major determinant of money incomes. Thus, the resource prices play a crucial role in determining the distribution of income in the community. The households supply the human and property resources and get money incomes in return.

Further, factor prices serve as a rationing device for the utilisation of the productive resources. These resources are allocated among various industries and firms through the mechanism of factor prices. These prices facilitate the shifting of resources from the less remunerative uses to the more profitable ones. Dynamic societies cannot function without this shifting.

From the point of view of the firms, the resource prices enable the application to production of the most economical or least cost combination of factors. To the firms factor prices are costs and these costs must be minimised in order to maximise profits.

Finally, since resource prices constitute incomes for the various sections of the society, a very important consideration is the attainment of equality of incomes. This trend in modern societies highlights the ethical and political significance of factor prices.

Theory of factor pricing is also called theory of distribution.

Meaning of Distribution

By "distribution" in the present context, we do not mean the distributive activities of traders and middle men. "The economics of distribution," says Chapman, "accounts for the sharing of the wealth produced by a community among the agents, or the

owners of the agents, which have been active in its production."¹

The theory of distribution is concerned with the evaluation of the services of the factors of production, a study of the conditions of demand for, and supply of, the units of these factors and the influences that bring about changes in their market price.

For instance, in the factor market, it is not acres of land which are being bought or sold but the services of land. Similarly, neither labour nor capital goods are being bought and sold but the services of labour and of capital. Thus, rent is not the price of land but the price of service or the use of land; wages, the price of service of labour; interest, the price of the use of capital; and profit, the reward for entrepreneur's services.

Functional vs. Personal Distribution

It may be pointed out that the distribution discussed here is **functional and not personal**. It is distribution not among individuals but among agents of production. The same person may represent in his person all the four agents, e.g., a peasant-proprietor. He is the entrepreneur, the labourer, the capitalist (for he has some capital of his own), and the landlord all rolled in one. Here we do not discuss how much he earns as an individual but the reward that he gets separately for supplying each factor of production. Thus, we study distribution in the form of rent, wages, interest and profits and not among the different individuals in the community.

It may also be understood that the prices of the factors of production are really the prices paid for them by the firms using them in producing a

1. *Outlines of Political Economy*, p. 278.

commodity. From the point of view of the firms, they are the costs of production. In other words, what is cost to a firm is income to the factors of production. Wages, rent interest, profit are the functional incomes respectively of the labourer, the landlord, the capitalist and the entrepreneur. The reward that each factor gets is the price paid for his service by the entrepreneur. Thus, from one angle it is income and from the other it is cost.

Role of Factor Prices

Price theory covers both product-pricing and factor-pricing. So far we have studied product-pricing and now we turn to factor-pricing. The factor prices play a key role in the working of the free-enterprise economy. Briefly:

(a) It is through factor prices that employment levels of the various productive resources are determined. That is, how much of a factor of production should be utilised in the process of production. A fall in the price of a factor will lead to increase in its demand and more of it being employed, and vice versa.

(b) The second important function of the factor prices is to allocate the productive resources among the various alternative uses. They signal the resources from the less important to the more important uses. This sort of discrimination will be simply out of the question in the absence of guide-lines furnished by the factor prices.

(c) As a corollary from the second function, the factor prices not only guide the individual firms regarding the use of resources, they also bring about the most efficient allocation of the resources of the community. Without pricing of factors, optimum utilisation of factors will not be possible.

(d) Finally, since we all are suppliers of one resource or another, factor-pricing determines the incomes of all of us, *i.e.*, our respective shares in the national output.

MARGINAL PRODUCTIVITY THEORY

Now let us see how the services of factors of production are evaluated. One theory put forward in this connection is the marginal productivity theory. We find references to marginal productivity theory in von Thunen's *Der Isolierte Staat* (1826), Longfield's Lectures on *Political Economy*, and in Henry George's *Progress and Poverty* (1879). But John Bates Clerk's name is most widely associated with its development. In 1880's and 1890's Jevons, Wicksteed, Marshall, Wood, Walras and others made important contribution to its development.

The entrepreneur buys the services of the various factors of production. Since he works for profit, he can only pay a price for a factor which he finds just

worthwhile. Obviously, he cannot afford to pay more than its marginal productivity. Since there is open competition, nobody will accept less than marginal productivity. That is how marginal productivity (not total productivity) determines the remuneration or the price of a factor of production.

The entrepreneur, in employing the various factors of production, acts on the principle of substitution. He substitutes one factor for another till the marginal productivities of all the factors he uses are equalised. In this way, he maximises his profit. What is marginal productivity?

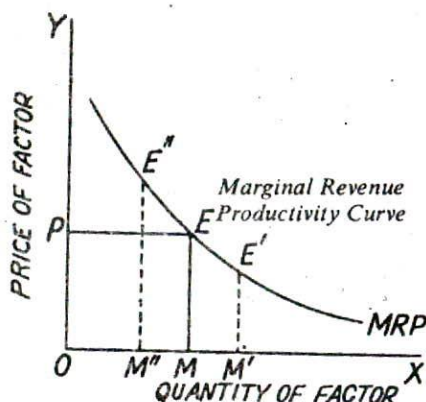
Marginal product may be **physical marginal product**, *i.e.*, the increase in output secured by an increase in the use of a factor. If we multiply increase in output by the prevailing price, we get the value of the marginal product or the revenue from the marginal product. This is called **marginal revenue product**, *i.e.*, the addition to the total revenue resulting from the use of one more unit of a factor of production. We use briefly marginal productivity in place of marginal revenue productivity.

We repeat that by the marginal productivity of a factor of production we mean the addition made to total production by the employment of the marginal unit, *i.e.*, the unit which the employer thinks just worthwhile employing. It may be distinguished from average productivity which is obtained by dividing the total product by the number of factor units employed.

At the margin of employment, the payment made to the factor concerned is just equal to the value of the addition made to the total production on account of the employment of the additional unit of a factor. If, for instance, the prevailing wage is less than the marginal productivity, then more labour will be employed. Competition among employers will raise the wage to the level of marginal productivity. If, on the other hand, the marginal productivity is less than the wage, the employers are losing and they will reduce their demand for labour. As a result, the wage will come down to the level of marginal productivity. In this way, by competition, wage tends to equal the marginal productivity. This applies also to the other factors of production and their reward.

This is illustrated in the Fig. 31.1 given on next page.

Here MRP is the marginal productivity curve. It is the demand curve for a factor of production. At the price OP, the quantity of the factor employed is OM, because at this level of employment, marginal productivity equals price. If he employed more, *i.e.*, OM', the marginal productivity E'M' will be less than the price he has to pay. He will thus be a loser. And if he employed less, *i.e.*, OM'', the marginal product E''M'' is more than the price, and he will add to his profit by employing more, till at OM when marginal revenue productivity equals price.



Marginal Productivity Theory

Fig. 31.1

Thus, the marginal productivity theory lays great emphasis on the close connection between the price of a factor and the price of the product that it produces. It seems that the factors of production earn that they get. Under perfect competition, a firm will be maximising its profit by equating marginal revenue product of each factor with its price. This indicates the extent to which the use of a factor of production will be pushed in production.

Also, factors of production tend to move from those uses in which their marginal productivity is low to those in which it is high. In this way, a given supply of a factor of production is distributed in such a way that its marginal productivity is equal in all the uses. Thus, we can say that the price of a factor of production is determined by its marginal productivity and this marginal productivity is the same in all its uses.

It also follows that in a competitive market assuming full knowledge of the various alternatives and of the right technology, paying factors of production according to their marginal productivity implies an efficient allocation of the community's productive resources. It gives a least-cost combination of the factors of production in a productive process.

Thus, in a position of competitive equilibrium: (i) the marginal productivity of a factor of production is the same in all its employments, (ii) the marginal productivity of a factor of production is measured by the price of the factor of production, and (iii) marginal productivities of various factors are proportional to their respective prices.

Hence, over the whole field of employment, therefore, each factor of production tends to be paid in proportion to its marginal productivity.

Thus, the distribution of National Dividend is not a scramble as the strikes or lock-outs make it appear to be. It is governed by a definite economic principle, viz., marginal productivity theory.

It should be noted that for any individual employer working under competition, the prices that he

has to pay for the factors of production are already determined. Since his demand for the factors of production is only an insignificant proportion of the total demand, his employing more or less of the factors does not appreciably affect their prices. What he does is to push the use of each of the factors to such a point as to make its marginal productivity equal to its price as already determined by the market forces.

Hence, the price of a factor of production is determined by the marginal productivity not of any particular employer but of employers in the aggregate.

Criticism of Marginal Productivity Theory

The theory, explained above, is true only under certain assumptions which make the theory unrealistic and render it inapplicable to actual conditions. It thus fails to explain the actual rewards earned by the factors of production. We give below the various grounds on which the marginal productivity theory is criticised:

(i) It assumes that all the units of a factor are homogeneous, so that any one unit is as good as any other. This is not actually the case. All labourers are not alike; they are of varying efficiency; nor are all the units of land similar. The capital equipment is also of different types. Thus we cannot talk of marginal productivity of a factor in general.

(ii) It is assumed that different factors are capable of being substituted for one another, so that, at the margin, it is possible to use a little more land, or a little more labour or capital, etc. If this substitution is not possible, marginal productivity of the various factors may remain unequal. Actually, it is not always possible to substitute labour for capital, and vice versa. Different factors of production are not close substitutes for one another.

(iii) It is also assumed that the amount of a factor used can be continuously varied, so that it is possible to apply a little more or a little less of the same factor. If this cannot be done, as is sometimes the case, the use of the factor cannot be pushed to the point at which its marginal productivity becomes equal to its price.

(iv) It is assumed that the factors of production are mobile as between various uses. We know that land lacks mobility; nor are labour and capital perfectly mobile. Human package is said to be the least portable. If a factor cannot be moved from one use or employment to another, its marginal productivity in the various employments may remain unequal.

(v) The theory is based on the law of diminishing returns as applied to the organisation of a business. This means that, other things being equal, a disproportionate increase in the supply of one factor increases total production at a diminishing rate. We

know, however, that in manufacturing industries, the operation of the law of diminishing returns is held in check.

It is under these assumptions that the reward of each of the four factors of production, *viz.*, rent of land, interest on capital, wages of labour and profits of enterprise, tends to equal the value of its marginal net product. These assumptions do not always hold good. The theory has been criticised on other grounds too.

(vi) The marginal productivity theory has been criticised by Keynes thus: One implication of this theory is that if employment is to be increased, wages should be lowered, so that more labour will be employed to make marginal productivity equal to the wage. This argument is fallacious. This may be true in the case of an individual industry or a firm. But it cannot apply to the economy as a whole. The total employment in a country depends on effective or aggregate demand, and not on the level of wages.

(vii) According to marginal productivity theory, marginal productivity determines the reward of a factor of production. In other words, the two are independent. This is not really the case. One affects the other. The marginal productivity or efficiency of a factor also depends on the reward it gets.

(viii) One common criticism is that a product is the result of the co-operative efforts of all the factors of production and that it is impossible to separate the share contributed by each. This criticism advanced by Taussig and Davenport is obviously based on a misreading of the concept of marginal productivity. As we have already explained, marginal productivity is not the net product solely due to the marginal factor. We merely impute that product to the factor on the margin of use. It is the net addition made to the total production by the employment of this additional factor or deduction caused in it if this factor were withdrawn.

(ix) Another attack is made by Hobson. It is held that if any particular factor unit is withdrawn, the whole business will be so disorganized that the loss to production will be much more than the productivity of the unit withdrawn. The criticism is also due to the wrong application of the theory. In this criticism, we are thinking of a small business organisation and large units of factors. If we conceive of a large business and small units of factors, it will be clear that withdrawing a unit at the margin will not appreciably affect the productivity of other factors.

(x) Then there is the view, according to which the sum of the marginal net products of all the factors will be less than the total product, the surplus being due not to any particular factor but to their co-operation. Wicksteed has answered this criticism. He assumes that the increase of all the factors will increase the quantity of the product in the same proportion. But this assumption, which implies that

the industry obeys the law of constant returns, is not always valid and introduces certain difficulties.

(xi) Another serious difficulty, which relates to the measurement of the marginal net product, has been pointed out by Joan Robinson,² Pigou and J. R. Hicks. It is argued that when there are economies of large-scale production, the marginal productivity of a unit of a factor to a particular firm will be considerably less than that to the industry as a whole. This is so because, when an additional unit is made available to an industry, it brings about a greater division of labour. But, when the industry has adjusted itself to the new supply, it is quite possible that the marginal productivity of a factor to an individual firm is less than that to the industry as a whole. This is so because its withdrawal will mean a much greater loss to the industry than to an individual firm. In such industries, therefore, marginal productivity of a factor is indeterminate.

(xii) Finally, to call marginal productivity theory a theory of distribution is a misleading statement. At the most it analyses the factors affecting the demand for a factor of production. The supply is taken as fixed. It is objected, however, that in actual practice, the reward enjoyed by a factor does affect its supply. The theory only approaches the problem from the side of demand. It is thus a one-sided explanation.

Conclusion. It should be remembered that the theory is valid only under the assumption of perfect competition. In real life, since competition is not perfect, actual rewards paid to the factors of production do not conform to their relative marginal productivities.

Moreover, this explanation of the determination of the shares of the various factors of production, in a capitalistic economy, should not be regarded as a justification, from the ethical point of view, of the system of distribution under such a system. The theory is essentially positive and not normative. It does not say that the reward of a factor according to marginal productivity is a just reward.

As Samuelson remarks, "It (marginal productivity theory) is not a theory that explains wages, rents, or interests; on the contrary, it simply explains how factors of production are hired by the firm, once their prices are known."

Commenting on the theory of distribution, Fraser remarks, "No economist would claim that the theory is as yet complete, even as a purely academic structure or framework. It has the defects of its qualities. Being simple and self-consistent it is abstract and impersonal. It is guilty of sins both of

2. *The Economics of Imperfect Competition*, p. 327.
3. *Economics of Welfare*.
4. *The Theory of Wages*.
5. *Economics*, p. 198, 526, 528.

omission and commission, its postulates are unduly rigid and narrow."⁶

FIRM'S EQUILIBRIUM IN FACTOR EMPLOYMENT

We have already seen in chapter 19 how a firm achieves equilibrium in regard to a factor combination. A firm will be in equilibrium in this respect when it is combining the various factors of production in such a way that the marginal rate of technical substitution between any two factors is equal to their price ratio. In this way, the producer will be using the most economical combination or the least-cost factor combination. Thus, the firm will be in equilibrium in regard to the employment of factors of production when $\frac{MP_a}{P_a} = \frac{MP_b}{P_b} = \dots = \frac{MP_n}{P_n}$.

This is as regards using the most economical combination of the factors of production.

But here the problem is different. Instead of finding the optimum factor proportion, the firm now wants to find out the absolute amounts of the factors of production which it should use or employ in order to be in equilibrium. In this connection, we may say that the various market situations are possible, viz., (i) perfect competition in both the factor market and the product market; (ii) perfect competition in the factor market but imperfect competition in the product market; (iii) monopsony in the factor market but perfect competition in the product market, and (iv) monopsony in the factor market and monopoly in the product market.

Since the entrepreneur is supposed to be rational, he will compare revenue and cost of utilizing an extra unit of a factor. If he finds that the marginal revenue (*i.e.* the additional income from the use of an additional unit of a factor) is greater than the cost of hiring it, he will use more of that factor. He will stop employing extra factor units when the marginal revenue product of the factor is equal to its marginal factor cost (*i.e.* $MRP = MFC$).

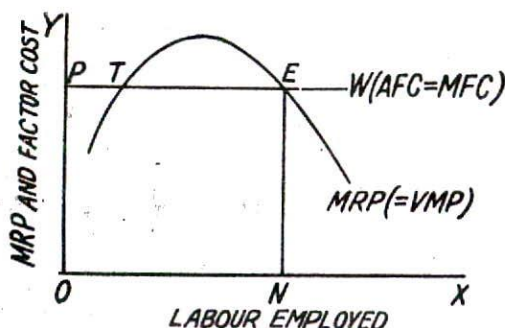
We shall now discuss the employment of a factor of production (a) under perfect competition, and (b) under imperfect competition.

Equilibrium Under Perfect Competition In The Factor Market

When there is perfect competition, an individual firm cannot influence the price of a factor, by buying more or less of it, because its own demand for the factor constitutes a small fraction of the total demand. Consequently, a firm has to accept the price of a factor, say, labour, prevailing in the market (*i.e.* wage) as given. Similarly, no individual supplier of a factor supplies an appreciable quantity thereof so as to be able to affect its price in the market. In this case, therefore, marginal factor cost

(MFC) and average factor cost (AFC) will tend to be equal and their curves will be the same or they will coincide, and will be a horizontal straight line as in the case of product prices. In the commodity market also, under perfect competition, marginal revenue (MR) and average revenue (AR) are equal and they are represented by a horizontal straight line.

The following diagram (Fig. 31.2) represents this state.



Equilibrium Under Perfect Competition
Fig. 31.2

MRP is the marginal revenue productivity curve and the same curve represents the value of marginal productivity (VMP), because we assume perfect competition in the product market. PW horizontal straight line represents both marginal factor cost MFC and average factor cost AFC. This firm will obtain maximum profit where marginal revenue productivity ($MRP = VMP$) is equal to the marginal factor cost (MFC). They are equal at the point E where the two curves intersect. At the point of equilibrium, the firm will employ ON quantity of the factor and pay OP (= EN) as the price or wage. This is the position of maximum profit.

In the diagram, it will be seen that these curves intersect also at an earlier point, T. But this cannot be the position of maximum profit, since at this point marginal revenue productivity (MRP) is still rising, whereas the factor cost is the same. Hence, the producer will not stop at T; he will continue to employ more labour till he reaches E where MRP is falling. It is thus clear that E is the real point of equilibrium. We can, therefore, lay down a principle that a purely competitive producer maximises short-run profits (or minimises losses) only when he employs each productive resource or a factor of production up to the point where the value of its marginal product equals the market price of the factor.

Thus, under perfect competition, since marginal factor cost (MFC) is equal to average factor cost (AFC), at equilibrium marginal revenue productivity (MRP) is equal to marginal factor cost, it is also

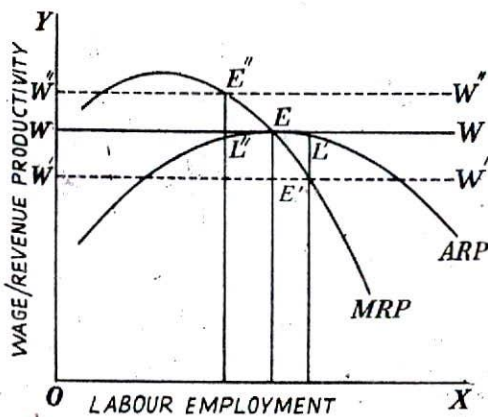
6. *Economic Thought and Language*, 1941 p. 354.

equal to average factor cost. This means that for a firm to be in equilibrium in a competitive factor market, two conditions must be satisfied, viz.

- (1) $MRP = MFC$
- (2) MRP curve must cut the MFC curve from above.

From the above diagram (Fig. 31.2), we cannot say whether the firm makes a profit or incurs a loss and if there is profit how much? This we shall be able to know from the following diagram (Fig. 31.3)

In this diagram ARP is the average revenue productivity curve and MRP is the marginal revenue productivity curve. MRP curve cuts the ARP curve at its highest point E. At the OW' wage level, the



Profit or Loss
Fig. 31.3

equilibrium will be at E' . Here there is extra profit per unit $E'L'$. This is however, a short-run situation.

In the long run, new firms will enter the industry because super-normal profit is being made. When the new firms enter, the demand for labour will increase; hence the wage level will go up to OW , where E is the point of equilibrium. Here supernormal profit has disappeared, because wage is equal to average revenue productivity. If on the other hand, the wage level is OW' , the equilibrium will be at E'' . In this case, the wage is higher than the average revenue productivity. Hence; there is loss instead of profit. In the long run, therefore, some firms will leave the industry; the demand for labour will decrease till the wage-level comes down to OW and the loss will disappear. The equilibrium at E will be restored and the wage will be OW . Here the wage is equal both to ARP and MRP.

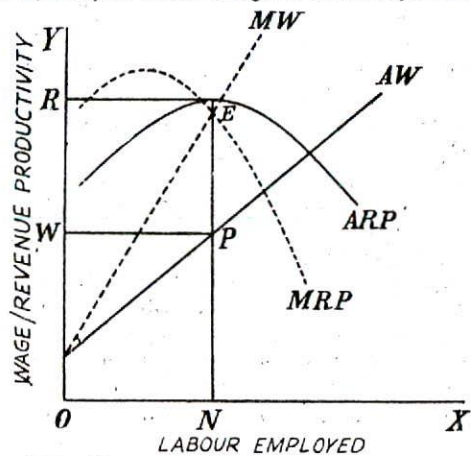
Thus, we conclude that, **under competitive conditions, wages tend in the long run to equal marginal revenue productivity as well as average revenue productivity.**

Equilibrium Under Imperfect Competition in the Factor Market

We have analysed above conditions of equilibrium in a factor market under perfect competition. But we know that in the real world, perfect competition does not prevail. The real world is of imperfect competition. We shall now see how a firm reaches a state of equilibrium under imperfect competition in respect of the quantity of the factor employed and the price paid for it.

We shall take here an extreme case of imperfect competition, i.e., monopsony, in which there is only a sole purchaser of a factor, i.e., a sole employer and no one to compete with him, i.e., there is none else who needs that factor. Unlike under perfect competition, here the employer has control over the price he pays for the factor, say, wage, i.e., he can raise it or lower it as it suits him, for there is no fear of competition from any other employer. If he needs more labour, he will have to pay a higher wage. Hence, the average wage curve AW will rise upward from left to right and the marginal wage curve MW will be above it.

In the diagram (Fig. 31.4). ARP is the average revenue productivity curve and MRP the marginal revenue productivity curve. In this case, the firm will be in equilibrium position where marginal wage (MW) is equal to the marginal revenue productivity



Equilibrium Under Imperfect Competition
in the Factor Market
Fig. 31.4

(MRP). This is at the point E, because it is here that MW and MRP curves intersect. In this equilibrium position, we see that the average wage $OW (= NP)$ is less than marginal revenue productivity MRP which is EN in this diagram. This means that the labour gives to the employer more than the wage the employer pays him. In other words, he is exploiting labour. This is known as **monopsonistic exploitation**. It is natural that, when there is no competition

and the employer enjoys a monopoly of purchasing a factor, he must pay as little as possible. Hence, under imperfect competition, labour will usually be exploited at the equilibrium position.

Conditions of Monopsony

A monopsonistic situation, inter alia, may arise in the following ways:—

(i) Units of a factory may become specialised to a particular use. Having become thus specific, they have no use for other purposes. For example, a labour of specialised skill may be developed to meet the requirements of a particular firm and no other firm has any use for it; or a big automobile manufacturer may be the sole purchaser of certain parts. In the film line, certain artists may be bound by agreement to work for a certain produce only and cannot work for any other during the currency of the agreement.

(ii) Monopsony may also arise owing to the immobility of a productive resource. For instance, a certain body of workers may be attached to a certain firm on account of emotional ties or caste considerations. Ignorance or dread of the unknown may keep them tied to a particular place or employer. They may be too poor to seek jobs elsewhere. Seniority, provident fund or pension rights earned already do not permit the employees to leave. There is also geographical immobility.

Under perfect competition, each firm maximises its profits by employing larger and larger quantities of the factor till the marginal revenue product of the factor is equal to the price paid for it. But a monopsonist maximises his profit by employing smaller quantity than would equalise marginal revenue to its price per unit. That is, the marginal product exceeds the price and the factor cost too exceeds the price. Hence, factor units are paid less than what any one of them contributes to the total receipts of the firm. In other words, the monopsonist restricts the quantity of the factor employed and keeps down the price that he pays. This is how the factor supplies are exploited by a monopsonist.

The Remedies. Is there any remedy to prevent this exploitation? Broadly speaking, two remedies are possible:—

(i) The factor suppliers may organise themselves and they or the government may fix minimum prices for the supply of the factor exploited. The firms using these resources must be made to pay these minimum prices. The minimum price will not only eliminate exploitation but, if fixed at a reasonable level, may also increase the quantity of the factor used. If it is fixed too high, it will create unemployment of the resource. Fixing a correct price may not

be so easy. Hence, collective bargaining on a firm-to-firm basis may be more appropriate.

(ii) Measures may be adopted to increase the mobility of the factor. In case of labour, where exploitation is general, steps can be taken to reduce geographical mobility by spreading information, improving transport or by subsidising migration; horizontal mobility may be increased by facilitating transfer from one industry to another and vertical mobility by imparting technical and general education.

MODERN THEORY OF DISTRIBUTION

The marginal productivity theory, which we have discussed above only tells us how many workers will an employer engage at a given wage-level in order to maximise his profit. It does not tell us how that wage-level is determined. We also saw that the marginal productivity theory approaches the problem of the determination of the reward of a factor of production from the side of demand only. It ignores the supply side. Hence, the marginal productivity theory is not an adequate explanation of the determination of the factor prices.

The modern theory of factor prices which provides a satisfactory explanation of factor prices is the **Demand and Supply Theory**. Just as the price of a commodity is determined by the demand for, and supply of, a commodity, similarly the price of a productive service also is determined by demand for, and supply of, that particular factor. We shall now analyse these two aspects, viz., demand and supply.

Demand Side

About the factor demand, we should remember two things:

- (a) Demand for a factor is not a direct demand but a derived demand and
- (b) the demand for it is a jointly interdependent demand.

The demand for a factor of production is not a direct demand; it is an indirect or derived demand. It is derived from the demand for the products that the factor produces. For instance, labour does not satisfy our wants directly. We want labour for the sake of the goods that it produces. The demand for a factor of production is, therefore, ultimately determined by the desires and preferences of the consumers for the final product. It follows, therefore, that if the demand for such goods increases, the demand for the factors which help to produce these goods will also increase. Also, if the demand for goods is elastic or inelastic, the demand for factors too will be elastic or inelastic. The higher the price of the commodity produced the higher the price of the productive service used.

No factor of production works alone. It is generally required to work in combination with others. The demand for a factor of production, therefore, depends on the quantity and the prices of the other factors required in the process. Generally speaking, the demand price for a given quantity of a factor of production will be higher, the greater the quantities of co-operating productive services. The factors of production are not only complementary to one another, they also sometimes compete with one another and are, therefore, substitutes for one another. Hence, the price of one factor affects the price of another. Thus, their cross elasticities are as important as their own elasticities.

If more of a factor of production is currently employed, the marginal productivity of the factor will fall and the lower will be the demand price for the unit of the productive service. This is another rule connected with the demand for a factor of production.

The demand price of a factor also depends on the value of the finished product in the production of which the factor is used. The demand price will generally be greater, the more valuable is the finished product in which the factor is used.

Also, the more productive the factor, the higher will be the demand price of a given quantity of the factor.

The demand for a productive service also depends upon technological changes. Improvement in technology makes a factor more productive and thus increases its demand.

On the demand side, we have also to consider the elasticity of a factor demand. In this connection, we may note that the elasticity of demand for a factor varies directly with the elasticity of demand for the final product. Also, the demand for a factor will be more inelastic, the smaller the cost of a given factor in the total cost of the final product. But the demand for a factor will be more elastic, the easier it is to substitute some other factor for it.

To Sum Up. The demand for a factor of production depends on (a) the demand for the commodity produced, (b) the price of the factor concerned, (c) the quantity and the prices of the co-operating factors, (d) the quantity and the quality of the factor currently employed and (e) changes in technology.

These are a few points connected with the demand for a productive service. We know that the demand curve of the industry is the sum-total of the demand curves of the various firms in the industry. By a similar summing up, we can have the demand curve of all the industries using a particular productive service.

The demand of the employer for a factor depends on its marginal revenue productivity (in short marginal productivity) and the quantity of the factor that a firm will employ will depend on the prevailing wage-level. That is, more labour will be employed if wages are low and less if the wages are higher.

Fig. 31.5 (a) is the diagram representing the position of a firm regarding the employment of a factor say, labour. When the wage is OW , the firm is in equilibrium at the point E and the demand for the factor is ON ; similarly, at OW' wages the demand is ON' and at OW'' the demand is ON'' . MRP (marginal revenue productive curve) is the demand curve for a factor of production by an individual firm.

But for the determination of price it is not the demand of the individual firm that matters. What matters is the total demand, i.e., the sum-total of the demands of all firms in the industry. The total demand curve is derived by the lateral summation of the marginal revenue productivity curves of all the firms in the industry. This curve DD is shown in Fig. 31.5 (b) It may be carefully noted that Y-axes in both curves are drawn to the same scale, but X-axes are drawn on different scales.

We have supposed that there are 100 firms in the industry. At OW wage, the demand of the individual

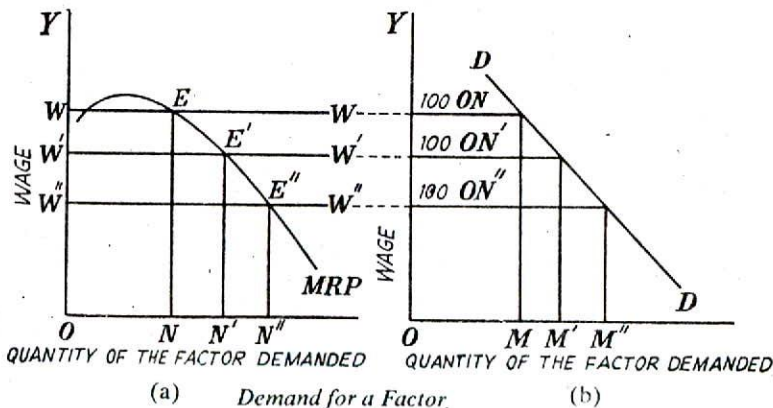


Fig. 31.5

firm is ON , but the demand of the whole industry at the same wages is OM which is equal to $100 ON$ (because the number of firms in the industry is 100). In the same manner, at OW' the demand of the firm is ON' but the demand of the entire industry is OM' , which is equal to $100 ON'$ and similarly at OW'' , the demand of the firm is ON'' and that of industry OM'' , which is equal to $100 ON''$.

It can be seen that the demand curve DD slopes downward to the right. The reason is that MRP curve whose summation is represented by DD also slopes down similarly to the right in the relevant portion. This means that according to the law of diminishing marginal productivity, the more a factor is employed, the lower is the marginal productivity.

Now we turn to the supply side.

Supply Side

The supply curve of a factor of production depends on the various conditions governing its supply. Take the case of labour—a very important productive service. The supply of labour depends on the size and composition of the population, its occupational and geographical distribution, labour efficiency, cost of education and training, costs of movement, the expected income, relative preference for work and leisure, and so on. In this manner, by considering all the relevant factors, it is possible to construct the supply curve of a productive service.

We may, however, add that the supply is a bit complicated affair. We generally say that the supply of land is limited. But the fact is that although for the community as a whole, land is limited, but for a particular firm or an industry, its supply is not limited. The supply can be increased, if higher rent is offered.

In the case of commodities, we see that generally an increase in price brings forth larger supplies. This, however, does not necessarily hold good in the case of the factors of production. It may happen in some cases that if wages go up, labour may be able to satisfy their needs by working for a short time. They may prefer leisure to work. In this case, when the price of a factor (or its remuneration) is increased, the supply is contracted. This peculiarity will be represented by a backward sloping curve after a stage.

Also, the supply of labour does not merely depend on economic factors; many non-economic considerations also enter.

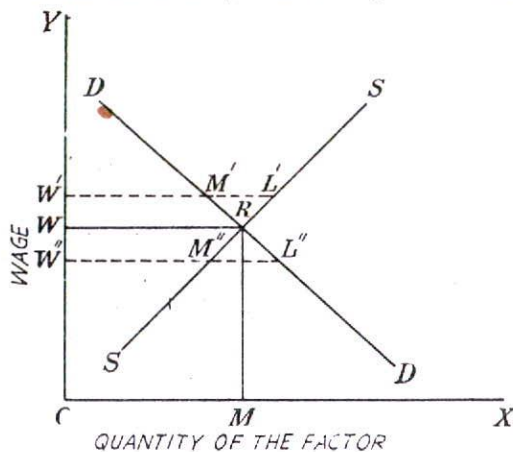
All the same, we can lay down a general rule that if the price of a factor increases, its supply will also generally increase, and vice versa. Hence, the supply curve of a factor rises from left to right upwards. This is shown in Fig. 31.6.

Equilibrium of Demand and Supply

Now we have worked our way to the demand curve and the supply curve of a factor of production.

Both these curves are needed for the determination of the price of a productive service. That price will tend to prevail in the factor market at which the demand and supply are in equilibrium. This equilibrium is at the point of intersection of the demand and supply curves.

In the diagram 31.6, they intersect at the point R and the price of the factor will be $OW (=MR)$. At OW' wage, demand $W'M'$ is less than the supply $W'L'$. Hence, the competition among the sellers of



Determination of a Factor Price
Fig. 31.6

the service will tend to bring down the price to OW . On the other hand, at OW'' price, the demand $W''L''$ is greater than the supply $W''M''$; hence price will tend to go up to OW at which the demand and supply will be equal.

This is how the price of a factor of production in the factor market is determined by the interaction of the forces of demand and supply relating to that factor of production. The modern economists consider this theory as a satisfactory theory of distribution.

TRANSFER EARNINGS

Modern economists make use of the concept of transfer earnings, in order to explain the remuneration of the factors of production. Take the case of a worker. Suppose in his present occupation he is earning Rs. 250. In case, he has to give up that occupation, suppose the next best alternative employment fetches him Rs. 200. This is his transfer earnings, i.e., what he will get on transfer to the next best alternative employment. This means that in his present employment, he must at least get Rs. 200 (his transfer earnings), otherwise he will give it up and move to his next best alternative employment where he can earn Rs. 200. In other words, the transfer earning is the minimum price which must be paid to a factor of production to induce it to stay in the industry or which is necessary to retain it

there. If this price is not forthcoming, the factor unit will transfer itself to its next best alternative use.

In Benham's words, "The amount of money which any particular unit could earn in its best paid alternative use is sometimes called its transfer earnings." To quote Mrs. Robinson, "The price which is necessary to retain a given unit of a factor in a certain industry may be called its transfer earnings or transfer price, since a reduction of the payment made for it below this price would cause it to be transferred elsewhere; and any particular unit of a factor may be said to be at the margin of transference, or to be a marginal unit, if the earnings, which it receives in the industry, where it is employed, are only just sufficient to prevent it from transferring itself to some other use."

Take the case of land. Suppose a land used for the cultivation of cotton earns a little more than when it is used for the growing of sugar-cane. If the profit in cotton cultivation falls a little or that in cane cultivation rises a little, the cotton land will be converted into cane land. Thus, fluctuations in prices of different crops may make the cultivation of one crop more profitable than that of another. There will then be a tendency for the land to be transferred from one use to another in search for higher earnings. If an agent is earning more than what it would earn when transferred to the best-paid alternative, the surplus is rent.

As Mrs. Robinson says, "Each individual unit of each factor will be fitted into the place where its earnings will be greatest; when its earnings in that use fall it will retreat to its next most profitable use and, if there is an appreciable difference between its actual earnings and its earnings in the next most profitable use to which it could be put, it will be receiving rent. If each productive unit is like its neighbour, both in respect to their efficiencies in the industry employing them and in respect to their efficiencies in alternative uses, there will be no rent."

The concept of transfer earnings has a great significance in the theory of value. The supply of a factor for a particular industry depends on its transfer earnings. The industry must pay a price for the use of a factor which is at least equal to its transfer earnings. This is how transfer earning affects the supply of a factor which is one of the two factors affecting the price of a factor. In this way, the concept of transfer earnings has an intimate bearing on factor pricing.

EQUITY IN INCOME DISTRIBUTION

We have studied above the general theory of

7. *The Economics of Imperfect Competition*, 1945, p. 104.

8. *The Economics of Imperfect Competition*, 1945, p. 107.

income distribution. Two questions arise: (a) Is the distribution equitable, fair and just? (b) Why is there inequality of incomes? The first question is ethical and involves value judgment and does not admit of any objective measurement. Also, no two opinions agree on the criterion of fairness or justice. It is futile, therefore, for economic theory to attempt to answer this question.

As for inequalities of income, they arise from two basic sources: (a) Differences in quality and quantity of resources owned by individuals and (b) differences in prices paid for the use of the productive resources.

The various causes responsible for the inequalities of income may be spelt out below:

(i) So far as differences in labour earnings are concerned, horizontal differences are due to differences in demand and supply conditions for a particular type of labour. Quantitative differences in the amount of work done and qualitative differences in the abilities and skills of the workers account for differences in earnings; vertical immobility also gives rise to differences in earnings of certain professions. Vertical mobility depends on the opportunities for training and for going up and it is also accounted for differences in social inheritance.

(ii) Differences in earnings also arise from differences in ownership of capital, e.g., business assets, land, factories and property of all kinds. This may, in turn, be due to inequalities of inheritance or gifts received, qualities of thrift and saving propensity and opportunities for investment or windfalls, chance or luck.

(iii) Differences in income may also arise from interference with the price mechanism. The owners of resources may manipulate the prices of the resources they own and the producers, the prices of their products, e.g., farmers getting high procurement prices fixed up, cane-growers getting minimum prices of sugar-cane. In an era of price control, all manufacturers endeavour to get minimum prices fixed for their products. How far the owners of resources are able to raise their distributive share will depend upon the elasticity of demand for the factor concerned. By restricting the resource supplies, the resource prices, and hence their income, can be raised. The existence of a monopoly or monopsony results in the manipulation of factor incomes.

Remedies

There are two broad remedies:

(a) **Via Administered Prices.** Prices may be fixed or manipulated in favour of the proper sections of society.

(b) **Via Resources Redistribution.** Labour resources can be redistributed by increasing vertical

mobility, by providing greater equality in educational and training opportunities, greater economic opportunities by reducing inequalities of capital ownership and breaking down barriers to entry into higher jobs and professions. Appropriate fiscal and taxation measures can be adopted not only to reduce economic inequalities directly by heavy taxation of the top classes but also to provide greater equality in capital accumulation and holdings of capital resources. Provision of more comprehensive social security and social welfare programmes will have the same effect.

Relation between Production, Value and Distribution Theories

According to the traditional division of Economics into different parts, value, production and distribution are considered its principal parts, but there exists an intimate relation between them so as to give Economics a unified look. Let us now consider this relationship.

Relation Between Value and Distribution Theories

Value means the price of a product and distribution determines the share of the factors of production in the national output. In the general theory of distribution, we say that a factor of production is remunerated according to the value of its marginal product under conditions of perfect competition. That is, the reward of a factor of production is determined by its marginal productivity.

But the value of the marginal product of a factor depends on the value of the product that it produces. In other words, the value of the marginal product of a factor is equal to the marginal physical product of the factor multiplied by the price of the product that it helps to produce. Obviously, higher the price of the product, the higher will be the value of the marginal product of the factor that produces it, and hence, greater will be its reward or the income. In case the value of the product goes up, the demand for the factor required to produce that product will rise and as a result the price or the income of the factor will go up. In this way, the demand for the factor is derived from the demand for the product.

Thus, the value of the products and the distribution of the income of the factors of production are very closely related to each other.

Relation Between Production and Distribution Theories

We have seen above how value and distribution (that is the share of the factors of production in national income) are intimately related to each other. Let us now see how production is related to distribution. Theory of distribution, as we have said above, determines the share that goes to a factor of

production out of the national output. It can be easily seen that the greater the share of a factor of production, the better will be his standard of living and greater his efficiency in production. The more efficient the factors of production, the higher will be the national income. In this way, the character of distribution in a country determines its productive efficiency, which in turn determines the level of output in the country.

Thus, we find that the theories of production and distribution are very closely related to each other.

We have already explained in Chapter 17 how a production function can be expressed. The formula is $X = F(a, b, c, d, \text{etc.} \dots)$ Here X is the output of a commodity and a, b, c, d, etc. are the various productive resources which go into the making of the commodity and F means the function of *i.e.* it varies with. From this equation, it is clear that production varies with the quantities of the various factors of production used. When we employ the factors of production, we do not merely consider their number but also their efficiency which depends on whether distribution of the national dividend is fair or uneven.

Marginal product of factors of production, say labour and capital, plays a vital role in the theory of distribution (*i.e.* marginal productivity theory). According to this theory, wages are equal to the marginal product of labour and interest equal to the marginal product of capital. The equality of factor rewards with their marginal productivities is based upon the assumption that the entrepreneurs aim at maximizing their profits. They combine labour and capital in a proportion which is most advantageous to them. The most advantageous combination will be where the marginal productivities of each factor are equalized. It is said that the forces are constantly at work which bring about an equality between the remuneration of the factors of production and their marginal productivities. This clearly brings out a synthesis between production theory and distribution theory.

Diminishing marginal returns, which is a special feature of production function, has great relevance for the theory of distribution. We know that the entrepreneur continues to employ labour or capital till its marginal product comes down to the level of wage or interest. Had the marginal product increased or remained constant instead of diminishing, the equality between the factor rewards and the marginal product could not have been achieved and the theory of distribution would have been nowhere.

The possibility of substitution between the factors of production is another feature of production function which has a bearing on the theory of distribution. If it were not possible to substitute one factor for another and if the factors were to be used in rigidly fixed proportions, their marginal productivities would have been zero, and zero marginal

productivity does not furnish any basis for the theory of distribution. It is only the possibility of substitution between factors which makes it possible to build a theory of distribution based on the concept of marginal productivity.

It was Prof. Hicks who introduced the concept of elasticity of substitution between factors and pointed out its importance for determining the distributive shares of the various factors of production. If there were a change in the ratio of the prices of factors of production, the cheaper factor would be substituted for the factor whose price is relatively higher. If for instance, wages rise relatively to interest, the entrepreneur will substitute capital for labour. This would naturally reduce the distributive share of labour and raise that of capital.

It may also be pointed that the possibility of substitution between the factors discourages unions from unduly pressing for higher wages. Any unreasonable behaviour on the part of unions will simply compel the entrepreneur to substitute capital for labour. The result would be that volume of employment will diminish and the labour itself would suffer.

Thus, we find that there is a very intimate relation not only between the theories of distribution and production but also of between value and distribution.

Product Exhaustion Problem or Adding-Up Problem

The marginal productivity theory of distribution

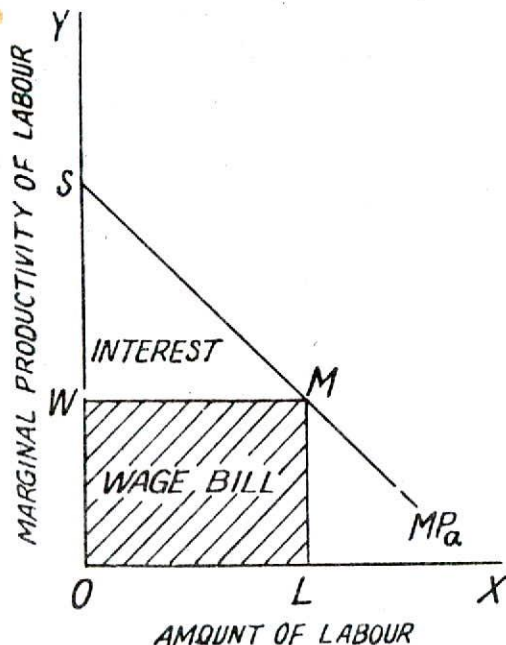


Fig. 31.7

states that each factor of production is paid remuneration equal to its marginal product. It has been urged out that if that were so, the total product would be just exhausted without leaving any surplus or deficit. This problem is called the adding-up problem.

The following figures numbering 31.7 and 31.8 illustrate the product exhaustion problem.

We assume that there are only two factors, labour and capital, used in production. Let A stand for Labour and B for Capital. The marginal product of a factor can be found by varying its quantity and keeping the other factor constant. The reward of a variable factor is equal to its marginal product when a certain quantity of it is used. Then the reward for the fixed factor can be shown as surplus (or residual income) of the total product over the marginally determined reward of the variable factor. In figure 31.7, labour (*i.e.* A has been treated as a variable factor and is shown on the X-axis whereas capital has been taken as a fixed factor. When OL is the equilibrium amount of the labour employed, its marginal product is LM and the wage rate determined is OW. The total wage bill is OLMW. The total product is the whole area under the marginal productivity curve *i.e.* OSML. After paying the total wage bill *i.e.* OLMW, the residual income SWM will go to capital as interest. In this way, when we have paid the total wage bill OLMW and interest SWM, nothing is left, *i.e.* the total product is exhausted.

In order to show that the total product is exhaust-

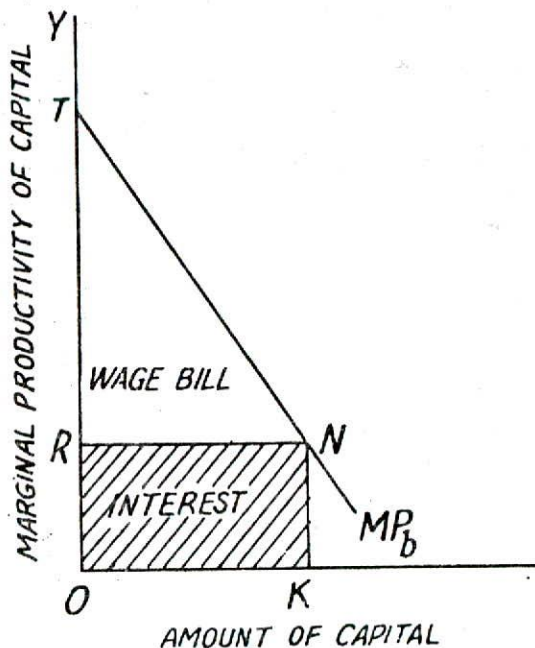


Fig. 31.8

ed, we may now take capital as the fixed factor and labour as the variable factor. This has been shown in figure 31.8. In this figure, capital gets ORNK as the marginally determined interest out of the total product OTNK. Labour receives the residue in the form of the wage bill *i.e.* RTN. Here again the total product is just exhausted by the rewards paid to labour and capital, leaving no surplus or deficit.

We have now to prove that the area OKNR in figure 31.8 is equal to the area WMS in figure 31.7 and area RNT in figure 31.8 is equal to OLMW in figure 31.7. Only in this way we can show that the payments made in accordance with marginal productivity to both labour and capital exactly exhaust the total product. We can thus show that the marginally determined reward of a factor is equal to the income of the factor determined as residual. In this way, we can illustrate the problem of product exhaustion, But illustration is not a proof.

Professor Wicksteed was one of the first economists who posed this problem of product exhaustion and also furnished a solution. This he did by applying a mathematical proposition known as the Euler's Theorem. He proved that the total product will be just exhausted if all the factors of production are paid according to their marginal product. Wicksteed's solution was criticised by economists like Walras, Edgeworth and Paréto. These writers say that production function is not homogeneous of the first degree, *i.e.*, returns to scales are not constant in the actual world.

The critics have pointed out that production function is such that it yields a U-shaped long-run average cost curve. This implies that increasing returns to scale occur upto a point and then there are diminishing returns to scale. If a firm is operating under increasing returns to scale, then payment to factors of production according to their marginal product would make the total factor rewards greater than the total product. If, on the other hand, the firm is working under diminishing returns to scale and all factors are paid equal to their marginal product, the total factor rewards would not fully exhaust the total product and there will be a surplus. It is pointed out, therefore, that Euler's Theorem does not apply and the adding-up problem does not hold good either when there are increasing returns to scale or decreasing returns to scale.

It is also pointed out as a defect in Wicksteed's solution that when there is a constant return to scale, the long-run average cost curve of the firm is a horizontal straight line which is incompatible with the perfect competition. But competition is essential to the marginal productivity theory and so to Wicksteed's solution.

Economists like Wicksell and Walras advanced more satisfactory solution after Wicksteed to the problem that marginally determined rewards would just exhaust the total product. They pointed out that

in the long-run, under perfect competition, the firm is in equilibrium and the minimum point of the long-run average cost curve is where the returns to scale are momentarily constant *i.e.*, constant within the range of a small variation of output. Thus, it was pointed out that the condition required for the marginally determined rewards to exhaust the total product (*i.e.*, the operation of constant return to scale) was fulfilled at the minimum point of the long-run average cost curve, where a firm operating under perfect competition is in long-run equilibrium. Hence, in the case of perfectly long-run equilibrium, the marginally determined rewards of the factors would exactly exhaust the total product.

From the above discussion, two solutions of the problems of product exhaustion emerge: (a) Wicksteed's solution which assumes the operation of the constant return to scale and (b) the solution by Wicksell and Walras which assumes that the firms operated at the lowest point of the long-run average cost curve.

IMPACT OF TECHNOLOGICAL PROGRESS ON RELATIVE FACTOR SHARES

Recent technical advance in agriculture, industry, trade and transport has raised the question of its effect on the relative remuneration of the factors of production, in particular of labour and capital. Technological progress implies the introduction of innovations *i.e.* new processes, new products or more economical methods of production.

The analysis of relative factor shares requires a detailed examination of the nature of technical progress as well as the underlying production function. In this analysis two basic concepts are involved *viz.*, the elasticity of substitution and the linear homogeneous production function.

Elasticity of substitution (a concept developed by J. R. Hicks in his book 'Theory of Wages') is measured by the ratio of the proportional change in a factor combination to a proportional change in the marginal rate of substitution between factors represented by the ratio of the marginal products of the two factors. If the production functions are of Cobb-Douglas type (*i.e.* the unitary substitution elasticity) the relative shares of the factors will remain constant. If elasticity of substitution between labour and capital is less than one, the share of labour will go up relatively to the share of capital, and if it is greater than one opposite will be the result.

