prefer a steady performance with 'satisfactory' profits to spectacular profit maximisation projects. If they realise maximum high profits in one period, they might find themselves in trouble in other periods when profits are less than maximum. Sixthly, large, growing sales strengthen the power to adopt competitive tactics, while a low or declining share of the market weakens the competitive position of the firm and its bargaining power *vis-à-vis* its rivals.

The desire for a steady performance with satisfactory profits, coupled with the separation of ownership and management, tend to make the managers reluctant to adopt promising projects which are risky. The top managers become to a certain extent risk-avoiders, and this attitude may act as a curb on economic growth. However, the desire for steady performance has a stabilising effect on economic activity. In general, large firms have research units which develop new ideas of products or techniques of production. The application of these projects is spread over time so as to avoid wide swings in the economic performance of the firm. Baumol seems to imply that the risk-avoidance and the desire for steady growth of the large corporations secure 'orderly markets', in the sense that they have stabilising effects on the economy.

II. INTERDEPENDENCE AND OLIGOPOLISTIC BEHAVIOUR

Although Baumol recognises the interdependence of firms as the main feature of oligopolistic markets, he argues that in 'day-to-day decision-making management often acts explicitly or implicitly on the premise that its decisions will produce no changes in the behaviour of those with whom they are competing . . .'. It is only when the firm makes 'more radical decisions, such as the launching of a major advertising campaign or the introduction of a radically new line of products, [that] management usually does consider the probable competitive response. But often, even in fairly crucial decisions, and almost always in routine policy-making, only the most cursory attention is paid to competitive reactions'.

This attitude towards competitors is attributed by Baumol to several reasons:

The complexity of the internal organisation of large firms renders decision-making a lengthy process: proposals originate from some sections, but final decisions are taken by top management after these proposals have passed through various levels of management and often from different departments. It is a characteristic inherent in the delegation of authority within the firm that each decision-maker will attempt to shift the responsibility on to others. Thus any reaction of competitors is bound to take place after 'a considerable time lag'.

Large organisations work to a 'blue-print' which includes a variety of rules of thumb, which simplify complicated problems such as pricing, size of advertising expenditure, level of inventories. Prices are set by applying a standard mark-up to costs, advertising expenses are determined by setting aside a fixed percentage of total revenues, inventories are determined as a percentage of sales, and so on. Such rules of thumb clearly do not automatically take into account the actions of competitors, and the adaptation of the 'blue-print' of a firm to a new environment takes time.

The desire of top management for a 'quiet life' has led large enterprises to some tacit collusion: firms depend on each other to behave in an 'orderly' way. They expect no 'breach of etiquette' in the established order in the industry as a whole.

However, the above reasons do not imply that businessmen are completely indifferent to actions of competitors. In particular, being sales maximisers and growth seekers, they are very alert to any change in their share of the market. Top management will ignore competitors only to the extent that their actions do not encroach on the firm's market and do not interfere with the desired rate of growth of the sales of the firm.

III. BAUMOL'S STATIC MODELS

The basic assumptions of the static models

1. The time-horizon of a firm is a single period.

2. During this period the firm attempts to maximise its total sales revenue (not physical volume of output) subject to a profit constraint. The firm in these models does not consider what will happen in subsequent periods as a result of the decisions taken in the current period.

3. The minimum profit constraint is exogenously determined by the demands and expectations of the shareholders, the banks and other financial institutions. The firm must realise a minimum level of profits to keep shareholders happy and avoid a fall of the prices of shares on the stock exchange. If profits are below this exogenously determined minimum acceptable level the managers run the risk of being dismissed, since shareholders may sell their shares and take-over raiders may be attracted by a fall of the prices of shares.

4. 'Conventional' cost and revenue functions are assumed. That is, Baumol accepts that cost curves are U-shaped and the demand curve of the firm is downward-sloping.

We will examine four models:

(1) A single-product model, without advertising.
(2) A single-product model, with advertising.
(3) A multiproduct model, without advertising.
(4) A multiproduct model, with selling activities.

Model 1: a single-product model, without advertising

The total-cost and total-revenue curves under the above assumptions are shown in figure 15.1. Total sales revenue is at its maximum level at the highest point of the TR curve, where the price elasticity of demand is unity and the slope of this TR curve (the marginal revenue) is equal to zero.

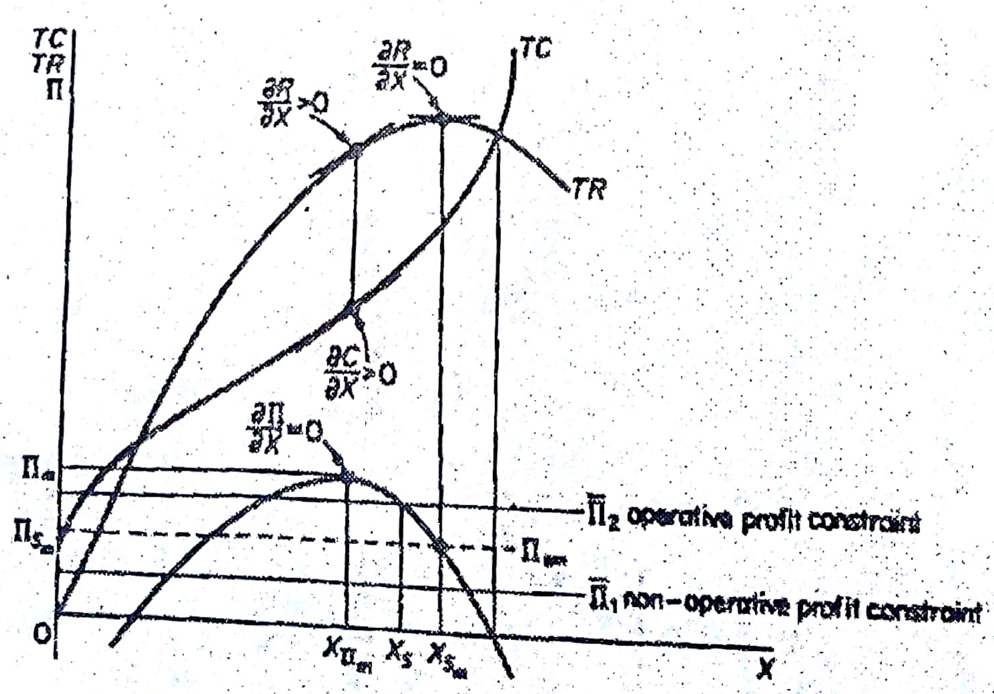


Figure 15.1

Proof

It has been established in Chapter 2 that

$$MR = P\left(1 - \frac{1}{e}\right)$$

At the point of maximum revenue the slope of the total revenue curve is

$$\frac{\partial R}{\partial X} = MR = 0$$

Therefore

$$0 = P\left(1 - \frac{1}{e}\right)$$

Given $P > 0$, we have

$$1 - \frac{1}{e} = 0$$

or

$$e = 1$$

Whether this maximum sales revenue will be realised or not depends on the level of the minimum acceptable level of profit which may act as a constraint to the activity of the firm. If the firm were a profit maximiser, it would produce the level of output $X_{\Pi m}$. However, in Baumol's model the firm is a sales maximiser, but it must also earn a minimum level of profit acceptable to shareholders and to those who finance its operations. If the minimum acceptable level of profit is Π_1, the firm will produce the level of output X_{Sm} which maximises its sales revenue. With this level of output (X_{Sm}) the firm earns profits Π_{Sm}, which are greater than the minimum required to keep the stockholders (and other interested parties) satisfied. Under these circumstances we say that the minimum profit constraint is not operative.

If the minimum acceptable profit is Π_2, the firm will not be able to attain the maximum sales revenue because the profit constraint is operative, and the firm will produce X_S units of output, which are less than at the level X_{Sm}.

In summary: '... two types of equilibria appear to be possible: one in which the profit constraint provides no effective barrier to sales maximisation (X_{Sm} units of output with a minimum acceptable profit of Π_1), and one in which it does (X_S units of output with a minimum acceptable profit of Π_2)'. (W. J. Baumol, *Business Behaviour, Value and Growth* (revised edn, Harcourt, Brace & World, Inc., 1967).) The firm is assumed to be able to pursue an independent price policy, that is, to set its price so as to achieve its goal of sales maximisation (given the profit constraint) without being concerned about the reactions of competitors.

Provided that the profit constraint is operative, the following predictions of Baumol's single-period model (without advertising) emerge:

The sales maximiser will produce a higher level of output as compared to a profit maximiser.

Proof

A profit maximiser produces the output $X_{\Pi m}$ defined by the equilibrium condition $MR = MC$ or

$$\frac{\partial R}{\partial X} = \frac{\partial C}{\partial X}$$

Given that the marginal cost is always positive ($\partial C/\partial X > 0$) it is obvious that at the level $X_{\Pi m}$ the marginal revenue is also positive ($\partial R/\partial X > 0$). That is, TR is still increasing at $X_{\Pi m}$, since its slope is still positive. In other words, the maximum of the TR curve (where its slope is $\partial R/\partial X = 0$) occurs to the right of the level of output at which profit is maximised. Hence $X_{Sm} > X_{\Pi m}$.

The sales maximiser sells at a price lower than the profit maximiser. The price at any level of output is the slope of the line through the origin to the relevant point of the total-revenue curve (corresponding to the particular level of output). In figure 15.2 the price of the profit maximiser is

$$P_{\Pi m} = \begin{bmatrix} \text{slope} \\ \text{of } 0A \end{bmatrix} = \frac{R_\Pi}{X_\Pi}$$

while the price of the sales maximiser is

$$P_{Sm} = \begin{bmatrix} \text{slope} \\ \text{of } 0B \end{bmatrix} = \frac{R_S}{X_S}$$

It is obvious that (slope $0A$) > (slope $0B$), that is, the price of the profit maximiser is higher than the price of the sales maximiser.

The sales maximiser will earn lower profits than the profit maximiser. In figure 15.2 the profit of the sales maximiser is $0\Pi_S$, which is lower than the profit $0\Pi_\Pi$ of the profit maximiser.