Now take the linear homogeneous production function called the Cobb-Douglas production function (See P. 130). This production function represents constant returns to scale.

Having studied the general theory of distribution, we are now in a position to take up the study of the determination of the share of the factors of production separately. We shall first take up wages followed by rent, interest and profit.

Meaning of Wages

The term 'wages' means payments made for the services of labour. "A wage may be defined," says Benham, "as a sum of money paid under contract by an employer to a worker for services rendered." A wage payment is essentially a price paid for a particular commodity, viz., labour services.

Is Labour a Commodity?

We have said that wage is the price of labour as if labour is a commodity. There is no doubt that labour has peculiarities of its own which differentiate it from other commodities (see next page) and labour protests against its being described as a commodity, yet it is true that from the aspect with which the economist is concerned, labour is regarded as a commodity. Like a commodity it is bought and sold. Labour participates in an exchange relationship, i.e., the worker gives his time or labour and, in exchange, gets money including some fringe benefits, if any.

Nominal Wages vs. Real Wages

According to the classical wage theory, labour supply was considered a function of real wages. But according to Keynes, the workers acted irrationally and generally bargained for money wages and they sharply reacted against any cut in money wages. That is, a rise in prices does not offend labour so much as a cut in the money wage. The money wage has also been called nominal wage.

The distinction between nominal wages and real wages may be clearly understood. Nominal wages

are wages paid or received in terms of money. But money wages alone may not give us a correct idea of what a worker **really** earns from his work. In order to ascertain real wages, which determine the standard of living of a person, the following factors have to be taken into consideration:

(a) **Purchasing Power of Money.** When comparing wages at different places and at different times, the changes in the purchasing power of money must be taken into account. The purchasing power of money varies inversely with the price level, i.e., higher the prices, the lower the purchasing power of money, and vice versa. A part of the high wages in England and America may be due to higher prices prevailing in those countries. Three hundred rupees in a village in India may provide a much more comfortable life than a similar amount in a town or vice versa, according to circumstances and tastes of the person concerned. One hundred rupees in 1950 had much greater purchasing power than in 1976. Even an increase of money wages thus leaves real wages at a lower level in 1976 as compared with 1950.

It is generally supposed that the prices rise faster than money wages during the times of rising prices and fall faster than money wages during the periods of falling prices. The result is that money wages decline in the former and rise in the latter case.

(b) **Subsidiary Earnings.** In addition to the regular money wage, an employee has extra earnings in the form of money or goods. For example, free board and lodging are provided to the domestic servants and may in some cases be given to peons. Professors earn additional income by marking examination papers or from tuition fees, and so on. Subsidiary earnings may also arise from opportunities of employment available to other members of the worker's family.

(c) **Extra Work Without Extra Payment.** If an

employee is required to do extra work without any compensation, his real wages are less by that extent. Peons are paid for doing their duty during working hours, but quite often they are required to work late. For such extra work they are paid nothing. This means that their real earnings are reduced to that extent.

(d) **Regularity or Irregularity of Employment.** Regular or more secure employments may give lower money wages, but the real wages may be higher than irregular and insecure employments which give high money wages. For instance, a person with Rs. 10 daily wage but whose employment is intermittent, may not be so well off as another, who earns regularly Rs. 5 a day. Hence, it is good to distinguish between wage rates and earnings.

(e) **Conditions of Work.** Some occupations are healthier than others, and in some the hours of work are shorter than in others; the work may be more pleasant or less pleasant; the employer may be more sympathetic or less, and so on. All these things should be taken into account in estimating a person's real earnings.

(f) **Future Prospects.** A low money income will be considered a high real wage if there are good prospects of a rise in the future. On the other hand, a high initial salary may not be considered as good in the absence of prospects for a further rise.

Why a Separate Theory of Wages

It is demand and supply relationship, or scarcity in relation to demand, which explains all values whether values of commodities or values of services of the factors of production. Why should we then have a separate theory of wages?

The ordinary theory of prices, *i.e.*, demand and supply theory, is not fully applicable to the determination of wages for the following reasons:—

(i) The demand for goods depends on their utility to the purchaser, whereas the demand for labour services depends on their productivity (and not utility) to the employer. Hence, demand for the labour needs a special treatment.

(ii) The human element in the labour market deserves special consideration. The labourer and the labour services are inseparably tied up; hence the importance of the human element.

(iii) The peculiar institutional and behavioural factors affect bargains in the labour market. For instance, in the rural areas of backward economies, wages are fixed by custom instead of competition.

(iv) Certain peculiarities of labour, which distinguish it from a commodity, also necessitate a separate theory of wages. These peculiarities are:—

(a) Labour has weak bargaining power as against the powerful employer. Hence, the actual wage is

less than what a worker may be entitled to by his productivity.

(b) Labour is most perishable. Hence, the worker has to accept a lower wage than to what he may be entitled. He has no staying power and must accept what is offered.

(c) The changes in the price of labour react rather curiously on its supply. It may be that when wages have gone up, the supply of labour may decrease instead of increasing, unlike what happens in the case of a commodity. The worker may prefer leisure to work. That is why supply of labour may be a backward bending curve.

(d) Another peculiarity is that the supply of labour cannot rapidly adjust itself to the changes in demand. It takes a generation to increase the supply of labour. When, therefore, demand increases, wages must go up. Nor can the supply of labour be decreased when there is unemployment.

Conclusion. In view of the above peculiarities of labour, it becomes necessary to formulate a separate theory of wages. All the same, it is well to remember that wages being a price (of labour services), the fundamental principle of pricing, *i.e.*, the interaction of the forces of demand and supply, applies to the wage determination also. The various theories of wages only differ in the assumptions regarding the basic conditions of demand and supply.

OLD THEORIES OF WAGES

Subsistence Theory¹

Several theories have been put forward to explain the general level of wages prevalent in a country. Take first the Subsistence Theory. This theory originated with the Physiocratic School of the French economists and was developed by Adam Smith and the later economists of the classical school. The German economist Lassalle called it the Iron Law of Wages or the Brazen Law of Wages. Karl Marx made it the basis of his theory of exploitation.

According to this theory, wages tend to settle at the level just sufficient to maintain the worker and his family at the minimum subsistence level. If wages rise above the subsistence level, the workers are encouraged to marry and to have large families. The large supply of labour brings wages down to the subsistence level. If wages fall below this level, marriages and births are discouraged and under-nourishment increases the death-rate. Ultimately, labour supply is decreased, until wages rise again to the subsistence level. It is supposed that the supply of labour is infinitely elastic. That is, its supply would increase if the price (*i.e.*, wage) offered rises.

Criticism. In backward countries, wages no doubt

1. See Dobb, M.—*Wages*, 1932, Ch. IV, Sec. 3, p. 100.

are to be found at or near the subsistence level. But the theory does not apply to advanced countries like England and America. The theory evidently is based on the Malthusian Theory of Population. But it is wrong to say that every increase in wages must inevitably be followed by an increase in birth rate. An increase in wages may be followed by a higher standard of living which in turn influences the wage level.

Ricardo, one of the exponents of the theory, stressed the influence of custom and habit in determining what was 'necessary' for the workers. But habits and customs change over time. Hence, the theory can hold good for only a limited period of time and cannot be true of all times, especially of a world characterised by fast changing habits. Ricardo, therefore, admitted that wages might rise above the subsistence level 'for an indefinite period in an improving society.'

Another criticism of the theory is that the subsistence level is more or less uniform for all working classes with certain exceptions. The theory, thus, does not explain differences of wages in different employments.

Further, it may be said that the theory explains wages only with reference to supply; the demand side is entirely ignored. On the demand side, the employer has to consider the amount of work which the employee gives him and not the subsistence of the employee.

Moreover, the fundamental weakness of the subsistence theory lies in its long-term character. It explains the adjustment of wages over the lifetime of a generation and does not explain wage fluctuations from year to year. As such it has little practical value.

Finally, the 'subsistence minimum' is a very vague term. Does it refer to the minimum requirements of a modern man or of a tribal savage? There is no rigidly fixed minimum and it is not independent of the wages ruling over a period of time.

Wages Fund Theory ²

This theory is associated with the name of J.S. Mill. "Wages," wrote Mill, "depend upon the demand and supply of labour, or, as it is often expressed, on the proportion between population and capital. By population is here meant the number only of the labouring classes or rather of those who work for hire, and by capital, only circulating capital and not even the whole of that but the part which is expended on the direct purchase of labour." Mill asserted: "Wages not only depend upon the relative amount of capital and population, but cannot, under the rule of competition, be affected by anything else."

According to this theory, therefore, wages depended upon two quantities, viz., (i) the wage fund or the circulating capital set aside for the purchase of labour, and (ii) the number of labourers seeking employment. Hence, the level of wages can be ascertained by means of a simple arithmetical operation: by dividing the wages fund by the number of workers. In other words, wages vary directly as the quantity of capital and inversely as the number of workers.

Wages, thus, cannot rise unless either the wage fund increases or the number of workers decreases. But since the theory takes the wage fund as fixed, wages could rise only by a reduction in the number of workers.

It would appear, therefore, that, according to this theory, the efforts of trade unions to raise wages are futile. If they succeeded in raising wages in one trade, it can only be at the expense of another, since the wage fund is fixed and the trade unions have no control over population. According to this theory, therefore, trade unions cannot raise wages for the labour class as a whole.

Criticism. In contrast to the subsistence theory, which represented a rigid view and attempted to provide deterministic long-term or static theory, the wages fund theory tried to explain movement of wages in a changing world. Instead of a single equilibrium to which wages must inevitably return, determined by 'cost of production of labour,' the wages fund theory provided a varying 'natural rate,' determined by varying ratio of capital to population.

The theory has been widely criticized and stands rejected now. Mill himself recanted it in the second edition of his "**Principles of Political Economy.**"

Mill thought that wages were paid out of circulating capital alone. Whether the source of wages is capital or the present products, has been the subject of a keen controversy in the past. The fact is that in some cases, where the process of production is short (e.g., final stages of the productive process), wages are paid out of the present production. On the other hand, when a process of production is long, the labourer obviously does not obtain wages from the product of his labour either directly or through exchange. In such cases, wages mainly come out of capital.

Mill argued that wages were paid out of a certain fixed proportion of capital set aside for this purpose. This also is not true. There is no fixed wages fund in this sense. The fund, if we can at all call it so, is elastic. Its volume changes according to the prospects of profits. The productivity of labour at a given time is an important factor in determining these prospects.

Further, the theory is a mere truism. It does not tell us about the sources of the wages fund and the method by which it is estimated. It simply tells us what is self-evident, namely, that the wage can be

² *Ibid.*, Ch. IV, Sec. 6. See also *Marshall's Principles*, Appendix (a).

ascertained by dividing the wages fund by the number of workers.

Again, the theory assumes a degree of antagonism between labour and capital that does not actually exist. According to this theory, wages can increase only at the expense of profits. But this is not necessarily so. In times of business prosperity, both wages and profits can go up.

It is also wrong to assume that the forcing up of the wages will drive capital abroad. Capital is not so sensitive, nor are the profits so inelastic. The theory fails to show why wages cannot be increased at the expense of rent and profit.

We cannot accept its corollary that the trade unions are powerless to increase wages and that any measures which hindered the accumulation of capital, e.g., heavy taxation, were bound to lower wages by reducing the wages fund. The theory is too unsympathetic to labour when it says that "the only hope of improvement for the workers lay in limiting the size of their own families and helping to increase the prosperity of their masters" (Dobb).

It is difficult to subscribe to the implication of the theory, *viz.*, that if any section of labour wrested a higher wage, it will be at the expense of other workers who must receive less or face unemployment. It also looks absurd that low wages paid to a certain section of workers would benefit others for it left larger wages fund available for them.

The theory ignores the principle of Economy of High-Wages. We know that a rise in wages would improve labour efficiency and increase the demand for labour to satisfy which the employers would set aside a larger fund for the purchase of labour. That is why the economists regard capital not as a fund but a flow.

Moreover, the wages fund theory does not explain why wages differ in different occupations. Besides, the wage rates prevalent in different countries do not correspond to the total amount of capital available there. In new countries, capital is scarce but the wages are high; the opposite is the case in the old countries.

Conclusion. In spite of the above criticism, it may be stated that the theory contains an element of truth. It may not apply to a highly industrialised country, but, in an under-developed country suffering from capital deficiency, wages cannot be increased unless national income is increased and capital accumulated through industrialisation.

Residual Claimant Theory

The Residual Claimant Theory has been advanced by the American economist, Walker. According to him, wages are the residue left over, after the other factors of production have been paid. Walker says that rent and interest are governed by contracts but profit is determined by definite princi-

ples. There are no similar principles, says the theory, operating as regards wages. According to this theory, after rent, interest and profit have been paid, the remainder of the total output goes to the workers as wages.

The theory admits the possibility of increase in wages through greater efficiency of labour. In this sense, it is an optimistic theory, whereas the subsistence theory and the wages fund theory are pessimistic.

Criticism. This theory also has been rejected by most economists. It has several defects: In the first place, it does not explain how trade unions are able to raise wages. Secondly, it ignores the influence of supply of labour on wages. Thirdly, one fails to understand why the same laws of supply and demand, that explain the remuneration of other factors of production cannot be applied to wages as well. Finally, it is not the worker who is the residual claimant but the entrepreneur, who undertakes to pay the other factors of production before he can expect to get anything.

Marginal Productivity Theory of Wages

Towards the end of the nineteenth century, the economists abandoned the Wages-Fund doctrine. Economists like Marshall treated demand for labour as derived from the demand for the products of labour and not from a pre-determined decision of the employers regarding the amount of capital they proposed to utilise in the purchase of labour. This view is represented by the marginal productivity theory.

While discussing the general theory of distribution (Chapter 31), we have given a detailed account of the marginal productivity theory and its criticism. There we explained the theory mostly as applicable to labour. This theory states that, under conditions of perfect competition, every worker of same skill and efficiency in a given category will receive a wage equal to the value of the marginal product of that type of labour.

We may repeat that the marginal product of labour in any industry is the amount by which the output would be increased if a unit of labour was increased, while the quantities of other factors of production employed in the industry remained constant. In short, it is the output attributable to a single unit of labour unaccompanied by any change in other factors of production. The value of the marginal product of labour is the price at which the marginal product can be sold in the market.

Under conditions of perfect competition, an employer will go on employing workers until the value of the product of the last worker he employs is equal to the marginal or additional cost of employing the last worker. Further, the condition of perfect competition implies that the marginal cost of labour is

always equal to the wage rate, irrespective of the number of workers the employer may engage. Every industry being ultimately subject to the law of diminishing returns, this marginal product must start declining sooner or later. Wages remaining constant, the employer stops employing workers at that point where value of the marginal produce of a worker is equal to the wage rate.

So far we have assumed that the quantities of other factors remain constant while that of labour alone increases. This, however, is not true because quantities of factors increase all round though this may not be true in the short run. To allow for this fact, the economists make use of the term "marginal net product of labour" instead of "marginal product of labour." The value of marginal net product of labour may be defined as being the value of the amount by which output would be increased by employing one more worker with the appropriate addition of other factors of production, less the addition to the cost of the other factors caused by increasing the quantities of other factors.

The theory may thus finally be re-stated as follows: Under conditions of perfect competition in the labour market and in the market for the products of the industry, and irrespective of the number employed, every worker will receive a wage equal to the value of marginal net product of his labour.

Limitations of the Marginal Productivity Theory. We have already studied in detail the various limitations and criticisms of the Marginal Productivity Theory as a general principle of distribution.³ With reference to its application to wages, we may repeat that the theory is true only under certain assumptions such as perfect competition, perfect mobility of labour from employment to employment, homogeneous character of all labour, constant rates of interest and rent and given prices of the product.

It is a static theory. The actual world is dynamic. All the factors assumed to be constant are in fact constantly changing. Competition is never perfect; mobility of labour is restricted for various reasons; all labour is not of the same grade; remuneration to other factors of production does not remain constant; and the prices of the products of labour vary. All these changes modify the theory when applied to actual conditions. The theory, however, as an assertion of a tendency, is true and is valuable in understanding the basic forces that determine wage rates.

In the real world, owing to the absence of the above assumptions, there is no single rate of wages that may be applicable to all labour of a particular type. Wages differ from place to place, from person to person and from employment to employment.

The following further points of criticism may now be noted:

Firstly, the theory has little applicability to reality. Labour is not perfectly mobile. Workers of the same skill and efficiency may not receive the same wages at two different places.

Secondly, though the condition of a large number of independent buyers of labour is fulfilled for a few industries of all countries and for most industries of some countries, the employers usually combine to the disadvantage of the worker. It is a case of monopsony, *i.e.*, one buyer and many sellers. The employers succeed in pulling down the wages below the value of the marginal net product of labour. If employees are also collectively organized, the wage rates may or may not be equal to the values of marginal net product of labour in the occupations or industries concerned. The wages are determined by the relative bargaining strength of the two parties, but cannot exceed the value of the marginal net product of labour.

Thirdly, the market for goods is in general characterised by imperfect competition. This also unsettles the theory.

Fourthly, the productivity of workers is also dependent on factors such as the quality of capital and efficient management. These factors are outside the control of workers.

Fifthly, it should be borne in mind that the marginal net product of labour depends not merely on the supply of labour but also on the supply of all other factors of production. If other factors are plentiful and labour relatively scarce, the marginal net product of labour will be high, and vice versa.

Finally, productivity is also a function of wages. Low productivity may be the cause of low wages, which may tell on the efficiency of the worker, lower his standard of living and ultimately check the supply of labour. The theory takes the supply of labour for granted.

Conclusion. In short, the marginal productivity theory ignores the effect of wage changes on the supply of labour, bargaining strength and monopoly conditions, *etc.*

Rejecting the marginal productivity theory, Marshall said: "This doctrine has been put forward as a theory of wages. But there is no valid ground for any such pretension . . . Demand and supply exert equally important influences on wages; neither has a claim to predominance; any more than has either blade of a scissors, or either pier of an arch . . . (but) the doctrine throws into clear light one of the causes that governs wages."⁴

3. See Chapter 31.

4. Marshall A., *Principles*, pp. 518, 538.

Taussig's Theory of Wages⁵

The American economist Taussig gives a modified version of the Marginal Productivity Theory of Wages. According to him, wages represent the **marginal discounted product** of labour.

Taussig thinks that the labourer cannot get the full amount of the marginal output. This is because production takes time and the final product of labour cannot be obtained immediately. But the labourers have to be supported in the meantime. This is done by the capitalist employer. The employer does not pay the full amount of the expected marginal product of labour. He deducts a certain percentage from the final output in order to compensate himself for the risk he takes in making an advance payment. This deduction, according to Taussig, is made at the current rate of interest.

Thus, wages equal the total product of labour on the marginal land or in the marginal firm minus the amount discounted as explained above. The present value of the product is ascertained by discounting its anticipated future returns.

Two weaknesses of the theory have been recognised by Taussig himself. First, that it is "a dim and abstract one, remote from the problem of real life." To this he replies that this weakness is common to all economic generalisations. Second, and a more serious, objection is that the joint product is discounted at the current rate of interest. But according to his own analysis, the rate of interest is a result of the process of advance to the labourers, because it depends on the excess of what the labourers produce in the future over what is advanced to them in the present. This would mean arguing in a circle. To meet this difficulty Taussig suggests that we determine the rate of interest independently of marginal productivity by the rate of time preference, and with the interest thus determined discount the marginal product of labour. This, however hardly solves the difficulty; it merely evades it.

Taussig's theory ultimately analysed is another version of the Residual Claimant Theory of Wages. He says, in fact, that wages are what is left after rent, interest and profits are deducted from the total output. As such, the theory is open to all the objections put forward against the Residual Claimant Theory.

MODERN THEORY OF WAGES

Although labour has certain peculiarities and cannot be regarded as an ordinary commodity, still wages are very largely determined by the interaction of the forces of demand and supply as in the case of an ordinary commodity. Thus, the Modern theory of wages is the demand and supply theory.

Demand for Labour

"The demand for labour by the individual firm is a function of both the productivity of labour and the money demand for the firm's product."⁶ In other words, the demand for labour reflects partly labour's productivity and partly the market value of the product at different levels of production.

The demand for labour is a **derived demand**. It is derived from the demand for the commodities it helps to produce. Greater the consumer demand for the product, the greater the producer demand for the labour required in making it. It may be observed that it is expected demand and not existing demand for the product that determines demand for labour. Hence, an **expected** increase in the demand for a commodity will increase the demand for the type of labour that produces this commodity.

Apart from the magnitude of demand (*i.e.*, big or small), we have also to consider its elasticity or responsiveness to change in wages. The elasticity of demand for labour depends on the elasticity of demand for its output. Also demand for labour will generally be inelastic if their wages form only a small proportion of the total wages. The demand, on the other hand, will be elastic if the demand for the commodity it produces is elastic or if cheaper substitutes are available.

The demand for labour also depends on the prices and the quantities of the co-operating factors. Suppose the machines are costly, as is the case in India, obviously more labour will be employed. The demand for labour will be more. Also, the greater the demand for the co-operating factors the greater will be the demand for labour.

Another factor that influences the demand for labour is the technical progress. In some cases, labour and machinery are used in definite proportions. For instance, the introduction of automatic looms reduces the demand for labour.

Thus, demand for labour is determined by (a) the nature of demand for the product of labour, (b) the proportion of the cost of labour to the total cost of the product, (c) its substitutability by other factors, and (d) supply of capital as determined by the ability and willingness of investors to save and invest.

After considering all relevant factors, *e.g.*, demand for the products, technical conditions, the prices of the co-operating factors, *etc.*, the employer is governed by one fundamental factor, *viz.*, marginal productivity. The demand for labour, under typical circumstances of a modern community, comes from the employer, who employs labour and other factors of production for making profits out of his business. The demand price of labour, therefore,

5. For a detailed study consult *Readings in the Theory of Income Distribution*, pp. 278-293.

6. Allen M. Carter, *The Theory of Wages and Employment*, 1959, p. 45.

is the wage that an employer is willing to pay for that particular kind of labour.

Suppose, the employer employs labourers one by one. After a point, the law of diminishing marginal returns will come into operation. Every additional labour employed will add to the total net production at a diminishing rate. The employer will naturally stop employing additional labourers at the point at which the cost of employing a labourer just equals (in fact it is little less than) the addition made by him to the value of the total net product.

Thus, the wages that he will pay to such a labour (the marginal unit of labour) will be equal to the value of this additional product or marginal productivity. But since all the labour units are supposed to be homogeneous, what is paid to the marginal worker will be paid to all the labourers.

The demand schedule of the employer for labour is like the demand schedule of the consumer for a commodity, *i.e.*, lower the wage, the larger the number of workers demanded. With the larger labour force, output will expand and each unit of output will have to be sold at a lower price. This is an additional reason why marginal productivity of industry as a whole declines, the other reason being a decline in the marginal output.

The change in wage rate determines the **direction** of change in the demand for labour but the **degree** of this change depends on the elasticity of demand for labour. In case of elastic demand, a small change in the wage rate will lead to a considerable change in demand for labour, and vice versa. Whether the demand for labour is elastic or not will depend on (a) the technical conditions of production and (b) elasticity of demand for the commodity which that labour produces. Generally, the short-term demand for labour is less elastic than the long-term demand. That is why the employers and the trade unions adopt a stiff attitude in wage negotiations.

In fact, for any particular employer, working under perfect competition, wages are already settled by the market forces. Each firm constitutes so small a portion of the entire industry that it cannot influence wages appreciably by employing more or less of labour. The supply curve of labour confronting each producer is perfectly elastic, *i.e.*, horizontal line at the level of the market wage rate. The individual demand curve is determined by marginal productivity. The individual employer hires as many labourers as will equate the marginal productivity of labour with the rate of wages in the market.

In Chapter 31, we have given two diagrams. Have a look at them again. Fig. 31.5 (a) shows the demand for labour of an individual firm and Fig. 31.5 (b) the demand of the industry. It is necessary to remember that it is the demand of the entire industry, not of an individual firm, which determines wages in the market. The individual firm has to accept the market rate of wages and adjust its

own demand for labour accordingly. The demand curve of the industry is derived from the lateral summation of the demand curves of individual firms.

Supply of Labour

The supply of labour depends on (a) the number of workers of a given type of labour which would offer themselves for employment at various wage rates and (b) the number of hours per day or the number of days per week they are prepared to work. The Supply of labour may mean three things: (a) Supply of labour to a firm, (b) Supply of labour to the industry and (c) Supply of labour to the entire economy.

To a given firm, the supply of labour is perfectly elastic because at the current wage rate, it can engage as many workers as it wants. Its own demand constitutes only a negligible fraction of the total supply of labour.

But for the industry as a whole, the supply of labour is not infinitely elastic. Hence, if it wants more labour it has to attract it from other industries by offering a higher wage. It can also work the existing labour force overtime. This in effect will mean an increase in supply. The supply of labour for the industry is subject to the law of supply, *viz.*, supply varies directly with price, which means low wage small supply and high wage large supply. Hence, the supply curve of labour for an industry rises upwards from left to the right.

The long-run supply curve of labour in any occupation is a horizontal straight line, if earnings in other occupations are constant. This is due to the mobility of labour among occupations. Even in the short-run, it will be approximately horizontal if the occupation is an old one and does not require much training. In case an occupation requires extensive training, the short-run supply curve will be inelastic.

The supply of labour for the entire economy depends on economic, social and political factors or institutional factors, *e.g.*, attitude of women towards work, working age, school and college age and possibilities of part-time employment for students, size and composition of the population and sex distribution, attitude to marriage, the size of the family, birth control, standard of medical aid and sanitation, *etc.*

A given supply of labour, under conditions of perfect competition, gets distributed in various employments in such a way as to make the marginal productivity of labour in all the employments the same. But if labour cannot move freely from one employment to another, the marginal productivity will be different in different employments, and consequently, wages will also be different even for the same kind of labour.

The supply of labour may be decreased by labourers refusing to work for a time. This happens

when labour is organised into trade unions. The labourers may not accept wages offered by the employer, if such wages do not ensure the maintenance of a standard of living to which they are accustomed. But, as we shall see, it is only when higher wages are justified by higher marginal productivity that high wages will have a chance to stay. Thus, labourers with low marginal productivity cannot demand high wages merely on the basis of their standard of living. In the long run, however, marginal productivity, wages and the standard of living tend to adjust to one another.

On the whole, we might say that, the number of potential workers being given, the supply of labour may be defined as the schedule of units of labour at the prevailing rates of wages. This depends on two factors: (a) the number of workers who are willing and able to work at different wages and (b) the number of working hours that workers are willing and able to put in at different rates of wages.

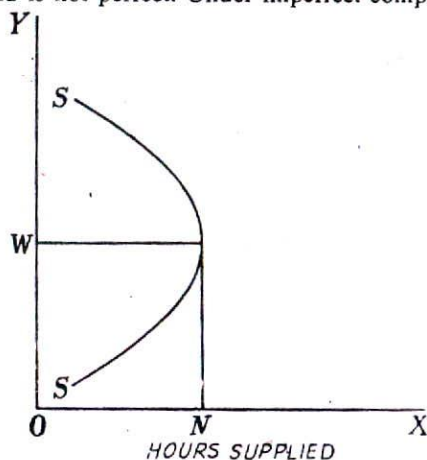
In case the workers have no staying power and the only alternative to work is starvation, the supply of labour in general will be perfectly inelastic. This means that wages can be driven down. Over a short period, reduction in wages may not cause any reduction in the supply of labour. But if wages are driven too low, competition among employers themselves will push them up. Even over a long period, the supply of labour is not very elastic.

There is a certain minimum wage below which labour will not work at all. Once this minimum is exceeded, supply of labour will increase as the wage rate goes up. But this will happen only up to a point beyond which wage increase will lead to a reduction in the supply of labour. Again, after a point, this tendency will be reversed when the worker thinks that he can move to a higher level of living. Hence, a rise in wages may lead to rise or fall in the supply of labour. It will depend on the worker's relative valuation of goods and leisure.

Thus, the supply of labour will depend on the elasticity of demand for income which will vary according to the worker's temperament and social environment. (We have already explained in Chapter 8 how indifference curves can be used to depict worker's preference for income and leisure). When the workers' standard of living is low, they may be able to satisfy their wants with a small income and when they have made that much, they may prefer leisure to work. That is why it happens that sometimes increase in wages leads to a contraction of the supply of labour. This is represented by a backward bending supply curve as in Fig. 32.1. For some time this particular individual is prepared to work longer hours as the wages go up (wage is represented on OY-axis). But beyond OW wage, he will reduce rather than increase his working hours.

Under perfect competition, the supply of labour,

as already mentioned, is perfectly elastic, since no single firm can influence the price of labour in the market as a whole. But actually competition in the world is not perfect. Under imperfect competition,



*Backward Bending Supply Curve
of Labour*

Fig. 32.1

we have an upward sloping supply curve. The particular slope of the supply curve will depend on a number of inter-related factors such as the existence or absence of unemployment in the neighbouring labour market, the reaction of the other firms to a change in the wage rate made by the firm in question, differences in the efficiency of the workers, and so on.

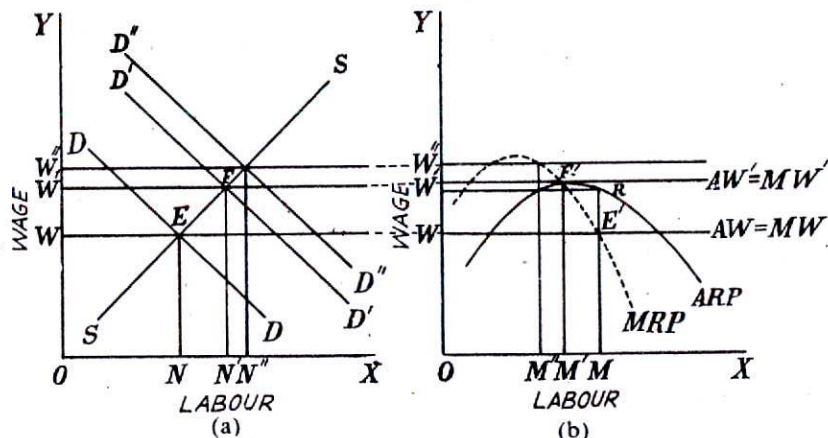
The supply of labour becomes less elastic as the margin of full employment is reached or if the particular firm accounts for a large share of total employment.

Interaction of Demand and Supply

Having analysed the demand side and supply side of labour, we shall now see how their interaction determines the wage level. This is shown by the diagrams given on next page.

In Fig. 32.2 (a) we take the case of an industry. The curve SS represents supply to the industry. DD is the demand curve. They intersect at E. The wage level, therefore, is OW (= EN).

Fig. 32.2 (b) represents the case of a firm. Under perfect competition, the firm has to accept the market wage OW settled by the industry above. From the OW level in Fig. 32.2 (a) draw a straight line towards Fig. 32.2 (b) representing the condition of the firm. We see that this extended line W-AW cuts the MRP (marginal revenue productivity) curve of the firm at F' in Fig. 32.2 (b). But at this level, the average revenue productivity (ARP) is MR which is greater than the wage OW. Hence, all the firms (this particular firm is representative of all firms in the



Interaction of Demand and Supply
Fig. 32.2

industry) are making supernormal profits at this wage level. This will lead to entry of new firms in the industry; the demand for labour will increase and the wage level will go up. Thus, the supernormal profits will be competed away in the long run by the entry of new firms.

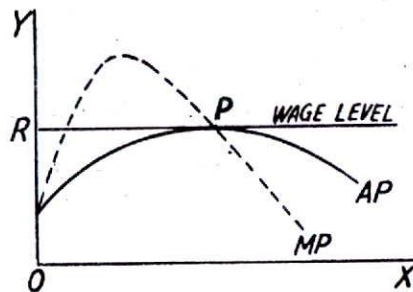
The new demand curve $D'D'$ cuts the supply curve SS in Fig. 32.2 (a) at F . The wage level in this situation will be OW' which is higher than the original wage level OW . This is how the interaction of demand and supply determines the wage. We can see that when the wage is OW , the firm is in equilibrium at E and when the wage rises to OW' , the equilibrium is at F . At this point, average revenue productivity (ARP) and marginal revenue productivity (MRP) are equal, and the average wage OW' is equal to both of them.

It can also happen that the occurrence of supernormal profits attracts some firms from outside, which may further increase the demand for labour to $D''D''$. The wage level will then be OW'' . Here the average revenue productivity (ARP) is less than the wage OW'' , i.e., the firms are suffering losses. The result will be that some firms will leave the industry and the wage will come down to the level of OW' . Here the wage, the marginal revenue productivity and the average revenue productivity are all equal.

Thus, under competitive conditions, wages are, in the long run equal to the marginal as well as average productivity of labour. If marginal productivity is greater than average productivity, it will be worthwhile to employ more labour till marginal productivity falls to the level of average productivity. On the other hand, if marginal productivity is less than average productivity, less labour will be employed and the marginal productivity will rise to the level of average productivity. Marginal productivity and average productivity thus tend to be equalised. Since wages are equal to marginal

productivity, they are also equal to average productivity.

This can be shown in a simpler diagram as in Fig. 32.3. Here OR represents the wage level, AP is the average productivity curve and MP the marginal



Wages Equal Marginal and Average Productivity
Fig. 32.3

productivity curve. They intersect at P which shows that wages are equal both to the marginal productivity and average productivity. When AP is rising, $MP > AP$, and when AP is falling, $MP < AP$. But when AP is neither rising nor falling, $MP = AP$. Hence,

$$\text{Wage} = MP = AP.$$

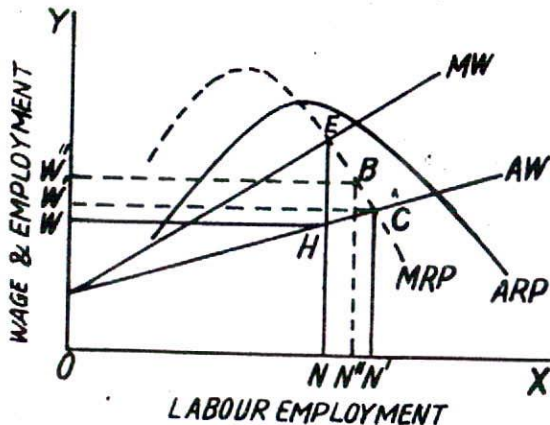
WAGES UNDER IMPERFECT COMPETITION

We have discussed above wage determination under perfect competition. But in the world of reality there is imperfect competition rather than perfect competition. Imperfect competition may result (a) when strong employers' associations are confronted with strong labour organisations. It is a case of Bilateral Monopoly, (b) when an industrial employer or a group of employers occupy a very strong monopolistic position as compared with labour to constitute what is called monopsony.

In bilateral monopoly, wage will be determined as price is determined under bilateral monopoly.

Monopsony. The usual case of imperfect competition so far as labour is concerned is that of monopsony where the employer embodies in his person concentrated monopoly power of being the sole purchaser of labour, whereas labour occupies a very weak position in comparison. Monopsony also occurs when a big employer employs proportionately a very large number of a given type of labour so that he is in a position to influence the wage rate. It may also occur when a group of big employers come to an understanding not to compete for labour and thus act as a single hirer of labour.

In the Fig. 32.4, the marginal revenue productivity



Wage Determination Under Monopsony
Fig. 32.4

curve (MRP) is the demand for labour. The supply curve of labour is AW (Average wage curve) which is rising to the right which shows that higher wages have to be paid to engage more labour. MW is the marginal wage curve corresponding to the average wage curve AW.

The two curves, viz., marginal wage curve MW and the marginal revenue productivity curve (MRP) intersect at E. It is at this point that the monopsonist will be in equilibrium. Here the marginal wage is equal to the marginal revenue product at the level of labour employment ON. In this situation, the wage is $NH (= OW)$.

It can be seen that this wage $NH (= OW)$ is less than marginal revenue productivity which is NE . Thus, each worker gets EH less than his marginal revenue product. This is the measure of labour exploitation under monopsony. This is called by Mrs. Robinson as 'monopsonistic exploitation.'

Thus, under monopsony, wage is lower and employment is less than under perfect competition in the labour market. Under perfect competition, the equilibrium would have been at C where the supply curve AW cuts the demand curve MRP. At

this point, the wage would have been higher at $OW (= NC)$ and the labour employed would have been larger at ON .

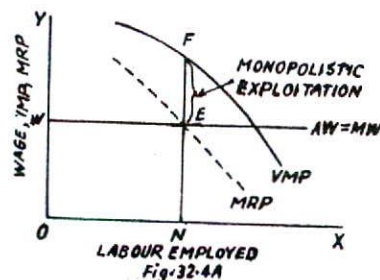
EXPLOITATION OF LABOUR

It is well-known that labour has weak bargaining power in relation to the employer; hence exploitation of labour. But what precisely do we mean by exploitation? As defined by Joan Robinson, labour is said to be exploited when the wage is less than the value of the marginal product (VMP), which is equal to the physical product multiplied by the sale price of the product. Obviously, there can be no exploitation of labour under perfect competition both in the labour market and in the product market. In this case, price and marginal revenue are equal ($VMP = MRP$) and labour is paid the wage rate equal to the value of its marginal product (see Fig. 32.3 where there is no exploitation of labour).

But in other market situations, labour exploitation occurs. We may distinguish three types of exploitation: (i) Monopsonistic Exploitation, (ii) Monopolistic Exploitation and (iii) Double Exploitation both monopsonistic and monopolistic. Now a word about these.

Monopsonistic Exploitation. We have already explained it in the preceding section. (see Fig. 32.4)

Monopolistic Exploitation. In this case, there is perfect competition in the labour market (average and marginal wage curves coincide) but imperfect competition in the product market, i.e. marginal revenue is less than the price of the product and therefore the marginal revenue product (MRP) is less than the value of the marginal product (VMP). In equilibrium, the firm will equate wage with marginal revenue product. This means that labour is paid less than the value of the marginal product which shows exploitation. This is shown in Fig. 32.4A below.



Double-Exploitation. It occurs when there is imperfect competition both in the labour market (monopsony) and in the product market (monopoly). Thus there is double exploitation of labour both monopsonistic and monopolistic and labour is exploited the most.

How to Stop Exploitation. The two weapons to

stop exploitation are (a) Government action and (b) Trade Union action. Since monopolistic exploitation arises from the monopoly power of the employer, it cannot be stopped either by the trade union or the government calling upon the employer to raise wages. If the trade unions compel the employers to pay high wages, they can reply by employing smaller number of workers so as to equate the new higher wage with the marginal revenue product of labour. Since the value of marginal product (VMP) will be greater than the marginal revenue product (MRP), labour will still be exploited. The only way to remove exploitation is to create conditions of perfect competition in the product market. The government can take some measures to remove monopolistic conditions. In the case of monopsonistic competition, the trade unions or the government can remove exploitation by raising wages.

TRADE UNIONS AND WAGE DETERMINATION

Establishment of a strong trade union introduces an element of monopoly in the labour market. Now, instead of each worker negotiating with the employer on the question of wages, the trade union bargains on behalf of them all. Here is a case of bilateral monopoly. Hence, determination of wages by collective bargaining is like determination of price under bilateral monopoly (p. 223).

The following diagram (Fig. 32.5) illustrates the determination of wage under collective bargaining on the assumption that the trade union seeks to achieve maximum wage regardless of its effect on employment.

In this figure ARP is the average net revenue productivity curve and MRP is the marginal net revenue productivity curve and IC_1, IC_2, IC_3, IC_4 and IC_5 are indifference curves showing different wage levels corresponding to the satisfaction of the trade union with the respective wage rate. The

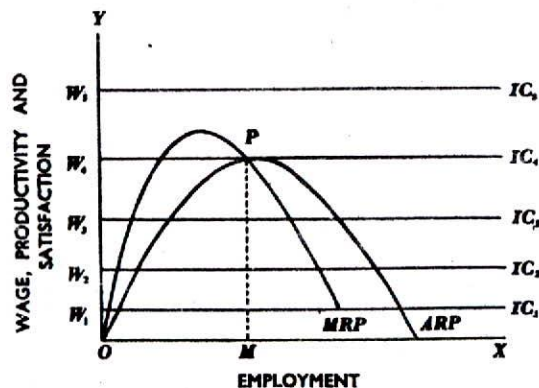


Fig. 32.5 The Two Limits in Collective Bargaining indifference curves show a constant rise in satisfaction between successive indifference curves. We can see that as the wages go up the gap between the

wage levels widens. This is due to the fact that as wages increase only a larger increase will yield constant satisfaction.

If the union had its way and could choose the wage rate unilaterally, it would choose that wage rate at which the corresponding indifference curve is a tangent to the average net revenue productivity curve (ARP). This happens at the point P in the figure given above and the wage rate will be set at $OW_4 (= PM)$, where the indifference curve IC_4 is tangent to ARP curve. The number of workers employed will be OM. A wage rate above OW being above the average net revenue productivity will put the employer to loss and hence will not be at all acceptable to him. He would rather stop production which would be detrimental to the workers who would lose their jobs. Thus, OW_4 is the upper limit.

But the employer would like to fix wages at as low a level as he could considering industrial conditions, elasticity of demand for the product, elasticity of substitution between labour and capital, prevailing wage rates, labour efficiency, their cost of living, etc.

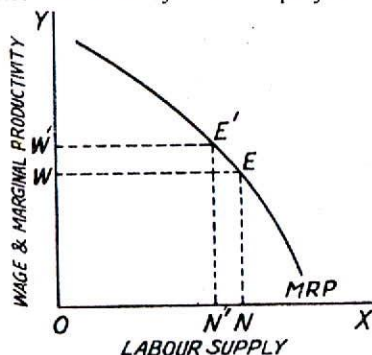
Can Trade Union Raise Wages?

For a long time, the economists held the view that the trade union action in increasing wages was futile and ineffective.

The classical economists argued that wages could be raised only at the expense of profits and a fall in profits, by reducing industrial activity, would reduce demand for labour and cause unemployment.

It is said that the wage level is determined ultimately by economic forces to which the strength of the contracting parties is irrelevant and which the bargaining power of the trade unions is powerless to bend.

Hubert Henderson held a similar opinion: "It is an illusion to suppose that the general level of wages can be appreciably and permanently raised by trade union action, except in so far as it increases the efficiency of workers or incidentally stimulates the efficiency of the employers."⁷



Futility of trade unions in raising wages is illustrated by the diagram 32.6 Here MRP is the

7. Supply and Demand, p. 145.

the Marginal Revenue Productivity curve which is the employer's demand curve. When ON is the supply of labour, equilibrium of demand and supply is at E and at $OW (=NE)$ wage, ON is the number of workers employed. But if the wages are forced up through trade union action to $OW' (=N'E')$, the number of workers employed is reduced to ON' i.e., NN' remains unemployed or through competition they will bring down the wages to OW , the previous level.

But, the modern economists do not subscribe to the above view. On the other hand, they strongly hold that trade unions can raise wages in a number of ways:

(i) **Stopping Exploitation.** The trade unions can ensure that labour is paid the full value of its marginal productivity. Under perfect competition, no doubt, wages tend to equal the marginal productivity of labour. But competition, in the real world, is not perfect. Hence, wages do not come up to the marginal productivity level due to the weak bargaining power of labour. By improving their bargaining power, the trade unions can raise wages up to the marginal productivity level and put an end to the exploitation of labour by powerful employers.

(ii) **Raising Marginal Productivity.** Trade unions can improve the marginal productivity of labour itself: (a) They can force the employer to use more up-to-date appliances and organisation. (b) They improve the efficiency of labour itself. This they do by fostering habits of sobriety, thrift and honesty and by helping the younger generation to acquire better education and training. (c) Trade unions may also increase the marginal productivity of a particular group of labourers by restricting its supply.

In Fig. 32.7 below: the trade union has raised the

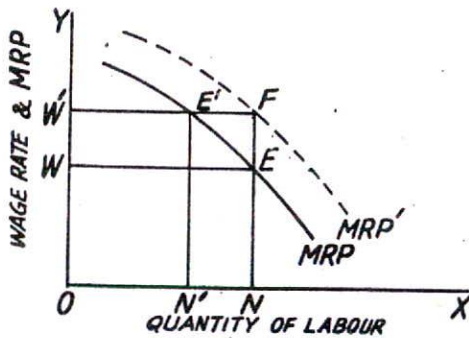


Fig. 32.7

marginal productivity from MRP to MRP' by improving workers' efficiency. In this way, they have raised the equilibrium wage from $OW (=NE)$ to $OW' (=N'E=NF)$ without reducing the number of workers employed. In this case, no unemployment has been created even though the wages have been forced up by trade union action.

(iii) **Restricting Labour Supply.** The trade unions usually adopt a number of restrictive devices, e.g., forcing the government to pass immigration laws, pressing for the reduction of working hours, long apprenticeships, restricting entry to the union and not permitting non-union labour to work, and so on. The aim clearly is to raise wages by reducing supply of labour when demand for it remains the same. When men workers get higher wages and are able to support the family, women workers may withdraw or the workers may work short-time preferring leisure to wages. In these ways, reduction in the supply of labour may raise the equilibrium wage rate.

There are special circumstances in which a particular set of workers can raise their wages by withdrawing their supply: (a) when the demand for that group of labour is inelastic, (b) when the wages of the said group form a small proportion of the total cost of production of the commodity concerned, and (c) when the other factors of production are "squeezable." In the long run, however, if the employers are forced to pay too high wages, there is a danger that they may adopt labour-saving devices and the demand for labour may fall, thus bringing down wages.

(iv) **Raising Standard Wage Rates.** Instead of putting restrictions on the supply of labour, modern trade unions fight for the raising of standard wage rates. This is a very common method of raising wages adopted by the unions today. Once certain standard wage rates are accepted by a representative body of the industry, individual firms quickly fall in line.

(v) **Increasing Labour Demand.** The trade unions also adopt ways and means to increase the demand for labour on the part of the employers. We know that demand for labour is a derived demand, i.e., derived from the demand for products that labour produces. The unions try to shift this derived demand curve upward. For instance, by improving labour productivity, they may enable the management to reduce cost of production and hence the price. This in turn will increase the demand. The unions may agitate for the protection of the industry so that domestic industry expands and employs more labour. They may influence the government to pay higher rates on public contracts for the goods made by the industry. They may help the employers to charge a high monopoly price and wrest a share for labour out of monopoly profits. Some unions have agitated for higher tariff in order to increase the domestic demand for goods they produce and consequently for labour used in their production.

By collective bargaining they can raise wages and increased wages mean higher marginal productivity. Higher marginal productivity means shifting of the demand curve upwards.

(vi) The trade unions, can raise wages, because a

large part of this rise can come about by squeezing the rent-element in the other factors of production and monopoly gains in other incomes.

(vii) It may also be argued that the raising of wages by the unions will not necessarily discourage investment. Today the bulk of investment comes not from individuals but from big corporations which usually maintain the level of investment but reduce dividend to shareholders when their income falls.

However, trade unions can raise wages if they are all-inclusive and it is difficult to import 'black-legs'.

Adverse Effect of Trade Union Action

We may note certain implications of wage determination by trade unions. By imposing a high wage, a trade union may prevent undue expansion of an industry when the only attraction to the new firms was the prevalence of a low wage. But it may also cause unemployment by insisting on a wage which is beyond the capacity of the industry to bear. Some employers will be forced out and others will contract output. In both cases, some labour will be discharged. The movement of the discharged labour into other industries will depress wages there too.

Conclusion

If a trade union can wrest monopoly profits from a monopolist, it is all right. The proper function of a trade union is to stop or prevent exploitation and when it goes beyond, it is bad. "The whole question of labour's bargaining power in a free enterprise economy boils down, therefore, to the question how far it can make inroads into capitalist consumption without endangering the capital substance on which its own long-run productivity and income depend."¹⁰

WAGE DIFFERENTIALS

Relative Wages

So far we were concerned with general wages. The problem of relative wages is different. Here we have to explain the causes of differences in wages in different employments or occupations or grades of employments and also between different persons in the same employment or grade. Wages everywhere tend to approximate to the marginal productivity of labour. But, the marginal productivity of labour is different in different employments and grades. It varies with the degree of scarcity of each kind of labour in relation to the demand for it, or ultimately in relation to the demand for the products of each kind of labour.

If there were free mobility of labour over the whole field of employment, real wage would tend to be in proportion to the relative efficiency of labour engaged in each kind of work. Real wages (not nominal wages) of workers of the same level of efficiency would tend to be the same. If workers in one employment were getting real wages more than

in proportion to their efficiency, labour would tend to move to that employment until increased supply would bring down its marginal productivity and wages. An opposite movement would take place if in an employment lower wages were paid than those justified by the relative efficiency of labour. Actually, however, labour cannot move freely from employment to employment especially in different grades. Different grades thus tend to become "non-competing groups."

We may summarise here the causes which create differences in wages in different employments, professions, and localities:

(i) **Differences in Efficiency.** These may be due to different inborn qualities, education, training, and conditions under which work is performed. When efficiencies are different, wages must be different.

(ii) **Existence of Non-competing Groups.** As explained above, these groups arise because of the difficulties in the way of mobility of labour from low-paid to high-paid employments. These difficulties may be due to geographical, social or economic reasons. They may arise from lack of transport facilities, existence of family ties or caste barriers, and lack of means for better training, etc.

(iii) **Difficulty of Learning a Trade.** The number of those who can master difficult trades is small. Their supply is less than demand for them, and their wages are higher.

(iv) **Differences in Agreeableness or Social Esteem.** Disagreeable employments must pay higher wages in order to attract labourers. If, however, disagreeable work can be performed by unskilled workers, who cannot do anything better (due to caste or other disabilities), wages may be quite low, e.g., sweepers in India.

(v) **Future Prospects.** If an occupation provides opportunities for future promotion, people will accept a lower start in it, as against another occupation offering higher initial rewards but no chances of rise in the future. The number of top prizes available in a profession also accounts for differences in wages.

(vi) **Hazardous and dangerous occupations** generally offer higher emoluments.

(vii) **Regularity or irregularity of employment** also exerts a strong influence on the level of wages. Regular employment generally carries a low salary.

(viii) **Collective Bargaining.** The differences in the strength and militancy of trade unions also account for differences in wages in different industries.

To Sum Up. We may sum up broadly the causes of differences in wages as workers' preferences, transfer costs, ranges of abilities and effects of collective bargaining.

It should be borne in mind that all these factors create differences in wages by affecting the adjustment of supply of labour to demand for it in various

10. Rothschild, K. W., *The Theory of Wages*, 1965, p. 110.

employments and grades. Wages are, in every case, determined by the degree of scarcity in relation to demand for labour, or by the marginal productivity of labour with respect to each kind of work.

Low Wages of Women

In most cases, women are paid lower wages for doing the same kind of work. There are several reasons for this: In the first place, it is due to long habit and custom. Until comparatively recently, a woman was regarded as a household drudge even in advanced countries. Even today women tend to crowd into occupations involving drudgery and depress wages there.

Secondly, since women do not make their work a life career, they do not equip themselves with proper education and training. Their aim normally is to get married, and after marriage most of them cease to earn independently.

Thirdly, for the same reasons, women do not organize themselves into trade unions to enforce higher wages for themselves. Their employment being only a stop gap between school/college and marriage, they do not try to improve their economic position.

Fourthly, women workers are prepared to accept lower wages because they have very limited obligations and responsibilities. In most cases, they do not depend merely on their own earnings, even when they are in employment. Husband, brother, or father, may and does, give them financial support.

Finally, male workers are supposed to be more capable of continuous and efficient work than female workers due to various reasons. Men are physically stronger and can undertake more strenuous work and can bear much greater nervous strain. Moreover, a woman, on account of biological reasons, is partly or wholly incapacitated for full work during certain periods of her life.

WAGES UNDER FULL EMPLOYMENT

When there is no full employment, pressure for increase in wage rates may lead to some reshuffling of demand for labour and hence of employment without upsetting the price system. For example, if as a result of increase in wages, there is increase in propensity to consume, level of investment remaining the same, output will increase to meet the increased demand and prices will not rise. If, on the other hand, investment declines, the national income will decline and prices, too, will fall in line. Thus, the wage increase will not affect prices if the productive resources of the community are not fully employed so far.

But under full employment, the conditions are not likely to remain stable. Since the resources are already fully employed, increase in wages will not be followed by increase in output; hence prices are bound to rise. The Government will make good by

public investment any falling off of private investment which might have depressed prices.

Also, the existence of full employment strengthens the bargaining power of labour. There can be no fear of unemployment as a result of rise in wages. This will only provoke labour to clamour for higher wages. The employers can maintain or increase production by offering higher wages, since there is no unemployed labour to draw upon. The employers, therefore, are likely to combine to resist an all-round increase in wages without their being able to get better supply of labour. But the labour will be in a stronger position and will succeed in securing wage increases. However, since, as explained above, prices are likely to rise, their real wages are not likely to go up.

Hence, under full employment, there will develop an inflationary spiral consequent on a rise in wages. This can be neutralized only by increase in the level of labour productivity. Further, insistent demand for higher wages may lead to the adoption of labour-saving and other technical improvements. Persistent rise in prices, wages and profits will lead to transfer of purchasing power from the fixed incomists to workers and entrepreneurs which is quite desirable socially.

Besides, inflationary situation and consequent instability, another consequence of permanent full employment is to rob the labour market of its flexibility. Changes in the industrial structure can no longer be met by redistribution of unemployed labour but by transfer of labour from one sector to another. In the absence of unemployment, helping in the process of adjustment, some new and more positive methods will have to be devised and adopted to ensure flexibility of the economic system. A national wage policy will have to be adopted which can be easily done in a planned system.

SHARE OF WAGES IN THE NATIONAL INCOME

Factors Affecting Share of Wages

Among the factors affecting the share of wages in the national income we may mention the following:—

(i) **Labour Productivity.** Other things being equal, the wages per worker would be higher where the labour productivity is higher and would increase with increases in labour productivity.

(ii) **Introduction of Labour-Saving Machinery.** With the progressive introduction of labour-saving machinery the share of wages in the aggregate national output would decline. It will be high where the work is mainly manual and low where it is highly mechanised.

(iii) **Condition of the Labour Market.** Labour's share in the total market, is subject to various

influences in the labour market. For instance, if labour is plentiful, their share will be low and it will be high if it is scarce. Similarly, collective bargaining, where there are strong labour unions, will increase labour's share.

(iv) **Mobility of Labour.** If there are no barriers to the entry of labour into the ranks of self-employed persons as small farmers, artisans, etc., labour will get a large share in the national income.

(v) **Growth of Monopolies.** The monopolists are in a position to exploit labour; hence labour's share in the national product would be reduced under monopoly.

(vi) **Export of Capital.** If capital migrates to other countries, where cheap labour and raw materials are available, there is a corresponding reduction in employment, and so in labour's share, in the exporting country.

(vii) **Cheaper Imports.** The real wages, and hence labour earnings, go up when cheap food and other consumable commodities can be imported from abroad.

From a continuous rise in wages, one is likely to get the impression that labour's share in the national income must have gone up. This, however, is not borne out by facts. The fact is that the share of wages in the national income has remained remarkably stable. This is so even in advanced countries where trade unionism is well established and very strong. In Great Britain, in the whole period 1880 to 1944, it fluctuated between 39 and 41 per cent.¹¹

According to an estimate by Sir Dennis Robertson, the share of wages in the net national income at factor cost varied from 40.3 per cent in 1938 to 43.6 per cent in 1953.¹² As regards the U.S.A., Dr. King's estimate showed that this share was just under 38 per cent in 1909 and just over 40 per cent in 1925.¹³ Thus, the share of wages in the national income has been remarkably steady.

There are two forces which seem to be responsible for keeping stable the labour's share in the national income: (a) The level of raw material prices, and (b) the degree of monopoly. These are the forces which in a capitalist economy determine the share of wages in the national income under conditions of imperfect competition.

In an economy, where productive capacity is not fully used, additional units of output can be obtained by combining more manual workers and raw materials with the given capital equipment without raising the wage and material costs of these additional units. Constant average wage and raw material costs are typical of the major part of the

economy. Other items of costs being negligible, the marginal costs of production approximate to the average cost of labour and raw materials.

The difference between the price and the marginal cost goes to the capitalists in the form of surplus profits or in the form of profit, interest, rent, depreciation, etc. The share of gross capitalist income and salaries in the total national income is determined by the average degree of monopoly; it increases or decreases if the degree of monopoly rises or falls. The rest of the national income is accounted for by wages and cost of raw materials.

Hence, the share of wages in the national income will depend on the cost of raw materials and degree of monopoly power. Therefore, a rise in the degree of monopoly power or in the cost of raw materials will reduce the share of wages and increase the gain of capital.

It should be clearly understood that the above analysis is applicable to an economy characterised by imperfect competition with some degree of monopoly power and not to an economy operating under perfect competition where the degree of monopoly power is zero. The world of imperfect competition is the world of reality.

Under perfect competition, the relative share of labour in the total output depends on the extent to which labour can be substituted for capital, and vice versa. The greater the substitution (technical or commodity substitution) of labour for capital, when the total output expands, the greater will be labour's share in the new output, and vice versa. This substitution will depend on elasticity of substitution. The relative share of labour will be greater or smaller according as the elasticity of substitution is greater or smaller than one. As Hicks observes, "An increase in the supply of any factor will increase its relative share, (i.e., its proportion of the National Dividend) if its elasticity of substitution is greater than unity."¹⁴

That the labour's share in the national income has remained stable is due to the fact that, on the one hand labour's attempts to raise wages have been baffled by their inability to reduce the degree of monopoly power; on the other hand, increase in the degree of monopoly power has been neutralised by a fall in the price of raw materials. During depression, the firms try to recover as much of their fixed expenses as possible and labour's share falls. But there is a greater fall in the prices of raw materials which checks the tendency of the wages to fall. In this manner, forces operating from opposite directions keep the share of wages in the national income more or less constant. This has been called a **law of capitalist economy**.

11. See Rothschild, K. W.—*The Theory of Wages*, 1965, p. 164.

12. Stamp Memorial Lecture for 1954, quoted in Maurice Dobb's *Wages*, 1956, p. 20.

13. Maurice Dobb—*Wages*, 1956, p. 20.

14. Hicks, J. R.—*Theory of Wages*, 1932, P. 117.

MEANING OF RENT

In ordinary speech, the term 'rent' is used in a wide sense to mean a hiring charge, *e.g.*, rent of a house, a tonga or a machine. But, in Economics, rent, or **economic rent** as it is called, is used in a special sense. It refers to that part of the payment by a tenant which is made only for the use of land, *i.e.*, free gift of nature.

Economic Rent and Contract Rent

The payment that an agriculturist tenant actually makes to the landlord is not necessarily equal to this economic rent. A part of this payment may consist of interest on capital invested in the land by the owner in the form of buildings, fences, drainage, wells, *etc.* That part of the payment which is made for the use of land **only is economic rent** and the total payment made by a tenant to the landlord is called **contract rent**.

Rent as Surplus

In Economics, the term 'rent' is being increasingly used in the sense of a **surplus**, *i.e.*, what a factor of production earns over and above what is essential to maintain its supplies in its present occupation. It can be easily understood that rent in this sense can arise only when the supply of a factor of production is less than perfectly elastic, and this is the case with most of the factors. In case the supply of a factor is perfectly elastic, it cannot earn any surplus over and above its supply price. Because whenever such a factor is found to be earning more than its supply price, more units of this factor will rush in and the surplus earning will disappear. This is so because, under perfect competition, the market price of a factor must equal its supply price.

But when the supply of a factor is not perfectly

responsive to changes in the reward of the factor, it can continue to earn more than what is necessary to call forth its supplies without any fear of new units of factors coming in to deprive it of extra reward. The supply of land in general, though not for a particular use, is absolutely inelastic, and, as such, its supply is independent of what it earns. That is a higher rent cannot attract more of it and a lower rent cannot drive it out. That is why it is said that land has no supply price. That is, no payment is necessary to call forth its supply; it is there already. Its supply price being zero, the whole of its earnings is rent in the economic sense.

Economic rent is the surplus which remains to the supplier of a factor after he has paid all the expenses of production and has remunerated himself for his own productive effort.

We have said above that, **from the social point of view**, the whole of the earnings of land (and of other free gifts of nature) can be termed rent because land has no supply price, or its cost of production is zero. Why then is any payment made for the use of such factors of production? This is simply due to the fact that they are scarce in relation to demand for them.

Rent and Transfer Earnings

Modern economists make use of the concept of transfer earnings in explaining economic rent. As was explained in a previous chapter (no. 31), the transfer earnings represent the amount which a factor can earn in its next best paid alternative use. Suppose a piece of land yields in its present use Rs. 500 a year and suppose further that if it is transferred to its next best use, it will yield Rs. 400. In its present use, it yields Rs. 100 more than its next best use. This sum of Rs. 100 is a sort of surplus that the land is yielding in its present use. This surplus is called rent.

We can look at it in another way. A payment of

Rs. 400 is essential in order to retain this piece of land in its present use, but actually it yields Rs. 500, *i.e.*, Rs. 100 extra. The price which is necessary to retain a given unit of a factor in a certain industry may be called its transfer earnings or transfer price. If, however, the factor is earning over and above its transfer earnings or transfer price, the surplus or excess earnings is economic rent.

In the modern sense, therefore, economic rent means surplus or excess over transfer earnings. This sort of surplus or economic rent is not peculiar to land, it can be found in other factors of production too.

RICARDIAN THEORY OF RENT

How Does Rent Arise?

More than a century ago, David Ricardo supplied the answer in the Theory of Rent associated with his name.

Ricardo defined rent as follows: "Rent is that portion of the produce of the earth which is paid to the landlord for the use of the **original and indestructible powers** of the soil." Economic rent, according to Ricardo, is the true surplus left after the expenses of cultivation as represented by payments to labour, capital and enterprise have been met.

Differential Rent

How this surplus arises may be illustrated by an example. Suppose in a country there are four kinds of land—A, B, C and D. Some pieces of land are more fertile than others and some areas are more advantageously situated as regards centres of population and means of transport, *etc.* But taking all the factors into consideration, let us suppose that we have four grades of land as mentioned above, so that land A is the most superior and B, C and D are 2nd, 3rd and 4th grade lands, respectively. Further, suppose that standard units of labour and capital called "doses" of labour and capital, when applied to these categories of land, produce wheat as given in the following table:

Doses of labour and capital	Return in quintals of wheat per acre			
	A	B	C	D
1st	10	9	8	7
2nd	9	8	7	6
3rd	8	7	6	5
4th	7	6	5	4
5th	6	5	4	3
6th	5	4	3	2

Suppose class A land is enough and to spare and it can meet the entire demand for food at the prevailing price. In this situation, land will command no rent. It will be like a free gift of nature.

Now suppose that population has increased to such an extent that the whole of the class A land is brought under cultivation, and still it is not enough to meet the growing demand for food. In order to meet the increased demand for food, more labour and capital will be put into lands of class A, and lands of class B will also be brought under cultivation. This will happen only when the price of wheat rises so much as to make it worthwhile putting one more dose (*i.e.*, two doses in all) of labour and capital into land A and putting first dose of labour and capital into land B.

In other words, according to our table above, 9 quintals of wheat must sell for as much as is the cost of second dose of labour and capital. The price of 9 quintals of wheat produced on land B is equal to the cost of labour and capital put in this land for producing this quantity. Since the price of wheat produced on land B just covers the cost of production on this land, there is no surplus and hence no rent. In other words, B is the no-rent land. But on lands of class A, two doses of labour and capital give a return of $10+9=19$ quintals of wheat. But only $9 \times 2 = 18$ quintals of wheat are enough to pay for the two doses. Hence, one quintal of wheat is the surplus on land A. Thus, cultivators of land can either cultivate B class lands free of rent and get 9 quintals of wheat per dose of labour and capital per acre, or they can pay one quintal of wheat (or its equivalent in money at the prevailing price) per acre to the owners of land A as rent. By applying one more dose of labour and capital per acre of land, they can obtain 19 quintals of wheat (*i.e.*, by applying two doses in all). The application of second dose on A class land yields 9 quintals of wheat per dose of labour and capital. If there is perfect competition, at this stage, this rent (*i.e.*, one quintal per acre) for A class lands is bound to be established, but B is no-rent land.

As the demand for food still grows and the price of wheat rises, this process will continue. More and more units of labour and capital will be applied to the superior lands on the one hand, and still inferior lands will be brought under cultivation, on the other. The available doses of labour and capital will be applied in such a way as to get equal returns at the margin of cultivation. For instance, if there are 10 doses available, 4 will be applied to land A, 3 to land B, 2 to land C and 1 to land D. In this way, the marginal or the last dose applied to each class of land will give the same return, *i.e.*, 7 quintals of wheat. The total production under these conditions will be $34+24+15+7=80$ quintals of wheat. No other arrangement will give more than this total.

These doses will be applied only if the price of wheat is such as to make only 7 quintals of wheat enough to meet the cost of application of doses of labour and capital.

Since the number of doses applied to each class of land is different, the output of these lands will also be different. On A class land, 4 doses of labour and capital will yield 34 quintals ($10+9+8+7=34$), B land with 3 doses will yield 24 ($9+8+7=24$) quintals, C land with 2 doses will yield 15 (*i.e.*, $8+7=15$) quintals and one dose applied to D class land will yield only 7 quintals which will just cover the cost of production there, yielding no rent.

It is clear that rent of each class of land is equal to the surplus output over and above the cost of production, which is equal to 7 quintals of wheat.

Under such circumstances, rent per acre of the various kinds of land will be:

Rent of A grade

$$= (\text{total produce}) - (\text{total cost}) \\ = 34 - 28 = 6 \text{ Qtls}$$

Rent of B grade = $24 - 21 = 3$ Qtls

Rent of C grade = $15 - 14 = 1$ Qtls

Rent of D grade = $7 - 7 = 0$ (zero) Qtls.

We have calculated the rent in terms of the produce. It can be easily converted into money value at the prevailing price of the produce.

Thus, rent arises on account of natural differential advantages enjoyed by a piece of land over the marginal land. The natural differential advantage may be due either to superior quality of land or its better situation.

Marginal or No-Rent Land

When marginal produce is 7 quintals, there will be no rent of D grade land. It is then the marginal land or land on the margin of cultivation. This is also called the "no-rent land". It produces no surplus over cost of production. Its produce is just enough to cover the expenses of production on it. The rent of all superior lands is measured upwards from, and with reference to, the marginal land.

Now, the marginal land may not be the poorest or the worst land. We have to look not to the quality of land but to its best alternative use. The marginal land, *i.e.*, the land on the margin of transfer, may be the best land. The land which is best for the cultivation of cotton may also be the best for the cultivation of wheat. If the price of cotton falls, it is this land which will be first diverted to wheat, and will, therefore, be marginal from the point of view of cotton. This may be considered the modern version of marginal land.

Diagrammatic Illustration of the Ricardian Theory of Rent.

The following diagram (Fig. 33.1) illustrates the Ricardian theory of rent. These four figures show how much rent each grade of land yields. The Fig (i) indicates A-grade land which is the most superior land, Fig. (ii) shows the B-grade land which is next best, Fig. (iii) shows the C-grade land and Fig. (iv) shows D-grade land which is the least productive land or which is most inferior. In all these figures, the curve AC is the average cost curve and MC is the marginal cost curve.

The D-grade land will be cultivated only when the price of agricultural output is equal to the average cost of production. It will be seen that the minimum average cost of production on D-grade land is LH. If this D-grade land is to be cultivated then the price of the produce (wheat) must be equal to OP. In other words, when the demand for wheat increases so much that price OP is determined, only then will this land be cultivated. Now suppose that the demand for wheat has increased so much that OP price is determined. In this situation OL quantity of D-grade land will be cultivated. In the figure (iv), the total cost of production is equal to the total value of the produce obtained (because price OP = Average cost of production LH). Thus, in this case of D-grade land, there is no surplus above the cost of production. In other words, there will be no rent on D-grade land. This is the no-rent land or marginal land.

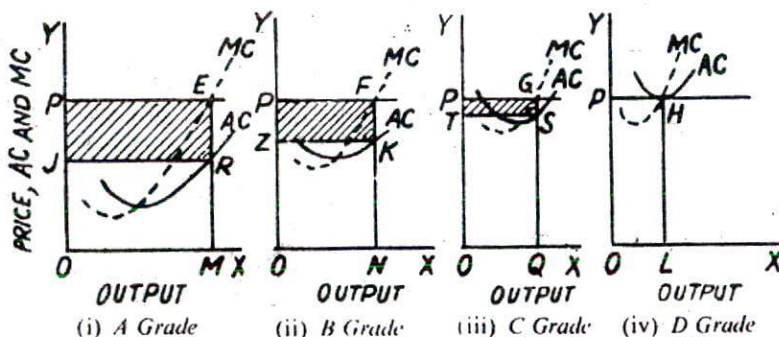


Fig. 33.1

Now take C-grade land (Fig. iii). In this case also, price is OP (because under perfect competition, price in the market must be the same whatever land may be producing that commodity), but the average cost is QS. Hence, there is GS per unit surplus or total surplus above total cost, is shaded area GPTS. This is the rent on the C-grade land. Similarly, the rent on B class land is represented by the shaded area FPZK and on A-grade land it is EPJR.

Does Marginal or No-rent Land Really Exist?

One would think that such a land cannot exist. Everybody will be keen on having such a land so that the owner may be in a position to demand some rent for it. But this is a superficial view. No rent land does really exist. The whole of the payment as rent for the use of land may be due to investment of capital, *i.e.*, interest. The land itself may be such as not to yield any surplus at all. For example, there may be some waste land lying useless and nobody may be willing to pay anything for its use. But a prospective-tenant may agree to pay something, if the owner sinks a well there. Obviously, the payment is for the capital invested in sinking the well, and not for the land which is really a no-rent land.

Sometimes the existence of no-rent land is concealed in a big farm, where the rent takes the form of so much money per acre. But in a farm there are mixed up good acres as well as bad ones. The latter, if let out alone, may fetch nothing.

Also, the rent paid may not be economic rent, but scarcity rent. This happens in an old country with a growing population. Then, even the marginal land pays some rent which is scarcity rent. The superior lands will pay, besides scarcity rent, economic rent, due to the natural differential advantages enjoyed by them.

Finally, the marginal land, with regard to an old country, like England, may exist in Australia or Canada supplying the same market.

Scarcity Rent

Besides economic rent, we have seen above, there is also scarcity rent. As the price of wheat rises, the worst land is also subjected to intensive cultivation and it yields a surplus over cost. This surplus is not a differential one compared to no-rent land, which does not exist. It is due to the scarcity of land as such. Hence, it is called scarcity rent.

The rent yielded by superior lands thus contains two elements: (a) differential surplus over the marginal land, and (b) payment due to scarcity of land as such. For instance, in our illustration, if cultivation is carried to a point where the worst land pays 2 quintals of wheat as rent, the superior land will pay a scarcity rent of 2 quintals in addition to the differential rent. The superior land will pay scarcity rent at the same rate as the worst land, but they will

also pay a differential rent. As Marshall says, "in a sense all rents are scarcity rents, and all rents are differential rents."¹ Differential rents arise because of scarcity of each particular grade of land.

Criticism of the Ricardian Theory

The Ricardian theory of rent has been widely criticized. First, it has been pointed out that there are no "original and indestructible powers of the soil." Good lands, after being constantly cultivated, lose their fertility to a large extent and get exhausted. To this may be replied that, if after exhaustion, good lands are manured equally with the bad ones, the former regain their productive power much more readily than the latter. It is also pointed out that, in an old country, where land has been constantly manured, the upper layer, which grows crops, is all man-made. There is nothing 'original' about it. But this is not correct. The climate, sunshine, air, situation, *etc.*, of a particular piece of land are all fixed by nature. They are all 'original and indestructible'.

Secondly, it is objected that Ricardo uses the term fertility of land in a vague manner. Apart from the factor of situation, fertility depends upon the ability of the farmers and the methods of cultivation used. Moreover, fertility is relative to the crops grown.

Thirdly, Ricardo's theory assumes that there exists a no-rent land which only repays the cost of cultivation. In most cases, it is true; there are lands which pay only a nominal rent. Such lands yield no true economic rent. The concept of scarcity rent can also explain this situation. For the substance of the theory, it is not necessary that there should exist a no-rent land. The concept of no-rent land is merely imaginary and theoretical and is not realistic.

Fourthly, according to the Ricardian theory, rent arises on account of natural differential advantages of superior lands over the marginal one. But even if all the land is of A-grade, rent will still arise. It will arise owing to the operation of the law of diminishing returns when land is intensively cultivated. The marginal unit of labour and capital applied must be compensated by the yield obtained. The earlier units will give surpluses over their costs (because their costs are less than the cost at the margin), which will constitute the rent.

The fact is that rent arises not on account of superiority or inferiority of land, but because land is scarce. If land, good or bad, were in a state of abundance, there would have been no question of paying or receiving rent. Even if the lands were homogeneous, rent will still arise owing to its scarcity. Ricardian Theory explains that superior

things have superior prices, but it does not explain why prices emerge.

Fifthly, as Carey and Roscher point out, it is historically wrong to assume that, in a new country, the best lands are cultivated first. In fact, lands that are first cultivated are not usually the best; they are only the most easily accessible. To this Walker replied that by the best land Ricardo meant not the most fertile land but that which was the best both from the point of view of fertility and situation.

Sixthly, criticism is levelled against Ricardo's corollary that since the marginal land pays no rent, and price is determined by cost of the marginal land, rent does not form a part of the price of the produce. The modern economists think that it is only from the point of view of economy as a whole that land has perfectly inelastic supply and earns a surplus or rent. This surplus is not included in cost and hence it does not enter into price. But, from the point of view of an individual farmer or industry, a payment has to be made to prevent land from being transferred to some other use. The payment called transfer earnings, is an element of cost and hence enters into price. For the individual farmer the whole of rent is cost.

"This concept of transfer earnings helps to bring the simple Ricardian theory—where transfer earnings are zero because it is the whole economy which is being studied—into a close relation with reality." (Stonier and Hague).

Finally, the most important criticism of Ricardo, however, comes from those who deny the necessity of explaining rent by a special theory not applicable to the rewards of other factors of production. They explain rent in the same way as wages, interest and profits. They deny its peculiar nature as contended by Ricardo. No specific and separate theory of rent is called for. The demand and supply theory which determines all values, also determines the rent of land. This point of view is explained below.

Further Appraisal of the Ricardian Theory

In the Ricardian theory, two facts stand out prominently: (a) That rent arises because some lands are superior and others inferior; and (b) that rent is measured from the no-rent margin.

Modern writers question both these contentions. In the first place, they assert that it is a matter of indifference to the general principle of rent whether the land is uniformly good, uniformly bad or gradable. The essential factor of rent is the relative scarcity of the products that land can yield. The scarcity of land is in fact derived from the scarcity of its products. If the problem is approached from this point of view, the necessity of assuming different grades of land disappears. The "differential principle" only explains why one particular acre of land commands a higher rent than the less fertile one; it does not explain why rent arises.

Fundamentally speaking, rent is paid because the produce of the land is scarce in relation to its demand. In the face of this scarcity, rent will arise even if all the lands in the country are exactly alike. The same is the case with wages, interest and profit. These payments arise because the products of the factors concerned are scarce in relation to the demand for them. Just as a superior labourer gets higher wages and a superior entrepreneur earns higher profits than the inferior ones, superior land also commands a higher rent. "Fundamentally, all that the Ricardian theory of rent amounts to is the truism that the better article will always command the higher price. A more fertile acre will be worth more than a less fertile one simply because they are different things. The same truism applies to wages."² Wicksell has pointed out that rent and wages are almost parallel cases.

Thus, there is no justification for placing rent in a special category. Land commands rent according to its marginal productivity just as labour commands wages or capital commands interest.

The second point of attack is the idea of the no-rent margin. This is the starting point of measuring rent according to the Ricardian theory. It is contended by modern writers that the no-rent margin may exist in some cases, but it is not fundamental to the emergence of rent. For instance, some lands may be fit only for a specific use, e.g., growing corn. If it is not profitable to grow corn on them due to fall in the price of corn, such lands may go out of cultivation, or they may just pay for the cost of the crop grown on them. Such lands may have significance from the point of view of rent but in a different sense than held by Ricardo. If such lands are cultivated, they tend to increase the supply of corn and thus lower rents, and if they go out of cultivation, rent rises due to a decrease in the supply of corn. The existence of such marginal land does not give any ultimate explanation of rent.

But when we do not refer to any particular crop, especially in a fully developed country, there is probably no land that cannot be put to some profitable use. Thus, the margin of cultivation may vary according to the use to which a particular land is put.

Conclusion

It is concluded, therefore, that the theory of no-rent margin and a series of differential rents created upon it, while true in particular cases, is a partial, not an ultimate, explanation of the phenomenon of rent. It does not reach the bottom of the problem. The real approach to rent, as to other phenomena of value, is through the principle of scarcity. The Ricardian Theory has, therefore, been rejected.

2. Briggs and Jordan, *op cit.*, p. 355.

Although the Ricardian theory has been rejected, but we should not ignore the elements of truth contained in it. Who can deny that with the increasing pressure of population, inferior land would be brought under the plough? It is true even today and not only historically true. The scientific improvements and new technology going under the label of 'green revolution' can only put off the process. Further, the theory exerted powerful influence on the contemporary economic and political thought. It highlighted the conflicting interests of the landlords and the rest of society. That is why abolition or socialisation of rent was advocated.

MODERN VIEW OF RENT OF LAND

Modern economists have furnished a better explanation of the rent of land. According to the modern view rent is a payment for the use of land. Rent, in this sense, is obviously determined by the demand for and the supply of land. Take first the demand side.

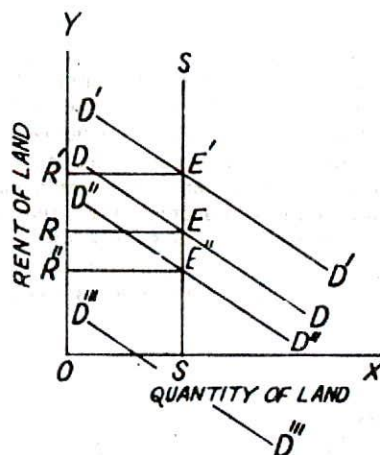
Demand for the Use of Land. The demand for land is a derived demand. It is derived from the demand for the products of land. If the demand for these products rises or falls, the demand for the use of land will correspondingly rise or fall leading to increase or decrease of rents. For instance, if population of a country increases, the demand for food will increase, resulting in increased demand for land and rise in its rent, and vice versa.

We have seen in a previous chapter (no. 31) that the demand for a factor of production depends on its marginal revenue productivity (or in short marginal productivity). The productivity is subject to the law of diminishing marginal productivity. That is why, as in the case of other factors, the demand curves shown in the Fig. 33.2 slope down from the left to the right.

Thus, on the side of demand, rent of land is determined by its productivity, not total productivity, but marginal productivity.

Supply Side. On the supply side, the supply of land is fixed so far as the community is concerned, although individuals can increase their own supply by acquiring more land from others or decrease its supply by parting with land. In spite of reclamation projects, the effect of which on the total supply is negligible, the supply of land remains practically fixed. It is a case of perfectly inelastic supply which means that whatever the rent (the rent may rise or fall), the supply remains the same. That is why it is said that **land has no supply price**. In other words, the supply of land in general is absolutely inelastic and as such its supply is independent of what it earns.

Interaction of Demand and Supply. We have analysed the demand and the supply side of land. The interaction of these forces is shown in the



Interaction of Demand and Supply
Fig. 33.2

diagram (see Fig. 33.2 given above). We assume that land is homogeneous and it is used for raising only one crop. Only then there can be one demand curve and one supply curve. We also assume perfect competition. *SS* supply curve, a vertical straight line, represents fixed supply. We start with *DD* as the total demand curve for land. These two curves intersect at *E*. In this *OR* (= *SE*) is the rent. If the rent is less than *OR*, say *OR''* (= *SE''*), the demand for land will increase; but the supply is fixed, hence rent will again rise to *OR*. If, on the other hand, rent rises above *OR* to *OR'* (= *SE'*), the demand for land will decrease and bring the rent back to *OR*.

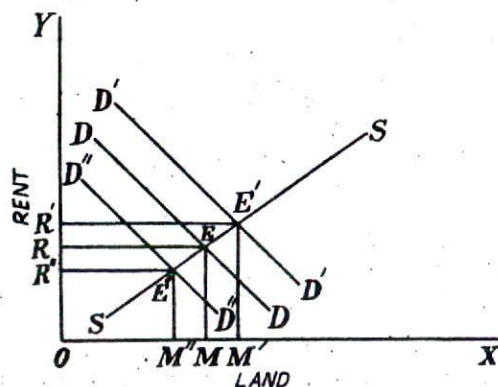
Suppose now that on account of increase in population or otherwise, demand for land has increased from *DD*, to *D'D'*. Supply curve is still the same *SS*. The new point of intersection will be *E'*, and, therefore, the rent will be *OR'* (= *SE'*). If demand falls to *D''D''* then the demand and supply curves intersect at *E''* and the rent will be *OR''* (= *SE''*). In case, the country is entirely new and land of good quality is surplus, then there will be no rent. This condition is shown by *D'''D'''*.

If the land is of different qualities, then each quality will have a separate demand curve and each quality of land will command different rent. Hence, the theory explains differential rent too.

Conclusion. Thus, the rent of land, like the remuneration of other factors of production, is determined by the equilibrium between demand for, and supply of land. In other words, it is scarcity in relation to demand that determines rent. Fundamentally speaking, rent is paid for land, because the produce of land is scarce in relation to its demand. The scarcity of land is in fact derived from the scarcity of its products. It is this scarcity which explains all values and rent is no exception.

Land for a Particular Use

We have analysed above the total demand and the total supply of land for the community as a whole. Let us now consider it from the point of a particular industry or use. For a particular use or industry, the supply of land cannot be regarded as fixed. By offering more rent, the supply can be increased; the supply will decrease if the rent in this particular case goes down. The supply is thus elastic and the supply curve will rise upward from the left to the right as is shown in the Fig. 33.3 DD is the



Land Rent for a Particular Use

Fig. 33.3

demand curve to start with E is the point of intersection: hence $OR (=EM)$ is the rent and OM is the land used.

Suppose the demand increases to $D'D'$. Now the two curves intersect at E' and the rent will be $OR' (=E'M')$ and the land used OM' . This means that, since for this particular use the rent of land has gone up, MM' land has been withdrawn from other uses and put to this use. If the demand for land decreases to $D''D''$, the rent will come down to $OR'' (=E''M'')$ and the quantity of land will come down to OM'' . This will mean that MM'' land has gone out of this particular use, since the rent has fallen.

MODERN THEORY OF RENT

Modern theory of rent does not confine itself to the determination of the reward of only land as a factor of production. Rent according to the modern sense can arise in respect of any factor of production. It is a surplus payment in excess of transfer earnings of that factor. We have already explained that transfer earnings means the amount of money which any particular unit of a factor could earn in its next best alternative use.

Economic rent of a factor of production is the excess over its transfer earnings, i.e., what a factor may be earning in its present employment over what it could earn in its next best employment. In other

words, transfer earnings of a factor mean what a unit of factor can earn in its next best alternative use, occupation or industry. We can also define transfer earnings as the minimum earnings which a unit of factor of production must be paid in order to induce it to stay in its present use or industry or occupation. If a factor is getting less than this minimum, it will give up its present employment and shift to its next best alternative employment. But if a factor in its present employment is earning more than the minimum necessary to keep it in that employment, the excess is called economic rent.

This concept of rent is applicable not merely to land but to all factors of production. That is, labour, capital and entrepreneurs too can earn economic rent in this sense.

Let us give some examples to explain this concept. Suppose a lecturer in economics is getting Rs. 600 per month as salary. Suppose further that his next best employment can be in a bank where he can get Rs. 500. If he cannot get Rs. 500 in a college, he will take up a job in a bank and earn that much. But since he is actually getting Rs. 600 as a lecturer in a college, he is earning Rs. 100 more than his next best alternative employment. That is, he is earning Rs. 100 as economic rent. Take another example. Suppose a piece of land is devoted to the cultivation of cane in which the owner of land is earning Rs. 1,500. If in the next alternative use, say, cultivation of cotton, it can fetch Rs. 1,200, then in its present use it is earning Rs. 300 more than its transfer earnings. This excess of Rs. 300 is surplus or economic rent. In the same manner, we can take the example of capital. Suppose one is getting 10 per cent on a certain investment of capital. Suppose further the next best investment is a fixed deposit in a commercial bank where one can get 7 per cent. This means that the present investment gives an excess of 3 per cent over its transfer earnings. This is economic rent.

From the above analysis it is clear that a unit of factor of production can earn more in its present use than its next best alternative use or the transfer earnings. Economic rent in such a case is the difference between the present earnings and transfer earnings. In Joan Robinson's words, "The essence of the conception of rent is the conception of a surplus earned by a particular part of a factor of production over and above the minimum earnings necessary to induce it to its work."³

How Economic Rent Arises

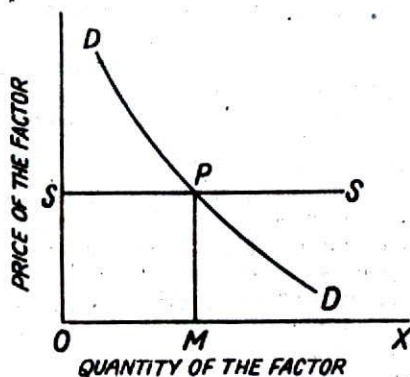
Now the question is how economic rent arises. Economic rent in the sense of surplus over transfer earnings will arise when the supply of the factor

3. *Economics of Imperfect Competition*, p. 102.

units is less than perfectly elastic or not perfectly elastic. From the point of view of elasticity of supply, there are three possibilities: (a) when the supply is perfectly elastic, (b) when it is less than perfectly elastic, and (c) when it is inelastic. Let us examine these possibilities from the point of view of rent.

(a) **When the supply of factor units is perfectly elastic.** In this case, there will be no surplus or economic rent and the actual earnings and transfer earnings will be equal. If there is any difference it will be competed away. When the supply of a factor is perfectly elastic, it means that at a given price, or remuneration, the entrepreneur can engage or employ any number of the factor units. It is obvious that, when the factor units are available at a minimum price or transfer earnings, their equilibrium price will be equal to that minimum price at which the present earnings are equal to the transfer earnings. Thus, no factor unit in such a situation will be able to earn more than its transfer earnings. That is, there will be no rent or surplus earnings.

This is shown in diagram 33.4 given below. Along OX is shown the quantity of the factor and along OY its price.



Perfectly Elastic Supply
Fig. 33.4

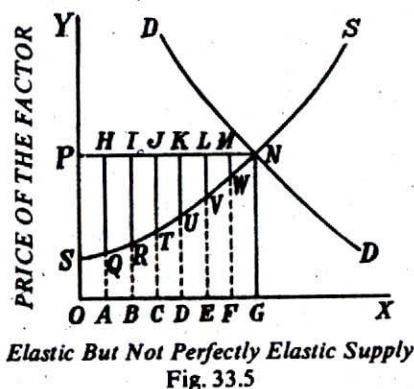
In this figure, the supply curve of the factor of production SS is perfectly elastic and is, therefore, shown as a horizontal straight line. This means that all factor units are available at the given price OS or in other words, the transfer earnings of each factor unit are also equal to OS. DD is the demand curve. The two curves intersect at P. OM is the quantity of the factor used. The price determined is OS (=PM). The total earnings are OSPM. But since transfer earnings are equal to the actual earnings, they are also equal to OSPM. There is no surplus and hence no rent. If this firm does not pay the price OS, the factor units will be shifted to some other use and earn there as much, because present earnings are equal to transfer earnings.

Thus, it is clear that if the supply of factor units is perfectly elastic for a particular use or industry, then no factor unit can earn surplus or economic rent.

Now let us have a situation when the supply of a factor is less than perfectly elastic.

(b) **Less than Perfectly Elastic Supply.** We have seen above that if the supply of a factor is perfectly elastic, there can be no rent. Now let us take a case when the supply is less than perfectly elastic, i.e., it is somewhat elastic. This means that the transfer earnings of all the factor units are not equal. As, in some industry or use, the price of the factor increases, more and more of the factor units will offer their services to this industry in use. Suppose that in a particular industry or use, a factor unit can earn Rs. 200 p.m. It is obvious that only such units of the factor will offer their services to this industry whose price in other alternative occupations is less than Rs. 200 or in other words the transfer earnings are less than the present earnings. In this manner, as the price paid for a factor in a particular industry or occupation increases, the supply of the factor will increase if the transfer earnings are less. It is clear that the supply of a factor of production depends on its transfer earnings.

This is shown in the diagram 33.5 given below.



Elastic But Not Perfectly Elastic Supply
Fig. 33.5

Along the X-axis is shown the quantity of the factor and along OY its price. SS is the supply curve sloping upwards to the right. It is somewhat elastic but not perfectly elastic as in the case (a). The supply curve SS indicates what quantity of the factor will be available at various prices. In other words, it shows the transfer earnings of different factor units. Thus, the transfer earning of A unit of the factor, is AQ whereas the price is OP. Therefore, surplus or rent is HQ. In the same manner, the other units earn surplus or rent. It is assumed that all factor units are equally useful for this industry. Hence, the price of all factor units in the industry will be the same. The supply curve cuts DD demand curve at N. In this case, OG is the quantity of the factor used. The rent or price per unit is OP (=GN). But the transfer earnings of each factor unit are less than the price.

OP. All units except the last G unit are earning more than their transfer earnings. That is they are earning economic rent. Economic rent or surplus will be different for different units because the transfer earnings are different, although the price is the same.

The total earnings are OGNP. But the transfer earnings are OGNs. If we take away the transfer earnings, we get PNS, the thick-lined area as surplus or rent.

(c) **Absolutely Inelastic Supply.** Now we come to a case when the supply of a factor is absolutely inelastic. The obvious example of this case is the supply of land for the community as a whole. We know that land for the community is fixed and it cannot be increased or decreased whatever the price offered. High price will not increase it or low price will not decrease it. That is why it is said that land has no supply price.

In the diagram 33.6, the supply curve SS shows an

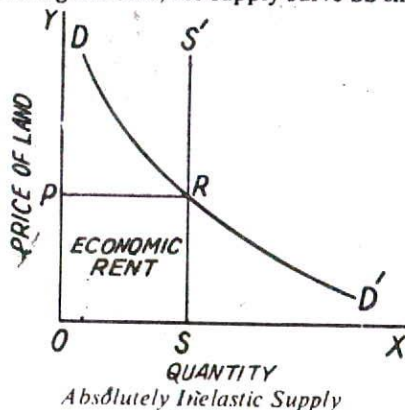


Fig. 33.6

absolutely inelastic supply which represents the supply of land for the community as a whole. Since the supply is fixed, the supply curve SS' is a vertical straight line. Even if the price of the use of land or rent falls to zero, the supply of land will remain the same. This means that from the point of view of the community as a whole, the transfer earnings are zero, since the land cannot be transferred to any other place.

Look at the Fig. 33.6 again $D'D'$ is the demand curve for the whole land. The supply curve SS' and the demand curve DD' intersect at R. In equilibrium, the price of land or rent is determined at OP and the total earnings of land are equal to OPRS area. Since in this situation the transfer earnings of land are zero, the entire earning of land, i.e., OPRS is rent. From the point of the community as a whole, there are two alternatives, viz., either the land should be cultivated or it should be kept idle. Hence, for the community, the transfer earnings of land are zero and whatever earnings are made in any use thereof constitute its rent or surplus.

Thus, it is clear that in case the supply of a factor is absolutely inelastic, its earnings are rent.

Conclusion. We may conclude that rent arises when the supply of a factor is less than perfectly elastic.

QUASI-RENT

The concept of Quasi-Rent was first introduced in Economics by Marshall. Quasi-rent, according to him, is the surplus earned by the instruments of production other than land. The term 'rent' is applied to income from land and other free gifts of nature, and quasi-rent to the income derived from appliances and machines which are the product of human effort. Quasi-rent stands for the whole of the income which some agents of production yield when demand for them has suddenly increased. It is earned during the period when their supply cannot be increased in response to an increase in the demand for them. Hence, it is a short-period concept.

The basis of distinction between rent and quasi-rent is the fact that while the supply of land, as a free gift of nature, is fixed for ever, that of the other instruments of production like building, machinery, etc., is fixed only temporarily and can be increased after some time.

Take the case of houses. During a war, the demand for houses in towns increases owing to increase in urban population. But the supply cannot be increased because of the scarcity of building materials. For the time being, their supply is as much limited as that of land. This abnormal increase in the return on capital invested in buildings is quasi-rent. It is not pure rent, because the supply of houses can be increased in the long run.

We may repeat that Quasi-rent is only a temporary surplus. With the increase in the supply of houses, as building materials become available, this surplus will tend to disappear. A similar surplus may arise in the case of other durable goods like machines, ships, etc. Similarly, quasi-rent may arise due to a temporary scarcity of a particular kind of skill which can be increased only if enough time is given.

The earnings from such durable goods like machines must, in the long run, equal the prevailing rate of interest. Temporarily, however, due to shortage, they may yield surplus earnings which are called quasi-rent.

We can apply the concept of transfer earnings also to explain quasi-rent. In the short run, a specialised machinery must remain in its present use. It cannot be transferred to any other use. This means that its transfer earnings are zero. Hence, in the short run, the whole of the earnings of machinery and capital equipment is a surplus over transfer earnings and is rent. But it is called quasi-rent because it is temporary. Their supply during the sort period is fixed, and

cannot be increased howsoever keen the demand may be. The surplus will, therefore, continue to be earned and cannot be competed away. But, in the long run, the supply can be increased and the surplus earnings will disappear. This cannot happen in the case of land whose supply is perfectly inelastic and is permanently fixed. Hence, rent of land will persist in the long run.

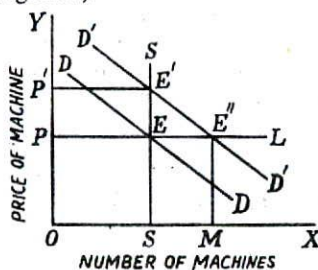
Quasi-rent has also been defined as the excess of total revenue earned in the short run over and above the total variable costs.

Thus,

Quasi-rent = Total Revenue — Total Variable Cost.

In the long run, all costs are variable and in the long run competitive equilibrium, total receipts are equal to total costs (including normal profit). There are no excess earnings over and above costs and hence no quasi-rent.

Quasi-rent can be illustrated by the following diagram (Fig. 33.7).



Quasi-Rent
Fig. 33.7

In this diagram (Fig. 33.7), SS is the absolutely inelastic supply curve of machines. It cuts the demand curve DD at E. That is, at OP (=SE) price, OS machines are supplied. If in the short run, demand increases to D'D', the price will go up to OP' (=SE') but the supply of machines remains OS. Since the number of machines is fixed in the short run, the transfer earnings are zero, the whole price is quasi-rent. It was OP before and it has become OP' now. But, in the long run, the supply is perfectly elastic and is represented by PL so that any number of machines will be supplied at OP. Supply having now increased to OM, the price comes down to E''M (=OP). Now the quasi-rent has vanished, because the price E''M just covers the supply price (or transfer earnings) which is OP.

It is thus clear that capital goods earn quasi-rent in the short run, because the supply is fixed, but it disappears in the long run, when supply becomes perfectly elastic.

Quasi-Rent and Interest

Quasi-rent, however, should be distinguished

from interest. The term 'interest' is applied to a return on free or floating capital. Quasi-rent is the return from specialised or sunk capital, *i.e.*, from old investments of capital.

It should be borne in mind that the distinction between rent and quasi-rent and, in fact, the return on any other factor of production, is only a matter of degree.

In the words of Marshall, "That which is rightly regarded as interest on 'free' or 'floating' capital or on new investments of capital, is more properly treated as a sort of rent—a quasi-rent on old investments of capital. And there is no sharp line of division between floating capital and that which has been sunk for a special branch of production, nor between new and old investments of capital; each group shades into the other gradually."⁴ Rent is thus "a leading species of a large genus."

Land, old investments and new investments they are all scarce in relation to demand. Differential surpluses arise in all of them. But since the supply of land in the absolute sense is limited in a larger degree than that of other factors, land is put in a separate category by economists. Fundamentally speaking, there is no justification for this distinction.

It is necessary to clear one misconception. Quasi-rent has been called "unnecessary profit," that is, it does not form part of costs. This misconception arises from not making a clear distinction between the long-run and short-run view. For the short period, quasi-rent may be regarded as unnecessary profit. It does not form part of price, because the appliances are already there in existence and no special costs have to be incurred. But, in the long run, several supplementary costs may have to be incurred and the businessman must be duly compensated for them. These will certainly form part of costs and must be regarded as 'necessary profit.'

Rent Element in Other Factors

The rent element is not peculiar to land. But it can be traced in other agents of production as well.

Rent Element in Profits. All entrepreneurs are not of the same ability. At one extreme are those who are just able to keep their heads above water earning low profits. But at the other extreme are those who, on account of their superior organising ability, are able to produce at a lower cost. They enjoy higher profits which are similar to the surplus enjoyed by superior lands. Corresponding to situational advantages possessed by land, certain entrepreneurs make large profits on account of birth or business influence. There is thus the rent element in profits. Profits have been called 'rent of ability'.

Rent Element in Wages. There are varying

4. *Principles of Economics*, 1936, p. 412.

degrees of efficiency to be found among the workers. Roughly speaking, wages correspond to personal efficiency. The more efficient workers enjoy a surplus or extra-wages as compared with the marginal worker. This is analogous to the rent of more accessible lands.

Rent Element in Interest. The prevailing rate of interest compensates the marginal investor, *i.e.*, one who is just induced to save. But there are others who are prepared to save even at lower rates. Such super-marginal investors enjoy a surplus (rent element). This corresponds to rent in the extensive form. But there is rent element in interest corresponding to the intensive form too. The last or marginal units of saving are just compensated by the rate of interest; the earlier units yield a surplus which is like rent.

In another sense also all factors of production yield rent. There is a minimum payment necessary to induce a factor to remain in a given employment. Any earning over and above this minimum amount is of the nature of rent. This can be traced in all factors.

Thus, rent is "a leading species" of a large genus."

RELATION BETWEEN RENT AND PRICE

According to Ricardo, rent does not enter into prices. But the modern theory of rent deviates from the Ricardian theory in this respect. Ricardo held that rent did not enter into the cost of production of the produce and hence did not affect its price. Rent rises because of the rise in price and not the other way round. We have seen that, according to this theory, rent is surplus over cost. Price is determined by the cost of production at the margin where there is no rent. Hence, rent does not enter into price. The marginal dose of labour and capital just pays for itself and leaves no surplus.

In fact the position of the margin, according to the Ricardian theory, is determined by price and not price by the position of the margin. Rent is thus not a part of price. Rent is price-determined and not price-determining.