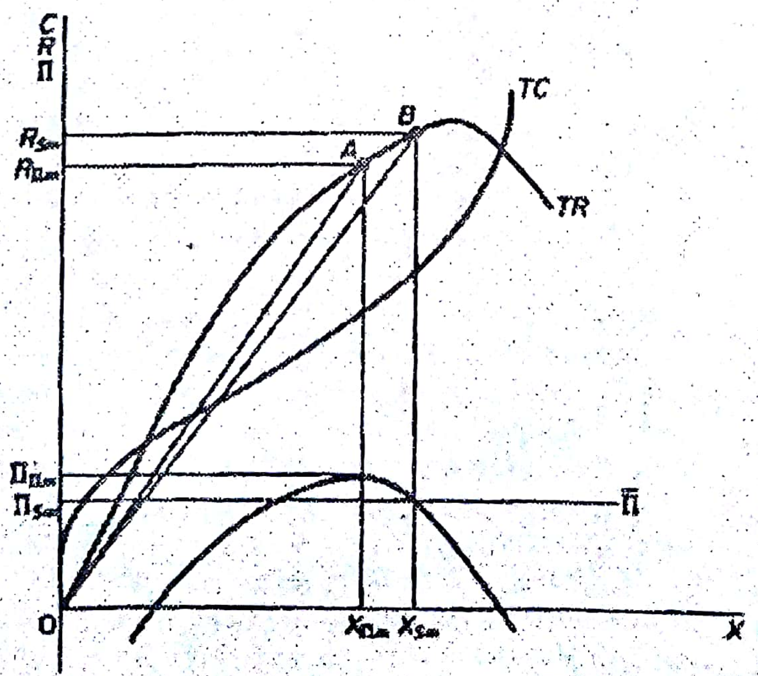


Figure 15.2

The sales maximiser will never choose a level of output at which price elasticity (e) is less than unity, because from the expression

$$MR = P\left(1 - \frac{1}{e}\right)$$

we see that if $|e| < 1$, the $MR < 0$, denoting that TR is declining. The maximum sales revenue will be where $|e| = 1$ (and hence $MR = 0$) and will be earned only if the profit constraint is not operative. If the profit constraint is operative the price elasticity will be greater than unity.

An increase in the fixed costs will affect the equilibrium position of a sales maximiser: he will reduce his level of output and increase his price, since the increase in fixed costs shifts the total-profit curve downwards. Subject to the profit constraint, the sales maximiser will pass the increase in costs to the customers by charging a higher price. This is shown in figure 15.3. The increase in fixed costs shifts the total costs upwards and the total-profits curve downwards (Π'). Subject to the profit constraint $\bar{\Pi}$, the firm will reduce its output (to X'_S) and will increase its price.

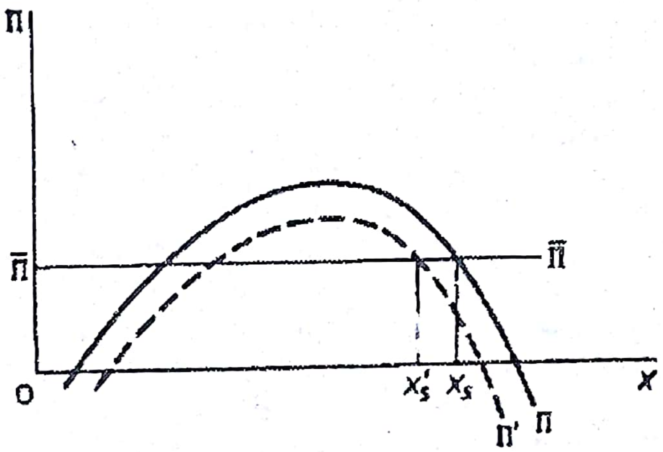


Figure 15.3

This prediction is contrary to the traditional hypothesis of profit maximisation. A profit maximiser will not change his equilibrium position in the short run, since fixed costs do not enter into the determination of the equilibrium of the firm. So long as the fixed costs do not vary with the level of output (and provided that the increase in TFC does not lead the firm to close down altogether) the change in the TFC will not lead the profit maximiser to change his price and output in the short run (see Chapter 6).

Baumol claims that firms in the real world do in fact change their output and price whenever their overhead costs increase. Thus he says that the sales-maximisation hypothesis has a better predictive performance than the traditional profit-maximisation hypothesis.

The imposition of a lump-sum tax will have similar effects. If the firm is a profit maximiser the imposition of the lump-sum tax will not affect the price and output in the short run: the profit maximiser will bear the whole burden of the lump-tax. If the firm is a sales maximiser, however, the lump-tax will shift the total-profit curve downwards and, given the profit constraint, the firm will be led to cut its level of output and increase its price, thus passing on to the consumer the lump-sum tax. Baumol argues that firms do in fact shift the tax on to the buyers, contrary to the accepted doctrine about the 'unshiftability' of the tax.

The imposition of a specific tax (per unit of output) will shift the profit curve downwards and to the left (figure 15.4). Given $\bar{\Pi}$, the sales maximiser will reduce his output

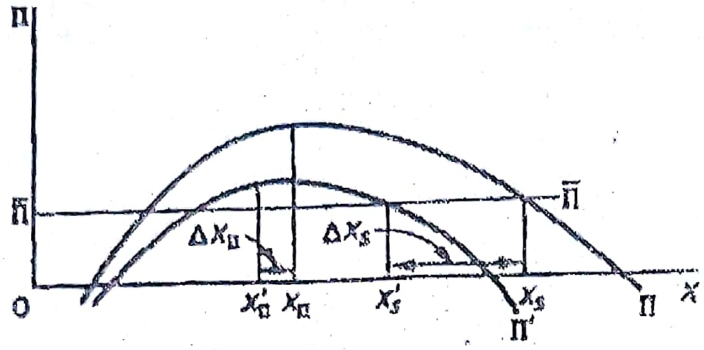


Figure 15.4

from X_S to X'_S and will raise his price, passing the tax to the buyers (at least partly). The profit maximiser will also reduce his output (from X_Π to X'_Π) and raise his price. However, the decrease in output will be larger than the decrease of the output of a profit maximiser.

A similar analysis holds for an increase in the variable cost. Both the sales maximiser and the profit maximiser will raise their price and reduce their output. The reduction in output, however, and the increase in price will be more accentuated for the sales maximiser, *ceteris paribus*.

A shift in demand will result in an increase in output and sales revenue but the effects on price are not certain in Baumol's model. Price will depend on the shift of the demand and the cost conditions of the firm. (See below.)

Model 2: a single-product model, with advertising

The assumptions of the model. As in the previous model the goal of the firm is sales revenue maximisation subject to a minimum profit constraint which is exogenously determined. The new element in this model is the introduction of advertising as a major instrument (policy variable) of the firm. Baumol argues that in the real world non-price competition is the typical form of competition in oligopolistic markets. The model presented by Baumol treats explicitly advertising, but other forms of non-price competition (product change, service, quality, etc.) may be analysed on similar lines.

The crucial assumption of the advertising model is that sales revenue increases with advertising expenditure (that is, $\partial R/\partial a > 0$, where a = advertising expenditure). This implies that advertising will always shift the demand curve of the firm to the right and the firm will sell a larger quantity and earn a larger revenue. The price is assumed to remain constant. This, however, is a simplifying assumption which may be relaxed in a more general analysis (Baumol, *Business Behaviour, Value and Growth*, p. 58).[1] Another simplifying assumption is that production costs are independent of advertising. Baumol recognises that this is an unrealistic assumption, since with advertising the physical volume of output increases and the firm might move to a cost structure where production cost is different (increasing or decreasing). But he claims that this assumption is simplifying and can be relaxed without substantially altering the analysis. (In fact

[1] An explicit geometric treatment of the interdependence between price and advertising has been presented by R. L. Sandmeyer, 'Baumol's Sales-Maximisation Model: Comment', *American Economic Review* (1964), and has been extended by R. Haveman and G. DeBartolo, 'The Revenue Maximisation Oligopoly Model: Comment', *American Economic Review* (1968) and by M. Kafoglis and R. Bushnell, 'The Revenue-Maximisation Oligopoly Model: Comment', *American Economic Review* (1970). The Haveman–DeBartolo model, as modified by Kafoglis and Bushnell, is presented on pp. 333–5 below.

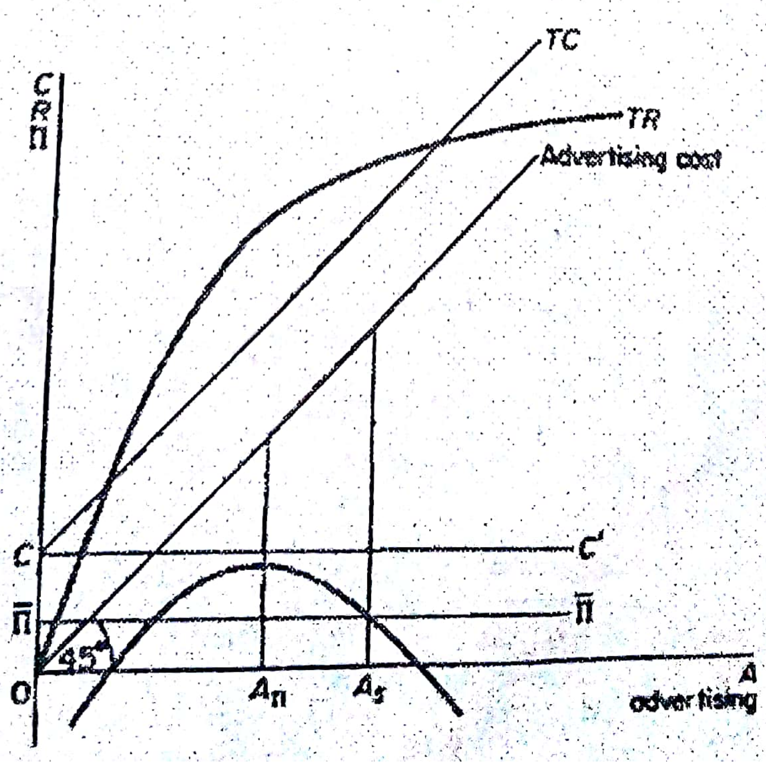
Baumol relaxes this assumption, as well as the assumption of constant price in the mathematical presentation of his model; see below.) From the above assumptions the following inferences can be drawn.

A firm in an oligopolistic market will prefer to increase its sales by advertising rather than by a cut in price. While an increase in physical volume induced by a price cut may or may not increase the sales revenue, depending on whether demand is elastic or inelastic, an increase in volume brought about by an increase in advertising will always increase sales revenue, since by assumption the marginal revenue of advertising is positive ($\partial R/\partial a > 0$).

With advertising introduced into the model, it is no longer possible to have an equilibrium where the profit constraint is not operative. While with price competition alone it is possible to reach an equilibrium (that is, maximise sales) where $\bar{\Pi}$ is not operative, with non-price competition such an unconstrained equilibrium is impossible. Unlike a price reduction, increased advertising always increases sales revenue. Consequently it will always pay the sales maximiser to increase his advertising expenditure until he is stopped by the profit constraint. Consequently the minimum profit constraint is always operative when advertising (or any other form of non-price competition) is introduced in the model.

The sales maximiser will normally have higher advertising expenditures than a profit maximiser. In any case advertising cannot be less in a sales-maximising model.

Baumol's single-product model with advertising is shown in figure 15.5. Advertising outlay is measured on the horizontal axis and the advertising function is shown as a 45° line. Costs, total revenue and profits are measured on the vertical axis. Production costs are shown as being independent of the level of advertising (curve CC'). If these costs are added to the advertising cost line we obtain the total-cost curve (TC) as a function of advertising outlay. Subtracting the total cost from the total revenue at each level of output we obtain the total-profit curve Π. The interrelationship between output and advertising and in particular the (assumed) positive marginal revenue of advertising permits us to see clearly that an unconstrained sales maximisation is (ordinarily) not possible. If price is such as to enable the firm to sell an output yielding profits above their minimum acceptable level, it will pay the firm to increase advertising and reach a higher level of sales revenue. The advertising outlay of the sales maximiser ($0A_s$) is higher than



that of the profit maximiser ($0A_\Pi$), and the profit constraint ($\bar{\Pi}$) is operative at equilibrium.