It is true that the differential aspect of land is not a cause of price. If land A gives more valuable service than land B, the extra payment yielded by land A would not affect the price of the produce. In fact, here we are considering two different things, each being paid according to its efficiency or marginal productivity. There can be a similar differential aspect of wages, and we can argue that such differential payments do not affect the price of the product of labour. But wages do enter into price.

In another sense also, it may be argued that rent does not form a part of price. Land is free gift of nature. No payment is necessary to maintain the total supply of land. In this sense, rent is not a part of the supply price of land and consequently of the

products of land. Since the supply of land is inelastic, it will always work for whatever it fetches under competition. Thus, the value of land completely depends on the value (*i.e.*, price) of the product, and not vice versa.

Suppose land is specific and it can be put to only one use and to none else, *e.g.*, when land can be used for production in one dominant industry only. It is again a case of inelastic supply to a particular industry, since land has no other use. Here also it will work for whatever it can get. In this case also rent will be price-determined and not price-determining.

When Rent Enters Price

When we are thinking, not of all the lands of the country, but of the land available for particular uses, rent does form an element of cost and hence affects price. This is clear from the conception of the opportunity cost. Most of the land is capable of being put to alternative uses. If it is put to one use, it is not available for another use. The minimum price that has to be paid for the use of land is the amount which this land could earn in its most profitable alternative use. This is called opportunity cost or transfer price. This payment for the use of land enters into price.

From the point of view of an individual farmer, prices paid for all factors (including the price for the use of land) must of course be included in the cost of production, and must, therefore, enter into price. If a farmer is using land belonging to someone else, the rent that he pays is obviously a cost for him. In the case of an owner-cultivator, too, the rent as a cost is there, only its presence is obscured. The payment that he could have received, if he did not cultivate himself, is the opportunity cost of this land.

The problem may be approached in another way. Prices are determined by the scarcity of the products in relation to demand. The rent that an entrepreneur pays is part of his cost of production. If the rent is high, the entrepreneur will tend to hire less land; and, conversely, he will use more land, if rent is low. If he uses more land the supply of land for other purposes is reduced; if he uses less land, the supply available for other purposes is increased. Thus, rent, by influencing relative scarcity of land for different uses; affects the prices of different products.

Thus, to use Samuelson's words, "What is a price-determined rent return to a factor which is inelastic in supply to the whole community or a dominant industry may, to each firm and to any small industry that is only one of many potential users, appear as a price-determining cost."⁵

Conclusion

In the last analysis, however, as Davenport points

5. Samuelson, P. A.—*Economics*, 1970, p. 540.

out, rent neither determines price nor is determined by price. Both price and rent are governed by the relative scarcities of the products of land. They both vary with the changes in this relative scarcity. The same principle applies to wages, interest and profits.

Rent and Economic Progress

How is rent affected by economic progress? Broadly speaking, economic progress may be indicated by (a) technical advances in methods of production, (b) improvements in transport, and (c) growth of population.

(a) Improved methods of agricultural production may affect all agricultural lands equally, or they may affect only the better lands or only the marginal lands.

If all lands are affected equally, the supply of the produce will increase, demand remaining the same, its price will, therefore, fall; marginal lands will go out of cultivation and, according to the Ricardian theory, rents of superior lands will fall.

If the improvements only affect the marginal land, according to the Ricardian theory, rents will again fall because the surplus of the superior land over the marginal land will be reduced.

If, on the other hand, the improvements affect only the best lands, their rents will rise, since their productivity, compared with that of the marginal land, will be higher. But if, in this latter case, the higher productivity of superior lands depresses the prices of the product, rents may fall again.

Thus, ultimately rents will depend upon the relation between supply and demand of the product of the land. **Agricultural improvements will affect rent not so much by affecting the margin of cultivation as the degree of scarcity of land or its products.**

(b) As regards transport improvements, according to the Ricardian theory, better transport will affect rent by affecting the factor of situation. Distant lands will become nearer the market as it were. Their rents will rise. The rents of more favourably situated lands will fall. This can also be explained in terms of scarcity. If improved transport increases scarcity (by leading to exports from the given region), rents will rise; in the contrary circumstances, the rents will fall.

(c) As for the growth of population, its result is to raise rents. According to the Ricardian theory, this will be due to more intensive cultivation of old lands and extension of cultivation to inferior lands, thus leading to a fall of the margin of cultivation. According to the more recent way of putting it, increase in population will increase the relative scarcity of land in relation to the demand for its products, and hence raise rents.

RENT IN A SOCIALIST STATE

In a socialist state, private ownership in land will

end, but a socialist state cannot do away with the concept of rent. Change of economic order from capitalism to socialism will not turn scarce land into an unlimited quantity. Rent as an index of productivity will help in the best allocation of the available land among its various uses. An accounting price-tag will have to be put on the various types of land, the good land being put a higher tag. As Samuelson puts it, "Only by putting a price upon inert sweatless land are we using it, and sweating breathing labour, most productively! The price or rent of land rises so as to ration its limited supply among the best uses." Thus, rent makes for the most efficient use of the land resources of the community. In this way, it solves the problems of FOR WHOM and HOW of the society. In countries, like India, where land is scarce and rents are high, farmers will take to cheaper labour-intensive production. On the other hand, as in the U.S.A., where labour is relatively more expensive, the farmers will use land-intensive methods. It is rent which indicates the relative scarcity and signals the best factor substitutions.

Under socialism too, the community must make the most efficient use of its resources.

Even a socialist state will have to direct land from one use to another so as to make its marginal productivity in all uses the same. The marginal productivity will be indicated by rent.

Thus, the existence of rent ensures a correct allocation of valuable human or non-human resources.

ECONOMIC RENT AS A BASIS FOR A SINGLE LAND TAX

Henry George, an American economist in his well-known book 'Progress and Poverty' (1879) advocated a single tax on land rent to replace the numerous other taxes which were then being levied. From the social point of view, land is a free gift of nature and as such land rent is not necessary to ensure its supply. Hence Henry George argued that taking away rents in their entirety would not affect the supply of land or impair the working of the economy. He justified such a tax on the ground that land rents went on rising as a result of social progress and not in no way due to the personal efforts of the land owners. The incidence of tax or rent is entirely on the land owners for whom this represents unearned income. A tax levied on a completely inelastically supplied good is absorbed entirely by the suppliers of that good. Hence Henry George argued that a single tax on land rent could take the place of numerous other taxes without producing any ill effects on the economy. Besides, it would not distort the allocation of land usage, because it will leave the supply and demand for land unaffected.

However, the proposal had little practical value.

What is Interest?

Perhaps interest has been the most controversial topic in the whole theory of distribution. Economists have differed regarding the nature of interest as well as how it is determined. In fact, the subject of interest is still an unsettled question of Economics. The very definition of interest depends on the interest theory which one accepts. Those, who believe in the classical or real theory, regard interest as payment for the use of capital goods. They also believe that interest is necessary to induce people to save. The followers of liquidity preference theory believe that interest is a price for surrendering liquidity preference. Still others who accept the loanable funds theory hold that interest is the price paid for the use of loanable funds.

Just as rent is a payment for the use of land, similarly interest is a payment for the use of capital. In Marshall's words, interest is "the price paid for the use of capital in any market."¹ Just as wage is the price of the service of labour, similarly, interest is the price of capital. It is expressed as a percentage return on capital invested after allowing for risks of investment.

Samuelson defines interest thus: "The market rate of interest is that percentage return per year which has to be paid on any safe loan of money, which has to be yielded by any safe bond or other type of security, and which has to be earned on the value of any capital asset (such as a machine, a hotel building, a patent right) in any competitive market where there are no risks or where all risk factors have already been taken care of by special premium payments to protect against risk."²

In this chapter, we shall discuss all these divergent views about the nature of interest and determination

of its rate, and shall attempt a synthesis among them.

Gross and Net Interest

Much of what is called interest is not the price of the service of capital as such. What people generally call interest is really gross interest. Pure interest is only a small proportion of gross interest as explained by this fact. Pure interest tends to equality in different employments, provided competition is free and perfect.

Gross Interest. The total amount which the debtor pays to the creditor is known as gross interest as already indicated. All that the borrower pays to the lender is not pure or net interest, *i.e.*, the price paid for the services of capital only. It is mixed up with so many other elements, of which Pure or Net Interest is only one. Gross interest consists of the following elements:—

(a) **Pure interest.** This is a payment only for the services of capital as such or for the money borrowed.

(b) **Insurance against Risk.** The lender is exposed to risk when he lends money. A certain amount must be paid to him to cover these risks. These risks are of two kinds: (i) personal risks due to the unreliable character of the borrower himself, and (ii) trade risks. Trade risks are due to the varying fortunes of the business in which the money is invested. Thus, the lender may fear that either the borrower may refuse to pay back the loan and the interest on it or he may lose the money in his business. The greater the risk of this kind, the higher will have to be the insurance money that the lender will expect before he is willing to lend.

(c) **Wages of Management.** A part of the payment may be due to the wages of management. The lender has to keep accounts and to arrange for new loans for short period. Money-lending is his whole-time job.

1. *Principles*, (8th Edition), p. 534.

2. Samuelson, P. A.—*Economics*, 1970, P. 595.

(d) **Return for Inconvenience.** A lender has to suffer certain inconveniences for which he seeks compensation in the form of interest. When a man lends money, he loses command over it for a period. He is unable to make use of it for himself, if he wants to. Favourable opportunities for its investment may also slip by. Inconvenience is of two types: The lender may not get back the money when he needs it and he may have to borrow from others on interest. Or, he may get the money when he cannot find for it some safe and remunerative investment so that the money lies idle and he loses interest. He must, therefore, compensate himself for such losses. Hence, he charges something extra over and above pure or net interest.

Keeping, these facts in mind, it will become clear why the village money-lender's rates of interest are high in spite of competition among the lenders. This is due to the fact that he faces greater risk and inconvenience. Similarly, the pawn-brokers seemingly charge very high rates when calculated on an annual basis. But they have to undergo a lot of trouble in their business in keeping small individual accounts, etc. This also explains why governments, especially those with sound financial traditions, can borrow at extremely low rates of interest. Here the risk and inconvenience to the lender are negligible and the interest paid is mainly "pure interest". The return on government securities is an example of pure rate of interest.

Differences in Interest Rates

There is an element of monopolistic competition in the money market inasmuch as different borrowers are charged different rates of interest. It may be noted, however, that **pure interest tends to be the same, if calculated over the same period of time in the same money market. The actual differences that prevail are differences in gross interest.** In other words, they are due to difference in the degrees of risk involved and inconveniences suffered by the lenders.

Differences in Pure Interest

Pure interest may be different in different investments, when the market is not the same owing to the following reasons:—

(a) **Differences Due to Distances.** People are usually more willing to invest their capital nearer home than at a long distance. This may create differences in supply and demand due to the comparative immobility of capital.

(b) **Differences Due to Time.** If people have to part with their money for a longer period, they expect higher rate of interest even though risks and other factors are the same. Of course, if money is lent for very short periods and has to be re-lent

again and again, the inconvenience of management will increase gross interest. But that will not be net interest.

(c) **Differences Due to the Amount of the Loan.** It is generally seen that the rate of interest varies inversely with the amount of the loan. The rate decreases as the amount of the loan increases, and vice versa.

(d) **Differences in Liquidity.** By 'liquidity' we mean the ease with which the loan given can be called back or the ease with which an asset like securities can be sold without monetary loss when its owner requires cash. The sale of some securities may involve delay, cost or capital loss, while others may be readily disposed of without such cost or monetary loss. The more liquid securities will carry a lower rate of interest and vice-versa.

(e) **Maturity Period.** Another reason for differences in interest rates is the maturity of the loan, i.e., the length of the loan. Other things being equal, long-term loans will carry higher rates of interest than do short-term loans, since the long-term lenders suffer greater inconvenience and possible financial sacrifice of foregoing alternative uses for their money for a longer period of time. On a long-term loan, there is a risk of drop in the value of the securities or a possible rise in the price level. The lender must, therefore, charge more to cover the possible loss. Hence, a long-term bond will have to offer a higher rate of interest than a short-term bond, other things being equal.

(f) **Market Imperfections.** Another factor which explains some of the differences in interest rates is the market imperfections or monopoly element. A bank in a small town, which has a monopoly in the local money market, may charge a higher rate of interest from the local people, because there the people find it very inconvenient to "shop around" at banks in distant cities. On the other hand a large corporation (i.e., joint-stock company) being not confined to a local market enjoys a more competitive position and can survey or shop around all possible lending houses. Therefore, a large nationally known corporation can sell its bonds on favourable terms.

Differences in Gross Interest

Differences in gross interest may, in addition to the causes already considered, be due to:—

(i) **Differences in Social Esteem.** A person with better reputation or integrity can borrow at lower rates. This is partly due to the element of lower personal risk already considered.

(ii) **Differences in Productivity.** Where capital can earn greater reward for the producer, he will be willing to pay higher interest. Such trades are usually more speculative and higher interest can be attributed to higher risks. Moreover, under perfect

competition, as we shall see, marginal productivity of capital, and hence the pure rate of interest, tends to equality. Higher rewards in particular employment, if not justified by greater risks or inconvenience, tend to disappear through force of competition. But if competition is imperfect, such differences may persist.

Productivity of capital may differ in different countries. In new States, for instance, the demand for capital is great, while the supply is very limited. Capital thus has a high marginal productivity, and interest is high. This is due to the immobility of capital over distance as already noted. Partly it is due to greater risk. In the latter case, it is gross interest which will be higher, and not net interest.

Thus, ultimately differences in interest rates can be reduced to differences in inconvenience or risks of lending except in cases where we are not dealing with the same market. **Pure interest tends to be the same in the same market.**

To sum up: "The interest rate charged on individual business is usually determined in personal negotiation between bank and borrower. It reflects such attributes as the borrower's size and general credit standing, his access to alternative credit sources, the size and maturity of the loan, the character of the borrower's business, the value to the bank of his deposit account and of other business relationships, and the nature of the security, if any, to be pledged. Certain other factors not related directly to the borrower of the loan but rather to the lending bank or perhaps the banking structure can also be shown to have some effect on interest rates charged for bank business loans. These are the size of the lending bank, the size of the centre in which the bank is located, and the area of the country where the loan was made."³

THEORIES OF INTEREST

A theory of interest should explain two things: (a) How interest arises and (b) how the rate of interest is determined. Some theories of interest explain only one of these aspects whereas others explain both. For instance, Productivity Theory, Abstinence or Waiting Theory, Agio Theory and Fisher's Time Preference Theory are all classical theories which explain only why interest arises or why interest is paid. On the other hand, the Classical Theory, the Loanable Funds Theory and Liquidity Preference Theory explain both how interest arises as well as how the rate of interest is determined.

We now start the discussion of the theories of interest with the first set of theories which explain how interest arises or why interest is paid.

3. Richard Youngdahl on "Structure of Interest Rates", *Federal Reserve Bulletin*, July 1947.

HOW INTEREST ARISES

Productivity Theory

The Theory. To take the Productivity Theory first, some older economists thought that capital was productive of goods in the same sense as land was productive of crops. They held that interest was paid because production with the aid of capital was greater than without it. Capital no doubt is productive; but not in the same sense as a flock of sheep or fertile land is productive. Capital is productive in the sense that labour, assisted by capital, produces more than labour without capital. A fisherman with a net can catch more fish than without it. A farm labourer with a tractor can produce more than without a tractor.

Criticism. Physical productivity of capital goods, however, does not explain interest. If people were willing to lend unlimited amounts of money without interest, business would expand up to a point where the falling price of the product would simply cover other charges (without any interest) in making thereof. **Interest would not be a cost.** But interest is a cost which every entrepreneur must reckon with. Hence, price, in the long run, must cover all costs including interest. Since demand for capital, when it is free of interest, must exceed the supply of it, interest is bound to emerge. Even if some lenders were willing to lend without interest, prevalence of competition would ensure that loans would be advanced only to those who would pay interest rather than to those who would not.

Hence, it is scarcity rather than productivity which explains interest.

Further, if interest depended merely on productivity (not marginal productivity), interest rates should vary in proportion to the productiveness of capital. In actual fact, pure rate of interest tends to be the same in the same market.

Moreover, if capital helps labour to produce more, it has to be seen how much of this extra production is due to capital and how much to labour, since capital without labour produces nothing.

Then, what about loans for consumption purposes? They are not productive, but interest has to be paid on them all the same.

Again, it may be pointed out that the productivity theory attempts to explain interest from the side of demand only. It ignores the supply side altogether. It is thus a one-sided explanation.

As we shall see later, the above objections can be overcome, if we talk in terms of marginal productivity and not mere productivity of capital.

Conclusion. Thus, it is clear that it is not productivity, as such which explains why interest is paid.

Abstinence or Waiting Theory

The Theory. Another theory of interest is the

Abstinence Theory. Whereas the Productivity Theory tries to explain interest from the side of demand; the Abstinence Theory approaches the problem from the side of supply. It was Senior who first pointed out that saving involved a sacrifice or "abstinence" as he put it. Saving was an act of abstaining from consumption. Since to abstain was painful, it was necessary to reward people for this act. This reward was in the form of the interest paid to those who saved, rather than consumed their incomes, or a part of their incomes.

Criticism. The idea of abstinence was widely criticised on the ground that it suggested positive discomfort, while the rich people, who are the main source of capital, save without the least inconvenience. That is why Marshall substituted the term "waiting" for "abstinence". Saving implies waiting. When a person saves, he does not refrain from consumption for all time; he merely postpones present consumption to a future date. Meanwhile, he has to wait. But since most people do not like to wait, an inducement is necessary to encourage this postponement of consumption. Interest is this inducement.

Some "waiting", however, may be forthcoming without any inducement in the way of interest payment. Other people will "wait" even with a negative rate of interest. But aggregate of savings thus made will not be adequate to meet the total demand for capital. Interest must be paid to induce many more people to save in the case of whom "waiting" does involve inconvenience. The rate of interest must be high enough to bring forth the marginal increment of saving or "waiting" in order to meet the aggregate demand for capital. The rate of interest will be fixed at a level at which the supply of "waiting" will be equal to the demand for it.

This theory has a considerable element of truth in it, but it does not clearly analyse forces acting on the side of demand for capital.

Austrian or Agio Theory

This theory is also called Psychological Theory of interest. First advanced by John Rae in 1834, this theory was given its final shape by Bohm Bawerk of the Austrian school of economists. It became popular later among some American economists, like Fisher, with slight modifications.

The Theory. The gist of the theory of Bohm Bawerk, is that interest arises because people prefer present goods to future goods, and that, therefore, there is an 'agio' or premium on present goods. One bird in hand is said to be worth two in the bush. The present gratification is attached greater importance than the future satisfaction.

In other words, future satisfactions when viewed from the present angle undergoes a discount. Interest is this discount which must be paid in order to induce people to lend money or postpone present

satisfactions to a future date. That is, interest is intended to equate future satisfaction with present satisfaction. To induce a person to part with Rs. 100 now, it is not enough to give him a promise of the return of only Rs. 100 after a year. The borrower must pay more, otherwise the lender will feel that he is a loser. The human mind is so constituted that there is commonly an under-valuation of the future purchasing power as compared with the present purchasing power.

Why do people prefer present satisfaction to future satisfaction? Bohm Bawerk gave three reasons for this fact. One is the "prospective underestimate for the future". The future is less clearly perceived than the present. It is uncertain. In the second place, present wants are felt more keenly than the future wants. The result is that the demand for present goods is greater than that for future goods. Thirdly, present goods possess "a technical superiority over future goods." This is so because the passage of time allows the use of more roundabout methods of production, which are more productive.

Another reason is that one may hope to improve his economic position in future as a result of which the marginal utility of his income will decline. He will, therefore, prefer to use his income at the present when the marginal utility of his income is high.

There is still another reason for a consumer to prefer present satisfaction to future satisfaction. He may expect growth of the economy making more and better goods and services available when the general standard of living will be higher indicating declining marginal utility of money.

These reasons, however, do not apply to all, at any rate, to the same extent.

Fisher's Time Preference Theory

Fisher emphasizes the fact of "time preference" as the central point in the Theory. Individuals prefer present satisfactions to equally certain future satisfactions. They are thus impatient to spend their incomes now. The degree of impatience depends upon the size of the income, the distribution of income over time, the degree of certainty regarding enjoyment in the future and the temperament and the character of the individual. Thus, people with larger incomes are likely to have their present wants more fully satisfied and will thus discount the future at a lower rate than poorer people.

As regards distribution of income over time, three kinds of situations may be imagined. The income may be uniform throughout one's life, or increase with age, or decrease with age. If it is uniform, the degree of impatience to spend (*i.e.*, the rate of discounting the future) will be determined by the size of the income and the temperament of the individual. If the income increases with age, it

means the future is well provided for, and the tendency will be to discount the future at a higher rate. If the income decreases with age, the converse will be true, *i.e.*, the future will be discounted at a lower rate.

As for the degree of certainty, it is clear that the greater the certainty of future enjoyment of income the smaller the degree of time preference or the rate of discounting the future, and vice versa.

Finally, the character of the individual will also influence this time preference. A man of forethought will discount the future at a lower rate compared with a spendthrift. The rate of time preference is also influenced by expectation of life. If a man expects to live long, his preference for purchasing power in the present will be comparatively low. Similar will be the case of a man who desires to leave behind a handsome patrimony.

The rate of time preference of an individual also depends on the proportion of his income that he is asked to lend. The larger the amount of loan, the higher must be the rate of interest, because it entails a greater sacrifice of liquidity on the part of the lender.

Thus, the rates of individual time preference, after having been determined in this way, tend to become equal to the rate of interest. An individual with a higher rate of time preference compared with the market rate of interests tends to borrow money in order to satisfy his more pressing wants. If his rate of time preference is lower than the market rate of interest, he will lend to the market and make a gain thereby. Thus, the individual will vary his income-stream by borrowing or lending. This process will tend to equalise the rate of interest with the rate of time preference.

Liquidity Preference Theory

According to Keynes, interest is not a reward for waiting, nor is it a payment for time preference. The rate of interest is a reward for parting with liquidity. This theory not only explains why interest arises, but it also explains how the rate of interest is determined. We shall discuss this theory in a later section of this chapter.

HOW RATE OF INTEREST IS DETERMINED

Among the theories which seek to explain the determination of the rate of interest we might mention:

- (i) Classical or Real Theory.
- (ii) Loanable Funds or Neo-Classical Theory, and
- (iii) Keynesian or Liquidity Preference Theory.

All these theories of interest seek to explain the determination of the rate of interest through the equilibrium between the forces of demand and

supply. In other words, all these three theories of interest are 'demand and supply theories' with rate of interest as the mechanism which brings about equilibrium between demand and supply. The difference among the theories of interest lies in the answer to the question: demand for what and supply of what?

According to the Classical Theory of interest, rate of interest is determined by demand for **saving to invest** and supply of **savings**. Loanable Funds theory seeks to explain the determination of the rate of interest through the equilibrium between demand for **loanable funds** and supply of **loanable funds**. Apart from current savings, loanable funds include other things also. Keynesian Theory of interest asserts that the **rate of interest** is determined by the demand for **money** to hold (*i.e.*, liquidity preference) and the supply of **money**.

We will now discuss these theories one by one.

CLASSICAL OR REAL THEORY OF INTEREST

There was no unanimity among the classical economists on the point as to how interest arises or why interest is paid. They had different views on this point. But they all agreed that the rate of interest is determined by the equilibrium of savings and investment.

The classical theory of interest also goes by the name of real theory as it seeks to explain the determination of the rate of interest by real factors like productivity and thrift, *i.e.*, productivity of capital goods and saving of goods. According to this theory, the rate of interest is a payment for saving. The rate of interest is, thus, determined by the demand for savings to invest in capital goods and the supply of savings. Let us explain these demand and supply sides.

Demand for Savings

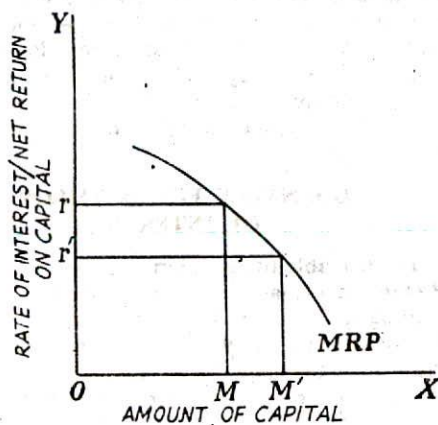
The demand for capital goods comes from firms which desire to invest, that is, to purchase or to make new capital goods. Capital goods are demanded because they can be used to produce consumer goods—because they have a revenue productivity like all other factors. For any given type of capital asset, *e.g.*, a machine, it is possible to draw a marginal revenue-productivity curve showing the addition made to the total revenue by an additional unit of a machine at various levels of the stock of that machine.

We have said that capital, like other factors of production, has marginal revenue productivity. But the marginal revenue productivity of capital is a more complex concept than that of other factors, because capital has a life of many years. A capital asset continues to yield returns for many years. Therefore, the entrepreneurs have to take into consideration the uncertainties of the future and

estimate the percentage yield or returns from capital after making allowance for maintenance and operating costs. In other words, they have to find out the **net expected return** of a marginal unit of capital expressed as percentage of the cost of the capital asset. The more capital assets of a given kind an entrepreneur has, the less revenue or income he will expect to earn by purchasing one more machine of the same kind. Therefore, the marginal revenue productivity curve of capital slopes downwards towards the right. (See Fig. 34.1 below).

Now, under perfect competition, it is profitable for a firm to purchase any factor up to the point at which the price of that factor equals its marginal revenue productivity. The price of the savings required to purchase the capital goods is obviously the rate of interest. Hence, the entrepreneur will demand capital goods or (which is the same thing) 'will demand savings to purchase capital goods' up to the point at which the expected net rate of return on the capital goods equals the rate of interest. Since the marginal revenue productivity curve of capital slopes downwards, it follows that, as the rate of interest falls, more capital goods will be demanded and also more money will be required to purchase these capital goods.

The way in which the demand for capital goods rises as the rate of interest falls is shown in Fig. 34.1,



Demand for Capital
Fig. 34.1

where MRP, is the marginal revenue productivity curve. On the Y-axis net rate of return on capital and the rate of interest are shown, while X-axis represents the amount of capital.

At O_r rate of interest, OM amount of capital is demanded. This is so because only at OM amount of capital, the falling net rate of return on capital becomes equal to the prevailing rate of interest O_r . Now, if the rate of interest falls from O_r to $O_{r'}$, the amount of capital demanded will increase from OM

to OM' , since at OM' the falling net rate of return equals the new interest rate $O_{r'}$.

Thus, it is clear that the marginal revenue productivity curve of capital shows the demand for capital and further that the demand curve for capital (or demand for savings to buy the capital) slopes downwards towards the right. This is true of individual firms, of individual industries and of the community as a whole.

Thus, we conclude that demand for individual capital goods and for capital goods in general will increase as the rate of interest falls.

Supply of Savings

According to this theory, the money which is to be used for purchasing capital goods is made available by those who save from their current income. By postponing consumption of a part of their current income, they release resources for productive purposes. Savings involve the element of waiting for the future enjoyment of savings. But people prefer the present enjoyment of goods and services to the future enjoyment of them. Therefore, if people are to be persuaded to save money and to lend it to entrepreneurs, they must be offered some interest as reward. More savings the people will do, the more consumption they will have to postpone, the higher must be the rate of interest they will ask to make such a postponement worthwhile. Thus, to induce people to save more, higher rates of interest must be offered.

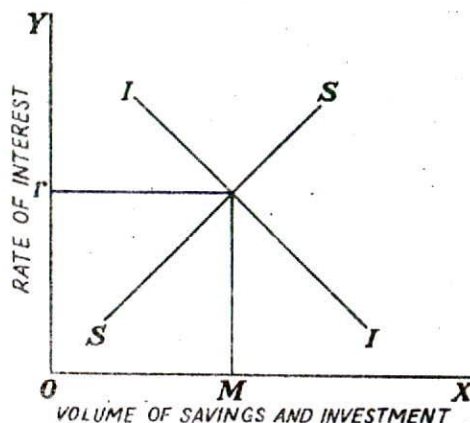
Moreover, higher rates of interest have also to be paid if savings have to come from those persons whose rates of time-preference are relatively more strongly weighed in favour of present satisfactions. The supply curve of capital will, therefore, slope upwards to the right.

Equilibrium between Demand and Supply

The rate of interest is determined by the interaction of the forces of demand for capital (or investment) and the supply of savings. The rate of interest at which the demand for capital (or demand for savings to invest in capital goods) and the supply of savings are in equilibrium will be the rate determined in the market.

How the rate of interest is determined by the interaction of demand for investment and supply of savings is shown in Fig. 34.2, where SS is the supply curve of savings and II is the demand curve of savings to invest in capital goods (II is also called demand curve for investment or simply investment demand curve). The demand for investment and supply of savings are in equilibrium at O_r rate of interest, where the curves intersect each other. Hence, O_r is the equilibrium rate of interest, which will come to stay in the market. In this equilibrium position, OM amount of money is lent, borrowed

and invested. If any change in the demand for investment and supply of savings comes about, the



Demand and Supply Equilibrium

Fig. 34.2

curves will shift accordingly, and, therefore, the equilibrium rate of interest will also change.

Criticism

The Classical Theory of Interest came in for serious criticism at the hands of Keynes. Firstly, it is pointed out that Classical Theory of interest is based upon the assumption of full employment of resources. In other words, it assumes that an increase in the production of one thing must mean the withdrawal of some resources from the production of other things. "Within the framework of a system of theory, built on the assumption of full employment; the notion of interest as a reward for waiting or abstinence is highly plausible. It is the premise that resources are typically fully employed that lacks plausibility in the contemporary world."⁴ If at any time in the country, unemployed resources are found on a large scale, there is no need for paying people to abstain from consumption, *i.e.*, to wait in order that more resources should be devoted to the production of capital goods.

Secondly, according to the Classical Theory of interest, more investment (production of capital goods) can take place only by curtailing consumption. Greater the reduction of consumption more are the savings and, therefore, more investment. But a decrease in the demand for consumer goods is likely to lessen the incentive to produce capital goods and therefore, will affect investment adversely.

Thirdly, by assuming full employment, the classical theory has neglected the changes in the income level. By neglecting the changes in the income level, the classical theory is led into error of viewing the rate of interest as the factor which brings equality of

savings and investment. As Keynes asserts, equality between savings and investment is brought about not by changes in the rate of interest but by changes in the level of income.

Fourthly, according to the Classical Theory, the investment demand schedule can change or shift without causing a change or shift in the savings curve schedule. For example, according to classical theory, if investment demand schedule or curve II shifts downwards, then the new equilibrium rate of interest will be determined where this new investment demand curve cuts the old savings curve which has remained unchanged. But this is wrong. As we know from Keynesian Economics, fall in investment leads to decrease in income and out of the reduced income, less is saved and therefore savings curve also changes. Thus, we see that classical theory ignores the effect of changes in investment on savings.

Lastly, the Classical Theory, as pointed out by Keynes, is indeterminate. Position of the savings schedule or curve depends upon the income level, that is, the position of the savings curve or schedule will vary with the level of incomes. There will be different savings schedules for different levels of incomes. As income rises, for example, the savings schedule or curve will shift to the right. Thus, we cannot know what the rate of interest will be unless we already know what the income level is. And we cannot know the income level without already knowing the rate of interest, since a lower interest rate will mean a large volume of investment and so via the multiplier, a higher level of real income. The classical theory, therefore, offers no solution and is indeterminate.

LOANABLE FUNDS THEORY OF INTEREST

The loanable funds theorists believed in the Time Preference explanation of how interest arises. According to Loanable Funds Theory, also called the Neo-classical Theory, interest is the price paid for the use of loanable funds. Like the Classical and Keynesian theories of interest, it is also a demand and supply theory. It asserts that rate of interest is determined by the equilibrium between demand and supply of loanable funds in the credit market. There are several sources of both supply and demand of loanable funds which we discuss below.

Supply of Loanable Funds

The supply of loanable funds is derived from four basic sources, namely, (a) savings, (b) dishoarding, (c) bank credit, and (d) disinvestment.

(a) **Savings.** Savings by individuals or households constitute the most important source of loanable funds. In the loanable funds theory, savings are looked at in either of these two ways, firstly, as

4. Dillard—*Economics of J. M. Keynes*, p. 162.

ex-ante savings, *i.e.*, savings planned by individuals at the beginning of a period in the hope of expected incomes and anticipated expenditures on consumption; or secondly in the Robertsonian sense savings of the difference between the income of the preceding period and the consumption of the present period. In either case, the amount saved varies at various rates of interest. Savings by individuals and households primarily depend upon the size of their incomes. But, given the level of income, savings vary at various rates of interest. More savings will be forthcoming at higher rates of interest, and vice versa.

Like individuals, businesses also save. A part of the earnings of a business concern is consumed as declared dividends; the undistributed part constitutes business or corporate savings. Such savings depend partly upon the current rate of interest. A high rate of interest is likely to encourage business savings as a substitute for borrowings from the loan market. But these business savings are often demanded for investment purposes by the firms themselves and, therefore, they do not enter the market for loanable funds.

(b) **Dishoarding.** This is another source of loanable funds. Individuals may dishoard money from the hoarded stock of the previous period. Thus, cash balances, lying idle in a previous period, become active balances in the present period and are available as loanable funds. At higher rate of interest, more will be dishoarded. At very low rates of interest, there is a greater tendency to hold on to money.

(c) **Bank Credit.** The banking system provides a third source of loanable funds. Banks, by creating credit money, can advance loans to the businessmen. Banks can also reduce the amount of money by contracting their lending. The new money created by the banks in a period adds greatly to the supply of loan funds. The supply curve of funds provided by banks is to some degree interest-elastic, *i.e.*, it varies with various rates of interest. Generally speaking, the banks will lend more money at higher rates of interest than at lower ones, other things remaining the same.

(d) **Disinvestment.** Disinvestment is the opposite of investment and takes place when, due to the structural changes or bad venture, the existing stock of machines and other equipment is allowed to wear out without being replaced or when the inventories are drawn below the level of the previous period. When this happens, a part of the revenue from the sale of the products, instead of going into capital replacement, flows into the market for loanable funds. "Disinvestment is encouraged somewhat by a high rate of interest on loanable funds. When the rate is high, some of the current capital may not produce a marginal revenue product to match this rate of interest. The firm may decide to let this

capital run down and to put the depreciation funds in the loan market."⁵ Thus, disinvestment adds to the supply of loanable funds. In Fig. 34.3, DI is the disinvestment curve and slopes upwards to the right.

By the lateral summation of the four curves, savings (S), dishoarding (DH), disinvestment (DI), and bank credit (BM), we get the total supply curve of loanable funds which slopes upwards to the right showing that a greater amount of loanable funds will be available at higher rates of interest, and vice versa.

Demand for Loanable Funds

The demand for loanable funds comes mainly from three fields: (i) investment, (ii) consumption and (iii) hoarding.

The bulk of demand for loanable funds comes from business firms which borrow money for purchasing or making new capital goods, including the building up of inventories. Demand for loanable funds for investment purposes by business firms is the most important constituent of total demand for loanable funds. The price of the loanable funds required to purchase the capital goods is obviously the rate of interest. It will pay businessmen to demand loanable funds up to the point at which the expected net rate of return on the capital goods equals the rate of interest. Businessmen will find it profitable to purchase larger amounts of capital goods, when the rate of interest (*i.e.*, the price of the loanable funds) declines.

Thus, the demand for loanable funds for investment purposes is interest-elastic and slopes downwards to the right. The demand for loanable funds for investment purposes is represented by curve I in Fig. 34.3 given on the next page.

The second big demand for loanable funds comes from individuals or households who want to borrow for consumption purposes. Individuals or households demand loanable funds when they wish to make purchases in excess of their current incomes and cash resources. Generally, the loans for consumption are demanded for buying durable goods like automobiles, refrigerators, radios, television sets; *etc.* Lower rates of interest will encourage some increase in consumer borrowing. Demand for loanable funds for consumption purposes is shown by the curve 'C' in Fig. 34.3, which is interest-elastic and slopes downward to the right.

Lastly, the demand for loanable funds may come from those who want to hoard money, *i.e.*, to satisfy the liquidity preference. Hoarding signifies the people's desires to hold their savings as **idle cash balances**. An important point to be noted here is that the one who supplies the loanable funds is the same person who demands the loanable funds for

5. Bober, M.M.—*Intermediate Price and Income*
1955, p. 271

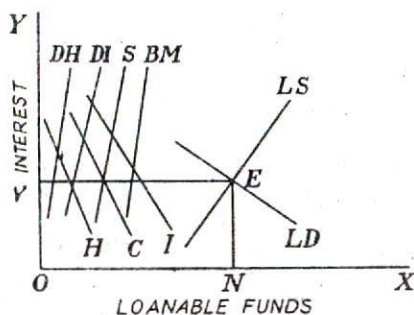
hoarding. A saver who hoards his savings can be said to be supplying loanable funds and also demanding them to satisfy his liquidity preference.

Demand for hoarding is shown by curve H in Fig. 34.3. The demand for hoarding money is interest-elastic and slopes downwards to the right. At higher rates of interest, people will hoard or hold less money, because much of the money will be lent to take advantage of the higher interest rates. Similarly, at lower rates of interest, people will hoard more money, because the loss incurred by hoarding in this case is not very much.

Equilibrium Between Demand for and Supply of Loanable Funds

The rate of interest will be determined by the equilibrium between the total demand for loanable funds and the total supply of loanable funds as shown in Fig. 34.3 below.

In this figure LS is the total supply curve of loanable funds, which has been derived by the lateral summation of the savings curve S, dishoarding



Demand and Supply Equilibrium
Fig. 34.3

ing curve DH, Bank credit curve BM and disinvestment curve DI. Total demand curve for loanable funds is LD which has been found out by the lateral addition of curves I, C and H, which show respectively the demand for loanable funds for investment purposes, consumption purposes and for hoarding. The curve LD of total demand for loanable funds and curve LS of the total supply of loanable funds intersect each other at the rate of interest $Or (= NE)$. At this rate, the loanable funds lent or supplied are equal to the loanable funds borrowed or demanded. Hence, Or is the equilibrium rate of interest which will tend to settle in the loan market.

Criticism

Most of the criticisms made against the classical theory are valid also in the case of the loanable funds theory. In fact, there is not much difference in

the classical and loanable funds theories. The difference lies only in the meaning of saving. In the classical theory, 'savings' is in fact the same thing as the "loanable funds" of the loanable funds theory. In classical theory, savings mean savings out of the income of the previous period. In the loanable funds theory, loanable funds consist of savings out of the income of the previous period plus borrowed bank deposits plus activated idle balances. In classical language, savings out of current income may well exceed the savings of loanable funds theory, because current income is increased by bank loans or the injection of idle balances. Thus, the supply schedule of savings of classical theory amounts to the same thing as the supply schedule of loanable funds of the loanable funds theory.

Further, the Loanable Funds Theory like the Classical Theory is indeterminate. According to this theory, the rate of interest is determined by the intersection of the demand curve for loanable fund with the supply curve. Now the supply of loanable funds is composed of savings plus Bank credit and dishoarding. But since the 'savings' part of the supply curve varies with the level of income, it follows that the total supply of loanable funds will also vary with income. Thus, this theory is also indeterminate.

LIQUIDITY PREFERENCE THEORY

In his epoch-making book, "The General Theory of Employment, Interest and Money", the late Lord Keynes gave a new view of interest. According to him, "interest is the reward for parting with liquidity for a specified period."⁶

A man with a given income has to decide first how much of this income he is going to consume and how much to save. The former will depend on what Keynes calls, the **propensity to consume**. Given this propensity to consume, the individual will save a certain proportion of his given income. He now has to make another decision. How should he hold his savings? How much of his resources will he hold in savings? How much of his resources will he hold in the form of ready money (cash or non-part with or lend. This latter decision will depend upon what Keynes calls his "**liquidity preference**."

Meaning. Liquidity preference means the demand for money to hold or the desire of the public to hold cash. In the words of Prof. Meyer, "Liquidity preference is the preference to have an equal amount of cash rather than of claims against others."

Factors Governing Liquidity Preference

Liquidity preference of a particular individual depends upon several considerations: The question

6. *Ibid.*, p. 167.

is: Why should the people hold their resources liquid or in the form of ready money, when they can get interest by lending such resources? The desire for liquidity arises because of three motives: (i) the transactions motive, (ii) the precautionary motive, and (iii) the speculative motive.

Transactions Motive. The transactions motive relates to the demand for money or the need for cash for the current transactions of individual and business exchanges. Individuals hold cash in order "to bridge the interval between the receipt of income and its expenditure." This is called the '**Income Motive.**' Most of the people receive their income by the week or the month, while the expenditure goes on day by day. A certain amount of ready money, therefore, is kept in hand to make current payments. This amount will depend upon the size of the individual's income, the interval at which the income is received and the methods of payments current in the locality.

The businessmen and the entrepreneurs also have to keep a proportion of their resources in ready cash in order to meet their current needs of various kinds. They need money all the time in order to pay for raw materials and transport, to pay wages and salaries and to meet all other current expenses incurred by any business. This Keynes calls the '**Business Motive**' for keeping money. It is clear that the amount of money held, under this business motive, will depend to a very large extent on the turn-over (*i.e.*, the volume of trade of the firm in question). The larger the turn-over the larger, in general, will be the amount of money needed to cover current expenses.

Precautionary Motive. Precautionary motive for holding money refers to the desire of the people to hold cash balances for unforeseen contingencies. People hold a certain amount of money to provide for the danger of unemployment, sickness, accidents and other more uncertain perils. The amount of money held under this motive will depend on the nature of the individual and on the conditions in which he lives.

The combined sum of balances held for transactions and precautionary motives, Keynes termed 'active balances' and labelled M_1 . The demand for active balances can be referred to as L_1 and thus $L_1 = f(Y)$. This relationship is shown diagrammatically in the following diagram. According to Keynes, L_1 is directly proportionate to the level of national income. OL_1 is the demand for active balances at OY_1 of national income. When the national income increases from OY_1 to OY_2 the demand for active balances goes up from OL_1 to OL_2 . It should be noted here that the demand for active balances, according to Keynes, is independent of the rate of interest or the demand for active balances with respect to the rate of interest is absolutely inelastic. OM is the demand for active balances which does

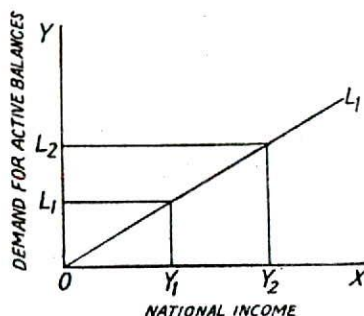


Fig. 34.4

not change, whether the rate of interest is 8% or 6%.

Speculative Motive. The speculative motive relates to the desire to hold one's resources in liquid form in order to take advantage of market movements regarding the future changes in the rate of interest (or bond prices).

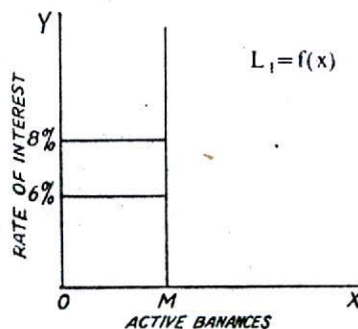


Fig. 34.5

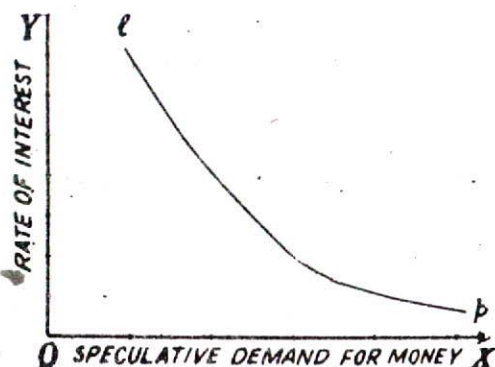
The notion of holding money for speculative motive is a new typically Keynesian idea. Money held under the speculative motive serves as a store of value as money held under the precautionary motive does. But it is a store of money meant for a different purpose. The cash held under this motive is used to make speculative gains by dealing in bonds⁷ whose prices fluctuate. If bond prices are expected to rise, which in other words means that the rate of interest is expected to fall, businessmen will buy bonds to sell when the price later rises. If, however, bond prices are expected to fall, *i.e.*, the rate of interest is expected to rise, businessmen will sell bonds to avoid capital losses. Nothing being certain in this dynamic world, where guesses about the future course of events are made on precarious bases, businessmen keep cash to speculate on the

7. All securities and other such papers as yield a fixed and known amount of interest over a period of time are known as bonds.

probable future changes in bond prices (or the rate of interest) with a view to making profits.

Given the expectations about the changes in the rate of interest in future, less money will be held under the speculative motive at a higher current or prevailing rate of interest and more money will be held under this motive at a lower current rate of interest. The reason for this inverse correlation between money held for speculative motive and the prevailing rate of interest is that at a lower rate of interest less is lost by not lending money or investing it, that is by holding on to money; while at a higher rate the holders of cash balances would lose more by not lending or investing.

Thus, the demand for money under the speculative motive is a function of the current rate of interest, increasing as the interest rate falls and decreasing as the interest rate rises. In other words, demand for money under this motive is a decreasing function of the rate of interest. This will be clear from the following diagram (Fig. 34.6).



Liquidity Preference Schedule

Fig. 34.6

Along OX-axis is represented the speculative demand for money (called inactive balances, by Keynes) and along OY the rate of interest. The liquidity preference schedule lp is a downward sloping curve towards the right signifying that the higher the rate of interest, the lower the demand for speculative motive, and vice versa. The schedule becomes more elastic towards the right end at very low rates of interest.

But the demand for money to satisfy the speculative motive does not depend so much upon what the current rate of interest is as on expectation of changes in the rate of interest. If there is a change in the expectations regarding the future rate of interest, the whole curve or schedule of liquidity preference for speculative motive will change accordingly. Thus, if the public on balance expect the rate of interest to be higher (*i.e.*, bond prices to be lower) in the future than they had previously supposed, the

speculative demand for money will increase and whole liquidity preference curve for speculative motive will shift upwards.

If the total supply of money is represented by M , we may refer to that part of M held for transactions and precautionary motives as M_1 , and to that part held for the speculative motive as M_2 . Thus, $M = M_1 + M_2$. The money held under the transactions and precautionary motives, *i.e.*, M_1 , is completely interest-inelastic unless the interest rate is very high. Also, the amount of money held as M_1 , that is for transactions and precautionary motive, is mainly a function of the size of income and business transactions together with the contingencies growing out of the conduct of personal and business affairs. On the other hand, money demanded for speculative motive, *i.e.*, M_2 , as explained above, is primarily a function of the rate of interest.

Determination of the Rate of Interest

According to Keynes, the demand for money, *i.e.*, the liquidity preference and supply of money, determine the rate of interest. It is, in fact, the liquidity preference for speculative motive which, along with the quantity of money, determines the rate of interest. We have explained above the speculative demand for money in detail. As for the supply of money, it is determined by the policies of the Government and of the Central Bank of the country. The total supply of money consists of coins plus notes plus bank deposits.

How the rate of interest is determined by the equilibrium between the liquidity preference for speculative motive and the supply of money is shown in Fig. 34.7.

In Fig. 34.5 (a), IPS is the curve of liquidity preference for speculative motive. In other words, IPS curve shows the demand for money for speculative motive. To begin with, OM_2 is the quantity of money available for satisfying liquidity preference for speculative motive. Rate of interest will be determined where the speculative demand for money is in balance or equal to the (fixed) supply of money OM_2 . It is clear from the figure that speculative demand for money is equal to OM_2 quantity of money at Or rate of interest. Hence, Or is the equilibrium rate of interest. Assuming no change in expectations, an increase in the quantity of money (via open market operations) for the speculative motive, will lower the rate of interest.

In Fig. 34.7 (a) when the quantity of money increases from OM_2 to OM'_2 , the rate of interest falls from Or to Or' because the new quantity of money OM'_2 is in balance with the speculative demand for money at Or' rate of interest. In this case we move down the curve. Thus, given the schedule or curve of liquidity preference for speculative motive, an increase in the quantity of money brings down the rate of interest.

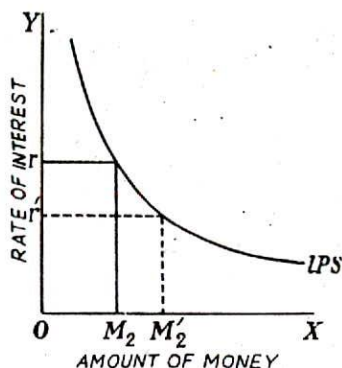


Fig. 34.7 (a)

Equilibrium Between Liquidity Preference and Money Supply

But the increase in the quantity of money may cause a change in the expectations of the public and thereby cause an upward shift in liquidity schedule or curve for speculative motive pushing the rate of interest up. But this is not certain. "New developments may only cause wide differences of opinion leading to increased activity in the bond market without necessarily causing any shift in the aggressive speculative demand for money schedule. If the balance of market expectations is changed, there will be a shift in the schedule. Central Bank policy designed to increase the money supply may, therefore, be met by an upward shift of speculative demand function leaving the rate of interest virtually unaffected."⁸

Thus, a large increase in the quantity of money may exert only a small influence on the rate of interest in certain circumstances.

It is worth mentioning that shifts in liquidity preference schedule or curve can be caused by many other factors which affect expectations and might take place independently of change in the quantity of money by the Central Bank. Shifts in the liquidity function may be either downward or upward depending on the way in which the public interprets a change in events.

If some change in events leads the people on balance to expect a higher rate of interest in the future than they had previously supposed, the liquidity preference for speculative motive will increase, which will bring about an upward shift in the curve of liquidity preference for speculative motive, and will raise the rate of interest.

In Fig. 34.7 (b), assuming that the quantity of money remains unchanged at OM_2 , with the rise in the liquidity preference curve from IPS to $I'P'S'$, the rate of interest rises from Or to Or'' because at Or'' , the new speculative demand for money is in equilibrium with the supply of money OM_2 .

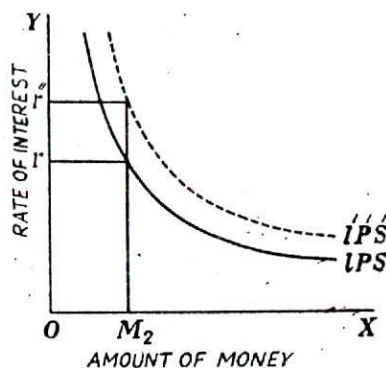


Fig. 34.7 (b)

It is worth noting that when the liquidity preference for speculative motives rises from IPS to $I'P'S'$, the amount of money held does not rise; it remains as OM_2 as before. Only the rate of interest rises from Or to Or'' to equilibrate the new liquidity preference for speculative motive with the available quantity of money OM_2 .

Thus, we see that Keynes explained interest in terms of purely monetary forces and not in terms of real forces like productivity of capital and thrift which formed the foundation-stones of both classical and loanable fund theories. According to him, demand for money for speculative motive together with the supply of money determines the rate of interest. He agreed that the marginal revenue product of capital tends to become equal to the rate of interest, but the rate of interest is not determined by marginal revenue productivity of capital. Moreover, according to him, interest is not a reward for saving or thriftiness or waiting but for parting with liquidity. Keynes asserted that it is not the rate of interest which equalises saving and investment. But this equality is brought about through changes in the level of incomes.

Criticism of Liquidity Preference Theory

Keynes theory, too, has met with criticism: Firstly, it has been pointed out that the rate of interest is not purely a monetary phenomenon. Real forces like productivity of capital and thriftiness or saving by the people also play an important role in the determination of the rate of interest.

Secondly, Keynes makes the rate of interest independent of the demand for investment funds. Actually, it is not so independent. The cash-balances of the businessmen are largely influenced by their demand for savings for capital investment. This demand for capital investment depends upon the marginal revenue productivity of capital. Therefore, the rate of interest is not determined independently

8. Hansen—*Guide to Keynes*, p. 133.

of the marginal productivity of capital or marginal efficiency of capital, as Keynes calls it.

Thirdly, liquidity preference is not the only factor governing the rate of interest. There are several other factors which influence the rate of interest by affecting the demand for and supply of investible funds.

Fourthly, the liquidity preference theory does not explain the existence of different rates of interest prevailing in the market at the same time. Owing to the perfect homogeneity of cash balance, the rates of interest have to be uniform. Actually it is not so.

Fifthly, Keynes ignores saving or waiting as a means or source of investible fund. To part with liquidity without there being any saving is meaningless.

Sixthly, the Keynesian theory only explains interest in the short run. It gives no clue to the rates of interest in the long run.

Seventhly, the borrower's intention is not so much to reward parting with liquidity as to get a return on investment.

Finally, exactly the same criticism applies to Keynesian theory itself on the basis of which Keynes rejected the classical and loanable funds theories. Keynes's theory of interest, like the classical and loanable funds theories, is indeterminate. According to Keynes, rate of interest is determined by the speculative demand for money and the supply of money available for satisfying speculative demand. Given the total money supply, we cannot know how much money will be available so satisfy the speculative demand for money, unless we know how much the transactions demand for money is. And we cannot know the transactions demand for money, unless we first know the level of income. Thus, the Keynesian theory, like the classical theory, is indeterminate.

"In the Keynesian case the supply and demand for money schedules cannot give the rate of interest unless we already know the income level; in the classical case the demand and supply schedules for saving offer no solution until the income is known. Precisely the same is true of loanable-funds theory. Keynes's criticism of the classical and loanable-fund theories applies equally to his own theory."⁹

KEYNESIAN AND OTHER THEORIES OF INTEREST COMPARED

Keynesian vs. Classical Theory

The Keynesian Theory differs from the Classical Theory in the following respects:—

(i) The classical theory applies to a situation of full employment and constant national income, whereas the Keynesian Theory assumes an equi-

librium with less than full employment, where both employment and income are fluctuating.

(ii) The second difference derived from the first is that since in the classical theory income is assumed to be constant, saving is also regarded as fixed and the rate of interest is determined by investment demand curve. It is the rate of interest which brings about equality between saving and investment. On the other hand, according to the Keynesian theory, there is a separate savings curve corresponding to each level of income and employment and the rate of interest is determined by the intersection of the liquidity preference and money supply schedule.

(iii) The classical theory regards the available funds as the supply factor, whereas Keynes treats them as a demand factor since they are determined by people's liquidity preference.

(iv) Keynesian saving is out of income. Therefore, according to Keynes, higher rate of interest will not increase savings as is the belief of the classical economists, because it will discourage investment and so decrease income out of which saving has to come. According to the view of the classical economists, saving automatically leads to investment. But Keynes held quite the opposite view, *viz.*, investment results in saving out of current income. In the classical view, investment could be increased by saving more, but according to Keynes it is investment which determines the volume of saving through the multiplier process. The classical confused the amount saved with the propensity to save (*i.e.*, thrift).

Keynesian vs. Loanable Funds Theory

We may note below a few points of difference between the Keynesian Theory and the Loanable Funds Theory:—

(i) Keynesian theory regards money as stock, whereas the loanable funds theory treats money as a flow. That is why Keynesian Theory explains the rate of interest at any given moment when the money stock is assumed to be fixed. On the other hand, the Loanable Funds Theory explains interest over a period of time when the supply of money is supposed to be fluctuating.

(ii) In the Keynesian theory, the quantity of money is regarded as an independent variable and is not affected by changes in the rate of interest. On the other hand, according to the loanable funds theory, the stock of money itself depends on the rate of interest.

(iii) Keynesian theory explains interest in a situation of less than full employment, while the loanable funds theory is valid in the long-run period when full employment of human and material resources has been realised. In this sense, the two theories may be regarded as complementary.

(iv) According to the Keynesian theory, it is the

9. Hansen—*Guide to Keynes*, p. 141.

demand for and supply of money which determine the rate of interest, whereas according to the loanable funds theory, the rate of interest is simply the price of credit and, as such, is determined by the demand for, and supply of, credit. Hence, the banking system can influence the rate of interest by expanding or contracting credit. In the Keynesian theory, supply of money is regarded as given.

(v) The loanable funds theory emphasises the crucial interdependence between the loan market and the commodity market, since money can be used for different alternatives like purchasing of consumption goods, investment in industries and for purchasing bonds. Keynes, on the other hand, ignores this inter-relationship.

(vi) Unlike Keynes, the loanable funds theory does not accept the notion that shifts in liquidity preference can affect the long-run equilibrium rate of interest.

MODERN THEORY OF INTEREST:

HICKS-HANSEN SYNTHESIS¹⁰

We have discussed above the various theories of the rate of interest put forward from time to time. But we have seen that all these theories suffer from various drawbacks and are indeterminate. The Keynesian theory considered only the monetary factors and the classical theory only the real factors as determining rate of interest. Modern economists have considered both types of factors, monetary and real. The economists like Prof. Hicks and Hansen have made a synthesis between these various theories and have given an adequate and determinate theory of interest. They are of the opinion that the classical and loanable funds theories amount to the same thing. The difference between these two theories, *i.e.*, classical and loanable funds theories, lies only in the meaning of savings. "The Pigovian supply schedule of savings amounts to the same thing as the Robertsonian or Swedish supply of loanable funds."¹¹

How the modern economists make synthesis between the classical or loanable funds theory on the one hand and the Keynesian theory on the other to give an adequate and determinate interest theory, we state in Prof. Hansen's own words:

"The classical or loanable fund formulation and the Keynesian formulation, taken together, do supply us with an adequate theory of the rate of interest. From the loanable-funds formulation we get a family of loanable-fund schedules (or saving schedules in the Pigovian sense) at various income levels. These together with the investment demand schedule give us the Hicksian IS curve. In other

words, the neo-classical formulation can tell us what the various levels of income will be (given the investment demand schedule and a family of loanable-funds schedules) at different rates of interest. But it does not tell us what the rate of interest will be.

From the Keynesian formulation, we get a family of liquidity preference schedules at various income levels. These together with the supply of money fixed by the monetary authority give us the LM curve (Fig. 34.7). The LM curve tells us what the

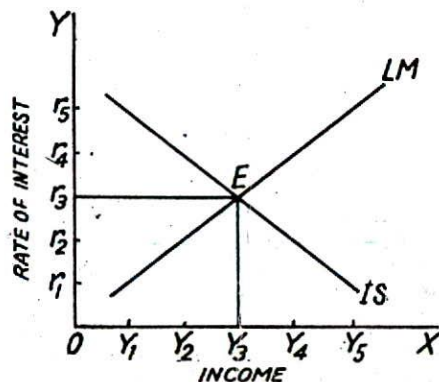


Fig. 34.7

various rates of interest will be (given the quantity of money and the family of liquidity preference curves) at different levels of income. But the liquidity schedule alone cannot tell us what the rate of interest will be. It is the intersection between the IS curve and LM curve which will determine the rate of interest.

In the above Fig. 34.7, the IS curve and the LM curve are curves relating to the two variables: (a) income and (b) rate of interest. Income and the rate of interest are, therefore, determined together at the point of intersection of these two curves, *i.e.*, the point E. Equilibrium rate of interest thus determined is O_r and the level of income determined is OY . At this point, the income and the rate of interest stand in relation to each other such that investment and saving are in equilibrium, and that the demand for money is in equilibrium with the supply of money. It should be noted that LM curve has been drawn by taking the supply of money as fixed.

Effect of Changes in the Variables

With the aid of Hicks-Hansen analysis, we can explain more satisfactorily the effect of changes in certain important economic variables such as desire to save, the supply of money, investment, liquidity preferences on the rate of interest. This is illustrated in Fig. 34.8 given on the next page.

Changes in the Supply of Money. Now suppose that the supply of money has been increased by the

10. See H. L. Ahuja, *Advanced Economic Theory*, 1975, pp. 944-95.

11. *Ibid.* p. 143.

action of the Central Bank. Given the liquidity preference schedule, with the increase in the supply

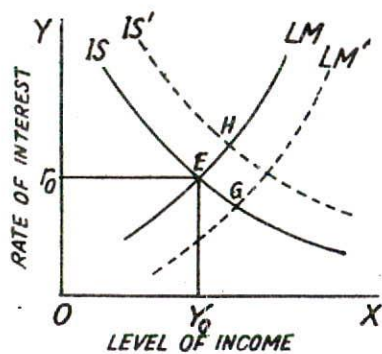


Fig. 34.8

of money, more money will be held for speculative motive at each level of income and the rate of interest will come down.

We see in the above Fig. that the LM curve has shifted to the right. When the LM curve has so shifted rightward, in the new equilibrium position, the rate of interest will be lower and the level of income higher than before. Point E is the intersection of LM and IS curves. When the supply of money has increased LM shifts to the dotted position LM' and with IS schedule remaining the same, the new equilibrium will be at point G corresponding to which the rate of interest is lower and the level of income greater than at E. On the other hand, if the Central Bank has reduced the money supply, less money will be available for speculative motive at each level of income. As a result, the LM curve will shift to the left of E; and the IS curve remaining unchanged, in the new equilibrium position, the rate of interest will be higher and the level of income lower than before.

Changes in the Desire to Save or Propensity to Consume. Let us now consider the effect of changes in the desire to save or propensity to consume. Suppose people's desire to save decreases, i.e., the propensity to consume goes up. As a result, the level of national income will go up at each rate of interest. The result will be that the IS curve will shift to the right. This is shown by IS curve shifting rightward to the dotted position IS'. With LM curve remaining unchanged, the new equilibrium position will be established at H corresponding to which the rate of interest as well as the level of income will be greater than at E. We find that decrease in the desire to save has led to the increase in both rate of interest and the level of income. If, on the other hand, the desire to save increases i.e., the propensity to consume decreases, the level of national income will fall for each rate of interest. As a result, the IS curve will shift to the left. With this shift and the LM curve remaining unchanged, the new equilibrium position

will be reached at some point to the left of E corresponding to which both the rate of interest and the level of national income will be lower than at E.

Changes in Investment and Government Expenditure. When investment and Government expenditure changes, the position of IS curve will be shifted. For instance, if private investment increases or government expenditure goes up, the national income will rise. This will shift the IS schedule to the right, and, given the LM curve, the rate of interest as well as the level of income will rise. On the other hand, if the private investment expenditure decreases or the government expenditure is reduced, the level of national income will go down. As a result, the IS curve will shift to the left, and, given the LM curve, the rate of interest will fall.

Changes in Liquidity Preference. The position of LM curve will also shift as a result of changes in the liquidity preference. If, for instance, the liquidity preference of the people rises, the LM curve will shift to the left, because the higher liquidity preference, supply of money remaining the same, will raise the rate of interest corresponding to each level of national income. When the LM curve has shifted leftward, given the IS curve, the new equilibrium rate of interest will be higher and the level of national income lower. On the other hand, if the people's liquidity preference decreases, the LM curve will shift to the right because, given the supply of money, a downward shift in the liquidity preference curve means that, corresponding to each level of income, there will be a lower rate of interest. When the LM curve has shifted rightwards, given the IS curve, the new equilibrium level of the rate of interest will be lower and the equilibrium level of the national income higher.

Thus, we see that changes in factors like propensity to consume (or desire to save) investment or government expenditure, the supply of money and the liquidity preference will bring about shifts in IS or LM curves. As a result, the rate of interest as well as the level of national income will change.

Hicks-Hansen integration of the classical and the Keynesian theories of interest shows clearly that the government is in a position to influence the level of economic activities or the level of national income by monetary and fiscal measures. We find, therefore, that Hicks-Hansen Theory of interest explained above has integrated the theory of money with the theory of income determination. In this way, the modern theory of interest has succeeded in synthesising the monetary and fiscal policies. Hence, both monetary and fiscal policies can play a useful role in regulating the rate of economic activity in the country.

Thus, a determinate theory of interest is based on: (1) the investment demand function, (2) the saving function (or conversely the consumption function), (3) the liquidity preference function and (4) the

quantity of money. Hence, according to modern economists both monetary and real factors, *viz.*, productivity, thrift, liquidity preference and the money supply play their part in the determination of rate of interest.

Critical Appraisal of Hicks-Hansen Synthesis Merits. Hicks-Hansen synthesis of the classical and the Keynesian theories of interest represents a significant advance in the theory of interest determination. It is a more general, inclusive and realistic approach to the determination of the rate of interest. Besides, it is an improvement in the theory of interest in another way. Hicks-Hansen integration synthesises the fiscal and monetary policies and the theory of income determination with theory of money.

Limitations. But the Hicks-Hansen synthesis of the theories of interest has its own limitation: (a) It is based on the assumption that the rate of interest is flexible, *i.e.*, free to vary and is not rigidly fixed by Central Bank. If the rate of interest were not flexible then the appropriate adjustment will not take place. (b) This synthesis of interest theories is also based upon the assumption that investment is interest-elastic, *i.e.*, investment varies with the rate of interest. If investment were not interest-elastic, then also the Hicks-Hansen synthesis breaks down because the necessary adjustment will not take place. (c) Don Patinkin and Milton Friedman have criticised the Hicks-Hansen synthesis as being too artificial and over-simplified. They think that the division of the economy into two sectors, *viz.*, monetary and real, is artificial and unrealistic. On the other hand, they are of the opinion that the monetary and real factors are interwoven and act and react on each other. (d) Patinkin also has pointed out that Hicks-Hansen synthesis ignores the possibilities of changes in the price level of commodities. He thinks that the various economic variables such as money supply, propensity to consume or save, investment, and liquidity preference not only influence the rate of interest and the level of income but also the prices of commodities and services. He, therefore, suggests a more integrated and general equilibrium approach which involves the simultaneous determination of not only the rate of interest and the level of income but also prices of commodities and services.

CAN THE RATE OF INTEREST FALL TO ZERO?

Theoretically, a zero rate of interest can be conceived. As pointed out above as time goes on, people's power to save and will to save tend to increase. The former because of the rising productive capacity and the latter because of the greater foresight and the tendency to discount the future at a low rate among the more advanced people. A

stage can be imagined in which capital accumulation may outstrip the demand for capital, thus lowering the marginal productivity of capital to zero, even making it negative; that is, people may even pay something for the care of their savings.

It is, however, extremely improbable that such a stage will be ever reached. In the first place, the demand for capital will increase with the increase in population and the increase in the variety of people's wants. Moreover, technical progress, including new inventions and discoveries, is constantly taking place which raises the net productivity of capital and so interest rates. It is true, of course, that new inventions may substitute methods, which economise capital; this may lead to a fall in the rate of interest. But the marginal productivity of capital will still be positive, and, hence there will always be a positive rate of interest. Moreover, the difference between the present gratification and future gratification cannot altogether be removed. The rate of interest is a measure of this difference:

Looking from the side of supply also, we come to the same conclusion. Some people may save even if the rate of interest is zero or negative. But the volume of savings would be seriously reduced if there were no compensation for postponing consumption, thus creating relative scarcity of capital.

Thus, there is no possibility of the rate of interest falling to zero. We can expect a zero rate of interest only if capital loses its character of scarcity. As it is, capital is scarce relatively to demand for its services. There are so many alternative uses to which capital can be put. It is the rate of interest which determines to what uses may be put.

That the zero rate of interest is improbable is also brought out by the Keynesian Theory of Interest, *viz.*, Liquidity Preference Theory. Look at the shape of the l_p curve in Fig. 34.6. It has three distinct parts: the upper reach which is steep, the middle one with a medium slope and the tail which is a horizontal straight line. The upper portion which is steeply falling indicates that, at a high rate of interest, the demand for money has low interest-elasticity. That is, even if the rate of interest falls, the demand for money to hold does not increase much. But when the rate of interest is neither very high nor very low (which is shown by the middle part of the curve), then the demand for money is interest-elastic, which means that even with a small fall in the rate of interest, the demand for money will increase considerably. But when you come to the flat portion, *i.e.*, the tail of the curve, the demand for money is infinitely elastic. In other words, the demand for money is perfectly elastic in respect of interest rate and as the liquidity preference reaches its maximum, *i.e.*, at the very low rate of interest, there is no limit to the amount of money which the people would like to hold. This is a position of Absolute

liquidity Preference. It is also called **Liquidity Trap** by some economists.

Now suppose that the position of liquidity trap or absolute liquidity preference is reached at 2 per cent rate of interest shown by that portion of the $1p$ curve which is a horizontal straight line. All increase of money now will be absorbed by the people and it will not lower the rate of interest. They will not now like to invest any portion of their money in bonds. The reason is that they do not expect the rate of interest to fall further or the price of bonds to rise. Since the price of bonds is not expected to rise, why should they take the risk of investing money in them? That is, they will like to keep the entire stock of money with themselves. In other words, at this stage no amount of increase in the money supply will lead to a reduction in the rate of interest.

Thus, from this characteristic of liquidity preference (being perfectly elastic in respect of low rate of interest) an important conclusion can be drawn, *viz.*, that the rate of interest is not likely to fall below a certain low figure. There is no way of depressing the rate of interest further even though such a fall may be in public interest. In other words, an important implication of perfect elasticity of liquidity preference at a certain low rate of interest is that the rate of interest cannot fall to zero.

Conclusion

We may conclude in Samuelson's words. "As long as any increase in time-consuming process could be counted on to produce any extra product and dollars of revenue, the yield of capital could not be zero. Also, as long as any land or other asset exists with a sure perpetual net income—and as long as people are willing to give only a finite amount of money today in exchange for a perpetual flow of income spread over the whole future—then we can hardly conceive of the rate of interest as falling to zero."¹²

SOCIAL IMPORTANCE OF INTEREST

Every society, socialist or capitalist, is faced with the problem of using its scarce resources in an optimum manner. Capital resources are scarce everywhere even in a socialist society.

Its Allocative Function. In all types of societies, therefore, all projects have to go through a screening process to find their place in the order of priorities.

Even in a socialist society interest cannot be abolished, though it may not be paid to private individuals. Even there, through its instrumentality, priorities regarding the use of scarce capital resources for various possible employments will be determined. The order of priorities in a socialist

society will probably be different from the one in a capitalist society, but the rate of interest will perform its function all the same. It has the same function as the function of price, *viz.*, it adjusts demand to the supply available. It enables capital to be apportioned between competing demands of alternative uses. It is through the rate of interest that those uses which promise the highest future returns receive the first consideration. Of course, the criterion of the highest future returns will differ in a capitalist society from that in a socialist society. In the former, expectations of profits for private entrepreneurs and in the latter the conception of welfare of the planning authority, will determine the priorities in investment.

Interest cannot be abolished; it can be socialised, if capital is socialised. Even if capital is socialised, the rate of interest will still be needed as an accounting device to give expression to the prevailing scarcity or abundance of capital for the guidance of the managers of public enterprises in a socialist economy.

Whatever the type of economic system, interest acts as a rationing device or performs an allocative function. Since capital is a scarce commodity, only those projects can be undertaken where the return on capital invested is greater than, or equal to, the rate of interest. Those projects whose prospective yield is lower than the rate of interest have to be dropped. Thus, the rate of interest directs the use of capital and the growth of productive capacity.

The interest rate is socially important because it determines both the level and composition of investment. According to Keynes' theory of investment, the level of investment is determined by marginal efficiency of capital and the rate of interest. The lower the rate of interest, the greater is the investment, and vice versa. However, the influence of the rate of interest depends on the interest-elasticity of investment. It may, therefore, be understood that raising interest rates as an anti-inflationary device will curb the rate of growth by reducing investment.

As pointed out above interest rate affects the composition of investment through its price-rationing function. Whether investment goes into consumer goods industries or the producing goods industries or into long-term or short-term investments depends on the rate of return in these investments.

There are, however, some limitations in its performance as an allocative mechanism, because government allocates a substantial portion of available capital to projects which it considers in public interest and oligopolistic firms can obtain capital on more favourable terms.

12. Samuelson, P. A.—*Economics* 1970, p. 579.

Nature of Profits

Profit is the reward of the entrepreneur, rather of the entrepreneurial functions. Profit differs from the return on other factors in three important respects:

(a) Profit is a residual income and not contractual or certain income as in the case of other factors. (b) There are much greater fluctuations in profits than in the rewards of the other factors. (c) Profits may be negative, whereas rent, wages and interest must always be positive.

In the ordinary speech, people understand by the term 'profit' all excess of income over costs and this includes the earnings of self-used factors, *i.e.*, entrepreneur's own land, capital and his own labour work called respectively implicit rent, implicit interest and implicit wage. But in economics, profit is regarded as a reward for the entrepreneurial functions of final decision-making and ultimate uncertainty-bearing.

Analysis of Gross Profits

The best way to understand the true nature of profits is to analyse the gross income of the entrepreneur into its various component parts. Gross profits stand for the total earnings of the entrepreneur, not necessarily for the entrepreneurial functions only. It is a mixture composed of several elements.

Prof. Walker was the first economist to draw the distinction between interest and profit in this mixture. According to him, the entrepreneur's income consists of two elements only, *viz.*, interest and profit.

There have been further refinements, in modern times, of the several items of which pure profit, however, is the return for the entrepreneurial functions only. From the total receipts of the business must be taken out what has to be paid out to the various factors of production engaged on a contract

basis. Thus, the rent of land, the wages of labour, and interest on capital have to be deducted out of the total income of the entrepreneur.

Apart from pure profit, the following are the main constituents of gross profit:—

(i) **Interest on Entrepreneur's Own Capital.** The entrepreneur could earn this interest by lending his capital. We must, therefore, make a deduction out of gross profit, for interest on the entrepreneur's own capital, at the current rate of interest in order to ascertain net profit.

(ii) **Rent of Land Owned by the Entrepreneur.** This rent cannot be counted as profits. The entrepreneur could earn this rent by giving his land to a tenant. A deduction must, therefore, be made, out of gross profit, for this rent at the current rate of the rent of land similarly situated and possessing similar other advantages for calculating net profit.

(iii) **Entrepreneur's Wages of Management or Superintendence.** This is the return for the work done by the entrepreneur as manager, and could have been done by him on a salary basis for another firm. This is also included in gross profit. The entrepreneur must be allowed a salary, which should be considered his wages rather than profits.

The above three elements, strictly speaking, are not the reward of the entrepreneur as such but he gets them as capitalist, landlord and manager. These he could earn without setting himself up as an entrepreneur. These are included in gross profit and must be deducted for finding out the net profit.

(iv) **Reward of The Entrepreneur as Risk-taker.** The function as a risk-taker must be performed by the entrepreneur himself. Certain risks, however, are insurable, *e.g.*, risks of accident, fire and marine, *etc.* But, as we shall see, many risks cannot be insured, and these must fall on the entrepreneur.

(v) **Gains as Superior Bargainer.** Certain gains

accrue to the entrepreneur when he bargains with labourers, capitalists, landlords, suppliers of raw materials and consumers—all those with whom he has dealings. These gains he makes owing to his superior skill in bargaining. They are also mixed up with gross profit.

(vi) **Monopoly Gains.** These gains are due to imperfect competition, which enables the entrepreneur to charge higher prices or to pay lower rewards to the factors hired by him and thus increase his profits. They, too, form part of gross profits.

(vii) **Conjunctural Gains.** Another element contained in gross profit are the conjunctural gains. They are also called "windfall profits." These are due to favourable circumstances or pure luck, e.g., outbreak of a war giving high profits to producers of essentials of war and even to other producers. During the second world war, many entrepreneurs made enormous profits due to war demands and high prices in India and elsewhere.

Sudden rise in prices is not the only source of windfall profits. They may also arise on account of sudden increase in demand due to non-monetary causes like changes in tastes or discovery of new uses of a product. Windfall profits will also emerge from a shift in population.

Pure or Net Profit

The tendency among modern economists is to accept the American view of profits as being the reward for purely entrepreneurial functions, i.e., functions which cannot be performed by paid employees. The income from risk-taking and gain from bargaining are regarded by some economists as "pure profits" of the entrepreneur as against the "gross profits" which include all the items enumerated above.

Pure or net profit is "the amount that accrues to the entrepreneur for assuming the risk inseparable from all business under the system of production in anticipation of demand." "It (pure profit) is a payment made exclusively for bearing risk. The essential function of the entrepreneur is considered to be something which only he can perform. This something cannot be the task of management, for managers can be hired, nor can it be any other function which the entrepreneur can delegate. Hence, it is contended that the entrepreneur receives a profit as a reward for assuming final responsibility, a responsibility that cannot be shifted on the shoulders of any one else."¹

THEORIES OF PROFITS²

Profit as Rent of Ability

One view of profits makes them analogous to rent.

1. Thomas, S.E.: *Elements of Economics*, 1939, pp. 293-94.

2. For the history of the theories of profit see F. H. Knight—*Risk, Uncertainty and Profit*, 1940, Ch. II.

The Rent Theory of Profit, as it may be called, was propounded by the American economist, F.A. Walker. He was the first to introduce a distinction between a capitalist and an entrepreneur. An entrepreneur need not be a capitalist. He is a person who may undertake a business without using any of his own capital.

The Theory. Walker regards profits as rent of ability. Just as there are different grades of land, there are different grades of entrepreneurs. The least efficient entrepreneur, who must remain in the field of production to meet the current demand, just recovers his cost of production. Above him are entrepreneurs of superior ability. Just as rent arises because of the differential advantages enjoyed by a superior land over the marginal land, profit also is the reward for differential ability of the entrepreneur over the marginal entrepreneur or the no-profit entrepreneur. **Profits are thus like rent and, like rent, they do not enter into price.**

Wages of management are not profit and the marginal employer only earns the wages of management, and no more. With a slight unfavourable turn of prices or costs, he may have to work as an employee rather than as an employer. Wages of management, thus, must be paid to maintain the given supply of entrepreneurs. Such wages thus enter into price.

Criticism. This theory has the same weakness as Ricardo's theory of rent. The employer, who will leave the business with a slight unfavourable turn of events, is not necessarily the least efficient. He may be higher up in the scale and may be attracted by more profitable alternative ventures.

The theory, moreover, does not explain the real nature of profits; it merely provides at best a measure of profits.

It is wrong to say, again, that profits do not enter into price. They may not enter in the short period, but they must be covered by price in the long run. The entrepreneur performs the essential function of risk-bearing and unless the price of the commodity is high enough to compensate the entrepreneur for this, the supply of entrepreneurs will decrease until the price rises high enough to pay for the risk-bearing service.

Some entrepreneurs may earn high profits and others may suffer heavy losses. When the average is taken over a long period, the so-called surplus tends to disappear.

Finally, the theory even fails to explain the size of the profit. The differential gain is due to scarcity of superior entrepreneurs; but why does this scarcity arise? In the case of land, scarcity is due to natural limitations. In the case of entrepreneurs, there are no such limitations. The theory of profit must explain the cause of such scarcity.

Thus, there is no doubt a differential element in profits as in rent, superior entrepreneurs earning higher profits. But the analogy ends here. There may

exist no-rent land, but there cannot exist no-profit entrepreneur. Rent may not enter into price but normal profits do. This is due to the fact that the supply of land is, there, rent or no rent. But the supply of entrepreneurs cannot be maintained unless profits are earned. Expected profits or prospective profits do enter into price, though the realised profits may not. But even realised profits enter indirectly into price in the long run by influencing expectations of profits.

Wages Theory of Profits

The connection between profits and wages can be looked at from two points of view: (i) The socialists regard profits as simply deductions from the produce of the workers' labour. According to this view, profits are not justified; because they are earned at the expense of the wage-earners. (ii) The second view is represented by Prof. Taussig. He regards profits as simply a particular kind of wages.

As for the socialist view, it may be pointed out that, under perfect competition, rates of wages for the same type of labour tend to equality in the same industry. The same wages are paid by the employer who earns no profits as by the one who earns high profits. The superior employer earns profits, not because he pays lower wage (this he cannot do under competition) but because, on account of his superior organising ability and uncertainty-bearing powers, he can produce at a cost lower than that of his inferior rival.

Now consider Taussig's view that profits are merely wages for a special kind of labour. "Profits," says Taussig, "are best regarded as simply a form of wages."

Criticism. The position taken up by Taussig cannot be accepted. There are fundamental differences between wages and profits:—

(i) Profits are essentially a reward for assuming risks of business. But the labourer takes no risk. His reward or wage is primarily for the work done by him. His risk is insignificant as compared with the risk taken by the employer.

(ii) There is much greater element of chance gains in profits than in wages. Wages thus are "earned incomes" in a much more real sense than profits.

(iii) Part of profits, and in some cases a major part, is due to imperfections of competition. Under imperfect competition, while profits tend to swell, wages tend to be depressed, for reasons already noted.

Marginal Productivity and Profits

Does the marginal productivity theory apply to profit? The answer is "in a way, yes." Marginal

productivity, as we have already explained, is the expression of relationship between scarcity and demand. The supply of entrepreneurs is short and it is not easy to increase it. The demand for the entrepreneurs' services, especially entrepreneurs of exceptional ability, on the other hand, is great in the modern conditions of production. Since the marginal productivity of entrepreneurs is high, therefore, profits are high. The greater the factor of uncertainty and the greater the scarcity of entrepreneurs of ability high enough to make a success of business, the higher the profits.

The only difference in the application of the theory of marginal productivity to entrepreneurs, as compared with other factors, is that here the forces of competition work directly, while in the case of the other factors they work through the employer. The ultimate substance is the same. It is the community's competing demands that have to be satisfied, and thus entrepreneurial ability has alternative uses. Forces of competition tend to equalise the profits of entrepreneurs of equal ability.

The theory of marginal productivity, however, cannot be applied in such a clear-cut way to profits as it applies to other phenomenon of value. This is due to the fact that the entrepreneur performs a complex set of functions and the element of uncertainty refuses to be standardised. Moreover, the entrepreneurial factor cannot be increased or decreased in minute doses, since it constitutes one large unit. Withdrawal of one unit may disorganise the whole business. It is thus difficult to measure the marginal net product of the services of the entrepreneur.

Marginal revenue productivity of the entrepreneur to an industrial firm cannot be measured as the marginal revenue productivity of other factor—land, labour, and capital—can be measured. You cannot compare the marginal revenue productivity of one entrepreneur with that of two or three or of two with three or four and so on. Contribution of an entrepreneur to a firm cannot be measured in any physical units. No set rules are available to measure entrepreneurial ability. That is why to try to calculate the marginal revenue productivity of the entrepreneur to an individual firm is unrealistic.

But the marginal revenue productivity of an entrepreneur to an industry can be measured in principle. This is so because, the number of entrepreneurs in an industry can vary and results of such variations or alterations can be studied. However, it will be necessary to assume that all entrepreneurs in the industry are homogeneous.

In Fig. 35.1, MRP is the marginal revenue productivity curve or the demand curve of the industry for entrepreneurs. SS is the supply curve of entrepreneurs. The equilibrium is at E, the point of intersection of the curves, where all entrepreneurs make normal profit OS (= EM) in the long run. In

the short run, the number of entrepreneurs being OM' instead of OM , they may make super-normal profit $E'M'$ ($=OS'$). But, in the long run, the

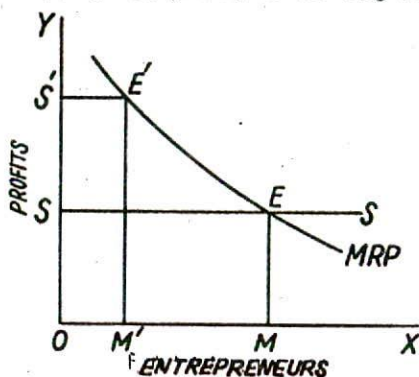


Fig. 35.1

number of entrepreneurs will increase to OM and the extra profits will be competed away so that the profits fall to the normal level EM ($=OS$).

Dynamic Theory of Profits

No Profit in Stationary State. This theory is associated with the name of the American economist, J. B. Clark. He says that in a static world, where the size of the population, the amount of capital, the quantity and quality of human wants, the methods of production, technical knowledge, the organisation of business, etc., remain the same, profits tend to disappear under the force of competition. Profits represent the difference between selling price and cost. It is surplus above costs. But if competition works in a frictionless manner, the surplus will vanish. Wherever there is a surplus, production will increase, bringing down the price. That is how the surplus will disappear. "In the static state each factor secures what it produces, and since cost and selling prices are always equal, there can be no profits beyond wages for the routine work of supervision." ³In a stationary state everything is known and knowable. There is no risk and no uncertainty; hence no profits.

Profits Arise in Dynamic World. But we are not living in a stationary state. Ours is a dynamic world and some changes are constantly taking place. The clever entrepreneur foresees these changes. He is a pioneer. Somehow by invention or otherwise he lowers the cost and makes profits. The changing world offers limitless opportunities to the far-sighted, daring and clever entrepreneurs to make profits by turning the facts of the situation in their favour. It is only because the world is dynamic that

it is possible for them to keep the lead and reap the profits.

That the world is dynamic is due to two sets of factors: (a) internal and (b) external. In other words, there are certain changes which the entrepreneur himself brings about such as innovations and there are other changes which are brought about by external forces. The external changes are of two kinds; (a) regular changes like trade fluctuations which affect all profits and (b) irregular as breaking out of fire, earthquake, floods, strikes, changes in tastes, changes due to government policies, war, etc.

Thus, profits arise out of changes. In the static world, there is no change; hence no profits. In an economy, where nothing changes, there can be no profit.

Prof. Knight, however, is of the opinion that only those changes, which cannot be foreseen and which cannot be provided for in advance, will yield profits and not others. He says, "It cannot, then, be change, which is the cause of profit, since if the law of change is known, as in fact is largely the case, no profits can arise. Change may cause a situation out of which profit will be made, if it brings about ignorance of the future."⁴

Thus, it is ignorance of the future or uncertainty, and not necessarily change, which, according to Knight, is the cause of profit.

Innovations Theory of Profits

In the dynamic changes, which give rise to profits according to the dynamic theory of profits, Joseph Schumpeter has singled out for special treatment the part played by innovations. The daring and the dynamic entrepreneurs continue to hit at one innovation or another, keeping their business ahead of others and thus make handsome profits. According to Schumpeter, the principal function of the entrepreneur is to make innovations and profits are a reward for performing this important function.

Schumpeter has given the term innovation very wide meaning. Discovery of a new material or a new technique of production resulting in the lowering of the cost of production or improving the quality of the product is an innovation. Any new measure or new policy initiated by the entrepreneur comes under innovation as Schumpeter uses the term. Innovations may be of two types: (a) Those which change the production function and reduce the cost of production, and (b) those innovations which stimulate the demand for the product, i.e., which change the demand or utility function. In the first type are included the introduction of new machinery, improved production techniques or processes, exploitation of a new source of raw material or a

3. Knight, F. H.—*Risk, Uncertainty and Profits*, 1940, p. 33.

4. *Ibid.*, p. 3.

new and better organisational pattern for the firm. The second type of innovations are those which are calculated to increase the demand for the product by introducing a new product or a new variety of an old product, new and more effective mode of advertisement, discovery of new markets, etc. Success of any of these innovations brings a handsome increase in profits. Profits increase because either the cost of production is lowered or the product fetches a higher price.

It may be pointed out, however, that profits owing to innovations are only temporary and tend to be competed away. Soon the innovations come to be imitated by the rivals and they cease to be innovations or lose their novelty. Only when the innovation is got patented can the originating entrepreneur continue to enjoy it. But in a dynamic world and progressive economy, the superior entrepreneurs continue to make innovations and enjoy the profits thereof. As Stigler observes, "these profits may exist for a considerable time because of the ignorance of other firms of their (innovations) existence or because of the time required for the entry of new firms. More important, the successful innovator can continuously seek new disequilibrium profits since the horizon of conceivable innovations is unlimited."⁵

We may also remember that profits are both the cause and effect of innovations. Profits serve as a necessary incentive for making innovations; hence profits are a cause of innovations. But since innovations result in profits, profits are the effect of innovations.

Criticism. Schumpeter's innovation theory can be criticised on the same ground as Clark's dynamic theory:

(a) Schumpeter also like Clark ignores uncertainty as a source of profit.

(b) He also denies that risk-bearing plays any role in the determination of profit.

Risk-bearing Theory of Profits

Most people do mind the risk which makes them hesitate to take a plunge in business. The greater the risk the higher must be the expected profit in order to induce them to start the business. All businesses are more or less speculative, and unless the risk-taker is going to be amply rewarded, business will not be started. As risk acts as a great deterrent, the supply of entrepreneurs is kept down, and those who do take the risk earn much more than the normal return on capital. Hence, profits are regarded as a reward for risk-taking or risk-bearing.

This theory of profits is associated with F.B. Hawley's name.⁶ He says profits is the reward for

risks and responsibilities that the undertaker . . . subjects himself to. Drucker mentions four kinds of risks: replacement, risk proper, uncertainty and obsolescence. Replacement, generally known as depreciation, is calculable and is counted as cost. Obsolescence is the least calculable but is also an item in the cost. Risk proper (*i.e.*, risk of marketability of the product) and uncertainty are not costs in the conventional sense, but are charges against profits. They may be called costs of staying in business. Physical risks like fire, accident, etc., can be provided against by insurance, and are, therefore, included in costs. There are, however, risks that cannot be foreseen, and hence cannot be provided against. It is for undertaking these risks that the entrepreneur is rewarded.

Criticism. As against this, there is the view that though profits do contain some remuneration for risk-taking, yet the high profits made by entrepreneurs cannot in their entirety be attributed to the element of risk. They are not, at any rate, in proportion to the risk undertaken. On the contrary, it is pointed out by Carver that profits arise not because risks are borne, but because the superior entrepreneurs are able to reduce them.⁷ We might say—though it may seem paradoxical—that profits are made not because risks are borne but because they are avoided. Still, it cannot be denied that a great deal of pure profit is the reward for risk-bearing.

Uncertainty-bearing Theory of Profits

The Theory. According to Prof. Knight, it is uncertainty-bearing rather than risk-taking which is the special function of the entrepreneur and leads to profit. We have seen that there are certain risks which are foreseen and provided against. Risks of death and of accident like fire and ship sinkings are statistically determinable. Their incidence is measurable. The insurance companies undertake these risks in return of premia paid to them. The payments of these premia are included in the cost of production. The entrepreneur gets no profit on account of these risks. Hence, risk-taking is not the function of the entrepreneur, but of the insurance companies.

But the genuine economic risks of the marketability of the product due to shifts in demand, are unforeseen and unpredictable. Knight will not call them risks but uncertainty. The term 'risk' is applied to those dangers which can be known and foreseen. The entrepreneur gets a remuneration for bearing uncertainties (unforeseeable risks) and nothing for the risks which have been foreseen, the incidence of which is on the insurance companies.

Just as waiting (capital) is a factor of production,

5. Stigler, G. J.—*The Theory of Price*, 1952, p. 182.

6. *Enterprise and the Productive Process*, 1907.

7. *Distribution of Wealth*, p. 273.

uncertainty-bearing has also been given the status of a factor of production. Like other factors of production, uncertainty-bearing has a supply price, *i.e.*, unless a certain return is expected, no entrepreneur will be induced to face uncertainty. The supply of this factor, uncertainty-bearing, depends on the temperament of the entrepreneur, the total resources at his command and the proportion of these resources he is inclined to expose to uncertainty. A rich entrepreneur of a bold and venturesome spirit, who has made up his mind to invest a big proportion of his wealth, can bear greater uncertainty. A greater gain is necessary to induce an entrepreneur to expose a large proportion of his capital than when only a small proportion is exposed.

Criticism. The theory of uncertainty-bearing, as a cause of profit, has been criticised on the following grounds:

(i) Uncertainty is not the only factor that limits the supply of entrepreneurs. Lack of funds, lack of knowledge, lack of opportunities and the presence of economic friction are some of the factors that restrict the supply of entrepreneurs.

(ii) Uncertainty-bearing is not the only function of the entrepreneur. The profit that he gets is also the reward for other services that he renders, *e.g.*, initiating, co-ordinating, *etc.*

(iii) Uncertainty-bearing cannot be elevated to the status of a factor of production. It is an element of real costs which means exertion, abstinence, sacrifice, *etc.*, as distinguished from money cost. Cost is not generally measured in terms of real cost. We know that capital is a factor of production but not abstinence that is needed to have capital.

(iv) Knight's theory does not seem to have much relevance to the real world. Businessmen continue to estimate profits *ex-ante* in defiance of this theory. This theory has been criticised as heavily insulated from empirical testing and empirical relevance.

Conclusion Regarding Theories of Profit

We have discussed above the various theories of profits. The question arises: which theory shall we accept? How do profits arise? Here we are thinking of not gross profit but net profit. The fact is that in the real world there are several causes which give rise to profits, but the principal cause is uncertainty. This uncertainty is due to the dynamic nature of the world. In this real world of ours, some or the other change is always taking place. But such changes only are the causes of profits as cannot be foreseen as we have read in Knight's theory.

However, in a static world, profits can arise in one way, *etc.*, owing to monopoly. The monopoly profits arise in the dynamic world also. Besides monopoly

profits can arise also from any other position of advantage.

In short, we can say there are two sources of profits: (a) uncertainty and (b) position of special advantage, monopoly or otherwise.

Normal Profits

Some writers introduce the concept of normal profits. As Prof. Knight points out, normal profits belong to the equilibrium state, or to the state in which changes are taking place which can be anticipated and calculated. Of course, it is difficult to imagine a world entirely devoid of change. But we can have societies, specially old established societies, in which conditions are relatively static and business methods have become more or less of a routine nature.

Absence of Uncertainty. In a static society (or state of equilibrium), resources are more or less fixed and have attained such a distribution among various industries that there is no motive for transferring them from one employment to another. There is no tendency for any firm to enter or leave the industry. Uncertainty, therefore, would be at a minimum and competition perfect. Under such conditions, pure profits will tend to disappear and the entrepreneurs will earn only wages of superintendence. Normal profits, therefore, practically amount to earnings of management. They are just sufficient to induce an entrepreneur to stay in the industry.

In a progressive state also, similar results can follow, provided competition is unrestricted and changes can be anticipated. But since changes are not uniform, entrepreneurs in some industries may be able to make higher profits than in others.

Surplus profits can arise, in the short run, in a dynamic world even though the factor of uncertainty is absent. But this will only be a temporary condition. In the long run, industries yielding profits higher than the normal will attract entrepreneurial ability and other factors of production. The rate of profits in such industries will tend to fall. In industries, which are being depleted of such factors, profits will rise, until the normal level of profits is established.

But normal profits in a dynamic world need not necessarily be equal to the wages of a hired manager, even though the factor of uncertainty is absent. The very fact that change exists, even though it can be anticipated, means heavier responsibility on the entrepreneur than in a purely static state. To keep up the supply of entrepreneurs, therefore, inducement must be given to them to take up such responsibility. But this additional payment will be kept within narrow limits by the forces of competition.

Existence of Uncertainty. In a dynamic world, with element of uncertainty in existence, even under

competition profits can be kept permanently above the normal level. Super-normal profits also can permanently arise, where the supply is in the hands of a monopolist and demand is inelastic. This will be so whether uncertainty exists or not.

It should be noted that uncertainty in its turn is not something constant. There are degrees of uncertainty. The greater the degree of uncertainty, the higher must be the profits to compensate the entrepreneurs for bearing its risks.

Monopoly and Profits

So far we assumed that the employer is working under conditions of competition. Under perfect competition, there are no profits in the long run. Profits must, therefore, be either temporary or monopoly profits. The monopolist is able to control output so that the price is not allowed to fall to the level of cost, as is the case under competition. By restricting entry of new firms into business by means of agreements and through the use of patent rights and similar devices, monopolists are able to reap monopoly profits. But the most common source of monopoly profits lies in monopolistic competition or product differentiation.

An element of monopoly profits can also be traced in what have been called **innovation profits** or pioneering profits. A firm which produces a new product, or is able to discover a new material or a cheap process or a new market, will always be able to make extra gains, till its rivals make an inroad into its business. Since competition is absent, partially or totally, gains arising out of innovation or pioneering may be termed as monopoly gains.

In the actual world, the typical cases are of imperfect competition, since absolute monopoly is rare. And, therefore, the element of monopoly gains is not as rare as one would think. In fact, monopoly element will be found almost in all profits.

A monopolist is able to make profits both in static and dynamic conditions. He is able to so fix the price of the product that he may make substantial profits by exercising his monopoly power. In order to make profits he raises the price by restricting the level of his output.

Monopoly is a matter of degree only. Monopoly power is exercised by a pure monopolist who produces a product which has no close substitute. Monopoly is also exercised, though to a somewhat lesser extent, by firms under monopolistic competition and oligopoly as mentioned above. Under various categories of imperfect competition, we know that the demand curve slopes downwards. Hence monopoly is associated with a downward sloping demand curve. Since the demand curve under conditions of pure monopoly, monopolistic competition and oligopoly slopes downwards, firm's

equilibrium (*i.e.* equality between marginal revenue and marginal cost) is achieved at a price which is higher than the marginal cost of production. Also, the price so determined is often higher than the average cost of production which, therefore, yields positive profits to the firm enjoying monopoly power. Since there is strong resistance to the entry of new firms into the industry, the firms working under the monopoly or monopolistic competition continue to make supernormal profits even in the long run.

Even under monopolistic competition, owing to product differentiation, entry into the industry by new firms is not wholly free since product differentiation gives a firm a certain degree of monopoly power so that it can set its own price. No new firm can produce exactly the same product as that of the existing firms. Since the entry of new firms is restricted, demand does not fall, even in the long run, to the tangency position with average cost curve. The result is that entrepreneurs working under monopolistic competition continue to enjoy positive profits by virtue of their monopoly power.

The monopolist enjoys profits owing to his monopoly power. In a previous chapter (No. 28), we have given Prof. A.P. Lerner's quantitative measure of the degree of monopoly power. This measure of the degree of monopoly power is based on the fact that the price set by the monopolist is higher than the marginal cost due to the demand curve sloping downwards. This measure of the degree of monopoly power is based upon the ideal market, *i.e.*, that is a market under perfect competition in which monopoly power is completely absent and, in equilibrium, price is equal to the marginal cost.

Some economists, especially Prof. M. Kalecki, have asserted that the greater the degree of monopoly, the greater the size of profits made by the firm. Kalecki thinks that the degree of monopoly power is the most important determinant of the value of profits; in fact, he thinks that it is the only determinant of the level of profits. It may be borne in mind that the degree of monopoly power of a firm (*i.e.*, its power to set the price above the marginal cost of production) depends upon the elasticity of the demand curve facing the monopolist. That is, smaller the elasticity of demand for the product, the greater is the power of the monopolist to set the price and hence to make profits. But elasticity of demand for a firm's product depends upon the extent to which the product can be differentiated. The greater the product differentiation, the less elastic is the demand for the product. The degree of monopoly power also depends upon the firm's share in the market for its product. The greater is its share, the greater is the monopoly power.

Thus, an entrepreneur is able to set the price for its product owing to monopoly power that it has gained by the extent of distinctiveness of its product

and its share in the total output or market. The entrepreneur earns profits based on his ability to set the price as high above the marginal cost as he can.