It should be stressed that the validity of this model rests on the crucial assumption that advertising always increases sales revenue. Baumol *assumes* that $\partial R/\partial a > 0$, but does not establish the implied positive relation between total revenue and advertising. In particular Baumol does not examine explicitly the interrelationship between advertising, price, cost of production and level of output.

If total production costs are independent of advertising, (that is, production costs remain constant after advertising takes place) as Baumol assumes, this implies that total output X will remain constant after advertising has taken place; consequently an increase in sales revenue R, given X, can be attained only if P is raised. This case is in fact implied in figure 15.5, reproduced from Baumol's book (p. 59).

However, this is inconsistent with what Baumol states elsewhere (p. 60), that 'unlike a price reduction a *ceteris paribus* rise in advertising expenditure involves no change in the market value of the item sold'. This statement implies clearly that advertising will *not* change the price. Hence Baumol implies that the increase in revenue will be attained from an increase in the volume X. But then production costs will increase, since MC is always positive.

In short, Baumol's graphical representation of his model is inconsistent with his statements. In particular the price implications of a change in advertising are not obvious in Baumol's analysis. Sandmeyer,[1] Haveman and DeBartolo,[2] and Kafoglis and Bushnell[3] have highlighted this deficiency of Baumol's model. They suggested that with advertising expenditures the TR curve will shift and in the new equilibrium revenue will be higher and advertising expenditure will be higher (consistent with Baumol). However, output may be lower and price higher in the new equilibrium, depending on the shift and the elasticity of the demand curve following advertising, as well as on the cost conditions of the firm. This situation has not been explicitly envisaged by Baumol, whose model has been interpreted as implying that all 'excess' profit will be devoted to advertising and that, therefore, the increase in revenue will accrue from an increase in output resulting from the shift of the demand curve following advertising. However, Baumol's mathematical model allows for the possibility of a change in price as well as of advertising and output (see below).

Haveman and DeBartolo have presented a model which they call 'generalised Baumol model'. In their model price, cost, output and advertising expenditure are all free to vary. We will first present graphically their model as modified by M. Kafoglis and R. Bushnell[4] and by C. J. Hawkins.[5] We will next present their model mathematically and will point out that actually their model is identical to Baumol's mathematical presentation of his advertising model.

The cost curves. It is assumed that: (*a*) Production costs vary proportionally with output. Thus the total production cost function is a straight (positively-sloping) line through the

[1] R. L. Sandmeyer, 'Baumol's Sales-Maximisation Model: Comment', *American Economic Review* (1964).

[2] R. Haveman and G. DeBartolo, 'The Revenue-Maximisation Oligopoly Model: Comment', *American Economic Review* (1968) pp. 1355–8.

[3] M. Kafoglis and R. Bushnell, 'The Revenue-Maximisation Oligopoly Model: Comment', *American Economic Review* (1970). Also R. Haveman and G. DeBartolo, 'Reply', *American Economic Review* (1970).

[4] M. Kafoglis and R. Bushnell, 'The Revenue-Maximisation Oligopoly Model: Comment', *American Economic Review* (1970)

[5] C. J. Hawkins, 'The Revenue Maximisation Oligopoly Model: Comment', *American Economic Review* (1970).

origin. (*b*) Advertising expenditure may change but is independent of the level of output. Thus a given level of advertising is presented by a straight line parallel to the X-axis. Higher advertising levels are shown by parallel lines which are further away from the X-axis. (*c*) The minimum profit constraint is exogenously determined and is denoted by a line parallel to the X-axis.

The total cost function is the summation of the production cost (C), the advertising expenditure (A_i) and the minimum profit constraint ($\bar{\Pi}$). Given the production cost function and the minimum profit constraint, a change in advertising (A_i) will generate a family of total-cost curves which will be upward-sloping (with their slope equal to the slope of production cost function). Such a family of total-cost curves is shown in figure 15.6.

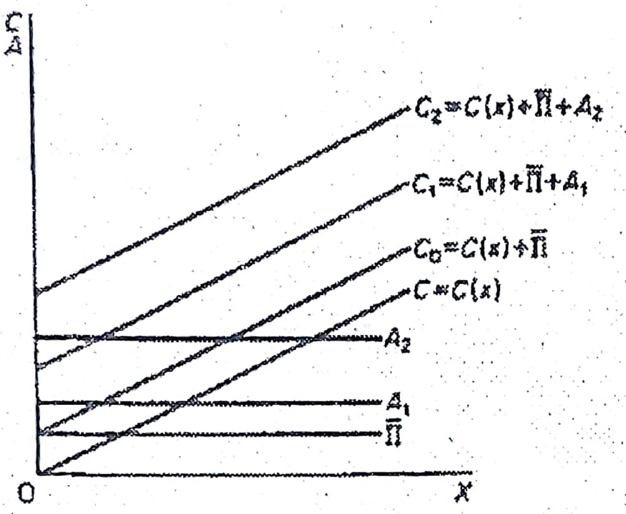


Figure 15.6

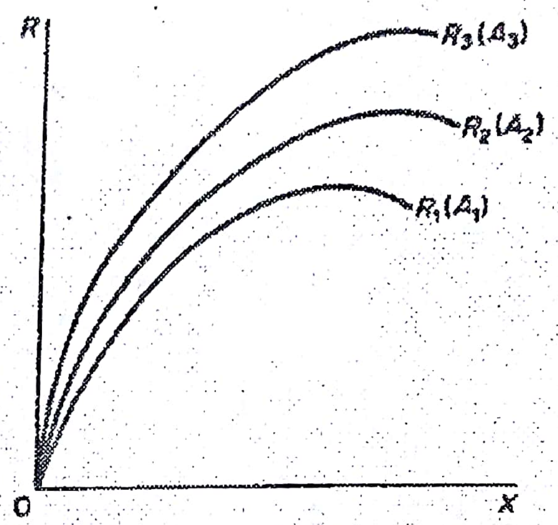


Figure 15.7

The revenue curves. The total-revenue curve has the usual shape, initially increasing but at a decreasing rate, reaching a maximum (where $\partial R/\partial X = 0$), and then decreasing (as $\partial R/\partial X < 0$).

The total-revenue curve shifts upwards as advertising is increased. Thus, by changing advertising we may generate a family of total-revenue curves, each representing the relationship of total revenue to output at different levels of advertising expenditure. Such a family of total-revenue curves is shown in figure 15.7. Curve R_1 is drawn on the assumption that advertising expenditure is A_1. Curve R_2 implies an advertising expenditure of A_2, and so on.

Equilibrium of the firm. If we superimpose figures 15.6 and 15.7 and join the points of intersection of total-cost and total-revenue curves corresponding to the same amount of advertising expenditure, we obtain a curve which is called by Haveman and DeBartolo the '$TC = TR$' curve. It is the dotted curve in figure 15.8. The firm is in equilibrium when it reaches the highest point of this curve. The equilibrium of the firm is at point a^*, with total costs C^*, total revenue R^*, output X^*, advertising A^*, and price equal to $0R^*/0X^*$.

It should be clear that two conditions must be satisfied for equilibrium: Firstly, the firm must operate on some point of the '$TC = TR$' curve. Secondly, $MC > MR$ at equilibrium. Thus at point a_3 the first condition is fulfilled ($C_3 = R_3$) but the second condition is violated, since at a_3 the two curves are tangent, implying $MC = MR$. Thus if the sales maximiser was producing at X_3, he would substitute production expenditure for advertising expenditure (a reallocation of resources from advertising to increased production) until output increased to X^*. In the process of adjustment price would fall,

Scanned by CamScanner

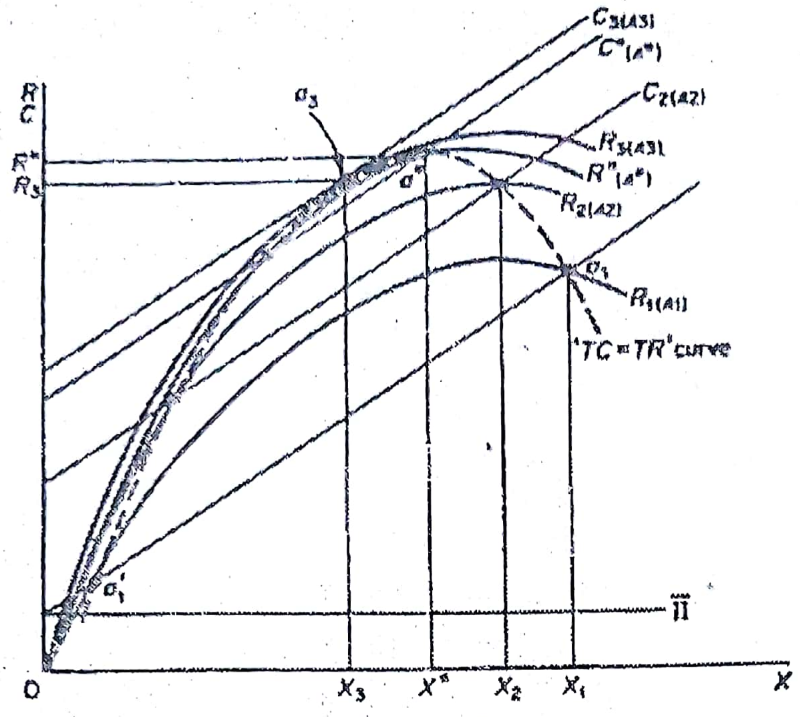


Figure 15.8

but the loss in revenue from this cause would be more than offset by the additional revenue from the increased output sold. In figure 15.8 we see that $R^* > R_3$.

A mathematical presentation of Baumol's Model 2

We define

$$R = f_1(X, a) = \text{total revenue function}$$
$$C = f_2(X) = \text{total production cost function}$$
$$\bar{\Pi} = \text{minimum acceptable profit}$$
$$A(a) = \text{total cost of advertising function}$$

The firm aims at the maximisation of

$$R = f_1(X, a)$$

subject to the minimum profit constraint

$$\Pi = R - C - A \geq \bar{\Pi}$$

It is assumed that

$$\frac{\partial R}{\partial a} > 0, \qquad \frac{\partial C}{\partial X} > 0, \qquad X > 0$$

Baumol's assumption

$$\frac{\partial R}{\partial a} > 0$$

that is, demand shifts always in response to advertising, ensures that the constraint is operative.[1]

[1] Alternatively the assumption $\partial R/\partial X > 0$, that is, demand is always elastic ($e > 1$) for any relevant output, would alone ensure that the constraint is operative.

16. Marris's Model of the Managerial Enterprise

1. GOALS OF THE FIRM

The goal of the firm in Marris's model[1] is the maximisation of the *balanced* rate of growth of the firm, that is, the maximisation of the rate of growth of demand for the products of the firm, and of the growth of its capital supply:

$$\text{Maximise } g = g_D = g_C$$

where g = balanced growth rate
g_D = growth of demand for the products of the firm
g_C = growth of the supply of capital

In pursuing this maximum balanced growth rate the firm has two constraints. Firstly, a constraint set by the available managerial team and its skills. Secondly, a financial constraint, set by the desire of managers to achieve maximum job security. These constraints are analysed in a subsequent section.

The rationalisation of this goal is that by jointly maximising the rate of growth of demand and capital the managers achieve maximisation of their own utility as well as of the utility of the owners–shareholders.

It is usually argued by managerial theorists that the division of ownership and management allows the managers to set goals which do not necessarily coincide with those of owners. The utility function of managers includes variables such as salaries, status, power and job security, while the utility function of owners includes variables such as profits, size of output, size of capital, share of the market and public image. Thus managers want to maximise their own utility

$$U_M = f(\text{salaries, power, status, job security})$$

while the owners seek the maximisation of their utility

$$U_0 = f^*(\text{profits, capital, output, market share, public esteem}).$$

Marris argues that the difference between the goals of managers and the goals of the owners is not so wide as other managerial theories claim, because most of the variables appearing in both functions are strongly correlated with a single variable: the *size* of the firm (see below). There are various measures (indicators) of size: capital, output,

[1] R. Marris, 'A Model of the Managerial Enterprise,' *Quarterly Journal of Economics* (1963). Also R. Marris, *Theory of 'Managerial' Capitalism* (Macmillan, 1964).

revenue, market share, and there is no consensus about which of these measures is the best. However, Marris limits his model to situations of *steady rate of growth* over time during which most of the relevant economic magnitudes change simultaneously, so that 'maximising the long-run growth rate of any indicator can reasonably be assumed equivalent to maximising the long-run rate of most others.' (Marris, 'A Model of the Managerial Enterprise'.)

Furthermore, Marris argues that the managers do not maximise the absolute *size* of the firm (however measured), but the *rate of growth* (=change of the size) of the firm. The size and the rate of growth are not necessarily equivalent from the point of view of managerial utility. If they were equivalent we would observe a high mobility of managers between firms: the managers would be indifferent in choosing between being employed and promoted within the same growing firm (enjoying higher salaries, power and prestige), and moving from a smaller firm to a larger firm where they would eventually have the same earnings and status. In the real world the mobility of managers is low. Various studies provide evidence that managers prefer to be promoted within the same growing organisation rather than move to a larger one, where the environment might be hostile to the 'newcomer' and where he would have to give considerable time and effort to 'learn' the mechanism of the new organisation. Hence managers aim at the maximisation of the *rate of growth* rather than the absolute size of the firm.

Marris argues that since growth happens to be compatible with the interests of the shareholders in general, the goal of maximisation of the growth rate (however measured) seems *a priori* plausible. There is no need to distinguish between the rate of growth of demand (which maximises the U of managers) and the rate of growth of capital supply (which maximises the U of owners) since in equilibrium these growth rates are equal.

From Marris's discussion it follows that the utility function of owners can be written as follows

$$U_{\text{owners}} = f^*(g_C)$$

where g_C = rate of growth of capital.

It is not clear why owners should prefer growth to profits, unless g_C and profits are positively related. At the end of his article Marris argues in fact that g_C and Π are not always positively related. Under certain circumstances g_C and Π become competing goals (see p. 364 below). Furthermore from Marris's discussion of the nature of the variables of the managerial utility function it seems that he implicitly assumes that salaries, status and power of managers are strongly correlated with the growth of demand for the products of the firm: managers will enjoy higher salaries and will have more prestige the faster the rate of growth of demand. Therefore the managerial utility function may be written as follows

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

where g_D = rate of growth of demand for the products of the firm
s = a measure of job security.

Marris, following Penrose, argues that there is a constraint to g_D set by the decision-making capacity of the managerial team. Furthermore Marris suggests that 's' can be measured by a weighted average of three crucial ratios, the liquidity ratio, the leverage-debt ratio and the profit-retention ratio, which reflect the financial policy of the firm. As a first approximation Marris treats 's' as an exogenously determined constraint by assuming that there is a saturation level for job security: above the saturation level the marginal utility from an increase in 's' (job security) is zero, while below the saturation level the marginal utility from an increase in 's' is infinite. With this assumption the managerial utility function becomes

$$U_M = f(g_D)\bar{s}$$

where $\bar{s}$ is the security constraint.

Thus in the initial model there are two constraints – the managerial team constraint and the job security constraint – reflected in a financial constraint. We will examine these constraints in some detail.

II. CONSTRAINTS

The managerial constraint

Marris adopts Penrose's thesis of the existence of a definite limit on the rate of efficient managerial expansion. At any one time period the capacity of the top management is given: there is a ceiling to the growth of the firm set by the capacity of its managerial team. The managerial capacity can be increased by hiring new managers, but there is a definite limit to the rate at which management can expand and remain competent (efficient). Penrose's theory is that decision-making and the planning of the operations of the firm are the result of teamwork requiring the co-operation of all managers. Co-ordination and co-operation require experience. A new manager requires time before he is fully ready to join the teamwork necessary for the efficient functioning of the organisation. Thus, although the 'managerial ceiling' is receding gradually, the process cannot be speeded up.

Similarly, the 'research and development' (R & D) department sets a limit to the rate of growth of the firm. This department is the source of new ideas and new products, which affect the growth of demand for the products of the firm. The work in the R & D department is 'teamwork' and as such it cannot be expanded quickly, simply by hiring more personnel for this section: new scientists and designers require time before they can efficiently contribute to the teamwork of the R & D department.

The managerial constraint and the R & D capacity of the firm set limits both to the rate of growth of demand (g_D) and the rate of growth of capital supply (g_C). (See section III below.)

The job security constraint

We said that the managers want job security; they attach (not surprisingly) a definite disutility to the risk of being dismissed. The desire of managers for security is reflected in their preference for service contracts, generous pension schemes, and their dislike for policies which endanger their position by increasing the risk of their dismissal by the owners (that is, the shareholders or the directors they appoint). Marris suggests that job security is attained by adopting a prudent financial policy. The risk of dismissal of managers arises if their policies lead the firm towards financial failure (bankruptcy) or render the firm attractive to take-over raiders. In the first case the shareholders may decide to replace the old management in the hope that by appointing new management the firm will be run more successfully. In the second case, if the take-over raid is successful, the new owners may well decide to replace the old management.

The risk of dismissal is largely avoided by: (*a*) Non-involvement with risky investments. The managers choose projects which guarantee a steady performance, rather than risky ventures which may be highly profitable, if successful, but will endanger the managers' position if they fail. Thus the managers become risk-avoiders. (*b*) Choosing a 'prudent financial policy'. The latter consists of determining optimal levels for three crucial financial ratios, the leverage (or debt ratio), the liquidity ratio, and the retention ratio.

The *leverage* or *debt ratio* is defined as the ratio of debt to the gross value of total assets of the firm

$$\begin{bmatrix}\text{Leverage}\\ \text{or}\\ \text{Debt ratio}\end{bmatrix} = \frac{\text{Value of debts}}{\text{Total assets}} = \frac{D}{A}$$

The managers do not want excessive borrowing because the firm may become insolvent and be proclaimed bankrupt, due to demands for interest payments and repayment of loans, notwithstanding the good prospects that the firm may have.

The *liquidity ratio* is defined as the ratio of liquid assets to the total gross assets of the firm

$$\begin{bmatrix}\text{Liquidity}\\ \text{ratio}\end{bmatrix} = \frac{\text{Liquid assets}}{\text{Total assets}} = \frac{L}{A}$$

Liquidity policy is very important. Too low a liquidity ratio increases the risk of insolvency and bankruptcy. On the other hand, too high a liquidity ratio makes the firm attractive to take-over raids, because the raiders think that they can utilise the excessive liquid assets to promote the operations of their enterprises. Thus the managers have to choose an *optimal* liquidity ratio: neither too high nor dangerously low. In his model, however, Marris assumes without much justification, that the firm operates in the region where there is a positive relation between liquidity and security: an increase in liquidity increases security.

The *retention ratio* is defined as the ratio of retained profits (net of interest on debt) to total profits

$$[\text{Retention ratio}] = \frac{\text{Retained profits}}{\text{Total profits}} = \frac{\Pi_R}{\Pi}$$

Retained profits are, according to Marris, the most important source of finance for the growth of capital (see section III below). However, the firm is not free to retain as much profits as it might wish, because distributed profits must be adequate to satisfy the shareholders and avoid a fall in the price of shares which would render the firm attractive to take-over raiders. If distributed profits are low the existing shareholders may decide to replace the top management. If the low profits lead to a fall in the price of shares, a take-over raid may be successful and the position of managers is thus endangered.

The three financial ratios are combined (subjectively by the managers) into a single parameter $\bar{a}$ which is called the '*financial security constraint*'. This is exogenously determined, by the risk attitude of the top management. Marris does not explain the process by which $\bar{a}$ is determined. It is stated that it is not a simple average of the three ratios, but rather a weighted average, the weights depending on the subjective decisions of managers.

Two points should be stressed regarding the overall financial constraint $\bar{a}$.

Firstly. Let

$$a_1 = \text{liquidity ratio} = \frac{L}{A}$$

$$a_2 = \text{leverage ratio} = \frac{D}{A}$$

$$a_3 = \text{retention ratio} = \frac{\Pi_R}{\Pi}$$

Marris postulates that the overall $\bar{a}$ is negatively related to a_1, and positively to a_2 and a_3. That is, $\bar{a}$ increases if *either* the liquidity is reduced, *or* the debt ratio is raised by increasing external finance (loans), *or* the proportion of retained profits is increased. Similarly, $\bar{a}$ declines if the managers increase the liquidity of the firm, or reduce the proportion of external finance (D/A), or reduce the proportion of retained profits (that is, increase the distributed profits), or a combination of all three.

Secondly, Marris implicitly assumes that there is a negative relation between 'job security' (s) and the financial constraint $\bar{a}$: if $\bar{a}$ increases (by either reducing a_1 or increasing a_2 or increasing a_3) clearly the position of the firm becomes more vulnerable to bankruptcy and/or to take-over raids, and consequently the job security of managers is reduced. Thus a high value of $\bar{a}$ implies that the managers are risk-takers, while a low value of $\bar{a}$ shows that managers are risk-avoiders.

The financial security constraint sets a limit to the rate of growth of the capital supply, g_C, in Marris model (see below).

III. THE MODEL: EQUILIBRIUM OF THE FIRM

The managers aim at the maximisation of their own utility, which is a function of the growth of demand for the products of the firm (given the security constraint)

$$U_{\text{managers}} = f(g_D)^1$$

The owners–shareholders aim at the maximisation of their own utility which Marris assumes to be a function of the rate of growth of the capital supply (and not of profits, as the traditional theory postulated)

$$U_{\text{owners}} = f^*(g_C)$$

The firm is in equilibrium when the maximum *balanced-growth rate* is attained, that is, the condition for equilibrium is

$$g_D = g_C = g^* \text{ maximum}$$

The first stage in the solution of the model is to derive the 'demand' and 'supply' functions, that is, to determine the factors that determine g_D and g_C.

Marris establishes that the factors that determine g_D and g_C can be expressed in terms of two variables, the diversification rate, d, and the average profit margin, m.