But it may be noted that the enjoyment of monopoly power is no guarantee that positive profits will be made. Much depends upon the demand-cost situation. If the demand-cost situation is unfavourable, which means that costs are higher and the demand or revenue is less, even under pure monopoly, monopolistic competition and oligopoly, under which the firms enjoy varying degrees of monopoly power, will be incurring losses. As Professor Bober observes, "He skates on thin ice who identifies profits with monopoly and monopoly with profits."⁸

But generally through the devices of advertisement, product differentiation and other sales promotion activities, the firms enjoying monopoly power make sure that the demand for the product remains above the marginal cost of production yielding a good deal of profits. They can not only exploit the consumers by setting a higher price and making profits but also by virtue of their monopsonistic power are able to exploit the various factors of production and underpay them.

It may be emphasised that in order to make their monopoly power lasting, the firms already in the market must be able to raise strong barriers to the entry of new firms into the industry. Hence, the ability of the monopolist to enjoy monopoly power and make profit depends ultimately on the restrictions they are able to impose on the entry of the new firms. By control over the supply of an essential raw material or legal restrictions such as patent rights, the existence of goodwill enjoyed by the existing firms, reputation of their brands, and trade marks of their products, economies of large-scale production *etc.*, the entry of new firms into the industry will be effectively barred and the existing firms will continue to earn monopoly profits.

Critical Evaluation of Monopoly Theory of Profits

It cannot be denied that monopoly is a very important source of profit. The firms operating under monopoly or monopolistic competition, with a downward sloping demand curve and so enjoying price-setting power, are able to make positive profits. But the assertion, as of Prof. Kalecki, that monopoly is the sole source or determinant of profit cannot be accepted. Dynamic changes, innovations by the bold entrepreneurs and uncertainty also contribute to profits. In any theory of profits, the influence of these factors cannot be ignored. It may also be mentioned that there is no contradiction between uncertainty and monopoly theories of profits as Kalecki thinks. In fact, market imperfections in

monopolistic competition increase uncertainty, which also makes the entry of new firms into industry difficult. Hence, the monopoly theory of profit only supplements the uncertainty theory of profits.

Kalecki's concept of the degree of monopoly power has also been criticized. For instance, under perfect competition, the degree of monopoly power is zero. Hence the rate of profits should also be zero and the share of labour and other factors should be 100 per cent which is really ridiculous. We must add that a great deal of criticism against monopoly theory of profits is really against Kalecki's measure of degree of monopoly power as a source of profits. There is no doubt that monopoly is an important cause and source of profit.

DO PROFITS TEND TO EQUALITY?

Here again, we cannot give a straight answer. It will depend upon the conditions that prevail. In a state of equilibrium, profit in the sense of wages of superintendence will be equalised. Pure profit will disappear.

Absence of Uncertainty. In a state of society, in which change is present but the factor of uncertainty is absent, profits will tend to equality about the normal level as already explained. Difference, however, will not be entirely absent because of the differences of ability. But these differences will be kept narrow by the force of competition.

On the whole, one may say that the greater the routine character of an industry, the greater the tendency of profits to equality, provided the period is fairly long and competition unrestricted. In the short period, however, inequalities can exist.

Existence of Uncertainty. But when we are dealing with a changing society, in which the factor of uncertainty is prominent, there is no tendency towards equality of profits even in the long run. Here profits may show considerable variations.

SOCIAL FUNCTION OF PROFIT⁹

It is commonly thought that the profit is for the entrepreneur; it is his concern, and it is at the expense of the consumer or the general public. It is supposed that there is an inherent conflict between the interests of the individual entrepreneur and social welfare or between private net product and social net product, as Pigou would put it.

But this view is a mere delusion. Far from there being a conflict, there is an underlying harmony between what the entrepreneur gets and the society gives. The conflict is more apparent than real. Whatever the form of society—capitalist, socialist, communist or fascist—profit performs a very essential and useful function.

8. Bober, M. M. *Intermediate Price and Income Theory*.

9. See P. F. Drucker on "The Function of Profit" in *Fortune*, March, 1949.

Philanthropically inclined entrepreneurs sometimes speak of their duty to promote the welfare of the society and of those who are associated with them as workers in the enterprise. They say their duty is not to make profits but to make the people happy. But the foremost duty of an economic enterprise is economic performance, which means the preservation of the resources entrusted to it. "The preservation of these resources is the corporate enterprise's first obligation to itself and to society. If it fails in this it not only weakens its own power to survive, but impoverishes the whole society." In order to keep its resources intact, the enterprise must avoid losses and work for the maintenance of normal rate of profit.

In its battle for survival, it is the duty of each individual enterprise to try to cover (a) current costs of business; and (b) the future costs of staying in business arising out of four kinds of risks already mentioned, (*viz.*, replacement, obsolescence, risk proper and uncertainty).

In addition to these, there are two more functions which a profit-making industrial enterprise must perform: (a) One is to fill what Drucker calls "dry holes," *i.e.*, a successful enterprise must from the social point of view, cover the losses of the unsuccessful ones. "Just as a productive oil well must pay for the pipe and labour that went into a dry hole, so the surviving company must pay for the economic loss of its competitor's failure." This is the insurance principle. Any individual enterprise may try to avoid these dry holes, but the society as a whole cannot.

(b) There is still another function which profit must discharge, *viz.*, to bear the social burden. The successful industrial ventures must bear the cost of social service or social security measures like education, health services, poor relief, old-age pensions, maternity relief, defence, civil administration, etc.

These four functions profit has to perform for the sheer maintenance and preservation of modern industrial economy. But no modern community could be satisfied with mere maintenance or survival. It must provide for a progressive growth and development. The economy must expand and increase productivity not only by the exploitation of virgin resources, but also by making a fuller and better use of existing resources through technological advances, so that they are made more productive than what they have been. All this has also to be achieved out of profits.

Finally, we may add that the rate of profit is an indicator of the directions in which the community's resources should move. It is a guide for the optimum allocation of the available resources. Whatever the economic system, no community can ignore the imperative necessity of deriving maximum benefit out of its resources. It can ignore it only at its peril. The punishment will be economic degeneration. Even a communist society is faced with scarce

resources and multiplicity of wants. The rate of profit, regardless of the pockets it goes into, will guide the investment. Broadly speaking, profits stimulate innovations by inducing businessmen to undertake new ventures. Further, desire for profits leads to efficient allocation of the country's productive resources. In a free enterprise or capitalistic economy, profit is the prime mover or energiser of all economic activity.

Thus, even under a socialist or communist organisation of society, profit will have an essential function to perform. Hence, whatever the form of our economy, the rate of profit guides the economy.

Here is an excellent summing up of the role of profit by Samuelson: "Profits and high factor returns are the bait, the carrots dangled before us enterprising donkeys. Losses are our penalty kicks. Profits go to those who have been efficient in the past—efficient in making things, in selling things, in foreseeing things. Through profits, society is giving the command over new ventures to those who have piled up a record of success." Again, "profits are the report card of the past, the incentive gold star for the future, and also the grubstake for new ventures."¹⁰

Thus, the most important function of profit is resource allocation. Occurrence of profit gives the signal for re-allocation of resources. It is, therefore, an important part of the signal—incentive mechanism which facilitates the functioning of the price-mechanism in a free enterprise economy. The rate of profit influences both the level of resource utilisation as well as the allocation of productive resources among alternative uses.

MACRO-THEORIES OF DISTRIBUTION

We have discussed so far micro-theories of distribution *i.e.* how wage is determined in an individual firm/industry, rent of a particular piece of land and profit of an individual firm or an entrepreneur. But we should also understand how national income is shared among the aggregative shares of rent, wages, interest and profits. This is macro theory. We shall now briefly notice some macro-distribution theories.

THE RICARDIAN OR CLASSICAL THEORY

The Ricardian theory makes use of two principles in income distribution *viz.* the 'marginal principle' and 'surplus principle'. With the marginal principle, the theory explains the determination of the share of rent in the aggregate national output and with the surplus principle the shares of wages and profits in the remaining national income. As Kaldor observes, rent is the difference between the product of labour on the marginal land and the product of average land.

10. Samuelson, P. A.—*Economics*, 1976, p. 602.

In the Ricardian Theory, the marginal product is assumed to be equal to the sum of wages and profits. The wage rate is said to be determined by the minimum subsistence level. The balance of the marginal product (i.e. after deducting the subsistence level wage) is the share of profits. Thus profits are a residual income (i.e. what is left over after paying rent and wages).

With economic growth the share of wages increases at the expense of profits. Thus in the Ricardian Macro economic model, there is a continuous tendency towards a declining rate of profit, with growth in output and employment.

THE MARXIAN THEORY

As Prof. Kaldor observes, the Marxian theory is an adaptation of Ricardo's surplus theory. According to Marx, the value of a commodity is determined by the labour-time necessary for its production. But labour produces more than the value of its labour power i.e. more than what is necessary for maintaining the minimum subsistence standard. Hence a surplus emerges which is expropriated by the capitalists in the form of profits. This is Marx's Theory of Surplus Value. Thus according to Marx profits represent exploitation of labour by the capitalists. The capitalists are able to exploit labour through the ownership of the non-human means of production (i.e. machines).

The notion of surplus value is crucial to the marxian theory: It is the amount of this surplus that determines the relative share of profits in the national income. Wages remaining constant at the subsistence level, the difference between total output and the subsistence output (i.e. the rate of exploitation) increases with technical progress. Hence under the capitalistic system the relative share of wages in the national income declines and the share of profits increases (quite the opposite of what Ricardo thought), though the rate of profit will go down as a result of capital accumulation. This is what Marx calls 'immiseration of the proletariat.'

But Marx has proved to be a bad prophet of future events. Workers' living conditions have continued to improve and the rate of profit has not fallen. Labour theory of value on which he bases his theory has been rejected.

KALECKI'S 'DEGREE OF MONOPOLY' THEORY

According to Kalecki, the distribution of national income into profits and wages depends upon the degree of monopoly in the economy. The degree of monopoly of a firm is measured by $(p-a)$ which is the difference between the price of the product and the average cost on manual labour and raw mater-

ials per unit of output. This difference is made up of entrepreneurial profits, interest, depreciation and salaries and thus represents gross capitalist income (inclusive of salaries) per unit of output. We can get his total income by multiplying it by the total number of units produced. Also, in order to get the gross capitalist income of the economy as a whole, we have to sum up the gross capitalist incomes of all firms which may be represented by the formula $Ex(p-a)$. If we divide it by T (aggregate turn over), we get

$$\text{macro-degree of monopoly} = \frac{\text{gross capitalist income}}{\text{aggregate turn over}}$$

Kalecki shows the dependence of labour's share in national income on the macro-degree of monopoly power. The relative share of wages in national income is given by the formula $\frac{\text{Wages}}{\text{National Income}}$. This share is inversely related to the degree of monopoly power. In other words, increase in the degree of monopoly power will reduce the relative share of wages (i.e. manual labour's share).

Labour's share in national income has remained constant by and large. This, according to Kalecki is due to the fact that increasing degree of monopoly power has been counter-balanced by a fall in the price of basic raw materials.

NEO-CLASSICAL THEORY

According to Neo-classical theory, the marginal principle can be applied to all factors of production by taking them as variable factors, and their rewards are equal to their marginal products. It is just like the micro theory of distribution. The total absolute share of a factor is determined by marginal products multiplied by the amount of the factor used. The absolute share of labour in the national income is determined by the amount of labour multiplied by its marginal products. The relative share of labour in the national income is the absolute share divided by the total national income.

KALDOR'S OR KEYNESIAN THEORY OF INCOME DISTRIBUTION

Kaldor has called his income distribution theory as Keynesian Theory since he employs Keynesian theoretical framework in its elucidation; Kaldor also divides the national income into two parts, viz., wages and profits. Since profits are defined as the incomes of property owners, it includes rents and interest besides ordinary profits.

In Kaldor's model, the share of profits in the national income is a function of investment-income ratio (I/Y) : the greater this ratio, the greater the share of profits, and vice versa. According to Kaldor's theory, a shift in the distribution of income in favour of capitalist class is essential if full employment equilibrium is to be maintained.

Having studied the pricing process under different market conditions, we should discuss briefly the role of the price system in a Modern Economy.

In order to appreciate the role of price in a modern economy, we should have some idea of what a modern economy is. A modern free enterprise economy is based on economic freedom, where every individual is free to take up any job he likes and give up the one he does not like. He is free to work or not to work. Further, every individual is free to own any property and use it as he likes. Also, every individual is free to enter into any contract or agreement, that he thinks most profitable for himself. Thus, freedom of occupation, freedom of owning and using property and freedom of contract are three main pillars of a modern capitalistic economy.

As a result, we find minute division of labour and specialisation creating millions of businesses, enterprises, millions of farmers, professional men, governmental and non-profit and philanthropic institutions all working in the economic system giving a fair degree of satisfaction to all concerned. In this system, every individual, including legal individuals like firms and corporations, decides for himself what contribution, in the form of goods and services, he will make to the economy and be sure that he will get the price acceptable to himself and the buyers thereof. Also, he is confident that he can get goods and services contributed by others at prices acceptable to them.

What is the instrument of this remarkable achievement? It is the 'invisible hand'—the price system—about which Adam Smith spoke so eloquently thus: . . . "every individual necessarily labours to render the annual revenue of the society as great as he can. He generally, indeed, neither intends to promote the public interest, nor knows how much he is promoting it . . . by directing that industry in such a manner as its

produce may be of the greatest value, he intends only his own gain, and he is, in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention."

PROBLEMS TO BE TACKLED BY PRICE SYSTEM

The price system seeks to tackle the following problems:

(1) The consumers constituting the household have a variety of unlimited wants; but the resources at their disposal for the satisfaction of these wants are limited. They seek to select such combination of goods and services which will give them maximum satisfaction. The consumers' demand schedule or curve indicates how much they would be willing to purchase at various prices.

(2) The businessmen or producers supply the goods and services to the households. Their supply schedule or curve communicates to the consumers the amounts they would be able to supply at various prices. But they have to select for production those goods and services and the quantities to be produced which will maximise their profits. For this purpose, they must choose the most economical or least-cost combination of productive resources.

(3) The producers will maximise profits and minimise costs by substituting low-priced resources for high-priced resources. The resources' demand curves will indicate to the resource suppliers how much of the resources the firms would be willing to employ at the various prices of the productive resources.

(4) The households, as resource suppliers, offer to the firms productive resources in the form of land, labour, capital and entrepreneurial talents. The decisions of the competing resource suppliers are communicated to the producers by means of resource supply curves. The resource suppliers tend to shift their resources from the less remunerative to

more remunerative employments. In this way, they seek to obtain the maximum return from the resources they supply.

Through the price mechanism, the product demand intentions of the consumers or the households and the product supply decisions of the firms result in series of product prices. In the same manner, the resource demand decisions of the firms and the resource supply decisions of the households will establish a series of resources prices. In short, "the competitive price system is a mechanism both for communicating the decisions of the consumers, producers and resource suppliers to one another and for synchronizing those decisions toward consistent production objectives."¹ This is series of inter-related decisions on the part of consumers, the resource suppliers and producers or resource users. These decisions or choices are based on the fact of scarcity of resources at the disposal both of consumers and producers and are influenced by considerations of substitutability of both products and resources, because both products and resources are capable of being put to alternative uses. The prices of products and resources reflect the relative scarcities and when the products and resources become scarce, their prices rise and it is a signal for substitution. Through substitutability, the society uses less of scarce resources and products and more of those which are comparatively less scarce.

Thus, the price system enables the community to tackle the fundamental problems of the economy, viz., (a) what to produce, (b) how much to produce, (c) how to produce, i.e., organising production and (d) for whom to produce, i.e., the distribution of the total output and (e) to adapt itself to change or flexibility.

Let us now study in some detail the operation of the price system.

FUNCTIONS OF THE PRICE SYSTEM

The price system performs the following useful functions:—

Coordination. The price system functions in such a manner that the adjustments in the economic system take place almost automatically without any direction or dictation from a central authority. Price is the co-ordinator both of production and consumption. The consumers are able to convey their preferences through the prices they are willing to offer. Similarly, the producers are able to indicate the scarcity or abundance of a commodity by means of the price they are willing to accept. If price rises, it checks demand and stimulates supply, and vice versa. If there is a greater demand for a commodity than the supply thereof, then the adjustment

between the two will be brought about through a rise in price. Conversely, if the supply is greater than the demand, the price will fall and bring about an equilibrium between the two.

The same type of dovetailing is effected in the sphere of distribution where we are concerned with the remuneration of services of the agents of production. Rent, wages, interest and profits are the prices paid for the factors of production. If, for instance, there is a large supply of labour, wages will fall and induce the employers to employ more and absorb the extra supply. If, on the other hand, the supply of labour has been curtailed by war or epidemics, the wages will rise and give a warning signal to the employers of this shortage. The price system collates and transmits to millions of consumers and producers the required information, as it were, regarding the availability of, and desires for, inconceivably large number and bewildering variety of goods and services and the industries are motivated to react appropriately.

Guides Economic Activity. The price obtained either for a commodity or a service constitutes an income which, through its purchasing power, determines the extent and the direction of economic activity. There seems to be every justification, therefore, for the present economic system to be designated as a "government by price."

Harmonizes Conflicting Interests. The price-mechanism is supposed to harmonize the interests of both the consumers and the producers. In Benham's words, "It tends to harmonize the desire of entrepreneurs for profits and the desire of consumers to satisfy their wants as fully as possible from the factors of production available."

This harmony, however, is not always realized in actual practice. We may find that either the consumers are being exploited or the entrepreneurs are suffering a loss. While the consumers enjoy the benefits of cheap production, they are sometimes deliberately defrauded, or they are at the mercy of the monopolist who does not hesitate to exploit them. On the other hand, the joint-stock company may suffer because of the inefficiency of the directors or because the directors are enriching themselves at the expense of the company. Thus, we generally find divergence rather than harmony between the interests of the producers and the consumers, for the price-mechanism does not often work smoothly and freely.

Allocation of Working Force. Everybody has his own preference for jobs. He may like to be a top business executive, a barrister, a top physician or a chief engineer, but he cannot. He must be put in his proper place. This is the unpleasant task of the price system to assign each individual a task which befits him. Not only that, the price system allocates the available labour force of a country numbering millions among thousands of trades, professions and

1. McConnel, Campbell, R.—*Economics, Fourth Edition*, p. 75.

occupations some of which are monotonous, irksome and hazardous. A change in wages and terms of employment matches the demand for, and supply of, each type of the workers. By adjusting the patterns of economic activity, the price system balances the relevant supply and demand. There is no favouritism and no injustice. No other alternative, e.g., job assignment by a central authority can do the trick except allocating jobs and workers by a free labour market.

Curbing Consumption. There may be times, for instance, when saving is essential for the growth of the economy or when shortage is feared and consumption has to be curtailed. In a free enterprise economy, a rise in the price will achieve the objective. Rationing cannot be tolerated for long and is out of place in normal times. An economy exists to provide the people with goods and services they want, unless they voluntarily withdraw their demand in the face of a forbidding price. The price system compels the consumers to express their preferences in a manner to match these preferences to the producers' costs. The price system thus makes an economy responsive to the desires of the consumers and the potentialities of the producers.

Allocation of Resources. The productive resources of the community, human and material, are automatically, as it were, allocated among the various uses in such a manner that each makes a maximum contribution to the total output. Any misallocation will be rectified by transfer of resources from one use to another through the price indicator. For instance, if land can make more valuable contribution, when devoted to the cultivation of food crops, the object will be accomplished by higher food prices.

Making for Adaptability. Unless an economy is flexible and can quickly adapt to changed situations, efficient administration of its resources over time will not be possible. It is the price system which provides this flexibility. As consumers' tastes and preferences change, new techniques of production are discovered and resource supplies undergo change, the particular allocation of resources will cease to be appropriate and efficient. Adjustment to change takes place through price-mechanism. For instance, if consumers' tastes change, this will be communicated to the producers through changes in prices of the goods concerned. If certain goods are in greater demand, the prices will rise giving supernormal profits to the existing firms and new firms will enter the industry. In this way, industry expands to meet increased demand. On the other hand, the industry, for whose goods the demand is decreased, contracts. Thus, the price system elicits appropriate responses from the firms as well as resource suppliers to a given change in consumers' preferences. This is its **guiding and directing function**. By and large, the competitive price system is conducive to

technological improvements and capital formation and other inter-related changes which result in greater productivity and higher level of material well-being.

CRITICAL APPRAISAL OF THE PRICE SYSTEM

Let us assess the functioning of price-mechanism in order to see whether the price system is an unflinching guide as to what is to be produced, how much to be produced, how production is to be organised and how total output is to be distributed. Opinions naturally differ on this point. There are people who are all praise for the competitive price system which is a characteristic feature of capitalism and there are those with whom it does not find favour. We shall briefly notice the case for and the case against the price system.

Case For

(i) In support of the price system, it is urged that it results in the most efficient allocation of the community's resources. All errors of judgment on the part of the producers are heavily penalised by losses. At the risk of being squeezed out, the producers have to employ the most economical techniques of production. Self-interest of the producers induces them to produce goods which consumers want the most. In short, the price system makes for maximum economic efficiency. The consumers get maximum satisfaction, the producers maximum profit and the community maximum welfare.

(ii) On the non-economic considerations, the proponents of the price system point out that it makes for economic freedom, i.e., freedom of enterprise for the businessmen and freedom of choice for the consumers. They are free to pursue their self-interest subject to rewards and punishments imposed by the price system itself.

Shortcomings

The price system is criticised on the following grounds:

(i) The economic freedom permitted by the price system kills wholesome competition which is beneficial to the consumers and the workers. Instead, the self-seeking and profit-greedy entrepreneurs combine to extract highest prices from the consumers and pay the lowest wages to the workers. This is sheer exploitation. "Combination, conspiracy, cut-throat competition, sheer productive efficiency are all means to the end of reducing competition and escaping its regulatory powers." (McConnel)² Attainment of maximum economic efficiency requires that

2. *Economics*, p. 86.

the number of producing firms should be small rather than large. Hence, rivals are ruthlessly squeezed out. The sovereignty of the producers kills the sovereignty of the consumers.

(ii) There is a clash between private interest and social welfare. The entrepreneurs are out to enrich themselves unabashed at the expense of the society. Social considerations are thrown overboard. Private individual interest runs amuck.

(iii) It is wrong to say that the price system leads to the production of goods which the consumer prefers the most. Rather, the consumers are hypnotised through high pressure salesmanship, to buy goods which the producers think most profitable to produce. The consumer's sovereignty is a myth. Besides, the competitive price system may not provide the consumers sufficient range of choice or for the development of new products. Pure competition entails product standardisation nor is it sufficiently progressive to develop new products.

(iv) The price system accentuates economic inequalities. It enables the big landlords and industrial magnates to acquire vast property which they can bequeath to their heirs and successors. Further, owing to differences in the quality and amount of human resources supplied by the households, there are galling differences in incomes.

(v) The critics also point out that the price system fails to register all costs and benefits. It shows only the costs of the producers, but not the social costs (*i.e.*, harm which a system of production may do to the society, *e.g.*, air and water pollution). Similarly, it only registers utility of purchase to the individual consumers but takes no notice of the benefits to the society, *e.g.*, purchase of an X-ray plant for a hospital.

(vi) The price system is an indicator of individual wants but not the collective wants of the community. For example, public parks, libraries, education and public health system cannot be purchased individually.

(vii) The price system does not help the economy of the country to adjust quickly to drastic changes in community's production targets. The actual pattern of resource allocation always lags behind the country's production targets. This is due to geographical and occupational immobility of the productive resources.

(viii) The price system has proved to be an imperfect mechanism for achieving full employment. Unemployment persists even in highly developed and prosperous societies.

(ix) The price system does not necessarily entail the use of most efficient productive techniques or the development of improved techniques. This is due to the fact that profit from innovations may be competed away; hence it is discouraged. A typical

competitive firm is too small to be able to finance research programmes.

Conclusion

While we cannot ignore the various criticisms levelled against the price system and some of them are reasonably valid, we have to concede that the functioning of the price system is conducive to economic efficiency, *i.e.*, efficient allocation of community's productive resources.

On the whole, it may be said that working of the capitalistic system is out of the question without price-mechanism. It is price-mechanism which brings about the necessary adjustments between the various parts of the economic system. It is really difficult to think how the economic system could function in the absence of price-mechanism. It is not intended to convey that the economic system functions justly or smoothly. But price-mechanism certainly enables it to work normally without any serious jolts and jerks.

PRICE SYSTEM IN A DEVELOPING ECONOMY

In a mixed economy like ours, where a substantial part of economic activity is in the private sector, the price-mechanism has an important role to play. Decisions regarding what to produce, how much to produce and how to produce are based on prices of products. This has been recognised in our plan documents.

One cannot, however, rely entirely on the free play of market forces. The prevailing price-relations in an under-developed economy are often the result of market imperfections and rigidities. Nor can the changes in relative prices by themselves always bring about the necessary re-allocation of resources quickly or adequately. In a developing economy, the basic trend of government operations in the fiscal and monetary fields is inevitably expansionary. Hence economic policy in a planned economy must influence the allocation of resources through direct controls, allocations, taxation and subsidies, so that it conforms more closely to the objectives of planning. Thus, the government has to play a positive role in using the price-mechanism for purposes of planning.

There is no aspect of economic policy which does not, in some way or another, have a bearing on prices. Therefore, we should not view price policy in isolation. It must be viewed as an integral part of general economic policy. We may conclude that within a broad framework of a properly conceived plan, the evolution and implementation of an appropriate price policy can be of assistance in securing the plan objectives.

BOOK TWO :
THEORY OF INCOME AND
EMPLOYMENT
OR
MACRO-ECONOMICS

Part I

Social Accounting

37

SOCIAL ACCOUNTING

Introduction: Importance

Social accounting or preparation of social accounts has assumed great importance in modern times. This is so because economic theory is being increasingly applied for the solution of practical problems. If study of economics is to be fruitful, knowledge of social accounts is absolutely essential. In the absence of a clear picture of the working of the economy, an economist is seriously handicapped in giving practical advice to the government or to businessmen. It is only with the help of social accounting that one can clearly trace the effects of changes in one section of the economy on its other sections. Nature of economic relationship is very complex in the present-day world, and therefore, no student can clearly grasp the principles of economics without having in his mind a clear picture of the economy which only social accounting can give.

What is Social Accounting

As we have said above, the nature of economic relationship has become very complicated in modern times. We have also said that it is the business of social accounting to make the understanding of this relationship easy and simple. But what precisely is social accounting? 'Social accounting' is a term which is applied to the description of the various types of economic activities that are taking place in the community in a certain institutional frame-work. In social accounting, we are concerned with statistical classification of the economic activity so that we are able to understand easily and clearly the operation of the economy as a whole. In the words of Stone and Murray, "The term social accounts is used in a general sense to denote an organized arrangement of all transactions, actual or imputed, in an economic system. In such a system distinctions are drawn between: (i) Forms of economic activity, namely, production, consumption

and accumulation of wealth; (ii) sectors or institutional division of the economy; and (iii) types of transactions, such as sales and purchase of goods and services, gifts, taxes and other current transfers, etc."¹

Here is another description of the field of social accounting, "The field of studies-summed up by the words 'social accounting' embraces, however, not only the **classification** of economic activity, but also the **application** of the information thus assembled to the investigation of the operation of the economic system."² Social accounting is thus concerned with the analytical as well as the statistical elements of the study of national accounts.

In social accounting, a transactor is supposed to keep a set of three accounts in which transactions are entered: (a) In the first account, incomes and outgoings relating to a productive activity of the transactor are brought together. The difference between the two indicates the profit or the gain. (b) The second account seeks to show how this profit and any other income that accrues to the transactor are allocated to different uses. The excess of income over outlay is the measure of savings. (c) The third account shows how this saving and any other capital funds are used to finance the capital expenditure or to give loans to other transactors. These accounts show the assets and liabilities of the person concerned at the beginning and at end of the accounting period. Since in the economy as a whole, the transactors are numerous, they are grouped into sectors. In the sector, accounts of a same type are consolidated. The sector accounts form the units in a system of social accounts.

1. Stone, Richard and Croft-Murray, G. *Social Accounting and Economic Models*, 1959, pp. 9-10.

2. Eday, H.C. and Peacock, A.T., *National Income and Social Accounting*, 1959, p. 11.

Social Accounting and Private Accounting Compared

We have all a fairly clear idea about accounting of private individuals. They make purchases and after incurring some incidental expenses sell the commodities they have purchased or manufactured. After debiting some necessary expenses, we are able to ascertain the amount of profit that they make. We also know that at the end of the year, they draw a profit and loss account and a balance sheet which sets out their assets and liabilities. But social accounting is not as simple as that. Social accounting, on the other hand, becomes complicated in several ways. We may set down below certain similarities and differences between private accounting and social accounting:

- (a) It is common knowledge with those dealing with accounts that private accounting is done on the method of double entry book-keeping. That is, each transaction is recorded twice in the books of the businessman. For instance, a cash sale will be entered once in the appropriate ledger account as a credit to the good sold, and it is again entered in the cash account as a debit in respect of the cash received. In social accounting, however, cash transactions are not separately presented. On the other hand, cash balances are recorded in the capital transaction account. This shows that social accounting also adopts the double entry method but the second entries are not recorded in detail.
- (b) Another difference between private accounting and social accounting is that private accounts relate to the individual businessmen. Each transaction is thus recorded from one point of view only. On the contrary, social accounts relate to a connected and closed network of all businessmen. There are no loose ends. Each transaction is recorded from the point of view of the two transactors connected with it.
- (c) Still another difference between private accounting and social accounting may be noted. As we have mentioned above, the accounts of private businessmen are usually presented in the form of a profit and a loss account which shows income and its allocation. There is also set out a balance sheet which shows the stock of assets and liabilities at the end of accounting period. The profit and loss account of a private individual resembles in social accounting to what is called as the appropriation account. The only difference is that in private accounting, the profit often includes some elements of costs such as depreciation on plant and machinery and fees paid to the directors of the company. On the other hand, in social accounting, these incomes are shown net. There is no counterpart at all of a balance sheet in social accounting since there are insuperable difficulties in collecting the necessary information completely and on a uniform basis

regarding assets and liabilities of all transactors in the economy.

Uses of Social Accounting

One might as well ask what after all is the object and purpose of preparing social accounts. To what uses are the social accounts put? We may give below in a summary fashion the purposes for which the social accounts are prepared or the uses to which they are put:

- (i) One purpose of the preparation of social accounts is to give the reader a clear picture of the economy as a whole. As we have mentioned above, in social accounting we find a classified account of the various transactions entered into in the various sectors of the economy. From these transactions, we can have a fairly clear view of the working of the entire economic system. These days we are not so much interested in the accounts of private businessmen, however big and prominent they may be. On the other hand, we are interested in the health of the economy and the way in which it functions. It is understood that a healthy economy is able to impart strength and health even to the economic affairs of private individuals. Intelligent citizens are keen on knowing how an economy is faring at a particular moment or in a particular period of time. This is clearly reflected in social accounting.
- (ii) We are all interested in the health, efficiency and stability of the economy since, as it has been mentioned above, a healthy and efficient economy results in the health and efficiency of individual businesses. We are all fairly familiar with the objectives placed before our Five-Year Plans. They all relate to the economy as a whole. We want economic growth with stability or we want to build up a self-reliant and self-generating economy. How far we are able to achieve these objectives can be very well found out from the description of various transactions and activities given in social accounting. Thus, if we want to promote efficiency and stability of our economy, preparation of social accounts is a 'must'.
- (iii) Measurement of economic welfare is another purpose of the preparation of social accounts. We have mentioned above our Five-Year Plans. At the end of each plan, we naturally like to know how far the masses have benefitted from the plan. In other words, we want to know to what extent economic welfare of the masses has been promoted. From social accounting and its study we can know at a glance to what extent the masses are better off than at the time when planning started.
- (iv) There is another important use which social accounting serves. From a study of social accounts, we are in a position to find out how the different sectors of the economy are inter-related to each other. We can, for instance, find out to what extent

the industrial sector depends on the agricultural sector, and vice versa. We can also know to what extent the growth of our export sector is conditioned by our industrial and agricultural growth. To the economist and to the person engaged in economic planning these inter-relationships are of very great use. In fact, the information that social accounting furnishes about the mutual relationship of the various sectors of the economy is indispensable, if planning is not going to be a leap in the dark.

(v) Social accounting serves a very practical purpose for the statesman, the government administrator and the politician. It is on the basis of social accounts that intelligent and effective government policies in fiscal, monetary and other economic spheres can be formulated and executed. In the absence of social accounts, such policies can well be mis-leading and may result in economic disasters. The national resources are limited and it will be criminal to fritter them away. It is, therefore, very necessary that every care is taken in the formulation of national policies. Only social accounts can give us proper guidance in this connection.

PREPARATION OF SOCIAL ACCOUNTS

We shall now see how social accounts are prepared (a) in a closed economy and (b) in an open economy.

Social Accounts in a Closed Economy

A closed economy is one which is self-contained and self-sufficient so that it has no economic dealings with the outside world. In a closed economy, the accounts may be prepared on a

uniform basis and classified as (i) transactions relating to productive activity, (ii) transactions relating to the use of gain from productive activity, and (iii) transactions relating to capital. By consolidating these accounts, prepared for each firm, household, government agencies and every other transactor, we get three consolidated national accounts: (a) Production (or operating) Account, (b) Consumption (or appropriation) Account and (c) Accumulation (or capital transactions) Account. These accounts can be illustrated for purpose of easy understanding, by a diagram given below. These transactions are represented as flows of money by arrows.

Let us now say a word about these accounts one by one.

Production (or operating) Account. In this account, we include all productive activities being carried on in the entire economy. It is a consolidated production statement relating to all firms operating in the economy. These firms manufacture commodities meant for consumption and capital goods and equipment for generating and accumulating wealth. The income obtained by these firms, therefore, comes through two channels, viz., partly by selling consumption goods and partly by selling capital equipment. The flow of this revenue is indicated by the direction of the arrows. In these accounts, purchases and sales by the firms from one another have obviously to be left out. They may be important for individual accounts but in the national or social accounts they cancel out.

In the above diagram, all proceeds of the sales of production are shown as being paid out as income to

Diagram 1

Simplified Flow Diagram for Closed Economy

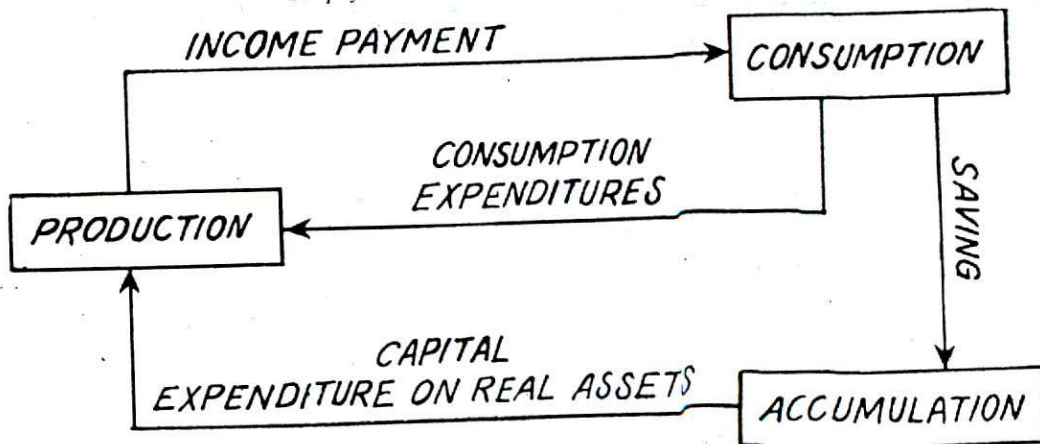


Fig. 37.1

3. See Richard Stone and Giovauna Croft Murrury, *A Social Accounting and Economic Models*, 1959, p. 12.

the factors of production, *i.e.*, to those who have taken part in the process of production. These payments take the form of wages to labour, salaries to the other employees, interest on capital, rent to the landlord, *i.e.*, the owners of land and business premises and profit going to the entrepreneur. The incomes are shown in the diagram as flowing from production to consumption by an arrow.

The incomes received by the factors of production may be either spent on consumption, *i.e.*, for the satisfaction of their immediate wants or they may be partly saved. The income spent on commodities is shown by an arrow going towards production from consumption and the part which is saved is shown by the arrow going from consumption to accumulation. Transactions of the consumers among themselves paying cash to one another are ignored just as we ignored the buying and selling activities of the producers among themselves.

We have seen that the whole of the sale proceeds of production go to consumption and consumption income is divided into spending and saving. It follows, therefore, that saving is equal to capital expenditure on real assets (or investments). We thus get the following two independent equations:

$$Y = C + S \quad \dots (1)$$

$$S = I \quad \dots (2)$$

In these equations Y stands for income paid to the factors of production, C stands for consumers' expenditure on commodities and services and S is savings (or income not consumed) and I is investment (or capital expenditures on real assets). If we substitute for S from equation (2) into equation (1), we get the following equation:—

$$Y = C + I \quad \dots (3)$$

As we shall see in the Keynesian theory discussed in the chapters that follow, these equations are the fundamental identities around which the Keynesian theory is built up. These relationships pertain to a simple closed economy which has no trading or financial relations with the rest of the world.

Social Accounts in an Open Economy

An open economy is an economy having economic dealings with other countries of the world. Simple representation of an economy given above seems to be unrealistic when we think of what happens in the economic systems actually. We find that in practice economic systems are seldom self-contained. On the other hand, they enter into extensive economic relationship with other economic systems. To put it into simple language we can say that no country in the world is self-sufficient. There may have been a time when various territorial units were ignorant of the existence of one another and led a self-contained economic life. It is often said that the

Indian villages were at one time economically self-sufficient, *i.e.*, they consumed what they produced and produced only for domestic consumption. Such a situation no longer exists. On the other hand, we find that there are extensive trading and capital transactions taking place between one country and another. Even China, a big communist country, is not self-sufficient. Thus, we find that economic systems of the world represented by individual countries are very closely inter-related. If social accounts are to be realistic, we have to introduce this new element in social accounts which may be represented by 'rest of the world'. Hence, for an open economy, we can classify the social accounts into (a) Production; (b) Consumption, (c) 'Rest of the world' and (d) Accumulation. The 'rest of the world' stands for the totality of other economies with which the economy in question is connected by virtue of economic relationship, *i.e.*, buying and selling, borrowing and lending, *etc.*

The diagram 2 given on next page represents the social accounting picture of an open economy.

It will be seen that this diagram is partly a reproduction of the earlier diagram. But we can also notice the additional items namely imports and exports of the country in question from and to the rest of the world. These are shown by flows connecting production with the rest of the world. In view of the fact that the value of the exports and imports is seldom equal to each other, we show in the diagram an arrow towards the rest of the world from accumulation. This flow is called 'net lending abroad'. This is a situation in respect of a country which has a favourable balance on current account which it lends abroad.

Like diagram No. 1, the flows into any given box in diagram 2 sum up to the same total as the flows which go out of that box. The total has to be same, it cannot be otherwise. We, therefore, get the following equations:

$$Y = C + S \quad \dots (4)$$

$$S = I + L \quad \dots (5)$$

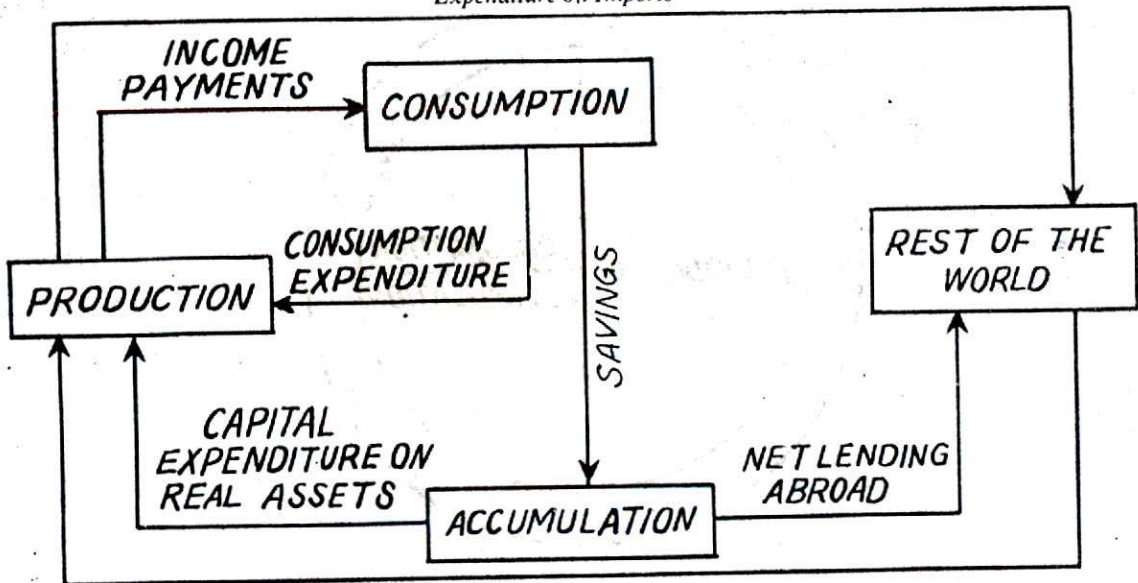
$$X = M + L \quad \dots (6)$$

In the above equations, as before, Y stands for income (paid or payable to the factors of production), C is consumers' expenditure on commodities, S the saving (or the income not consumed) and I the investment (or capital expenditure on real assets). The new letters X , M and L stand respectively for exports, imports and the net amount lent by a country in question to the countries in the rest of world.

Coming to the equations we can see that equation (4) is the same as equation (1). It is with equation (5) that we come to the difference made by introducing the 'rest of the world'. This equation shows that saving is no longer equal only to domestic investment as in diagram (1); but it is equal to domestic

Diagram 2

*Simplified Flow Diagram for an Open Economy
Expenditure on Imports*



Revenue from Exports
Fig. 37.2

investment plus foreign investment. That is, saving of a country finances investment (or capital expenditure on real assets) both in the country itself and also includes the amount which it lends abroad. In the new equation (6), showing relationship between transactions with the rest of the world, value of exports of the country in question is equal to the value of its imports plus the amount which it lends to the foreign countries. By substituting for L from equation (6) into (5) and then for S from (5) into (4), we get the equation

$$Y = C + I + X - M \quad \dots(7)$$

This equation shows that income payments are equal to consumption expenditures plus capital expenditures on real assets (i.e., investment) plus the excess of the value of exports over the value of imports.

There is no doubt that the second diagram is more realistic than the first one, but even this diagram is not sufficiently detailed as to give a complete picture of an economic system actually in operation. It does not contain provision for depreciation which is very necessary since buildings and machinery do not last for ever. Provision for depreciation is taken directly to 'accumulation' to serve as a source for funds meant for the replacement of real assets. It has, therefore, to be represented by an arrow pointing from production to accumulation.

Secondly, we have to take into account taxes, both direct and indirect, levied by the government and also subsidies given by it. Indirect taxes are paid from 'production' into 'consumption' which includes 'consumption' by government also. Actually, all subsidies paid by the government are set off against the indirect taxes received by them and the resulting flow is labelled 'indirect taxes (net)'.

We have also to take note of the fact that the number of transactions between the economy in question and 'the rest of the world' is much greater than described here. For instance, a firm belonging to the country may have set up branches abroad from which profits flow in. Also, the citizens of the country may earn dividends from investments abroad. These flows of income payments, from and to the rest of the world, have also to be taken into account. These are shown as received by and paid out by 'consumption'. Then there are gifts to and from abroad and also inter-governmental grants.

Thus, the actual situation to be represented by social accounts becomes very much complicated. If the social accounts are to be realistic, the flow called 'consumption expenditures' going from 'consumption' to 'production' should include not only the consumption expenditures of individuals and of non-profit institutions but also public consumption, i.e., of public agencies. It may be more convenient to route these expenditures abroad through the domes-

Diagram 3

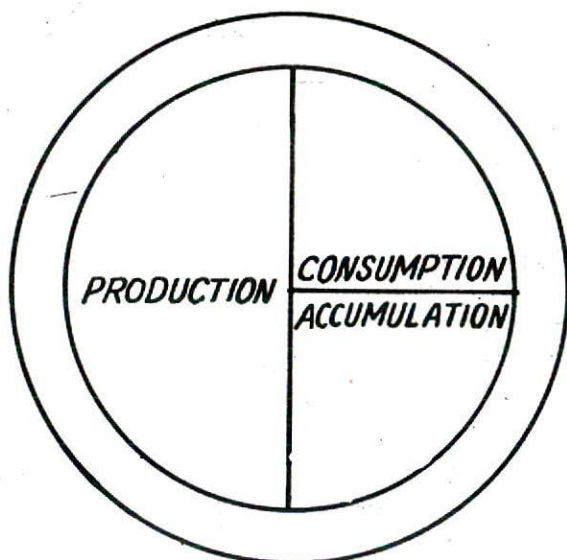
Regions and Boundaries

Fig. 37.3

tic productive system which will, therefore, appear both as imports and as purchases from production. Similarly, the flow of savings going from consumption to accumulation should include not only the savings of individuals, non-profit institutions and business firms but also savings by government agencies. In case the government has a deficit on current account showing negative saving, the flow of saving will represent the excess of positive saving by the private sector over the negative saving of the public sector.

Look at the above diagram. The outer circle represents the world economy and the inner circle domestic economy. The domestic economy is divided into two equal parts of which one is production and the second includes consumption and accumulation. This diagram can give some idea of the economic activities in a domestic economy and its relation with the world outside.

ITEMS INCLUDED IN THE VARIOUS SECTORS OF THE ECONOMY

We have classified the various sectors of the economy as production, consumption and capital transactions account. We shall now say a word about the various items included in the social accounts relating to each sector.

Items included in the Production Sector

The production sector has two sides: (a) the in-

comings (or revenues) and (b) outgoings (or costs). The sale proceeds of commodities is the principal source of revenue or income of the firms. The goods are sold to other producers to enable them to carry on current production, to consumers consisting of households, private non-profit institutions or departments of central and local government. Then, there are sales to other producers for the purposes of maintaining or adding to their capital equipment; and finally, there are sales to the rest of the world. The second source of revenue consists of the value of increase in stocks. The stocks increase when production exceeds sales. Still another source of revenue is subsidies given to industries, agriculture, etc.

As far the costs, they are incurred on the intermediate products used in production which are bought from domestic producers or from abroad. Depreciation is another item of costs. It represents that value of a fixed asset which is deemed to have been used up in current production. The third item on the cost side is the indirect taxes or taxes on sales.

When we balance the costs mentioned above against the revenue of sale proceeds of the goods, we shall get what the business has gained and which is represented by excess of revenue over costs. Out of this profit only that part will accrue to the entrepreneur which is left after paying wages and salaries to the employees and interest to the bond-holders.

An account is prepared on above lines for every producer in the country and all accounts are added together. Sales and purchases of intermediate products will be equal and are, therefore, to be

4. Stone, R. and G, *National Income and Expenditure*, 1961, p. 29.

cancelled out. Subsidies may be treated simply as reduction in costs instead of revenue and hence subtracted from indirect taxes. Conversely, imports instead of being considered as costs may be treated as reduction in revenues from exports. In this way, we can get a total revenue consisting of the proceeds of final sales to the domestic product. From this if we deduct depreciation, we get the net domestic product at market prices. Further, if we deduct indirect taxes net, we get net domestic product at factor cost or domestic revenue.

Items Included in Consumption

We have seen above the various items included in the 'production' sector of the economy. We now turn to consumption and examine its two sides, *viz.*, income and outlay.

The consumers' income consists of the income which accrues to a country from its productive activity at home and abroad. In other words, it is the national income. Besides, there is some income from indirect taxes, which is counted as usual net of subsidies. As far as consumers' expenditure, we may say that most of the income that a country derives from its productive activity is spent on the purchase of goods and services. The second item of outgoings of the consumer is what are called current transfers abroad. They take the form of private remittances and government grants given to foreign countries. After taking into account similar remittances and grants received from foreign countries, we can show these transfers as net. These two items, however, namely consumers' expenditure on purchase of goods and services and current transfers abroad, do not exhaust the consumers' income. There is another item in consumers' outlay that is savings. Saving represents the excess of consumers' income over expenditure and their transfers abroad. Saving then flows from consumption into accumulation and thus finances capital expenditure.

Item Included in Accumulation Capital Transaction Account

The incoming stream into accumulation *i.e.* capital transaction account comes from two directions (a) depreciation coming from production and (b) saving coming from transactions. These are the incomings.

As for the outgoings, we may notice three types: domestic investment which is made up of expenditure on fixed assets (minus sales of existing assets) and investment in stocks. Then there are foreign investments, *i.e.*, the net increase in the country's claims against other countries of the world. After deducting depreciation from domestic investment, we get net domestic investment which means net addition to the country's tangible wealth. By adding to it foreign investments, we can get the addition made to the country's wealth at home and

abroad. In addition to the two outgoings mentioned above, there is the third, *viz.*, capital transfers abroad. After deducting similar transfers received from abroad they are shown net.

Rest of the World: Balance of Payment Account

Let us now bring in 'rest of the world' and see the balance of payment account of a country with respect to the totality of the other countries of the world. We have to take notice of the mixture of income flows and transfer flows and current and capital transactions. The total income of country consists of the income derived from production, from transfers abroad both current (net) and capital (net) and borrowings from abroad. Against these, we set off the total outgoings consisting of imports and income from productive activity.