The instrumental variables

The firm will first determine (subjectively) its financial policy, that is, the value of the financial constraint $\bar{a}$, and subsequently it will choose the rate of diversification d, and the profit margin m, which maximise the balanced-growth rate g^*.

The following are policy variables in the Marris model:

Firstly, $\bar{a}$ implies freedom of choice of the financial policy of the firm. As we shall see in section IV, the firm can affect its rate of growth by changing its three security ratios (leverage, liquidity, dividend policies).

Secondly, the firm can choose its diversification rate, d, either by a change in the style of its existing range of products, or by expanding the range of its products.

Thirdly, in Marris's model *price is given* by the oligopolistic structure of the industry.

[1] At this stage it is assumed that there is a saturation level of utility from security, so that s becomes a constraint. At a later stage this assumption is relaxed and a general managerial utility function is used in which U_M increases as both g_D and s increase.

Hence price is not actually a policy variable of the firm. The determination of the price in the market is very briefly mentioned in Marris's article. He argues that eventually a price structure will develop in which the market shares are stabilised. This equilibrium will be arrived at *either* by tacit collusion, *or* after a period of war during which price competition, advertising, product variation or all three weapons are used. The length of time involved and the level of price and the number of firms which will remain in business is uncertain, due to 'imperfect knowledge of the competitors' strength, determination, and skill', and from the unpredictability of games containing chance moves. From this line of argument it seems that Marris is not concerned with price determination in oligopolistic markets, but rather takes it for granted that a price structure will eventually develop. Thus Marris seems to treat price as a parameter (given) rather than as a policy variable at the discretion of the firm. Similarly, Marris assumes that production costs are given.

Fourthly, the firm can choose the level of its advertising, A, and of its research and development activities, R & D. Since the price, P, and the production costs, C, are given, then it is obvious that a higher A and/or R & D expenditures will imply a lower average profit margin and, vice versa, a low level of A and/or R & D implies a higher average profit rate. Implicit in Marris's model is the average-cost pricing rule

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

where $\bar{P}$ = price, given from the market
$\bar{C}$ = production costs, assumed given
A = advertising and other selling expenses
R & D = research and development expenses
m = average profit margin

Clearly m is the residual

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

Given $\bar{P}$ and $\bar{C}$, m is negatively correlated with the level of advertising and R & D expenditures. Thus m is used as a proxy for the policy variables A and R & D.

In summary, all the policy variables are combined into three instruments:

$\bar{a}$, the financial security coefficient
d, the rate of diversification
m, the average profit margin

The next step is to define the variables that determine the rate of growth of demand, g_D, and the rate of growth of supply, g_C, and express these rates in terms of the policy variables, $\bar{a}$, d and m.

The rate of growth of the demand: g_D

It is assumed that the firm grows by diversification. Growth by merger or take-over is excluded from this model.

The rate of growth of demand for the products of the firm depends on the diversification rate, d, and the percentage of successful new products, k, that is,

$$g_D = f_1(d, k)$$

where d = the diversification rate, defined as the number of new products introduced per time period, and k = the proportion of successful new products.

Diversification may take two forms. Firstly, the firm may introduce a completely new product, which has no close substitutes, which creates *new demand* and thus competes with other products for the income of the consumer. (Marris seems to narrow his analysis to firms producing consumers' goods.) This Marris calls *differentiated diversification,* and is considered the most important form in which the firm seeks to grow, since there is no danger of encroaching on the market of competitors and hence provoking retaliation. Secondly, the firm may introduce a product which is a substitute for similar commodities already produced by existing competitors. This is called *imitative diversification*, and is almost certain to induce competitors' reactions. Given the uncertainty regarding the reactions of competitors the firm prefers to diversify with new products. The greater d, the higher the rate of growth of demand.

The proportion of successful new products, k, depends on the rate of diversification d, on their price, the advertising expenses, and the R & D expenditure, as well as on the intrinsic value of the products:

$$k = f_3(d, P, A, \text{R \& D, intrinsic value})$$

Regarding the intrinsic value of the new product Marris seems to adopt Galbraith's[1] and Penrose's[2] thesis (rather far-fetched)[3] that a firm can sell almost anything to the consumers by an appropriately organised selling campaign, even against consumers' resistance. He implicitly combines intrinsic value with price, that is, price is associated with a given intrinsic value. Price is assumed to have reached equilibrium in some way or another. Thus price is taken as given, despite the fact that the product is new.

k depends on the advertising, A, the R & D expenditures and on d. The higher A and/or R & D, the higher the proportion of successful new products and vice versa. Marris uses m, the average profit margin as a proxy for these two policy variables. Given that m is negatively related to A and R & D, the proportion of successful new products is also negatively correlated with the average profit margin.

Finally, k depends on d, the rate of new products introduced in each period: if too many new products are introduced too fast, the proportion of fails increases. Thus, although the rate of growth of demand, g_D, is positively correlated with the diversification rate (d), g_D increases at a decreasing rate as d increases, due to the rate of introduction of new products outrunning the capacity of the personnel involved in the development and the marketing of the products. There is an optimal rate of flow of 'new ideas' from the R & D department of the firm. If the research team is pressed to speed up the development process of new products there is no time to 'research' the product and/or its marketability adequately. Furthermore, top management becomes overworked when the rate of introduction of new products is high, and the proportion of unsuccessful products is bound to increase.

In summary

$$g_D = f_1(d, m)$$

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

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

The g_D function is shown in figures 16.1 and 16.2.

[1] See J. K. Galbraith, *The New Industrial State* (Houghton, Mifflin, 1967).

[2] See E. Penrose, *The Theory of the Growth of the Firm* (Blackwell, 1959).

[3] See H. Townsend, 'Competition and Big Business', in T. M. Rybczynski (ed)., *A New Era in Competition* (Blackwell, 1973).

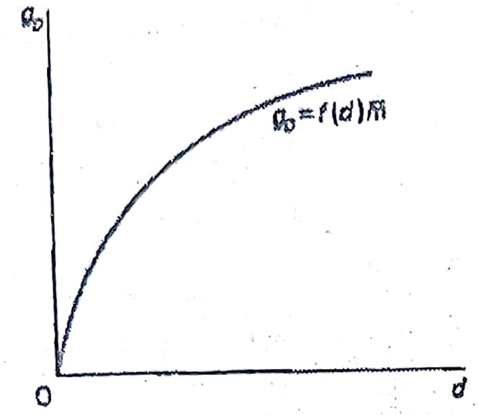


Figure 16.1 $g_D = f(d)$ given m

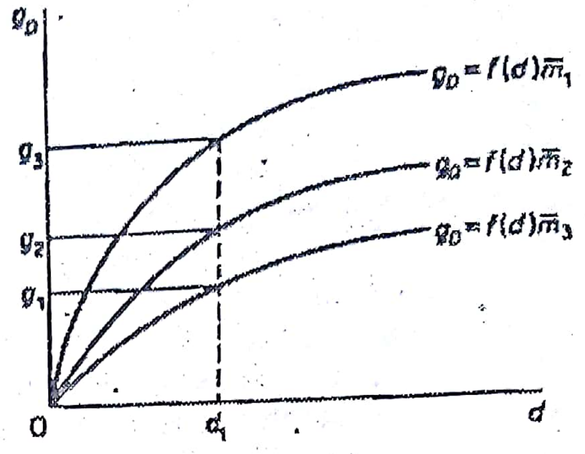


Figure 16.2 $g_D = f(d, m)$

The average rate of profit is constant along any g_D curve. But the curve shifts downwards as m increases ($\bar{m}_1 < \bar{m}_2 < \bar{m}_3$). This is due to the negative relationship between g_D and m. With a given rate of diversification (for example, at d_1 in figure 16.2) and given the price of the products, the lower m, the larger the A and/or the R & D expenses, and hence the larger the proportion of successful products and the higher the growth of demand ($g_3 > g_2 > g_1$). Of course the monotonic positive relationship between d and A (and R & D), which is implied by Galbraith's and Penrose's hypothesis and is adopted by Marris, is highly questionable on *a priori* and empirical grounds.[1]

The rate of growth of capital supply: g_C

It is assumed that the shareholder–owners aim at the maximisation of the rate of growth of the corporate capital, which is taken as a measure of the size of the firm. Corporate capital is defined as the sum of fixed assets, inventories, short-term assets and cash reserves. It is not stated why the shareholders prefer growth to profits in periods during which growth is not steady.

The rate of growth is financed from internal and external sources. The source of internal finance for growth is profits. External finance may be obtained by the issue of new bonds or from bank loans. The optimal relation between external and internal finance is still strongly disputed in economic literature.[2]

Marris takes the position that the main source of finance for growth is profits, on the following grounds. Firstly, the issue of new shares as a means of obtaining funds is, for prestige and other reasons, not often used by an established firm. Secondly, external finance is limited by the security attitude of managers, that is, from their desire to avoid mass dismissal. Financial security is achieved by setting an upper limit to the debt/assets ratio (leverage) and a lower limit to the liquidity ratio in the long run.

Although profits are the main source of finance for growth, the top management cannot retain as much profits as it would like. There is an upper limit to the 'retention ratio' set by the desire of managers to distribute a satisfactory dividend, which will keep shareholders happy and avoid a fall in the prices of shares. Otherwise the selling of shares, or a successful take-over raid, would endanger the position of managers.

The three security ratios are subjectively determined by the managers through the security parameter $\bar{a}$, which is a determinant of the retained profits, and hence a determinant of the rate of growth of capital.

[1] See H. Townsend, 'Competition and Big Business'.

[2] J. Meyer and E. Kuh, *The Investment Decision* (Harvard University Press, Cambridge, Mass., 1957).

Under Marris's assumptions the rate of growth of capital supply is proportional to the level of profits

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

where $\bar{a}$ = the financial security coefficient
Π = level of total profits

The security coefficient $\bar{a}$ is assumed constant and exogenously determined in this model. This assumption is relaxed at a later stage. It should be stressed, however, that so long as $\bar{a}$ is constant, growth, g_C, and profits, Π, are *not* competing goals, but are positively related: higher profits imply higher rate of growth.

The next step is to express g_C in terms of the policy variables d and m.

The level of total profits depends on the average rate of profit, m, and on the efficiency of the performance of the firm as reflected by its overall capital output ratio, K/X:

$$\Pi = f_4\left(m, \frac{K}{X}\right)$$

It is intuitively obvious that Π and m are positively correlated (an increase in the average profit margin results in an increase in the total profits)

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

The relationship between Π and the capital/output ratio is more complicated. The capital/output ratio is claimed to be a measure of efficiency of the activity of the firm, given its human and capital resources. The overall K/X ratio is not a simple arithmetic average of the capital/output ratios of the individual products of the firm, but is a function of the diversification rate d

$$(\bar{K}/X) = f_5(d)$$

Given $\bar{K}$, the relation between X and d is up to a certain level of d positive, reaches a maximum, and subsequently output declines with further increases in the number of new products: the overall output increases initially with d due to a better utilisation of the team in the R & D department as well as of the skills of the existing managerial team. Output reaches a maximum when the d is at its optimum level allowing the optimal use of the managerial team and the R & D personnel. Beyond that point, the total output X decreases with further increases in d, and the efficiency of the firm falls: the R & D personnel are overworked and the decision-making process becomes inefficient, as there is not enough time allowed for the development of new products or for the study of their marketability. Hence the success rate for new products falls and efficiency declines.

Substituting for K/X in the profit function we obtain

$$\Pi = f_4(m, d)$$

The relationship between Π and d is initially positive, reaches a maximum, and then declines as d is further accelerated.

We next substitute Π in the g_C function

$$g_C = \bar{a}.[f_4(m, d)]$$

The rate of growth of capital is determined by three factors: the financial policies of the managers, the average rate of profit and the diversification rate.

Marris assumes in his initial model that $\bar{a}$ is a constant parameter exogenously determined by the risk-attitude of managers, while there is a positive relation between g_C and m

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

The relationship between g_C and d is not monotonic. The rate of growth of capital, g_C, is positively correlated with d up to the point of optimal use of the R & D personnel and the team of managers; but g_C is negatively correlated with d beyond that point: a higher d implies hastening up of the diversification process $\longrightarrow$ inefficient decisions $\longrightarrow$ fall in the overall profit level $\longrightarrow$ low availability of internal finance and consequently a lower rate of growth g_C.

The relation between g_C and d, keeping $\bar{a}$ and m constant, is shown in figure 16.3. If we allow both d and m to change, while keeping $\bar{a}$ constant, we obtain a family of $g_C = f_?$ (d, m) curves (figure 16.4). The average profit rate is depicted as a shift factor of the $g_C = f(d)$ curve. The higher the average profit rate, the further from the origin the g_C curves will be ($m_1 < m_2 < m_3$). These curves are drawn under the assumption that $\bar{a}$ is constant. (The effects of a change in $\bar{a}$ are discussed in section IV below.)

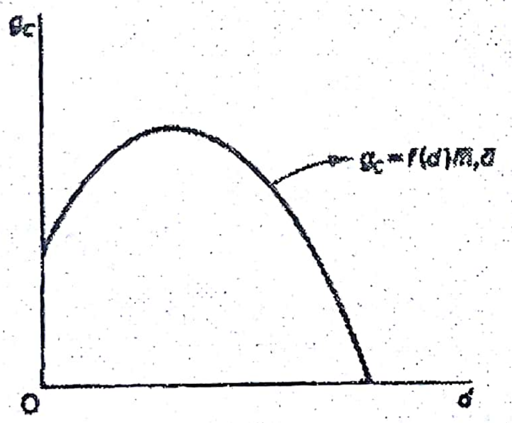


Figure 16.3 $g_C = f(d)$, given $\bar{m}$ and $\bar{a}$

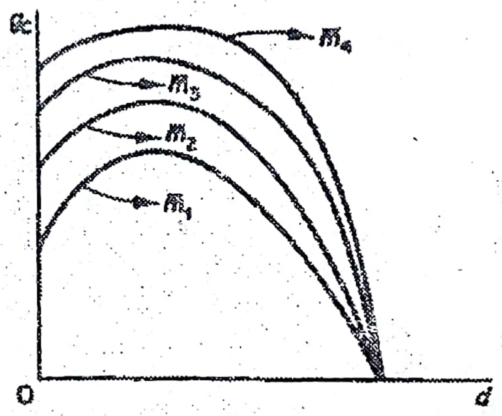


Figure 16.4 $g_C = f(d, m)$, given $\bar{a}$

Summarising the above arguments, we may present Marris's model in its complete form as follows.

$g_D = f_1(m, d)$	(demand-growth equation)
$\Pi = f_4(m, d)$	(profit equation)
$g_C = \bar{a}.[f_4(m, d)]$	(supply-of-capital equation)
$\bar{a} \leq a^*$	(security constraint)
$g_D = g_C$	(balanced-growth equilibrium condition)

$\bar{a}$ is exogenously determined by the risk-attitude of managers. The level of profit Π is endogenously determined. The variables m and d are the policy instruments. Given the balanced-growth equilibrium condition, we have in fact one equation in two unknowns (m and d, given $\bar{a}$)

$$f_1(m, d) = \bar{a}.[f_4(m, d)]$$

Equilibrium of the firm

Clearly the model cannot be solved (is underidentified), unless one of the variables m or d is subjectively determined by the managers. Once the managers define $\bar{a}$ and one of the other two policy variables, the equilibrium rate of growth can be determined.

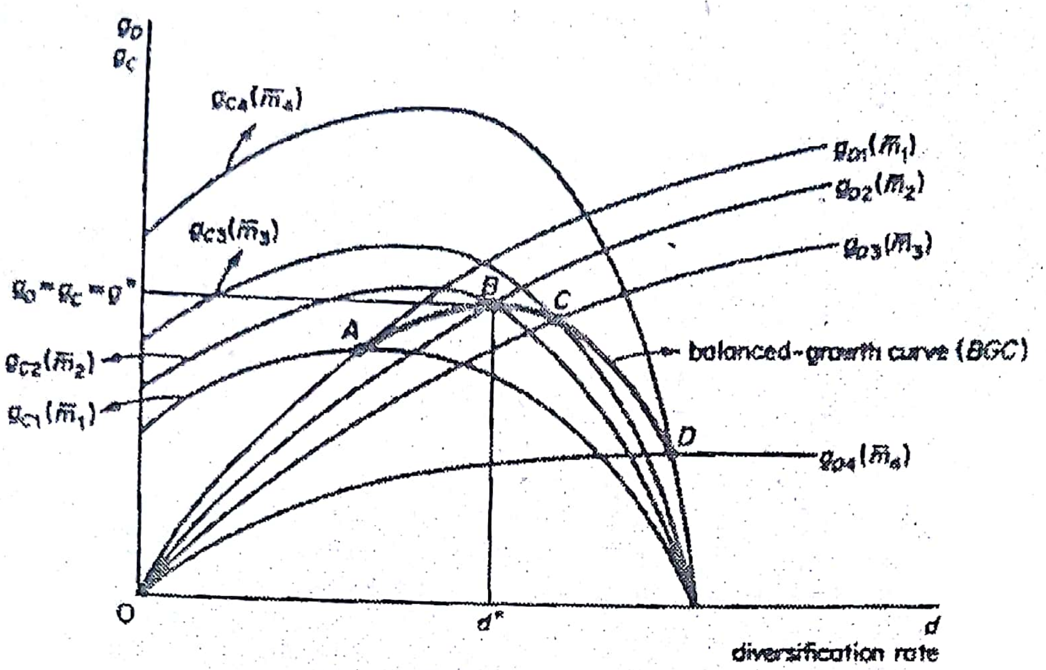


Figure 16.5

The equilibrium of the firm is presented graphically in figure 16.5, formed by superimposing figures 16.2 and 16.4. Given their shapes, the g_D and g_C curves associated with a given profit rate intersect at some point. For example, the g_D and g_C curves corresponding to m_1 intersect at point A; the g_D and g_C curves associated with m_2 intersect at point B, and so on. If we join all points of intersection of g_D and g_C curves corresponding to the same level of m we form what Marris calls the *balanced-growth curve* (*BGC*), given the financial coefficient $\bar{a}$.

The firm is in equilibrium when it reaches the highest point on the balanced-growth curve. The firm decides its financial policy, denoted by $\bar{a}$. It next chooses subjectively a value for either m or d. With these decisions taken, the firm can find its maximum balanced-growth rate, consistent with $\bar{a}$ and with the chosen value of one of the other two policy variables. In figure 16.5 the *BGC* corresponding to $\bar{a}$ is *ABCD*. The balanced-growth rate g^* is defined by the highest point B of this *BGC*. This g^* rate is compatible with a unique pair of values of the policy variables, m^* and d^*. If the firm chooses d^*, then m^* is simultaneously determined; alternatively, if the firm chooses m^*, then d^* is simultaneously determined from the function

$$g^* = f_1(m^*, d^*) = \bar{a}\cdot[f_4(m^*, d^*)]$$

Substituting m^* and d^* in the profit function

$$\Pi = \bar{a}[f_4(m, d)]$$

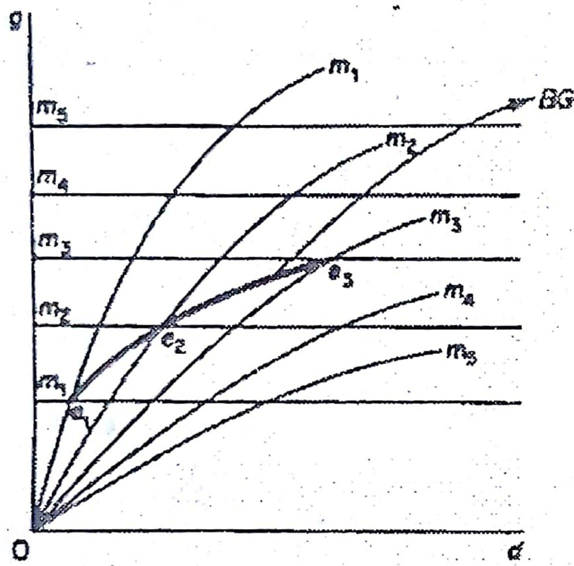
we find the level of profit, Π^*, required to finance the balanced-growth rate, g^*. Thus profit is endogenously determined in Marris's model. Furthermore, growth and profit are *not* competing goals (so long as $\bar{a}$ is constant). From the g_C function

$$g_C = \bar{a}\cdot(\Pi)$$

it is obvious that higher profit implies higher growth rate. However, if the financial coefficient $\bar{a}$ is allowed to vary, then profits and growth become competing goals (see below).[1]

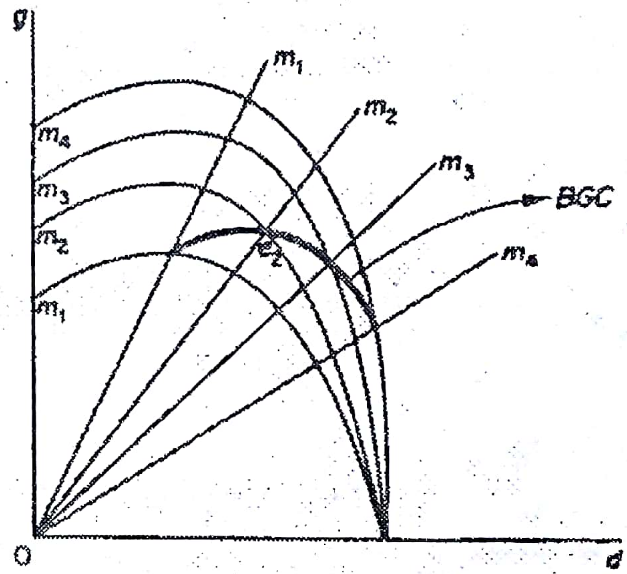
[1] If the endogenously determined profit Π^* is very low and is not adequate to satisfy the demands of the shareholders (for dividends), the managers will have to reduce the retention ratio, that is, reduce the $\bar{a}$ coefficient. The model will then be solved again with the new (smaller) $\bar{a}$. The solution will yield a lower growth rate, but a higher profit level Π. See section IV.