Conclusion

This briefly is the description of the process of the preparation of national or social accounts of a country. Social accounts show the structure of an economic system in the simplest form possible. The accounts are classified as production, consumption and accumulation. It may, however, be noted that it relates only to flows. This means that it only shows the additions to the country's wealth during a given period, but it leaves out the accumulated wealth already existing at the commencement of a period. That is, it does not give a balance sheet.

PROBLEMS OF CLASSIFICATION AND DEFINITION⁵

The description and classification of social accounting given in the previous sections seems to be a very simple affair. But it is not so. Actually when we come to define the various concepts and make an attempt to classify the accounts of an economy, we encounter several difficulties some of which appear to be formidable and insurmountable. We discuss below the main problems which arise from classification and definitions adopted in social accounting.

(i) Classification into Sectors and Accounts

One problem that has to be faced in the preparation of social accounts relates to classification into sectors and accounts. The relationships between production, consumption, investment and saving are easy to identify in hypothetical examples. But when we consider the actual accounting behaviour, we may not find it possible to ignore the various types of transactors operating in the economy. For instance, the directors of the company will not act as the owners of private businesses will do. The

5. Edey, H. C. and Peacock, A. T. *National Income and Social Accounting*, 1959, p. 63-78.

decisions of both are different from those of the government administrators running trading and manufacturing corporations. Hence it may be necessary to divide and sub-divide further the firms or production accounts corresponding to different types of business organizations. For example, it may be proper to sub-divide the government accounts into accounts for central and local government. Similarly, the accounts of the firms can be split into parts like 'production' or 'operating account' and the 'operating account' and the 'appropriation account'. It may be found necessary to have separate operating account and a separate appropriation account for each sub-division.

(ii) Transactions Recorded

Social accounts which are usually prepared do not give a complete picture of an economy. There are several items which are not included in the accounts. These items relate to the flow of value resulting from the production of new goods and services during the accounting period, the incomes accruing to the factors of production in respect of a particular product, current transfers of wealth between the various factors and to and from abroad in the forms of gifts, grants, taxes, interest payments, *etc.*, and capital transfers. The social accounts also exclude domestic transactions in second-hand goods. A common difficulty relates to the inclusion of the value of paid services of wives and other members of the family.

The payments of interest and dividends also raise a conceptual problem. We have to see to what extent should such payments be deemed to have been made in respect of current services and therefore, treated as part of the aggregate gross national product and to what extent should they be treated as mere transfers of income between sectors. Business payments of interest and dividends to shareholders also raise some difficulties. We have to decide whether such payments are to be considered as part of factor income or as transfers for the purpose of social accounting.

There is still another problem which arises in dealing with financial intermediaries such as banks, investment trusts, *etc.* The payments that these organizations receive for services rendered in the form of commissions, *etc.*, the payments that they make to the factors of production (*e.g.*, salaries and rents and payments for goods and services purchased) are similar to what the other firms do. The difficulty is that these institutions make profit from the excess of interest and dividend received over interest paid. If interest received and payments are treated as transfers, it may give a negative figure for their gross product which does not reflect the true state of affairs.

(iii) Principles of Measurement

The items included in social accounts are

measured in terms of exchanges at market values. When payments are made in terms of money, the valuation problem does not arise. But difficulty arises when we come to measure transactions which are not to be settled in money terms. These problems arise in the case of, for instance, valuation of farm products consumed by the farm household or the imputation of rental values to owner-occupied property. Difficulty particularly arises in the case of the valuation of investment in the form of both fixed capital formation and stock changes including the problem of depreciation. The difficulty also relates to the valuation of semi-manufactured and unsold finished stocks held by the firms. No market transactions may be available as guidance for setting value on them. Deductions may also have to be made for obsolescence, *etc.*, in the case of stocks. To make allowance for depreciation is not easy because capital equipment undergoes changes both in quality and quantity. The wornout capital equipment may not be replaced by identical equipment.

(iv) Collection of Data

Difficulties arise in connection with collection of data also. The difficulties relate to method of collection as well as the reliability of the data collected. Theoretically, if we independently calculate the sum of the net output of individual industries, sum of factor incomes and the sum of purchases of final output, it would furnish a check on each other in respect of different estimates of national income. But actually independent estimates can be made only to a limited extent, partly because some components of each aggregate are available from one source only and partly because of the practical difficulties in finding adequate reliable statistics.

As for the reliability of the data, it can be ensured if all economic units are required to keep accurate accounts and classify them according to official definition and reveal the information required correctly and completely. But the accounts actually kept are defective in all these respects. Even if the business units cooperate with the authorities, errors of computations may still arise and errors of classification may also be committed. Hence statistical discrepancies may arise when we aim to achieve consistency in the estimation of the national income from different sources.

Conclusion

There is no doubt that the task of preparing social accounts is beset with difficulties, but as the data improve both qualitatively and quantitatively, we may succeed in preparing national accounts in which errors are reduced to the minimum and the accounts become more reliable and dependable.

Introduction

In the foregoing chapters, we were mainly concerned with the principles determining the pricing of products and of productive resources. The theory of value and distribution is mainly concerned with these principles. Although prices are very important yet of much greater importance is the volume of activity and employment in the economy as a whole. This subject, however, did not receive sufficient attention till recently. The Great Depression of the 'thirties of this century led economists to think much more seriously about this question and Keynes' "General Theory of Employment, Interest and Money" was the pioneering attempt in that direction.

The subject has further gained in interest and importance as a result of the quest for rapid economic development in the under-developed areas of the world. The factors on which the state of employment in a country depends deserve serious consideration. We, therefore, now enter a new and much more extensive territory. Our interest now shifts from the constituent parts of an economy (micro-economics) to the economy as a whole. We referred to the term macro-economics in the second chapter of the book while discussing the scope of economics. We repeat that macro-economics is concerned with aggregates such as national income, total output and employment, total consumption, saving and investment, aggregate demand, aggregate supply and the general level of prices.

The problem which we have now to investigate is: What determines the level of total output and employment in the economy and fluctuations in this level from one time period to another? Thus, the problem has two aspects: (a) the determination of aggregate output and employment at a moment of time; and (b) the analysis of causes of fluctuations in this level over a period of time. The second aspect

opens out into the theory of industrial fluctuations more commonly known as the theory of trade cycle. Output and employment in a country depend on the size of national income (in fact they depend on each other) to the study of which we now turn.

Dr. Alfred Marshall defines National Income (or National Dividend) thus: "Labour and capital of a country acting on its natural resources, produce annually a certain 'net' aggregate of commodities, material and immaterial, including services of all kinds. The word 'net' means that from the gross value of the output depreciation of capital must be deducted. In Pigou's words, national income "is that part of the objective income of the community including, of course, income derived from abroad, which can be measured in money." Prof. Fisher bases his idea of national income on consumption instead of production. According to him national income refers "solely to services received by ultimate consumers, whether from their material or human environment."

The concept of national income has three interpretations. It represents a receipts total, it represents an expenditure total and it arises out of the fact every expenditure is at the same time a receipt (expenditure by one is receipt by another); and if goods or services bought are valued at their sales prices, we have three-fold identity that the value received equals the value paid, equals the value of goods and services given in exchange.

Definition of National Income

Keynes' concept of national income is somewhere between G.N.P. and N.N.P. (discussed below). From G.N.P. he subtracts only the 'User Cost', i.e., reduction in the value of capital equipment actually used and not full depreciation.

According to present ideas, National Income may be defined as the aggregate factor income (i.e., earning of labour and property) which arises from



Circular Flow of Income
Fig. 38.1

the current production of goods and services by the nation's economy. The nation's economy refers to the factors of production (i.e., labour and property) supplied by the normal residents of the national territory.

There is a circular flow which can be shown in the above diagram:—

To explain the above idea let us take an economy, where there are only two sectors: households and firms. Firms are required to produce goods. To produce them, they require services of the factors of production. Thus, incomes of these factors arise in the course of production. The sales value of net production must equal the sum total of payments made by the firms to the factors of production in the form of wages, rents, interest and profits. These incomes in turn become the sources of expenditure.

Thus, income flows from firms to households in exchange for productive services while products flow in return when expenditure by the households takes place.

Thus, there are three measures of national income of a country:—

(a) as the sum of all incomes, in cash and kind, accruing to factors of production in a given time period, i.e., the total of income flows;

(b) as the sum of net outputs arising in several sectors of the nation's production;

(c) as the sum of consumers' expenditure, government expenditure on goods and services and net expenditure on capital goods.

The total of income flows, net outputs and final expenditures will be the same but the significance of each arises from the fact that they reflect the total

operations of the nation's economy at the level of three basic economic functions—production, distribution and expenditure.

The discussion of the various concepts of national income will make the meaning of national income clear.

CONCEPTS OF NATIONAL INCOME

We study below the five important concepts of national income, viz., the Gross National Product, Net National Product, National Income, Personal Income, Disposable Income.

Gross National Product (G.N.P.)

This is the basic social accounting measure of the total output or aggregate supply of goods and services. Gross National Product is defined as the total market value of all final goods and services produced in a year. It is a measure of the current output of economic activity in the country.

Two things must be noted in regard to gross national product:

(i) It measures the market value of the annual output. In other words, G.N.P. is a monetary measure. There is no other way of adding up the different sorts of goods and services produced in a year except with their money prices. But in order to know accurately the changes in physical output, the figure for gross national product is adjusted for price changes by comparing to a base year as we do when we prepare index numbers.

(ii) For calculating gross national product accurately, all goods and services produced in any given year must be counted once, but not more than once. Most of the goods go through a series of production

stages before reaching a market. As a result, parts or components of many goods are bought and sold many times. Hence, to avoid counting several times the parts of goods that are sold and resold, gross national product only includes the market value of final goods and ignores transactions involving intermediate goods.

What do we mean by final goods? Final goods are those goods which are being purchased for final use and not for resale or further processing. Intermediate goods, on the other hand, are those goods which are purchased for further processing or for resale. The sale of final goods is included in gross national product, while the sale of intermediate goods is excluded from gross national product. Why? Because the value of final goods includes the value of all intermediate goods used in their production. For instance, the value of cloth includes the value of cotton used in the making of cloth. The inclusion of intermediate goods would involve double counting and will, therefore, give an exaggerated estimate of gross national product.

Another important thing to be borne in mind while calculating the G.N.P. is that non-productive transactions should be excluded. These are purely financial transactions or transfer payments like old-age pensions or unemployment doles which are merely grants or gifts or transactions relating to existing shares or second-hand shares.

Net National Product (N.N.P.)

The second important concept of national income is that of net national product. In the production of gross national product of a year, we consume or use up some capital, i.e., equipment, machinery, etc. The capital goods, like machinery, wear out or depreciate in value as a result of its consumption or use in the production process. This consumption of fixed capital or fall in value of capital due to wear and tear is called depreciation. When charges for depreciation are deducted from the gross national product, we get net national product. It means the market value of all final goods and services after providing for depreciation. Therefore, it is called 'national income at market prices.' Thus,

Net National Product

or

National Income at Market Prices

= Gross National Product -- Depreciation.

National Income or National Income at Factor Cost (N.I.)

The difference between 'national income at market prices' and national income at factor cost may be clearly understood. National Income at factor cost means the sum of all incomes earned by

resource suppliers for their contribution of land, labour, capital and entrepreneurial ability which go into the year's net production. In other words, national income (or national income at factor cost) shows how much it costs society, in terms of economic resources, to produce net output. It is really the national income at factor cost for which we use the term 'National Income.' The difference between national income (or national income at factor cost) and net national product (national income at market prices) arises from the fact that indirect taxes and subsidies cause market prices of output to be different from the factor incomes resulting from it.

Suppose a metre of mill cloth sold for Rs. 5 includes 25 P. on account of the excise and the sales tax. In this case, while the market price of the cloth is Rs. 5 per metre, the factors engaged in its production and distribution would receive only Rs. 4.75 P. a metre. The value of cloth at factor cost would thus be equal to its value at market price less the indirect taxes on it. On the other hand, a subsidy causes the market price to be less than the factor cost. Suppose handloom cloth is subsidized at the rate of 20 P. a metre and it sells at Rs. 2.80 P. Then, while the consumer pays Rs. 2.80 per metre, the factors engaged in the production and distribution of such cloth receive Rs. 3 per metre. The value of the handloom cloth at factor cost would thus be equal to its market price plus the subsidies paid on it.

Thus, national income (or national income at factor cost) is equal to net national product minus indirect taxes plus subsidies.

National Income
or

National Income at Factor Cost

= Net National Product (National Income at Market prices) -- Indirect Taxes + Subsidies.

Personal Income (P.I.)

Personal Income is the sum of all incomes actually received by all individuals or households during a given year. National income, that is, income received, must be different for the simple reason that some income which is earned—social security contributions, corporate income taxes and undistributed corporate profits—is not actually received by households and, conversely, some income which is received—transfer payments—is not currently earned. (Transfer payments are old-age pensions, unemployment doles, relief payments, interest payment on the public debt, etc.).

Obviously, in moving from national income, as an indicator of income earned, to personal income, as an indicator of income actually received, we must subtract from national income these three types of

incomes which are earned but not received and add incomes received but not currently earned. Therefore,

Personal Income = National Income – Social Security Contributions – Corporate Income Taxes – Undistributed Corporate Profits + Transfer Payments.

Disposable Income (D.I.)

After a good part of personal income is paid to government in the form of **personal taxes** like income tax, personal property taxes, etc., what remains of personal income is called **disposable income**.

Concepts Summarised. The following chart summarises the various concepts of National Income.

Disposable Income = Personal Income – Personal Taxes.

Disposable Income can either be consumed or saved. Therefore,

Disposable Income = Consumption + Saving.

MEASUREMENT OF NATIONAL INCOME

We have already pointed out that there are three possible measures of national income: (a) total income flows, (b) Net outputs, and (c) final expenditures. We have also said that all these methods arrive at the same result. Which of these methods is adopted in actual practice in calculating the national income of a country depends on the nature and condition of its economy as well as the purpose of undertaking this exercise. We discuss below briefly these methods.

Production or Output Method

This method approaches national income from the output side. According to this method, the economy is divided into different sectors such as agriculture, mining, manufacturing, small enterprises, commerce, transport, communication and other services. Then, the gross product is found out by **adding up net values of all the production** that has taken place in these sectors during a given year.

National Income Concepts Summarised

GNP	GNP	NNP	NI	PI	DI
Expenditure Approach	Income Approach	National Product	National Income	Personal Income	Disposable Income
Personal Consumption Expenditure	Wages	Wages	Wages	Wages	Consumption
	Rent	Rent	Rent	Rent	
	Interest	Interest	Interest	Interest	
Government Purchases	Dividends	Dividends	Dividends	Dividends	Savings
	Income of the unincorporated business	Income of the unincorporated business	Income of the incorporated business	Income of the unincorporated business	
	Corporate Income Taxes	Corporate Income Taxes	Corporate Income Taxes	Subsidies	
	Social Security contributions	Social Security contributions	Social Security contributions	Transfer Payments	
Gross Private Domestic Investment	Undistributed Corporate Profits	Undistributed Corporate Profits	Undistributed Corporate Profits	Subsidies	Personal Taxes
	Indirect Business Taxes	Indirect Business Taxes	Indirect Business Taxes		
Net foreign Investment	Depreciation				

In order to arrive at the net value of production of a given industry, the purchases of the producers of this industry from producers of other industries or sectors are deducted from the gross value of production of that industry. The aggregate or net values of production of all the industries and sectors of the economy plus the net income from abroad will give us the **Gross National Product**. By subtracting the total amount of depreciation from the figure of gross national product, we get the net national product, or national income.

This method of estimating national income enables us to trace the origin of the national income aggregate to the different sectors of the economy. Therefore, this is called **national income by industrial origin**.

This method can be used where there exists a census of production for the year. In many countries, figures of production of only important industries are known. Hence, this method is employed along with other methods to arrive at the national income. The one great advantage of this method is that it reveals the relative importance of the different sectors of the economy by showing their respective contribution to the national income.

Income Method

This method approaches national income from the distribution side. In other words, this method measures the national income after it has been distributed and appears as income earned or received by individuals of the country. Thus, according to this method, **national income is obtained by summing up of the incomes of all individuals in the country**. Individuals earn income by contributing their own services and the services of their property such as land and capital to the national production. Therefore, national income is calculated by adding up the rent of land, wages and salaries of employees, interest on capital, profits of entrepreneurs (including undistributed profits of joint-stock companies) and income of self-employed people.

This method of estimating national income has the great advantage of indicating the distribution of national income among different income groups such as landlords, capitalists, workers, etc. Therefore, this is called **national income by distributive shares**.

Expenditure Method

This method arrives at national income by adding up all the expenditure made on goods and services during a year. Income can be spent either on consumer goods or investment goods. Thus, we can get national income by summing up all consumption expenditure and investment expenditure made by all individuals as well as the government of a country during a year. Hence, the gross national product is found by adding up—

(a) what private individuals spend on consumer goods and services. This is called **personal consumption expenditure**;

(b) what private businesses spend on replacement, renewals, and new investment. This is called **gross domestic private investment**;

(c) what the foreign countries spend on the goods and services of the national economy over and above what this economy spends on the output of the foreign countries, i.e., exports minus imports. This is called **Net foreign investment**; and

(d) what the government spends on the purchase of goods and services, i.e., **government purchases**.

We have explained above the three alternative methods of estimating national income. The best way to arrive at national income will be to employ all these three methods so as to permit their cross-checking ensuring greater accuracy and throwing more light on details.

Identity of the Methods

It can be shown that the results obtained by any of these methods are similar to those obtained by any other method. But allowance will have to be made for all errors of omission and commission and all discrepancies removed before they can be found to be exactly the same. Hence in actual practice more than one method is used in combination with one another so as to provide a cross check.

By way of a hypothetical illustration we may take total GNP as 18,564 (figures in crores of rupees) by the income method composed of Rental Incomes 692, wages, salaries and supplements 9,564, Net interest earnings 844, Dividends of Companies 760, undistributed profits 240, Profit Tax Liability 292, Inventory Valuation adjustment 144, Indirect Taxes 1,872, Business Transfer Payments 92, Subsidies 96, Depreciation (Capital consumption) 1,620, Income of Unincorporate enterprises 2,256, Statistical Discrepancy 92.

By the Expenditure Method, we can get the same result, i.e., total GNP 18,564 consisting of Personal Consumption Expenditure 13,488, Gross Private Domestic Investment 1,800, Government Purchases of Goods and Services 3,096, Net Foreign Investment 176.

Difficulties of Measurement

There are some conceptual problems that crop up when we start measuring the national income of a country. Some of these problems are enumerated below:—

(i) The first problem relates to the treatment of non-monetised transactions such as the services of housewives to the members of their families and farm output consumed at home. On this point, the general agreement seems to be to exclude the services of housewives while including the value of farm output consumed at home in the estimates of national income. This, however, gives rise to certain

anomalies. For example, if a man employs a maid-servant for household work, payment to her will appear as a positive item in national income. If next day, the man were to marry the maid-servant, she would be performing the same services as before but without payments. In this event, the value of national income would go down though the real amount of goods and services performed remains the same as before.

(ii) The second difficulty arises with regard to the treatment of the government in national income accounts. On this point, the general viewpoint is that, so far as the administrative functions of the government like justice, administration and defence are concerned, they should be treated as giving rise to final consumption of such services by the community as a whole, so that the contribution of general government activities will be equal to the amount of wages and salaries paid by the government. As regards capital formation by the government, this is treated at par with capital formation by any other enterprise.

(iii) The third major problem arises with regard to the treatment of income arising out of activities of the foreign firms in a country. Should their income form a part of the national income of the country in which they are located or should it belong to the national income of the country owning the firms? On this point, the IMF viewpoint is that production and income arising from an enterprise should be ascribed to the territory in which production takes place. However, profits earned by foreign branches and subsidiaries are credited to the parent concern.

Special Difficulties of Measurement in Under-developed Countries. In under-developed countries like India, we face some special difficulties in estimating national income. Some of these difficulties are given below:—

(i) The first difficulty arises because of the prevalence of non-monetised transactions in under-developed countries like India, so that a considerable part of output does not come into the market at all. Agriculture, still being in the nature of subsistence farming in these countries, a major part of output is consumed at the farm itself. The national income statistician, therefore, has to face the problem of finding a suitable measure for this part of output:

(ii) Because of illiteracy, most producers have no idea of the quantity and value of their output and do not keep regular accounts. This makes the task of getting reliable information from a large number of petty producers all the more difficult.

(iii) Because of under-development, occupational specialisation is still incomplete so that there is a lack of differentiation in economic functioning. For instance, an individual may receive income partly

from farm ownership, partly from manual work in industry in the slack season, etc.

(iv) There is a general lack of adequate statistical data and this adds to the difficulties of estimating national income.

SIGNIFICANCE OF NATIONAL INCOME STATISTICS

There are several important uses of national income statistics and, therefore, there is great need for their regular preparation. National income estimates provide not only a single figure showing the national income, but also supply the detailed figures in regard to the various components of the national income. It is both the figure of national income and the details regarding its various components that throw light on the functioning and performance of the economy.

The following are some of the important uses of national income estimates:

First, national income estimate reveals the overall production performance of the economy, as it seeks to measure the level of production in a year. Per capita income, which is found by dividing the total national income by the population, gives us an idea about the average standard of living of the people. Economic welfare depends to a considerable degree on the level of national income and the average standard of living of the people. Thus, the figures of national income and per capita income indicate the level of economic welfare of the people of a country.

Secondly, by comparing national income estimates over a period of time, we can know whether the economy is growing, stagnant or declining. If the national income increases over years, it means that the economy is growing. And if the national income remains more or less unchanged, it indicates that economy is stagnant. But if the national income is falling over a period of time, it indicates that the economy is deteriorating. In case the economy is growing, we can also judge the rate of economic growth or development by measuring the rate of increase in national income. Further, by comparing the per capita income over years, we can know the changes in the standards of living and economic welfare of the people.

Thirdly, the national income estimates show the contribution made by the various sectors of the economy, such as agriculture, manufacturing industry, trade, etc., to the national income. Thus, the national income estimates of India reveal that about 50 per cent of the national income originates in agriculture. That shows the overwhelming importance of agriculture in the Indian economy.

Fourthly, national income estimates throw light on the distribution of national income among different categories of income such as wages, profits,

rents and interest. The distribution of national income between wages on the one hand and profits, interest, rent on the other, is of special significance, since inequality in personal incomes depends to a large extent on the share of the working classes (*i.e.*, wages) and the share of property owners (*i.e.*, rents, profits and interest).

Fifthly, the national income estimates also contain the figures of consumption, saving and investment in the economy. Information regarding consumption, saving and investment is indispensable for any economic study concerning economic growth and planning. It is the rate of saving and investment in the economy that determines the rate of economic growth. Further, consumption plus investment constitutes the level of aggregate demand on which depends the level of income or employment in a country.

Sixthly, with the help of national income estimates of various countries of the world, we can compare the standards of living and the levels of economic welfare of the people living in those countries. For this purpose, we have to adjust the national income figures for differences in production and in price levels. In other words, by the figures of the 'real' national income per capita, we can compare the standards of living or levels of welfare in different countries of the world. Moreover, developed and under-developed countries are usually classified on the basis of per capita income.

Finally, national income estimates are a valuable guide to economic policy, especially in these days of development planning and active government intervention in the economy. By looking at the national income statistics, the government can decide if the economy or its various sectors need any stimulus or regulation. From the national income estimates we can see the part played by the government in the national economy.

In fact, no development planning is possible without national income estimates. National income estimates prove very useful for formulating plans and fixing targets of production. Preparation of plans depends very much on the availability of data regarding national income, consumption, saving and investment which are all provided by national income estimates. Further, we can assess and evaluate the achievements or otherwise of the development targets laid down in the plans from the changes in national income and its various components.

Conclusion. We may conclude in Samuelson's words thus: "By means of statistics of national income, we can chart the movements of a country from depression to prosperity, its steady long-term rate of economic growth and development, and finally, its material standard of living in comparison with other nations."

LIMITATIONS OF THE GNP APPROACH

Is Growth of GNP an Unfailing Index of a Country's Economic Progress ?

We are accustomed to judging a country's economic progress from the rate of growth in GNP (Gross National Product). The GNP represents the total value of goods and services produced in the country in that year. It also includes the amounts set aside as savings or investment, expenditure incurred by Government and by the people.

But this concept of national income is not altogether valid in a country like India, where everything is not commercialised. For instance, the value of the services rendered by a son in the firm or a wife on the farm are not generally evaluated and paid for on the basis of the ruling wage rates. Besides, much of the work done in institutions like panchayats, village co-operatives, etc., is honorary. Hence tremendous amount of national economic effort is not recorded in GNP. Nor is a good deal of production in villages like foodgrains, milk, vegetables by self-employed small and marginal farmers reckoned in GNP.

It has, therefore, been suggested that it would be more appropriate or sensible for the developing countries to measure economic progress in terms of improvement in the conditions of the people, *i.e.*, in terms of increase or decrease in human suffering. That is, we can measure it by the size and magnitude of poverty and unemployment. Recently, a few economists in America, having recognised the unreliability of GNP as an instrument to measure a backward country's progress, have come up with a new index which they call the 'quality of life' to measure a developing country's progress.

The quality of life criterion consists of three factors, *viz.*, literacy, life expectancy and infant mortality. Since this index does not say anything about the distribution of income or relative level of poverty (things which are vital for judging a people's economic conditions), it may not be considered very sound. Hence it is necessary to take into account the indices of consumption by people belonging to economically weaker sections of the people.

Production of 'use values' is, suggested as another criterion. 'Use values' is defined as basic goods and services necessary for a person's survival. Thus, the total output of a country can be classified as 'basic goods or use values' and non-basic goods or 'exchange values.' In India 'use values' could include foodgrains like rice, wheat, pulses, oils, vegetables, milk, sugar, salt, clothing, bricks, timber, etc.

Thus, the inadequacy of GNP as a yard-stick to measure the progress of developing countries is sought to be replaced by 'quality of life' and the 'production of use values.'

Part II

Theory of Income Determination

39

THEORY OF EMPLOYMENT

Having familiarised ourselves with the national income and its various concepts, we are now in a position to study the theory of Income and Employment. It is also called simply 'income theory' or 'employment theory.' In macro-economics, 'income' and 'employment' are interchangeable terms, since in the short run national income depends on the total volume of employment or economic activity in the country.

The problem of employment has exercised the minds of the economists from time to time. We shall first study the theory of employment as propounded by the classical economists and then the modern theory of employment, the Keynesian theory. In his book "**General Theory of Employment, Interest and Money**", the Late Lord Keynes is said to have ushered in a revolution in economic thinking and "opened up new vistas and new pathways to a whole generation of economists." It has also been called 'New Economics' or "Keynesian Revolution." There have been also subsequent modifications and refinements in the theory of employment introduced by later economists. The classical economists assumed the prevalence of full employment. They thought that there might occasionally be unemployment, but there was a tendency in the economic system towards full employment all the same. But the Great Depression of 1929-1934, engulfing the entire world in widespread unemployment, low output and low national income for about five years, upset the classical theory. Keynes showed convincingly that the classical doctrine was untenable on theoretical grounds while it had been proved untrue in the practical world.

CLASSICAL THEORY OF EMPLOYMENT:

The term 'classical economists' was first used by Marx to describe economic thought of Ricardo and his predecessors including Adam Smith. But by classical economists' Keynes meant the followers of

David Ricardo including John Stuart Mill, Alfred Marshall and A.C. Pigou. The term 'classical economics', as used by Keynes, refers to the traditional or orthodox principles of economics which had come to be accepted, by and large, by the well-known English economists since the time of David Ricardo. They were so widely accepted and well established for over more than a century that they were labelled as 'classical'. Being a student and disciple of Marshall, Keynes had himself accepted and taught these classical principles. But he repudiated the doctrine of laissez-faire which the classical economists strongly advocated.

The classical economists did not formulate any specific theory of employment as such. They only laid down certain postulates. The two broad features of the classical theory of employment were: (a) the assumption of full employment of labour and other productive resources and (b) the flexibility of prices and wages to bring about full employment.

Assumption of Full Employment. The classical economists assumed that labour and other resources were always fully employed. According to them, the general over-production, and hence general unemployment, is impossible. There may be lapses from full employment at times but these were regarded as temporary and abnormal. The normal situation is stable equilibrium at full employment. If at any time unemployment persists for a long time, according to classical economists, it is because of the interference by government or private monopoly with the free play of market forces or wrong calculations of businessmen or artificial resistances in the economic system. In the view of classical economists, thus, in a free competitive capitalist economy, the persistence of general unemployment is unlikely and that there always exists full employment or a tendency towards full employment.

The classical economists believed that the policy of laissez-faire guaranteed normal full employment.

According to them, the laissez-faire capitalism was self-adjusting. They had great faith in free and perfect competition, efficacy of the profit motive and price mechanism to remedy the temporary ills of the economic system and ensure full employment.

The classical theory does not explain what determines the level of employment. Instead, it assumes full employment and tries to explain the allocation of given resources in production and the distribution of income among the participating resources. The market forces of demand and supply allocate the resources as well as determine their rewards. It is the general relations of demand and supply which determine the relative values of commodities and services. The pricing system serves as the planning mechanism.

Flexibility of Prices and Wages. Apart from full employment of resources, the classical economists believed that it is the flexibility of prices and wages which automatically brings about full employment. For instance, if there is general overproduction resulting in depression and unemployment, prices would fall as a result of which demand would increase, prices would rise and productive activity will be stimulated and unemployment would tend to disappear. Similarly, unemployment would be cured by cutting down wages which would increase the demand for labour and would stimulate activity. As Prof. Pigou says, "With perfectly free competition, there will always be at work a strong tendency for wage rates to be so related to demand that everybody is employed."

Thus, if the prices and wages are allowed to move freely without any let or hindrance from the Government or monopoly interests, unemployment would disappear and full employment level restored in course of time.

Further, the classical economists treated money as a mere medium of exchange. They ignored its role in affecting income, output and employment. In other words, they took note only of the transactions motive for holding money and overlooked the precautionary and speculative motives.

Say's Law

Say's Law is the foundation of Classical Economics. Assumption of full employment as a normal condition of a free market economy is justified by classical economists by a law known as '**Say's Law of Markets.**' It was this law on the basis of which classical economists thought that general overproduction and hence general unemployment were impossible.

Statement of the Law. According to J.B. Say, an early nineteenth century French economist, '**supply creates its own demand.**' In Say's words: "It is production which creates market for goods; for selling is at the same time buying and more of production; more of creating demand for other

goods. Every producer finds a buyer." In other words, every supply of output creates an equivalent demand for output, so that there can never be a problem of general over-production. Say's law, thus, denies the possibility of the deficiency of aggregate demand.

Say's law so conceived describes an important fact about the working of the free exchange economy that the main source of demand is the sum of incomes earned by the various productive factors from the process of production itself. The employment of hitherto unutilised labour and other resources pays its own way, because it enlarges the market demand for goods by an amount equivalent to the income created and the value of output produced. A new productive process, by paying out income to its employed factors, generates demand at the same time that it adds to supply. It is thus production which creates market for goods. It is the cause and sole cause of demand. David Ricardo, the chief among the classical economists, said: "No man produces but with a view to consume or sell, and he never sells but with an intention to purchase some other commodity which may be useful to him or which contributes to future production. In the words of James Mill; "Consumption is co-extensive with production." Say's law was faithfully followed by the neo-classical economists like Marshall and Pigou.

Thus, supply creates its own demand not only at the same time but also to an equal extent. Demand is generated simultaneously through the act of supply because supply creates income in the form of wages, interest and profit.

Suppose 1,000 metres of cloth are produced. The value of cloth has been distributed in the form of wages, rent, interest and profit as reward to the participating factors of production. The purchasing power so generated will be spent either on purchasing the cloth or some other commodity. The factors of production producing the other commodity will receive purchasing power as reward which may be spent on the purchase of cloth or again some other commodity. Thus, the circle of production as well as of purchase goes on widening till the supply of no commodity remains unsold in the market. Hence, the total or aggregate supply of commodities in the economy would be exactly equal to aggregate demand. There being no deficiency of demand, general overproduction is out of the question. It may be that at any given time, the supply of commodity may exceed the demand for it, but it will be only a temporary disequilibrium; ultimately demand will equal supply and the entire production will be taken off the market, provided, of course, there is no interference in the working of the free market forces.

"In brief, Say's law of markets is a denial of the possibility of general overproduction, that is, a

denial of the possibility of a deficiency of aggregate demand. Therefore, the employment of more resources will always be profitable and will take place to the point of full employment, subject to the limitation that the contributors of resources are willing to accept rewards no greater than their physical productivity justifies. There can be no general unemployment, according to this view, if workers will accept what they are 'worth.'¹

Equality of Savings and Investment. According to Say's law, there will always be a sufficient rate of total spending so as to keep all resources fully employed.² Most of the income, which is earned by resources or factors participating in the productive processes, is spent on consumer goods and a part of it is saved. But according to classical economists, savings are spent automatically on investment goods. Savings and investment are interchangeable terms and are equal to each other. "Since saving is just another form of spending, according to the classical theory, all income is spent, partly for consumption and partly for investment. There is no reason to expect a break in the flow of income stream and, therefore, supply creates its own demand."

Equilibrating Rate of Interest. If there is any gap between saving and investment, the rate of interest brings about equality between the two. In McConnell's words, "As classical economists saw it, the economy was analogous to a gigantic bath-tub wherein the water-mark measured the level of output, and employment. Any leakage down the drain of saving would be returned to the tub through the tap of investment. This had to be so because the interest rate, connected the drain pipe and the tap."²

Pigou's Modification

According to Prof. Pigou, the unemployment which exists at any time is because of the fact that changes in demand conditions are continually taking place and that frictional resistances prevent the appropriate wage adjustment from being made instantaneously. Thus, according to classical theory, there could be small amounts of **frictional unemployment** attendant on changing from one job to another, but there could not be **involuntary unemployment** for a long period of time. All people who sought employment would fairly quickly find it, if the wage rate is perfectly elastic. If all people seeking jobs at the current wage rate could not obtain them, then the wage rate would be bid down. A decrease in the wage rate would reduce the cost, shifting individual product supply curves to the right, thereby lowering the price of the products. With given demand curves, a larger quantity would

be purchased at lower prices; then more people would be employed to produce the larger output. In short, Pigou applied to the labour market what Say said about the commodity market.

Thus, according to the classical analysis, if people were unemployed, wages would fall until all seeking employment were in fact employed. Involuntary unemployment which was found at times of depression was because of the fact that wages were kept too high by the actions of labour unions and government. Therefore, Prof. Pigou advocated that a general cut in money wages at a time of depression would increase employment. Hence, according to Prof. Pigou perfectly elastic wage policy would abolish fluctuations of employment and would ensure full employment of labour.

Basic Assumptions of Say's Law

It is clear from the statement of the law given above that it is based on certain assumptions. The main assumptions are:

(i) The law can operate only in a free-exchange economy where there is perfect freedom for the buyers to buy and sellers to sell. There prevails perfect competition and there are no restrictions imposed either on the producers or on the consumers and there is no price control.

(ii) There is free flow of money incomes. As these incomes are received they are immediately spent. Even savings must be invested and spent on acquiring producer's goods.

(iii) Savings are equal to investment and this equality is brought about by flexible interest rate.

(iv) The government follows the policy of *laissez-faire* and does not interfere in any manner with the operation of the market forces.

(v) The size of the market is limited by the volume of production; only then will demand equal supply or supply create its own demand.

Implications of Say's Law

From the above study of Say's Law some conclusions logically follow. The following are its main implications:

(i) One of the implications of Say's Law is that the economic system is self-adjusting and functions automatically without being directed by any controlling authority. If there is any disequilibrium, it will be only temporary and there is a persistent tendency for the equilibrium to be restored. For instance, if there is over-production, prices will fall, demand will increase and the extra supply will be cleared. Similarly, if there is unemployment, wages would fall and it would become worth while to employ more labour so that unemployment disap-

1. Dillard, D.—*Economics of J. M. Keynes*, p. 19.

2. McConnell, C. R.—*Economics*, 1969, p. 208.

pears. That is how there is automatic adjustment in the economic system. There is built-in flexibility.

(ii) An important policy conclusion that follows from Say's Law is that the governments should act on the policy of *laissez-faire* (let alone) or of non-interference in economic activities. Any interference by the government in the automatic working of the economic system will simply create imbalances and disequilibria. In the absence of government interference, the disequilibrium will be temporary and will tend to be rectified by the free operation of market forces. Hence, government should not raise barriers in the way of smooth working of the economy and economic forces. Owing to built-in flexibility in the economic system, prices, wages and interest rates and the volume of production keep changing as the economic situation may require and there is no need for the government to interfere.

(iii) Another important conclusion that follows from Say's Law is that general over-production is not possible. This is so because as soon as phenomenon of over-production appears, prices will fall and increased demand under the impact of fall in prices will clear the surplus stock. At times, there may be over-production in a particular industry as distinguished from general over-production. But this would also be temporary because automatically adjustments would come about. But, in a free and fully competitive system, general over-production is simply out of the question. Whatever the amount of the annual produce, it can never exceed the amount of annual demand.

(iv) Similarly, under free and perfectly competitive economic system, general unemployment is impossible. It follows from Pigovian formulation of Say's Law that a general reduction in wages would create enough demand for labour to remove unemployment. Only there should be no wage regulation by government and no trade union pressure to resist reduction in wages. There will thus be persistent tendency to full employment. There may be sometimes unemployment in a particular industry but no general unemployment is possible. Whatever the state of demand, there will always be, via wage adjustment, tendency towards full employment.

Say's argument can be explained by taking into account the money flows in an economy, based on the following simplifying assumptions:

1. All income received by households is immediately spent on consumption.
2. There is no government activity of any nature, *i.e.*, no government expenditure, taxation, subsidies, price control, *etc.*
3. It is a closed economy.

The following figure shows the money flows between the firms and the households in the econ-

omy. Firms pay for factor services in money-wages, rent, interest and profit which in turn is the income of households.

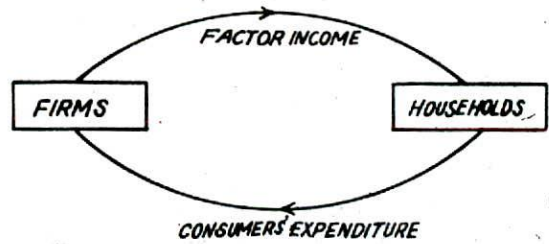


Fig. 39.1

Households spend their income on goods and services. There is thus a circular flow of money from firms to households and from households to firms.

The assumption that all income is spent is of course unrealistic. However, it does not make any difference with the working of this law even if we assume that some income is not spent, *i.e.*, part of income is saved.

The real question then is what happens to the unspent income or saving. Suppose we have an economy which has been in equilibrium with income = Rs. 1,000 and consumers expenditure Rs. = 1,000. Households now decide to save 1% of their income. The result will be firms' receipts fall to Rs. 900. Profits will fall and firms will tend to react by reducing output and hence reducing employment and income. But it should be pointed out here that in the above discussion we have assumed that firms only produce consumption goods and services, which is not true because they also invest in factories, machinery and stocks of raw materials and unsold goods. Provided that the firms wish to borrow and spend exactly the same amount that households wish to save, the circular flow of income can be maintained and so can employment. This is shown in the following figure, when savings are channelled to firms through banks.

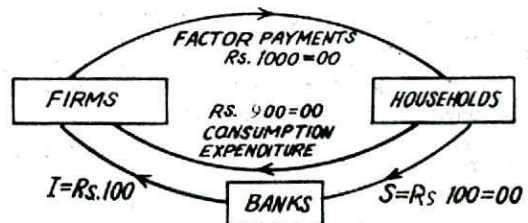


Fig. 39.2

(v) It also follows from Say's Law that it always pays to employ the unemployed or under-employed productive resources in the economy. The incomes earned by them will be sufficient to purchase their output and provide profit to the entrepreneurs. Thus, according to Say's Law, it would be unwise to

let the resources remain idle. They can be made to pay their way.

(vi) Since supply is supposed to create its own demand, there is no limit to productive activity in the economy. In other words, economic development can be carried out to any extent; for there can be no deficiency of aggregate demand. This opens out unlimited possibilities of economic development in under-developed countries.

Criticism of Classical Theory and Say's Law

Keynes in his *General Theory* made a vigorous attack on the classical theory of employment; mainly, it was two-pronged attack: (a) He bitterly criticised Say's law, "supply creates its own demand" and (b) in particular the Pigovian assertion that money wage-cuts would help to promote employment at a time of depression and unemployment.

As explained above, according to Say's law every supply of output creates an equivalent demand for output, so that there can never be a problem of general over-production and hence of general unemployment. Now, it is true that supply does create demand for goods and services, because various factors of production earn their incomes in the process of production by helping to create additional supply of output. When factors of production are employed, for instance, to produce cloth, they get their reward in the form of wages, rents, interest and profits.

But, from this, it does not follow that the entire supply of national output will always be demanded by them. The incomes of the factors of production are necessarily equal to the value added in the productive process, but it does not mean that the entire income will be automatically spent on goods and services created in a given time period. A part of income will be saved so that this part of income is not available to create demand for goods and services. Saving thus causes a break or leakage in the income stream and obstructs the income-expenditure flow. Unless investors are willing to invest to an equivalent extent of intended savings, the total effective demand which consists of demand for consumers' goods and producers goods will not be sufficient to absorb the entire available supply of output. And, if it happens like this, there will be over-production and the producers will not be able to sell their entire output, their profits will fall and they will reduce their production and this will create unemployment. Thus, supply does not necessarily create its own demand.

In a given time period, consumers are planning to spend a given part of their income and save the rest. Similarly, entrepreneurs are planning to invest in factories, machines, etc., to a given extent. The total effective demand is the sum of consumption and investment demand. Savers are saving for reasons

different from the investors and, in a free enterprise economy, there is no mechanism to ensure that what savers are planning to save is just equal to what investors are planning to invest. If the planned investment expenditure is not enough to fill up the gap of savings, then the present level of income and employment will not be maintained and, therefore, there will be a fall in income and employment.

Hence, the basic weakness of Say's law arises because of lack of any agency to ensure that intended savings are just equal to intended investment; and since savings and investment are undertaken by different persons and for different reasons, a discrepancy between the two is bound to arise, and, when it arises, the necessary mechanism to correct it is through changes in the volume of employment and income.

Thus, according to Keynes, there is no inherent reason to believe that investment expenditure plus consumption expenditure would always be equal to the cost of any given output; there is thus no assurance that demand would equal any given supply. Savings are determined primarily by income. But investment demand depends mainly, in the short run, on marginal efficiency of capital and rate of interest, and, in the long run, on factors like changes in technology and population growth. Therefore, investment demand, so determined, will not necessarily fill the savings gap between the income and consumption at the level of full employment and thus unemployment will be the result.

In sharp contrast to the classical view that full-employment equilibrium is the normal situation, Keynes in his *General Theory* firmly held the view that in a free private enterprise economy, there are more chances for the equilibrium to be established at less than full employment level.

Further, Keynes strongly opposed the Pigovian view that unemployment would disappear, if a general cut in money wages was applied. A general cut in wages, according to Keynes, will fail to bring about increases in employment, because it will mainly cause a reduction in aggregate demand. No doubt, the costs in all industries would be reduced as a result of a general wage-cut, but that in itself would not increase demand for the products, because the purchasing power in the hands of workers would have been reduced by cutting down their wages. A general wage-cut, by bringing about decline in aggregate demand, may actually decrease the volume of employment, and thus deepen the depression. As Hansen observes, "demand determines employment and employment determines marginal productivity or real wage, not the other way round". Employment cannot be increased by manipulating wages.

Besides, workers' organisations are too strong to permit general wage cuts. Thus, Pigou's theory has no relevance as a guide to policy. Hence, the

classical theory of employment must be rejected both on theoretical and practical grounds.

The fundamental fallacy in Say's Law is that partial equilibrium analysis which could apply to a particular industry, has been extended to the economy as a whole. Lowering of wage rate in a particular industry may increase employment there without decreasing demand. But if wages are reduced all round, it will reduce income and so effective demand and the volume of employment.

Besides, capitalism is not a self-regulating system as it is supposed to be. It cannot be depended to run itself.

The difference between the Pigovian and Keynesian views is fundamental. While Pigou is of the view that employment depends upon the level of money wages, and can be increased by lowering wages, Keynes contends that the volume of employment is determined by the level of aggregate effective demand, which may be adversely affected by the cuts in money-wages. In Keynes' view, even if wages rates were perfectly flexible, unemployment could still exist, if the aggregate demand was deficient. Hence, it is wrong to assert, as the classical economists did, that wage adjustment ensures full employment and interest rate adjustment tends to solve the saving-investment problem.

Summing Up. We can give the main points of criticism in a summary fashion as under:

(i) Supply may not create its own demand when a part of the income is saved. Aggregate demand is not always equal to aggregate supply.

(ii) Employment in the economy as a whole cannot be increased by means of a general wage cut though it may be possible in a particular industry. It is wrong to apply micro-economic principles to macro-economic activities or situations.

(iii) The classical economists looked at wages only from the employers' point of view, *i.e.*, the cost aspect and ignored the income aspect of wages. There is no direct relationship between wages and employment, nor is unemployment due to wage rigidities or artificial resistances.

(iv) **Interest-rate** adjustment cannot solve savings-investment problem. Savings and investment are not interest-elastic.

(v) The economic system is not so self-adjusting as it is supposed; hence government intervention in the economic sphere becomes necessary. Wages and prices are not so flexible as was supposed.

(vi) Assumption of free and perfect competition is not realistic.

(vii) It is **wrong** to suppose that money is a mere medium of exchange and has no role in affecting output and employment.

(viii) Say's Law cannot explain the occurrence of trade cycles.

(ix) The classical theory does not explain how the

level of employment is determined. It evades the problem by assuming full employment.

KEYNESIAN THEORY OF EMPLOYMENT

We have discussed above the classical theory of employment. We have also subjected it to detailed criticism. This criticism is mainly based on Keynes. In his epoch-making book, "**General Theory of Employment, Interest and Money**", Keynes has not only pointed out the shortcomings of the classical theory but he has also propounded his own theory of employment which is widely accepted by modern economists. In fact, as already pointed out, Keynesian Economics has been called 'New Economics' and 'Economic Revolution.' While propounding his theory of employment Keynes has invented and used new tools of economic analysis such as 'consumption function' or 'propensity of consumer', 'multiplier', 'marginal efficiency of capital', 'liquidity preference', 'effective demand', *etc.* We shall discuss these concepts in their proper place. Here we shall give in brief the Keynesian theory of employment.

At the outset, it may be pointed out that Keynesian theory is based on a short-run view. In the short run, it is assumed that capital equipment, population or manpower, technical knowledge, labour efficiency remain constant. That is why, according to Keynesian theory, volume of employment depends on the level of national income and output. Because if capital, working force and technology, labour efficiency remain fixed, the national income can be increased only by employing more labour (which was lying idle before). Hence, in Keynesian short-run, the increase in national income would mean increase in employment. The larger the volume of employment, the larger the national income and smaller the volume of employment the smaller the national income, and vice versa. That is why the Keynesian theory is called the theory of employment determination and also theory of income determination.

However, to simplify the analysis in this chapter, we shall explain Keynesian theory in terms of employment determination and the diagrams given in this chapter are intended to clarify the determination of the level of employment. In the next chapter, we shall explain the Keynesian theory regarding national income determination. But the factors that determine level of employment and national income are the same; only the diagrammatic representations are different.

PRINCIPLE OF EFFECTIVE DEMAND

The basic idea underlying or starting point of Keynesian theory of employment is the Principle of Effective Demand. According to Keynes, the level of

employment in the short run will depend on aggregate effective demand for goods in the country. Greater the aggregate effective demand, the greater will be the volume of employment, and vice versa. Total employment depends on total demand and unemployment is the result of a deficiency of total demand.

Effective demand represents the total money spent on consumption and investment. The total national expenditure is equal to total national income which is equal to national output. Effective demand being equal to total output as well as total expenditure, is also equal to national income.

Thus:

$$\begin{aligned} \text{Effective Demand} &= \text{National Income (Y)} \\ &= \text{National Output (O)}. \end{aligned}$$

Hence, effective demand holds the key at the same time to the volume of employment in the economy, to the level of national income and to the total national expenditure. Since effective demand determines the volume of employment in the economy at a particular time, the deficiency of effective demand results in unemployment. The deficiency of effective demand is due to the gap between income and consumption. As income increases, consumption also increases but in a smaller proportion than the increase in national income. Since consumption is less, the demand is less. This gap must be filled up by increasing investment and hence effective demand, in order to maintain employment at a high level. Thus, it is increase in effective demand which results in increase in employment or total output or national income.

In terms of expenditure, effective demand means the total expenditure of the community at a particular level of employment. This expenditure is just equal to the economy's aggregate supply price, i.e., the total cost of production of goods and services at that level of employment. In short, effective demand is the aggregate or total demand of the community both for consumption and investment.

Determination of Effective Demand

Keynes used two terms: Aggregate Demand Function or Price and Aggregate Supply Function or Price to explain the determination of effective demand. These are the two Keynesian 'blades of scissors' like Marshall's blades of scissors of demand and supply. Aggregate Demand Price and Aggregate Supply Price together determine effective demand which in turn determines the level of employment in the economy at a particular time.

An entrepreneur will employ that number of workers in his firm which gives him maximum profit. In the economy as a whole, the level of employment depends on the decisions of all individual employers, added together, as to how many workers should be employed so that it is most profitable. The volume of employment in an econ-

omy will be determined by aggregate supply price and aggregate demand price. Let us understand these two concepts.