The question is: does the *BGC* have a maximum? Marris argues that so long as either (or both) of the g_C or g_D curves flattens out or bends, there will always be a maximum point on the *BGC* curve. Furthermore, depending on the shape of the g_C and the g_D curves, the *BGC* may be *platykurtic*, that is, have a flat stretch which indicates that there are several optimal solutions: the g^* may be achieved by a large number of combinations of the values of the policy variables m and d (given $\bar{a}$ is already chosen). It is only if the g_C curve is parallel to the d-axis ($g_C = f(m)$ but $g_C \neq f(d)$) *and* the g_D curves are straight upwards-sloping curves (implying that $g_D = f(d, m)$, but $k \neq f(d)$ and hence the g_D curve does not flatten out) that the *BGC* increases continuously, never reaching a maximum. This situation is, however, improbable given the capacity for efficient decision making of the managerial team and the capacity for well-explored new products of the R & D department of the firm.



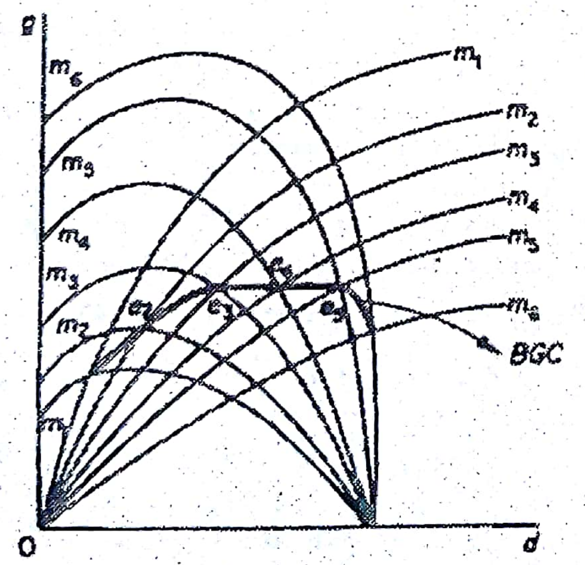
$g_D = f_1(d, m)$
$g_C \neq f(d)$
$g_C = f(m)$

Figure 16.6



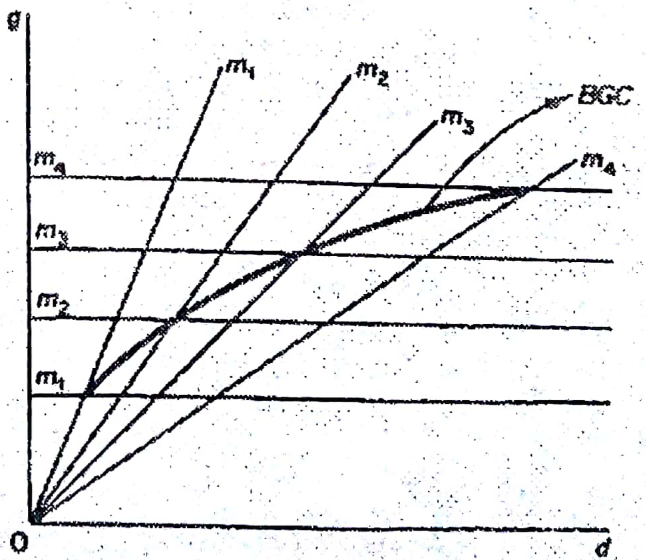
$g_D = f(d, m)$
(with $\partial g_D/\partial d$ constant)
$g_C = f(d, m)$

Figure 16.7



$g_D = f(m, d)$
$g_C = f(m, d)$

Figure 16.8



$g_D = f(m, d)$, $\dfrac{\partial g_D}{\partial d}$ constant

$g_D \neq f(d)$, $g_C = f(m)$

Figure 16.9

firms will have different time-horizons, different risk attitude and will form different assessments of uncertain future events. Consequently firms will respond differently to the same (given) conditions of the environment. The interactions among uncertainty, risk-aversion and the time-horizon of the entrepreneurs are not dealt satisfactorily with by the probabilistic approach adopted in the traditional theory of the firm.

The static nature of traditional theory

Time enters into the traditional theory in three respects.[1] Firstly, the distinction between the short and the long run implies time considerations. However, the theory does not really answer the question: how long is the long run in actual decision-making. Secondly, it is assumed that the firm has some time-horizon over which it attempts to maximise its profits. The discounting of future costs and revenues implicit in the long-run profit maximisation clearly involves the time element. But again the length of the time-horizon and its interaction with uncertainty and risk-aversion is not adequately dealt with. Thirdly, the analysis of the timing of demand relative to the production flow implies a period analysis. Considerations of the gestation period of investment and the final product also imply time considerations.

However, the traditional theory is basically static. The time-horizon of the firm consists of *identical* and *independent* time-periods. Decisions are treated as temporally independent, and this is probably the most important shortcoming of the traditional theory. It is clear that decisions are temporally interdependent: decisions taken in any one period are affected by the decisions in past periods, and will in turn influence the future decisions of the firm. This interdependence is ignored by the traditional theory, which postulates that the long-run profits are maximised as the firm maximises its short-run profits in any one period, by equating its marginal cost to its marginal revenue ($MC = MR$).

Entry considerations

In the traditional theory entry considerations differ, depending on the type of market structure. The common features of the treatment of entry in traditional theory are two: firstly, only actual entry is considered; secondly, entry is assumed to take place only in the long run: it is a long-run phenomenon.

In pure competition entry is free. The same is assumed in the model of monopolistic competition. In both models entry can occur only in the long run.

In monopoly entry is blockaded by definition.

The traditional models of oligopoly are implicit, cryptic or silent so far as entry is concerned.[2] The classical duopoly models are 'closed' models in that they do not allow for entry. These models can be extended to larger numbers of sellers, but the numbers remain unchanged in the final market equilibrium as they were at the initial situation. In the theory of cartels it is implicitly assumed that the entrants, if any, will join the cartel. Without this implicit assumption the inherent instability of cartels becomes even greater. Similar assumptions are made in the traditional price-leadership models: the entrant is assumed to be either a small firm which can be coerced to follow the leader, or is assumed to accept the *status quo* of the established leader.

[1] See Ira Horowitz, *Decision Making and the Theory of the Firm* (Holt, Rinehart & Winston, 1970) p. 332.

[2] See J. Bain, *Barriers to New Competition* (Harvard University Press, Cambridge, Mass., 1956) p. 6.

Potential entry and its effects on decision making are not dealt with in traditional theory.[1]

The marginalist principle

The behavioural rule postulated by the traditional theory in actual decision-making is described by the so-called 'marginalist principle'

$$MC = MR$$

In each period the firm maximises its (short-run) profit by setting its output and price at the level defined by the intersection of the MC and MR curves. Given the temporal independence of decisions, such short-run profit maximisation implies also long-run profit maximisation.

This behavioural rule has been attacked on several grounds. One line of argument is that although the goal of the firm is long-run profit maximisation, this is not necessarily attained by equating the short-run marginal cost ($SRMC$) to the short-run marginal revenue ($SRMR$). Another line of attack centres on the goal of profit maximisation as the single goal of the firm. This has been briefly examined in Section B above, and will be discussed in detail in the various chapters which deal with theories postulating other goals (managerialism, behaviourism, limit-pricing).

In the rest of this chapter we will examine briefly several arguments arising from the above lines of attack. These arguments constitute the essence of what has become known as the 'marginalist controversy.'

II. THE HALL AND HITCH REPORT AND THE 'FULL-COST' PRICING PRINCIPLE

In 1939 Hall and Hitch published some results[2] of research undertaken at Oxford and aiming at the investigation of the decision process of businessmen in relation to government measures.[3] Their study covered 38 firms, out of which 33 were manufacturing firms, 3 were retail trading firms and 2 were building firms. Of the 33 manufacturing firms, 15 produced consumer goods, 4 intermediate products, 7 capital goods, and 7 textiles. The sample was not random, but included firms which may well be expected to belong to 'efficiently managed enterprises'.[3]

The most startling results of the studies of 'The Oxford Economists Research Group' reported by Hall and Hitch were that firms did not attempt to maximise their profits, that they did not use the marginalist rule $MC = MR$, and that oligopoly was the main market structure of the business world. Up to then the theory of monopolistic or imperfect competition of Chamberlin and Joan Robinson had been generally accepted as typical or relevant. The firms were assumed to be able to act atomistically, ignoring their rivals' reactions and pursuing their short-run (and long-run) profit maximisation by equating marginal cost to marginal revenue in each time period. The tangency solution in the long run implied normal profits, but excess capacity (unexhausted

[1] See P. W. S. Andrews, *On Competition in Economic Theory* (Macmillan, 1964).

[2] Hall and Hitch, 'Price Theory and Business Behaviour', *Oxford Economic Papers* (1939), reprinted in P. W. S. Andrews and T. Wilson (eds.), *Oxford Studies in the Price Mechanism* (Oxford University Press, 1952).

[3] R. Barback, *The Pricing of Manufactures* (Macmillan, 1964).

economies of scale), which was the main source of criticism of the regime of monopolistic competition. (See Chapter 8.)

Hall and Hitch's findings may be summarised as follows.

Firstly, the firms do not act atomistically. Firms are continuously conscious of the reactions of their competitors. This behaviour, obviously in contradiction to the postulates of monopolistic competition, suggested that oligopoly was much more widespread than had been thought up to that time. Oligopolistic interdependence could not be dealt with within the framework of the traditional theory. Duopoly theory, based on assumptions of constant reaction patterns of competitors, also seemed inadequate to cope with oligopolistic interdependence and the ensuing uncertainty regarding the demand for the product of oligopolistic firms.

Hall and Hitch found that firms do not attempt to maximise short-run profits by applying marginalistic rules ($MC = MR$), but aim at long-run profit maximisation. Firms set their price on the *average-cost principle*.[1] That is, firms do not set their price and output at the levels determined by the intersection of the MC and MR curves, but they set a price to cover the average variable cost, the average fixed cost and a 'normal' profit margin ('usually 10%')

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

The reasons given by Hall and Hitch for the breakdown of marginalism may be summarised as follows. (*a*) Firms do not know their demand curve nor their marginal costs, hence the application of the marginalist rule ($MC = MR$) is impossible due to lack of relevant information. (*b*) Firms believed that the 'full-cost price' is the 'right' price, since it allowed a 'fair' profit and covered the costs of production when the plant was 'normally' utilised.

Hall and Hitch reported that firms' main preoccupation is price, and not output as the traditional theory implied. Thus the firms would set their price on the basis of the above average-cost principle and they would sell at that price whatever the market would take.

Although the firms in general would adhere to the average-cost pricing rule, they would be prepared to depart from it if they wanted to secure a big order, or if they thought that they could not set this price without damaging their goodwill and endangering their future position, in view of their rivals charging a lower price.