Aggregate Supply Price

At any given level of employment of labour, aggregate supply price is the total amount of money which all the entrepreneurs in the economy, taken together, **must receive** from the sale of the output turned out by that number of workers which it is just worthwhile employing them. In other words, the aggregate supply price, when any given number of workers is employed, is the total cost of production of the output at a certain level of employment. It is the sum total of all payments made by entrepreneurs to all the factors of production producing that output. Obviously, the entrepreneurs will not employ that number unless they can recover the total cost of employing them. A certain minimum amount of proceeds is necessary to induce employers as a whole to offer any given aggregate amount of employment. "This minimum price or proceeds, which will just induce employment on a given scale, is called the aggregate supply price of that amount of employment." (Dillard).

If the output of these workers does not fetch sufficient price so as to cover the cost, the entrepreneurs will employ less number of workers. In this way, aggregate supply price will be different for different number of workers employed. That is, we can prepare aggregate supply price schedule according to the total number of workers employed in the economy and we can have a corresponding aggregate supply price curve or **Aggregate Supply Function**. Thus, aggregate supply price is a schedule of the minimum amounts of proceeds required to induce varying quantities of employment. The greater the amount of proceeds, the greater the amount of employment offered by the employers, taken together, to the workers in the economy. Since more workers will be employed only if the employers in the aggregate expect to be paid more money for the larger output produced by these workers, the aggregate supply price curve will slope upwards to the right.

Aggregate Demand Price

Aggregate demand must be distinguished from the demand for the products of an individual firm and individual industries. The demand for a firm or industry means a schedule of various amounts of a commodity which will be purchased at a series of prices. Since, the output of the entire economic system (as distinguished from the output of an individual firm or individual industry) cannot be expressed in terms of a unit of a commodity, it is expressed by Keynes as the quantity of labour employed. "The aggregate demand price for the

output of any given amount of employment is the total sum of money or proceeds which is expected from the sale of the output produced when that quantity of labour is employed. In other words, the aggregate demand price at any level of employment is the amount of money which all the entrepreneurs in the economy taken together **do expect** that they will receive if they sell the output produced by this given number of workers. The aggregate demand price represents the expected receipts when a given volume of employment is offered to workers. The aggregate demand curve or **Aggregate Demand Function** represents a schedule of the proceeds expected from the sale of the output produced by different amounts of employment. The greater the number of workers employed, the larger the output. That is, the aggregate demand price increases as the amount of employment increases, and vice versa. The aggregate demand curve rises upwards to the right showing that demand increases with an increase of employment. This is quite the opposite of demand curve for the product of a firm or an industry which slopes downwards to the right showing that the quantity sold will increase as the price falls.

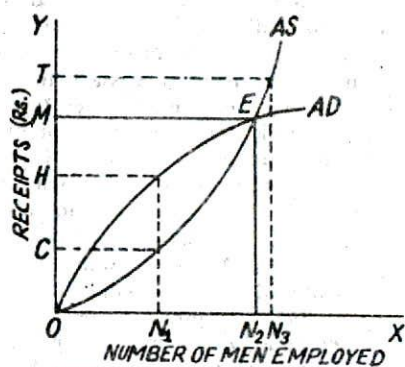
The student should carefully see the difference in the wording in the definitions of aggregate supply price and aggregate demand price since the wording seems to be similar. In the case of aggregate supply price, it is the amount of money which all employers taken together **must expect to receive** to induce them to offer certain amount of employment and in the case of aggregate demand price, it is the amount of money which they **do expect to receive** by the sale of the output produced by varying number of men. One is necessary payment, the other is an **expectation**.

Determination of Equilibrium Level of Employment

After having understood the concepts of Aggregate Supply Price and Aggregate Demand Price, we are now in a position to explain the determination of the level of employment with the help of the Aggregate Supply (AS) and Aggregate Demand (AD) curves.

In Fig. 39.3, AS (Aggregate Supply) and AD (Aggregate Demand) curves relate to an imaginary economy. Both AS and AD curves slope upwards to the right, but they do not follow the same course. For each level of employment there will be both an aggregate demand price and an aggregate supply price. There will be some levels of employment for which the receipts expected exceed the receipts necessary to induce given volumes of employment and at certain other levels of employment for which the receipts expected will not be sufficient to induce that amount of employment. In between these two limits there will be some level of employment for

which the expected receipts are just equal to what is necessary to make the employment worth while to the employers. This will be a point of intersection of the Aggregate Supply Curve and the Aggregate Demand Curve. This point of intersection determines the actual level of employment at any time. This is the point E in the Fig. 39.3.



Determination of Level of Employment
Fig. 39.3

In Fig. 39.3, the number of men employed are shown along the X-axis and receipts or proceeds of the goods and services produced by all the entrepreneurs are shown up the Y-axis, i.e., the total amount spent by the community on the output made by the entrepreneurs.

Take first AS curve. It shows the different total amounts which all the entrepreneurs taken together **must receive** to induce them to employ a certain number of men. For instance, if the entrepreneurs are convinced that they must receive OC amount of money, then they will employ ON_1 men. Now take AD curve. It shows the different total amounts which all the entrepreneurs, taken together, expect to receive at different levels of employment. For instance, if they employ ON_1 men, then they expect to receive OH amount from the sale of the total output. This means that at ON_1 level of employment, the total amount expected is greater than the amount necessary to be paid; hence there is no equilibrium at this point.

Look at the shape of AS curve. At first it rises slowly and then after a point it rises sharply. It means that in the beginning as more and more men are employed, the cost of output rises slowly. But as the amount received by the entrepreneurs increases they employ more and more men. So much so that all men who offer for work are given employment. In Fig. 39.3 ON_3 men, altogether seek employment and as soon as the entrepreneurs start getting OT amount, they will be prepared to employ all of them.

Now look at the shape of AD. In the beginning, it rises sharply, but flattens towards the end. This shows that in the beginning as more men are

employed, the entrepreneurs expect to get sharply increasing amounts of money from the sale of the output. But after employment has sufficiently increased, the expected receipts do not rise sharply.

The aggregate demand and aggregate supply for any community together determine the volume of employment which is actually offered by the entrepreneur. If the situation is such that the total amount of money which all the entrepreneurs, taken together, expect to receive from the sale of output exceeds what they consider necessary to receive, there will be competition among the entrepreneurs to offer more employment to the workers and employment will increase. For instance, on the left of N_2 (i.e., till the employment level reaches ON_2). Aggregate Demand (AD) is greater than Aggregate Supply (AS), i.e., the amount expected to be received is greater than the amount considered necessary. It is natural that there will be competition among entrepreneurs to employ more men.

On the other hand, if the number of men employed is greater than ON_2 , AD curve lies below the AS curve, that is, the amount expected by the entrepreneurs in the aggregate from the sale of the total output is less than what is considered necessary by them, and consequently they are losers. They will, therefore, reduce the number of persons employed. The retrenchment will continue till ON_2 level of employment is reached. At ON_2 level, AS and AD intersect. This is the point of equilibrium of the level.

Thus, employment in the economy as a whole will be in equilibrium only when the amount of proceeds which entrepreneurs **expect to receive** from providing any given number of jobs is just equal to the amount which they **must receive** if the employment of that number of workers is to be worth while for the entrepreneurs.

Further Discussion on Effective Demand

At the outset of the Keynesian theory of employment, we said that the Principle of Effective Demand is the starting point of the theory and greater the effective demand, the larger the volume of employment, and vice versa. After familiarising ourselves with the concepts of Aggregate Demand Price and Aggregate Supply Price, we are in a position to understand more clearly what 'effective demand' means and how it determines the level of employment.

We have seen in the above discussion that when in an economy, the Aggregate Demand and Aggregate Supply are equal to each other then employment in the economy is in equilibrium. When both aggregate demand and aggregate supply change but again are equal to each other, the equilibrium level of employment also changes. But in the short run, when aggregate demand remains the same, then it is

a short-run equilibrium of aggregate demand and aggregate supply. There is, in the economy, a schedule of aggregate demand (or aggregate demand price) which shows the aggregate demands at varying levels of employment. But of these varying levels of employment that aggregate demand is called effective demand which, at that level of employment, is also equal to the aggregate supply and which shows the short-run equilibrium of aggregate demand and aggregate supply. In other words, **effective demand is that aggregate demand price which becomes 'effective' because it is equal to aggregate supply price and thus represents a position of 'short-run' equilibrium.** There are several other points on the aggregate demand schedule but what distinguishes effective demand from all these points is that at this point aggregate demand price is equal to aggregate supply price. On all other points, aggregate demand price is either more or less than aggregate supply price.

It is clear that employment in the economy in the short run is determined by effective demand. The higher the level of effective demand, the greater is the volume of employment, and vice versa. Unemployment is due to the deficiency of effective demand and the basic remedy to remove this unemployment is to raise the level of effective demand. The classical economists believed that effective demand was always large enough to ensure full employment. But Keynes proved that it was not so and that is why the phenomenon of unemployment was common in all economies.

Effective demand means the total amount of money, actually spent on goods and services produced in the economy. Since money can be spent both on consumption goods and investment goods, effective demand will equal national expenditure both on consumption goods and investment goods. Thus, the main determinants of effective demand are total expenditure on consumption goods and on investment goods.

In the **Post-Keynesian sense**, government expenditure is also added to the total expenditure on consumption and investment on private account. Also, since all the money which all the entrepreneurs, taken together, receive from the sale of goods and services must be paid out in the form of wages, rent, interest and profit, effective demand will equal national income—the receipts of all the people in the country. Effective demand also represents the value of national output because the value of national output is nothing else but the total amount of money received by the entrepreneurs from the sale of goods and services. And money received by the entrepreneurs from the sale of goods is equal to the money spent by the people on these goods. Hence, effective demand equals national income or value of national output or national expenditure.

Thus,

Effective Demand = National Income.
 = Value of National Output.
 = National Expenditure.
 = Expenditure on consumption goods + expenditure on investment goods.

Importance of Effective Demand

The concept of effective demand occupies a very important place in Keynesian theory of employment. We have seen that the total volume of employment in the community depends on the state of effective demand. It is the deficiency of effective demand that results in unemployment in the country.

The concept of effective demand explains under-employment equilibrium as we shall show presently. We learn from the principle of effective demand that whatever is produced in the economy is not automatically consumed. That is, supply does not create its own demand as Say's Law said. It also shows that a general wage-cut is no remedy for unemployment.

Thus, the principle of effective demand demolishes the main notions of the classical theory. It brings out the crucial importance of investment in filling the gap between total income and total consumption.

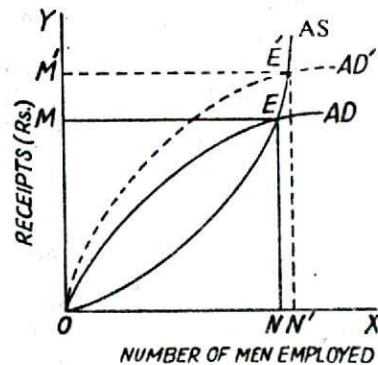
Equilibrium Not Necessarily at Full Employment

It is not necessary that the equilibrium level of employment is always at full employment. Equality between aggregate demand price and aggregate supply price does not necessarily indicate the full employment level. It can be in equilibrium at less than full employment or an under-employment equilibrium. The classical economists assumed full employment. Actually, however, there is always some unemployment even in the most developed countries. Keynes demolished the classical thesis of full employment both on theoretical grounds and by illustrations from real life. Full employment is important only as a limiting case. It is the level of employment beyond which further increases in effective demand do not increase output and employment.

At the point of intersection of AS and AD, the entrepreneurs are maximising their profits. The profits will be reduced if volume of employment is more or less than at this point. But there is no reason to assume that this point represents full employment. Aggregate demand and aggregate supply will be equal at full employment only if investment demand is sufficient to cover the gap between the aggregate supply price corresponding to full employment and the consumption expenditure out of income at full employment. The view of Keynes is

that the typical investment demand falls short of this gap. Hence, the aggregate demand schedule and aggregate supply schedule will intersect at a point less than full employment. Unless there is some external change there will be no tendency to depart from this under-employment equilibrium. Although with given aggregate demand and supply curves, there will normally be one position of equilibrium (where the two curves intersect), this need not be at the level of full employment. It is possible that the aggregate demand curve and the aggregate supply curve intersect each other at such a point at which the number of workers actually employed is less than those seeking employment. All the seekers of employment have not been able to secure employment and thus remain unemployed.

Look at the diagram (Fig. 39.4). Employment is



Employment Equilibrium at Full and Less than Full Employment Levels
 Fig. 39.4

measured along OX-axis and income or money receipts along OY-axis. AD curve shows expected incomes at various levels of employment. It shows that expected income increases as volume of employment increases. It has a diminishing slope and it becomes horizontal to X-axis after a certain level of employment showing that further employment will not increase the expected income. AS curve representing total cost of production rises steeply. This means that due to rising costs the entrepreneur will increase employment only if the maximum income that they must receive rises faster. It represents two situations, one of full employment and another of less than full employment. AS, the aggregate supply curve and AD, the aggregate demand curve, intersect at E. This means that ON men are employed and the entrepreneurs expect to get from this level of employment OM amount of money. In this situation of aggregate demand and aggregate supply in the economy, ON represents equilibrium level of employment. But as is shown in this diagram, in this situation of aggregate supply (AS), ON' number of men were seeking employment, whereas only ON number could secure employment. In this situation

there is equilibrium level of employment but it has not yet reached the full employment level, because NN' men are still unemployed. This unemployment will be removed only if on account of some favourable circumstances, aggregate demand increases so much (*i.e.*, it should increase from OM to OM' and there is shift from AD to AD') that the entrepreneurs now find it worthwhile to employ ON' men so that none who seeks work remains unemployed. In this diagram only at point E' there is full employment.

The situation in which an economy is in equilibrium at the level of full employment is called the optimum situation. In any other situation, it is in a state of less than full employment or there is unemployment. The classical economists denied that there could be an equilibrium at less than full employment. But we have seen in the above diagram that when AD and AS intersect at the point E , there is equilibrium level of employment but at less than full employment or there is under-employment equilibrium. There is unemployment indicated by NN' , *i.e.*, NN' men are seeking employment but are not actually employed.

Why Under-employment Equilibrium

According to Keynes, the root cause of the under-employment equilibrium is the deficiency of Aggregate Demand. This deficiency is due to the fact that there is a gap between income and consumption. As income increases, consumption increases but not proportionately, *i.e.*, marginal propensity to consume (to be discussed later) is less than unity. Coupled with this is the cause of insufficiency of investment. If investment increases sufficiently to cover this gap, there can be full employment, otherwise not. Investment is low owing to low marginal efficiency of capital (to be discussed later), *i.e.*, profit expectations are low.

Hence, the gap between income and consumption and insufficiency of investment to fill this gap are responsible for under-employment equilibrium. According to Keynes, under-employment (or unemployment) equilibrium is a normal feature of private free enterprise economy. To increase employment, government must step up its own investment and not leave things to the private investor.

Factors Determining Aggregate Supply and Aggregate Demand

We have seen above that we have to use Aggregate Supply Price curve and Aggregate Demand Price curve to determine the equilibrium level of employment. Let us know a little more about them. This will enable us to understand more clearly the principle of effective demand.

First take the Aggregate Supply Price curve. The Aggregate Supply Price Schedule or curve of the

economy depends in the last analysis on physical and technical conditions. The physical and technical conditions of production remain constant in the short period. Hence, given these technical conditions, production can be increased only by increasing employment. But production and employment can be increased only if the expenditure on production is increased. Production may be subject to the law of increasing, diminishing or constant returns, additional expenditure must be increased to increase production or employment. Hence, more men will be employed only if the entrepreneurs expect to increase their income from the sale of output produced by these men. That is why the aggregate supply curve slopes upwards to the right. The slope of the aggregate supply curve will depend on the physical and technical conditions of production in the economy. But when all the men seeking employment have been provided employment; then the AS curve will rise vertically, because even though the entrepreneurs may expect to receive larger income, employment cannot be increased since there are no unemployed men left. This has been shown in the diagram 39.4. As soon as ON' men are employed, the AS curve rises vertically.

We have said that the Aggregate Supply Schedule is determined by the physical and technical conditions prevailing in the economy, that is, the nature and quantity of labour, capital and raw materials available in the economy. As these change, or as productive resources are improved, AS curve will also change. But in the analysis of the problem of unemployment, it is unnecessary to pay more attention to aggregate supply schedule or curve. This is for the simple reason that at that time the main problem is how to employ idle resources to increase production and not that the techniques of production be improved. That is, AS curve need not be changed. The need of the moment is to increase employment and to increase aggregate demand for the purpose. When aggregate demand is increased the aggregate demand curve will rise upwards and it will intersect the AS curve more on the right, *i.e.*, the number of men employed will increase.

That is why Keynes assumes the AS curve to be constant and he pays greater attention to the factors determining aggregate demand.

When, however, the full employment level has been reached, then employment cannot be further increased by increasing aggregate demand. It will only result in inflation. In such a situation, it becomes necessary to pay attention to aggregate supply conditions and such production techniques have to be adopted which may increase productive efficiency. This will contain the inflationary pressure. In other words, we have to pay attention to aggregate supply only when full employment level has already been achieved and the economy is involved in inflation.

Since normally there is less than full employment (*i.e.*, there is some unemployment), aggregate demand has greater importance in the theory of employment than aggregate supply. Whereas aggregate supply depends on technical factors, aggregate demand schedule depends more on psychology of the people than on technology. The aggregate demand schedule shows the varying amounts of money that the entrepreneurs expect to receive at different levels of production or employment. In other words, it shows the varying amounts of total expenditure—which all the purchasers of goods and services are prepared to incur at different levels of employment. The purchasers of goods spend money on the purchase of two types of goods: (a) consumption goods and (b) investment goods. Thus, the shape and position of aggregate demand curve at each level of employment depends on the expenditure of the members of the community on consumption on the one hand and on investment on the other.

Summary of Keynesian Theory of Employment and Income

From the theory of employment explained above, we learn that in the short run, employment is determined by effective demand and the effective demand is equal to the total expenditure on consumption goods and investment goods. Hence, it is clear that employment depends on the total expenditure incurred on consumption and investment. In other words, if at any time owing to some reasons either consumption increases or investment increases, it would mean that effective demand has increased which in turn would increase employment in the country, and vice versa. This means that the key to employment and income determination lies in consumption and investment.

Keynes has introduced several new concepts in the discussion of consumption and investment. In the succeeding chapters, we shall discuss these concepts and principles such as propensity to consume and marginal efficiency of capital. We shall see how consumption expenditure is determined by

propensity to consume and investment by marginal efficiency of capital and the rate of interest. Marginal efficiency of capital depends on (a) future expectations of profit from investment and (b) present cost of replacing capital equipment. Obviously, higher the expected rate of profit, the greater the inducement to invest. Also, lower the replacement cost, the greater is the inducement to invest. If the expectations of profit remain the same and the rate of interest falls, then investment would increase. Keynes introduced the concept of liquidity preference to explain the determination of the rate of interest.

Thus, the two determinants of investment are the marginal efficiency of capital and the rate of interest.

After we have learnt fully the tools and principles used by Keynes in the theory of employment, we shall be able to understand clearly Keynesian theory of employment and income. We give below this theory in summary fashion:

- (1) Total income depends on the volume of total employment.
- (2) Total employment depends on total effective demand and in equilibrium aggregate demand is equal to aggregate supply.
- (3) Aggregate supply depends on physical and technical conditions of production, and, in the short run, these do not often change; hence, it is the changes in the aggregate demand that bring about changes in income and employment.
- (4) Effective demand is made up of (a) consumption demand and (b) investment demand.
- (5) Consumption demand depends on consumption function or propensity to consume, and, in the short run, consumption function is relatively stable.
- (6) Investment demand depends on (a) the marginal efficiency of capital, and (b) the rate of interest.
- (7) The marginal efficiency of capital depends on (a) the expectations of profit yields, and (b) replacement cost of capital assets.
- (8) The rate of interest depends on (a) the quantity of money, and (b) the state of liquidity preference.

In the previous chapter, we have given Keynes' theory from the point of view of the determination of the volume of employment. As we explained there, Keynes gives a short-run view in which the amount of capital, labour efficiency, techniques of production, the system of business organisation, etc., are assumed to be constant. In such a situation, the volume of employment and the level of national income increase and decrease together. That is, the determinants of income and employment are the same. They are both determined by aggregate demand and aggregate supply.

In this chapter, we shall explain the Keynesian theory in terms of the determination of national income. Since the determining factors of income and employment are the same, some repetition of what we have given in the previous chapter is unavoidable. The diagrammatic representation, as we shall see, will be different.

How the Level of National Income is Determined

As we have mentioned above, the level of national income, in the short period, is determined by aggregate demand and aggregate supply. The supply of goods and services in a country depends on the productive capacity of the community. But during the short period this productive capacity does not change. But it is not necessary that the actual production or aggregate supply should be always equal to productive capacity. The total output or aggregate supply will correspond to aggregate demand. If aggregate demand increases, output will also increase and the level of national output (i.e., national income) will rise. On the other hand, if aggregate demand decreases, the national output or national income will also decrease. It follows that the equilibrium level of national income is determined by aggregate demand since aggregate capacity remains more or less the same during the short period. We assume here that the national output (or income) will be as much as the effective demand.

There are two components of effective demand: (a) consumption demand, i.e., demand for consumption goods and (b) investment demand, i.e., demand, for investment goods or producers' or capital goods. Hence, by effective demand, we mean how much total expenditure the government and the people are willing to incur on consumption goods and investment goods.

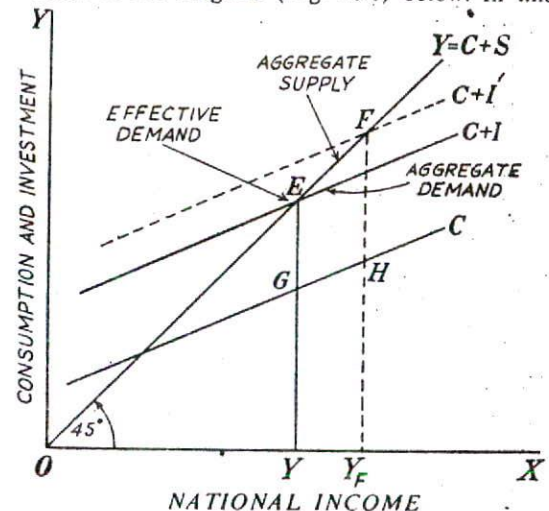
Thus,

Aggregate Demand = Consumption Demand + Investment Demand, i.e., $AD = C + I$.

(Here C is consumption demand, and I is investment demand).

So far as consumption demand is concerned, it depends on propensity to consume (consumption function) and income. Given propensity to consume, as income increases, the consumption demand will also increase.

Look at the diagram (Fig. 40.1) below. In this



Determination of National Income
Fig. 40.1

figure, national income and national output (G.N.P.) are shown along the X-axis and along the Y-axis are shown the consumption demand (C) and investment (I). A straight line making an angle of 45° is drawn along the X-axis. This 45° line represents aggregate supply curve and it is also called Income Line. This 45° line shows two things: (a) total output or aggregate supply (consumption goods + investment goods) and (b) national income in terms of money. Actually, national output (G.N.P.) and national income are two different names of the same thing.

In this diagram, C curve represents propensity to consume. It will be seen that this curve C rises upwards to the right which means that as income increases consumption (or consumption demand) also increases. Since 45° line also represents national income, the distance between the curve C showing propensity to consume and 45° line represents saving because a part of income is consumed and the rest is saved. Thus, national income = consumption + saving or $Y = C + S$.

(Here Y is national income, C is consumption and S is saving).

It will be seen in the diagram that the distance between the national income line Y (i.e., 45° angle line) and C (propensity to consume) goes on increasing which means that as income increases, the amount of saving increases.

One noteworthy thing about propensity to consume is that it remains stable or constant during the short period. This is because propensity to consume (or shape of C curve) depends on the tastes and needs of the people and these do not change in the short period. When we say that propensity to consume remains stable, it does not mean that there is no change in the consumption demand. As we have said above, consumption demand increases with increase in income. What we mean by the constancy or stability of propensity to consume is that the total consumption schedule or the shape of propensity to consume curve C does not change in the short period. Since consumption is more or less stable, variation in national income depends on variation in investment.

The second component of aggregate demand, viz., investment is a very important determinant of national income. Investment depends on two things: (a) marginal efficiency of capital and (b) the rate of interest. Of these two, interest rate is comparatively stable. Hence, change in investment is largely determined by change in marginal efficiency of capital. The marginal efficiency of capital means expectation of profit from investment. That is, expected rate of profit is called marginal efficiency of capital. The marginal efficiency of capital, too, depends on two things: one is the replacement cost of the capital goods and the other, profit expectations of the investors. Of these two, again, profit expectations are

a more important determinant of investment. This means that when a country wants to increase its national income or employment then it should create conditions in which profit expectations of investors and businessmen go high.

At any time in a country, keeping in view, the rate of interest and marginal efficiency of capital, the entrepreneurs wish to invest a certain amount of capital or there is a certain demand for investment goods. We assume that investment demand does not increase with increase in income. Actually, as people's income goes up, then their demand for goods will increase and as a result the entrepreneurs' profit expectations go up. When profit expectations go up, the marginal efficiency of capital goes up which will result in increase in investment. But it is obvious that the amount of investment does not directly depend on income. That is why in our diagram, we do not show investment demand rising with increase in income.

As we have said above, at any given time, in a country, the entrepreneurs are desirous of investing a certain amount of capital. If we join this investment demand with the curve C of propensity to consume, we get aggregate demand curve C+I in which C represents consumption and I investment. The distance between propensity to consume curve C and aggregate demand curve C+I is equal to investment (I). If propensity to consume curve is given, then higher the investment, the higher will be the aggregate demand curve (C+I).

What level of national income will be determined in any country at any time? The level of income will be determined at a point at which the aggregate demand curve (C+I) of the country and the aggregate supply curve, i.e., 45° angle line intersect each other, i.e., where the aggregate demand and aggregate supply are in equilibrium. This has been shown in the Fig. 40.1 on the previous page.

We see in this figure that aggregate demand curve C+I intersects the aggregate supply curve (i.e., 45° angle line) at E. Hence, the equilibrium level of income is OY. It will be seen in this figure that when income is either more or less than OY, then the aggregate demand and aggregate supply are not in equilibrium. You will see that if income is more than OY, then total output or aggregate supply is greater than aggregate demand C+I so that the entire output cannot be sold out. The result will be that output will be decreased and income will decrease. On the contrary, if income is less than OY, then the total output or aggregate supply will fall short of aggregate demand. As a result, output will be increased and national income will increase. It is only when income is OY that aggregate demand and aggregate output (supply) are equal so that there will be no tendency for output or income to decrease or increase. Hence, OY income will be determined

We have seen how the equilibrium level of income is determined by the interaction of aggregate demand and aggregate supply. The equilibrium level of employment (discussed in the previous chapter) is also determined by the interaction of aggregate demand and aggregate supply. This is as it should be because national income, output and employment are interchangeable terms.

Equilibrium Not Necessarily at Full Employment

We may repeat here that it is not necessary that equilibrium level of national income may be achieved at the point of full employment. This view of Keynes is altogether opposed to the classical view. According to the classical economists, the economy is always in a state of full employment. The lapses from full employment are strictly temporary, since there is persistent tendency in the economic system to restore the state of full employment. Keynes completely demolished this view and established both on theoretical grounds and with reference to reality the possibility of under-employment equilibrium.

Look at the figure 40.1 again. Suppose OY_F is the full employment level of national income. But in this diagram, OY is the equilibrium level of income and OY is less than OY_F which represents full employment level. It is obvious that the equilibrium level of income is at less than full employment. That is, it is an under-employment equilibrium. The equilibrium will be established at full employment income only when investment demand is sufficiently large so as to fill the saving gap between income and consumption corresponding to full employment.

It will be seen in Figure 40.1 that at OY_F level of income, which corresponds to full employment, HF is the saving. Hence, when investment demand increases so as to cover this saving only then the equilibrium level of income will be at full employment. But there is no guarantee that investment demand will be equal to saving at the full employment level of income. This is for the simple reason that saving is done by different people and investment by different people. Besides, the factors which determine saving are different from those that determine investment. People save money for the education and marriages of their children; they save to meet future contingencies or to build houses or for their old age. But investment at any time depends on the marginal efficiency of capital and the rate of interest. And it is not necessary that investment should be equal to saving at the full employment level of national income. When investment is less than saving at the full employment level of income, as it usually is, the equilibrium will be established at less than full employment.

EQUILIBRIUM LEVEL OF INCOME: EQUALITY OF SAVING AND INVESTMENT

We have seen how equilibrium level of national income is determined by the interaction of aggregate demand and aggregate supply. The equilibrium level of national income is established at the point where aggregate demand equals aggregate supply. But there is also another method for the explanation of the determination of national income. This alternative method explains the determination of national income directly by saving and investment.

Look at the Figure 40.1 again. In this figure, at the equilibrium level of national income OY , saving and investment are equal; they are both equal to GE . Given the aggregate demand curve $C+I$, the amount of saving at income more than OY is more than investment and for income less than OY , investment is more than saving. It is obvious that saving and investment are equal only at the equilibrium level of national income and when saving and investment are not equal, then the national income will not be in equilibrium. Let us see why it is so and how national income is determined by saving and investment.

When at a certain level of national income, intended investment by the entrepreneurs is more than intended savings by the people, this would mean that aggregate demand is greater than the total output or aggregate supply. This would induce the firms to increase production raising the level of income and employment. The result will be that national output will be increased on account of which national income will increase. Hence, when at any level of income, investment is greater than saving, there will be a tendency for the national income to increase. On the contrary, when at any level of national income, the investment demand is less than saving, it means that aggregate demand is less than aggregate supply. That is, the entrepreneurs will not be able to sell their entire output at given prices. The result will be that output will be reduced which will result in the reduction of national income.

Hence, when, at any level of national income, investment demand of the entrepreneurs is less than the intended savings of the people, the national income will decrease. It will come down to the level at which investment spending is just equal to the planned savings by the community. But, when at any level of national income, intended investment demand on the part of the entrepreneurs is equal to intended saving of the people, it means that aggregate demand is equal to the total output or aggregate supply as a result of which the national income will be in equilibrium.

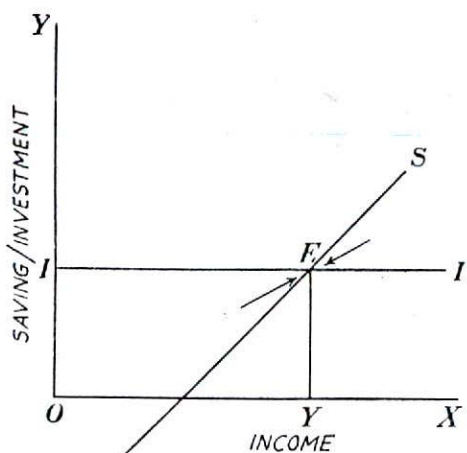
Hence, equilibrium level of national income will

be established at the level at which the amount of intended investment by the entrepreneurs is equal to the amount of intended saving by the people.

We can explain the determination of national income by saving and investment in another way. Saving is withdrawal of some money from the income stream. On the other hand, investment is injection of money into the income stream. Now if intended investment is more than intended saving, it means that more money has been put into the income stream than has been taken out of it. This would mean that the income stream, *i.e.*, national income would increase. On the contrary, if investment is less than intended saving, it means that less amount of money has been put into the income stream and more has been taken out of it. The result would be that national income would decrease. But when investment is just equal to saving, it would mean that as much money has been put into the income stream as has been taken out of it. The result will be that the national income will neither increase nor decrease, *i.e.*, it will be in equilibrium. It is thus clear that the equilibrium of national income will be established at the level at which the intended investment is equal to intended saving.

The determination of national income by investment and saving is illustrated by Fig. 40.2.

In this figure (40.2), income is shown on the X-axis and saving and investment on the Y-axis. SS is the saving curve which shows intended saving at different levels of income. II curve shows investment



Determination of National Income by Saving and Investment
Fig. 40.2

demand, *i.e.*, intended investment. The II investment curve has been drawn parallel to the X-axis. This is done on the assumption that in any year the

entrepreneurs intend to invest a certain amount of money. That is, we assume that investment does not change with income.

The saving curve SS and the investment curve II intersect each other at E. That is, the intended investment and intended saving are equal at the OY level of income. Hence, OY is the equilibrium level of income.

It will be seen in the Fig. 40.2, that at the level of income less than OY, the amount of intended investment is more than intended saving. As a result income will increase. On the contrary, at the level of income greater than OY, the amount of intended investment is less than the intended saving with the result that income will decrease. The decrease in income will continue till it becomes equal to OY. At the level of OY income, intended investment and intended saving are equal, so that there is neither the tendency for income to increase nor to decrease. Hence, OY national income is determined. It is thus that national income is determined by investment and saving.

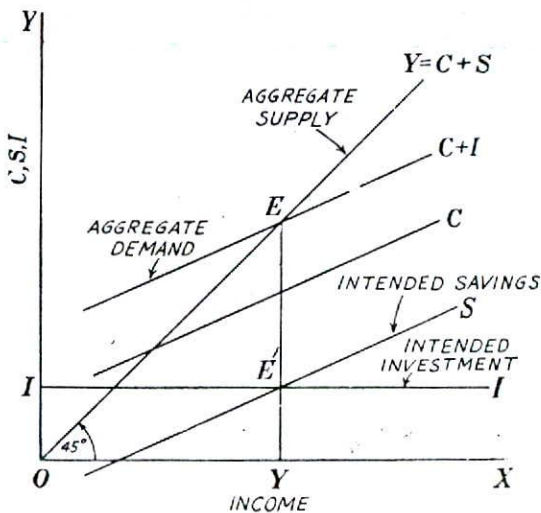
Conclusion

The determination of national income has been explained above by two methods: The equilibrium level of national income will be determined where two conditions are fulfilled:

- (i) **Aggregate Demand = Aggregate Supply, and**
- (ii) **Intended Investment = Intended Saving.**

In reality, the equality between Aggregate Demand and Aggregate Supply and between Intended Investment and Intended Saving mean the same thing.

This is illustrated in diagram 40.3 below.

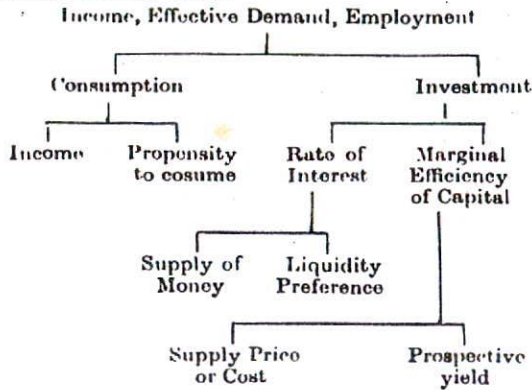


Similarity between Equality of Aggregate Demand and Aggregate Supply and of Intended Investment and Intended Saving
Fig. 40.3

Keynes' Theory in Outline

According to the Keynesian theory, employment is a function of income; greater the national income, the greater the volume of employment. Both income and employment are determined by the level of effective demand.

The Keynesian theory may also be presented in a tabular form as under:—



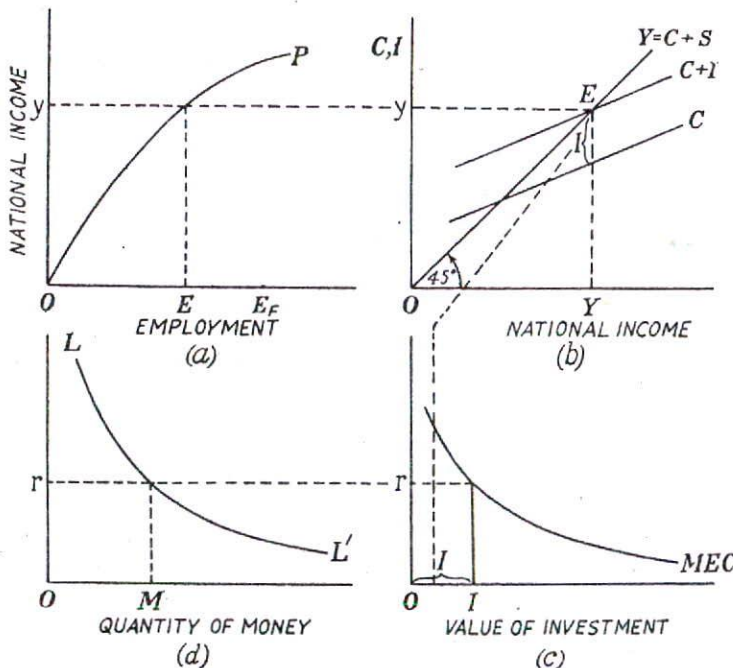
We have already explained how the various elements of the above chart combine and interact together to determine the equilibrium level of income and employment.

Keynes's Model of Income-Employment Determination: Graphical Representation. The curve OP

in diagram (a) of Fig. 40.4 shows the relationship between income and employment; the greater the national income the greater the employment. The diagram (b) shows that the aggregate demand (consumption plus investment spending) determines the equilibrium level of income. Diagram (c) shows the way in which marginal efficiency of capital and the rate of interest determine the volume of investment and diagram (d) shows the determination of the rate of interest by the liquidity preference curve LL' and the quantity of money OM.

Now given the liquidity preference curve LL' and the quantity of money OM, rate of interest Or is determined [see diagram (d)]. Given the schedule of marginal efficiency of capital MEC and the rate of interest Or, the volume of investment OI, will be determined [see diagram (c)]. Now given the investment of OI and the propensity to consume as represented by curve C in diagram (b), national income of OY will be determined, that is, economy will be in equilibrium at OY level of income. Now diagram (a) indicates that OY level of national income will generate OE volume of employment. Now, if OEF is the level of full employment, it is clear that the equilibrium is established at less than full employment level.

If the quantity of money in the economy is increased, it can be shown by the above diagram that the equilibrium income and employment will rise.



Income-Employment Determination Model

Fig. 40.4

INFLATIONARY AND DEFLATIONARY GAPS

We know that equilibrium is not always at the full employment level. Rather it can be at a level when there is less than full employment. Thus no particular virtue need be attached to what we call the equilibrium level of national income or employment. On the other hand, equilibrium level may involve much unemployment and waste of national resources, if the investment is not sufficiently high to ensure full employment. Hence, the only level of equilibrium that may be considered desirable is that which provides full employment or near full employment. This level can be reached if investment opportunities happen to match full employment saving. In actual practice, therefore, we may have inflationary gap or deflationary gap.

Inflationary Gap

Inflationary gap arises when consumption and investment spending together are greater than the full employment GNP level. This means that people are demanding more goods and services than can be produced. In other words, the implication of inflationary gap is that national income, output and employment cannot rise further. The only consequence of increased demand for goods and services on the part of people will be to raise the price level. Or, we may say that there will be an inflationary gap if scheduled investment tends to be greater than full employment saving. In a situation like this, more goods will be demanded than the economic system can produce. The result will be that the prices will begin to rise and an inflationary situation will emerge. Thus, if full employment saving falls short of scheduled investment at full employment (which means that peoples' propensity to spend is higher than the propensity to save), there will be an inflationary gap

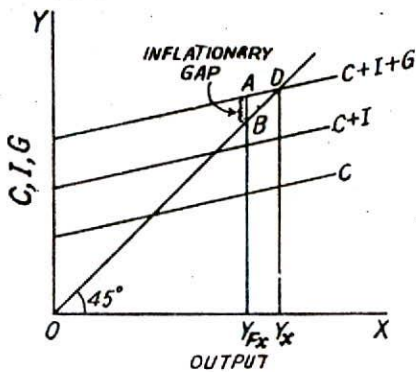


Fig. 40.5

The inflationary gap can be explained with the help of the above diagram. C, I, G stand for the

consumption, investment and Government expenditure respectively. $(C+I+G)$ line shows the total expenditure on demand in the economy. At this level, Y_x is the total real output, as shown by the intersection, point D, with the 45° line. Y_{fx} represents a full employment limit on real output Y_{fx} . Real income of the economy, obviously cannot reach Y_x . At Y_{fx} , total demand $(C+I+G)$ exceeds total output, leaving a gap AB, which is the inflationary gap in the Keynesian sense.

Deflationary Gap

Similarly, we can show with the help of a diagram "deflationary gap". This would come into existence, if total aggregate demand is insufficient to

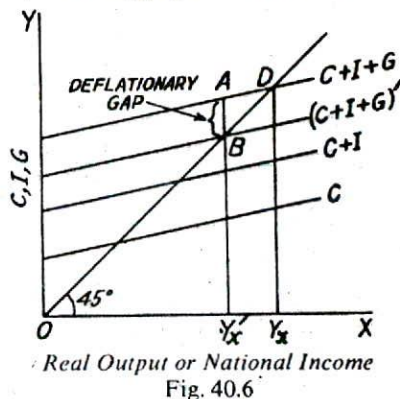


Fig. 40.6

create full employment. Y_x is the total output at full employment. Let us assume that the total demand is $(C+I+G)$ which cuts the 45° line at B, with real output Y_{fx} . AB then is the deflationary gap.

SAVING-INVESTMENT CONTROVERSY

To a beginner in the Keynesian Theory; this controversy is one of the perplexing ones. Are savings equal to investment or are they unequal? A layman may say that a man may not invest all his savings; how can then saving be equal to investment? Let us clear this matter.

The first thing is to understand the meanings of these two terms, saving and investment. Investment in Keynesian Theory means the net addition to the stock of capital goods like machinery, equipment, factories, etc. It also includes inventories. Investment in this sense does not mean the total stock of capital goods in existence but the net addition thereto in a certain period. Thus, the term investment is different from the term capital. Capital means the total stock of capital goods in existence, whereas investment means what has been added to this stock say in a year.

Saving means that amount which a man saves out of income after he has incurred his consumption

expenditure. Saving thus means the income which is not consumed. It has been defined as the excess of income over consumption expenditure. This even a layman can understand. Suppose a man is earning Rs. 5,000 per annum. Out of it he spends Rs. 4,000 on food, clothes, housing accommodation and on other household needs. His saving amounts to Rs. 1,000. Suppose he is a farmer and spends Rs. 500 on irrigation, agricultural implements or in repairing the farm house. We should not then think that his saving is Rs. 500. The saving will still be Rs. 1,000, for Rs. 500 spent on irrigation, agricultural implements, etc., is called investment and cannot be included in consumption expenditure. For finding out saving, we should deduct from income only domestic or consumption expenditure and not investment expenditure. This point needs to be remembered.

Why Unequal? Having defined the terms saving and investment, we now turn to the controversy whether they are equal or unequal. Economists before Keynes generally were of the view that the two are not equal, unless they reach an equilibrium position. They argued that savings are made by one set of people and investment by another set of people. That being so, their equality can only be by coincidence. Another argument put forward was that the entrepreneurs also borrowed money from banks for investment purposes. This being in addition to their own savings and other people's savings, investment must exceed savings.

Why Equal? Keynes, in his book "General Theory", has established that savings and investment are equal. The national income, he says, is derived from the production and sale of (a) consumers' goods and (b) investment goods. We therefore get the equation

$$\begin{aligned} \text{Income} &= \text{Consumption} + \text{Investment} \\ \text{or } Y &= C + I \end{aligned} \quad (1)$$

Y stands for income, C for consumption expenditure and I for investment expenditure.

This is the production and sales side of income. But income has also another side, i.e., spending side. From this point of view, income is exhausted in consumption and saving and we get the equation

$$\begin{aligned} \text{Income} &= \text{Consumption} + \text{Saving} \\ \text{or } Y &= C + S \end{aligned} \quad \dots (2)$$

Y stands for income, C for consumption expenditure and S for saving out of the income.

Bringing equations (1) and (2) together we get

$$\begin{aligned} \therefore C + S &= C + I \\ \therefore S \text{ (Saving)} &= I \text{ (Investment)}. \end{aligned}$$

We find that saving and investment are equal by the very definition we have given. In Keynesian

Economics, the terms 'income', 'consumption', 'saving' and 'investment' have been so defined as excess of income over consumption expenditure; such is also the definition of investment. The two must therefore be equal. Hence, both saving and investment are equal, because they have been equated to each other by the definition of both as excess of income over consumption expenditure.

We may remind the reader that we are now studying Macro-Economics and not Micro-Economics. The words 'savings' and 'investment' are used in the aggregate sense and not individual sense. For instance, saving does not mean here savings of an individual but that of the community. Similarly, income means aggregate income and consumption aggregate consumption. In the case of an individual, saving may not be equal to investment, but for the community or the economic system as a whole, saving must be equal to investment as defined by Keynes.

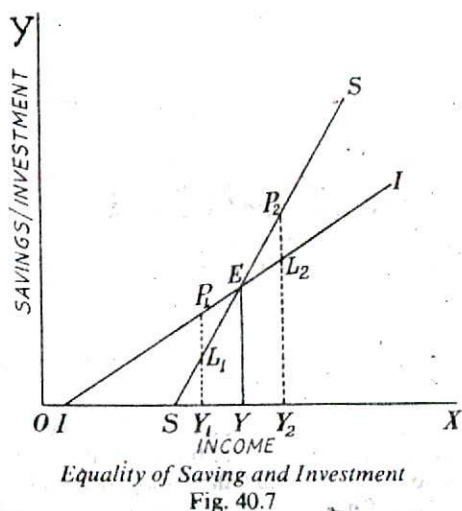
It may happen sometimes that the nation cuts down consumption expenditure in response to appeals by leaders in a national emergency. Savings will then increase. It may be asked how can saving be equal to investment in this case. The answer is that, if consumption has been decreased, the inventories or finished goods are lying idle; they have not been sold to the consumers. And those inventories are included in investment. Hence, investment is as large as savings.

In What Sense Unequal? Is there any sense in which saving and investment may not be equal to each other? The answer is that as the words have been defined in Keynesian Economics, they must be always equal. In this controversy, it is essential for the student to understand two words: ex-ante and ex-post. Ex-ante means anticipated or intended and ex-post means actual or realised. The former means what we expect in the future and the latter means what has already been realised by a past process.

Now ex-post (i.e., realised or actual) saving and investment are always equal, but ex-ante (i.e., intended or anticipated or expected) savings and investment may differ. Ex-ante (or intended or anticipated) saving and investment are equal only at the equilibrium level of income.

Suppose, in the sense of ex-ante (or anticipated or intended), investment is greater than saving. Greater investment means greater income which in turn will tend to increase saving and the two will become equal at the equilibrium position. If, on the other hand, saving exceeds investment, income will decrease, leading in turn to a decrease in saving so that saving and investment will tend to be equal.

This can be shown by the diagram (Fig. 40.7). SS is the saving curve and II is the investment curve. The point of intersection of the two curves E represents the equilibrium level of income. Here



both saving and investment are equal. At the OY_1 level, $Y_1 P_1$ is the ex-ante or intended investment and $Y_1 L_1$ is the ex-ante or intended saving, which is less than intended investment. Hence, income must increase and reach the OY level where the intended saving and investment are equal. At the OY_2 level, intended saving $Y_2 P_2$ is more than intended investment $Y_2 L_2$. Hence income must decrease and tend to reach OY level where again the two equal.

It is clear that ex-ante (or intended or anticipated) saving and investment are equal at the equilibrium level of income, and not at any other level of income (i.e., more or less). It is just like demand and supply, which are equal at the equilibrium price, and not at any other price. Just as price goes up and down and equates demand and supply at the equilibrium level, similarly, income increases or decreases to reach the equilibrium level and thus equates saving and investment.

SOME MACRO ECONOMIC TERMS

Stocks and Flows

The variables in macro-economics may be called 'Stocks' or 'Flows'. There is no time dimension in stocks but flows have a time dimension. That is, a stock must be specified at a point of time, but the magnitude of stock has no time dimension. On the other hand, a flow can be expressed only in terms of time units. It is clear that money is a stock but expenditure is a flow. Similarly, wealth is a stock but income is a flow.

We can trace stock-flow relationship in certain economic problems. Stocks change only through flows. It is the fast or slow changes in stock due to current rates of flow that certain macro-economic problems arise.

Macro Statics

In macro static, attention is given to the state of the economy at a given moment of time. That is, we confine our attention to equilibrium positions. We explain certain aggregative relationships in an equilibrium but we do not show how this equilibrium position was reached. In other words, we do not show the process of adjustment implied in reaching the equilibrium position.

For example, when we say $Y = C + I$, we only show that to maintain a certain level of income, the expenditure of the consumers and businessmen must be $C + I$, but we do not show how the total income reaches that particular level of Y . That is, there is no room for time analysis in macro statics. It is looking at a still picture and not a movie.

Macro Comparative Statics

As against static analysis, the comparative statics involves a comparison between two equilibrium positions. That is, we compare the state of economy at one particular time with that of another. In other words, we compare two levels of income or employment achieved by an economy at two different times.

Macro-Dynamics

In macro dynamics, we investigate the process by which the economy moves from one equilibrium position to another. For instance, the national income of India at constant prices rose from nearly Rs. 9,000 crores in 1950-51 to Rs. 22,000 crores nearly in 1975-76. The macro dynamics will investigate the changes in the economy which raised the income to that level in 25 years. It describes the time paths of macro variables and aggregate relationships. Thus it shows us a movie picture of the progressive changes that occurred in the economy during this period. For instance, it shows present savings depending on past rates of interest, level of consumption and cumulative total of past savings and dissavings.

Obviously, it is a more realistic description of complex phenomenon of cyclical fluctuations and secular growth.

The dynamic analysis is of particular importance in macro economics since it is widely used in the study of national income, principle of acceleration, trade-cycle theories and growth economics.