Finally, it was found that prices of manufacturers were fairly sticky, despite changes in demand and costs. Traditional theory predicted a change in price and output (at least in the short run) in response to changes in demand and costs. This prediction was not observed in the real world, in which 'stickiness' of prices was a general phenomenon. To explain this stickiness of prices Hall and Hitch introduced the Chamberlinian apparatus of the 'kinked' demand curve. The kink implies the following pattern of expected reactions of competitors, witnessed by the firms that were studied. Businessmen held the belief that if they raised their price their competitors would not follow, so that they would lose a considerable number of their customers; while if they cut their price their competitors would follow suit, with the result that their sales would increase by an insignificant amount. The price was set equal to the average cost and the kink would occur at that price. The firms would arrive at this average-cost price independently (without collusion). Firms reported that they did not enter any collusive agreement in order to increase the market price, because of fear of potential entrants who might endanger the long-run position of established firms. Similarly, they did not think that a collusive reduction in the price would be profitable to the 'group', because they believed

[1] This principle is often called, although but mistakenly, the *full-cost principle*, due to the use of this terminology by Hall and Hitch.

that the market demand was price inelastic. Furthermore businessmen were aware of the fact that frequent price changes were disliked by their customers. Thus firms prefer to keep their price constant, except in cases of general cost increases which affect all firms and lead to an increase of price by all. The 'kink' in the individual-demand curve, implying that above the given P the demand is very elastic while below P demand is inelastic, provides an explanation of the firms keeping their price constant.

Various writers have interpreted the device of the kinked-demand curve as a 'theory' of oligopolistic behaviour. This view has become so firmly established that the kinked-demand 'theory' has found its place in all textbooks of microeconomics. However, as we saw in Chapter 9, the kinked demand cannot be considered as a theory of pricing and output decisions in oligopolistic markets, since it cannot explain the level of the price, i.e. the location of the kink. In the framework of Hall and Hitch's approach the kinked demand makes sense as a device for explaining the 'stickiness' of the prices, but not their levels. Hall and Hitch suggest that the firms set their price independently of one another on the basis of the average-cost principle. And these prices, once set, tend to be sticky because of the expectations of firms about competitors' reactions to changes in the firm's price, which are such as to create a kink at the level of the average-cost price.

It does not seem that Hall and Hitch have developed a determinate theory of pricing in oligopoly markets, and although they mention several elements relevant to oligopolistic behaviour (long-run profit considerations, goodwill of firms, potential and actual competition and interdependence of firms, average-cost pricing procedures instead of marginalistic computations), they did not combine them into a comprehensive theory of oligopoly. However, their conclusions constituted a serious attack on marginalism. The reported results of the study of their sample suggest the following. Firstly, short-run profit maximisation was rarely stated by businessmen to be their goal. Most firms reported that they aimed at a 'fair' level of profit and that they had also other goals, such as building up their goodwill, being fair to their competitors, etc. Secondly, the demand curve and its price elasticity, on which marginalism so heavily relies, are unknown in practice, because neither consumers preferences nor competitors' reactions are known with certainty. Most firms are oligopolists who are faced with uncertainty concerning their customers and their competitors. Given this uncertain environment the demand schedules cannot be taken as known, hence marginal revenue schedules are unknown to the businessmen. Thirdly, marginal costs are also unknown in multi-product firms, which are typical in the modern business world. Fourthly, even if MC and MR were known, and firms aimed at the maximisation of their (short-run) profits, the adherence to this equality would require continuous changes in the price in view of the continuous changes in costs and demand. Such frequent changes in prices are not desirable, and prices have exhibited considerable stickiness despite changes in short-run costs and demand.[1]

III. GORDON'S ATTACK ON MARGINALISM

In 1948 Gordon[2] joined the marginalist controversy by attacking the assumptions and postulates of the traditional theory of the firm. His line of attack may be summarised as follows.

[1] This finding has been challenged by subsequent empirical studies. See G. J. Stigler, 'The Kinky Oligopoly Demand Curve and Rigid Prices', *Journal of Political Economy* (1947). Also J. L. Simon, 'A Further Test of the Kinky Oligopoly Demand Curve', *American Economic Review*, (1969).

[2] R. A. Gordon, 'Short-Period Price Determination in Theory and Practice', *American Economic Review* (1948).

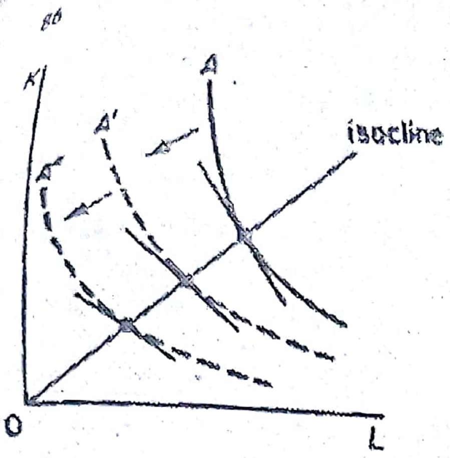


Figure 3.29 Capital-deepening technical progress

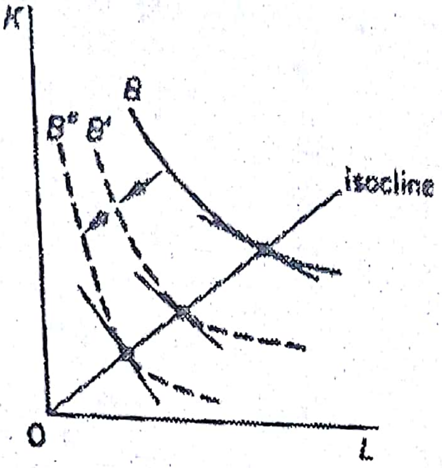


Figure 3.30 Labour-deepening technical progress

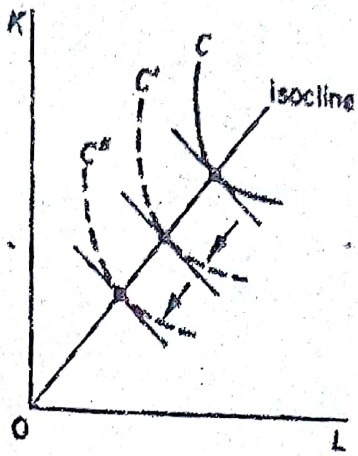


Figure 3.31 Neutral technical progress

Neutral-technical progress

Technical progress is neutral if it increases the marginal product of both factors by the same percentage, so that the $MRS_{L,K}$ (along any radius) remains constant. The isoquant shifts downwards parallel to itself. This is shown in figure 3.31.

IV. EQUILIBRIUM OF THE FIRM: CHOICE OF OPTIMAL COMBINATION OF FACTORS OF PRODUCTION

In this section we shall show the use of the production function in the choice of the optimal combination of factors by the firm. In Part A we will examine two cases in which the firm is faced with a single decision, namely maximising output for a given cost, and minimising cost subject to a given output. Both these decisions comprise cases of constrained profit maximisation in a single period.

In Part B we will consider the case of unconstrained profit maximisation, by the expansion of output over time.

In all the above cases it is assumed that the firm can choose the optimal combination of factors, that it can employ any amount of any factor in order to maximise its profits. This assumption is valid if the firm is new, or if the firm is in the long-run. However, an existing firm may be coerced, due to pressure of demand, to expand its output in the short-run, when at least one factor, usually capital, is constant. We will examine this case separately.

In all cases we make the following assumptions:

1. The goal of the firm is profit maximisation – that is, the maximisation of the difference $\Pi = R - C$ where

$$\Pi = \text{profits}$$
$$R = \text{revenue}$$
$$C = \text{cost}$$

2. The price of output is given, $\bar{P}_x$.
3. The prices of factors are given:

$\bar{w}$ is the given wage rate

$\bar{r}$ is the given price of capital services (rental price of machinery).

A. SINGLE DECISION OF THE FIRM

The problem facing the firm is that of a constrained profit maximisation, which may take one of the following forms:

(*a*) Maximise profit Π, subject to a cost constraint. In this case total cost and prices are given ($\bar{C}, \bar{w}, \bar{r}, \bar{P}_x$), and the problem may be stated as follows

$$\max \Pi = R - \bar{C}$$
$$\Pi = \bar{P}_x X - \bar{C}$$

Clearly maximisation of Π is achieved in this case if X is maximised, since $\bar{C}$ and $\bar{P}_x$ are given constants by assumption.

(*b*) Maximise profit Π, for a given level of output. For example, a contractor wants to build a bridge (X is given) with the maximum profit. In this case we have

$$\max \Pi = R - C$$
$$\Pi = \bar{P}_x \bar{X} - C$$

Clearly maximisation of Π is achieved in this case if cost C is minimised, given that $\bar{X}$ and P_x are given constants by assumption.

The analysis will be carried out first by using diagrams and subsequently by applying calculus.

For a graphical presentation of the equilibrium of the firm (its profit-maximising position) we will use the isoquant map (figure 3.32) and the isocost-line(s) (figure 3.33).

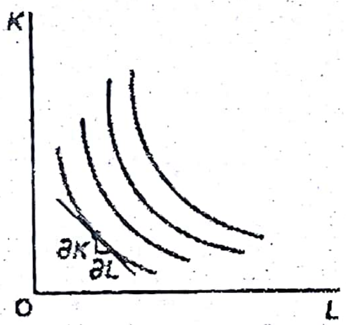


Figure 3.32

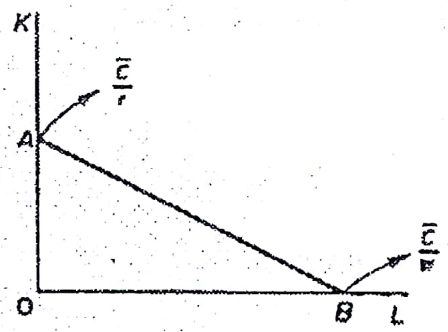


Figure 3.33

The isoquants have been explained in section I, where it was shown that the slope of an isoquant is

$$-\frac{\partial K}{\partial L} = MRS_{L,K} = \frac{MP_L}{MP_K} = \frac{\partial X/\partial L}{\partial X/\partial K}$$

The *isocost line* is defined by the cost equation

$$C = (r)(K) + (w)(L)$$

where w = wage rate, and r = price of capital services.

The isocost line is the locus of all combinations of factors the firm can purchase with a given monetary cost outlay.[1]

The slope of the isocost line is equal to the ratio of the prices of the factors of production:

$$\text{slope of isocost line} = \frac{w}{r}$$

[1] There is a close analogy between the consumer's budget line (Chapter 2, figure 2.10) and the firm's isocost line.

Proof

Assume that the total cost outlay the firm undertakes is $\bar{C}$. If the entrepreneur spends all the amount $\bar{C}$ on capital equipment, the maximum amount he can buy from this factor is

$$0A = \frac{\bar{C}}{r}$$

If all cost outlay is spent on labour the maximum amount of this factor that the firm can purchase is

$$0B = \frac{\bar{C}}{w}$$

The slope of the isocost line is

$$\frac{0A}{0B} = \frac{\bar{C}/r}{\bar{C}/w} = \frac{w}{r}$$

It can be shown that any point on the line AB satisfies the cost equation ($C = r \cdot K + w \cdot L$), so that, for given prices of the factors and for given expenditure on them, the isocost line shows the alternative combinations of K and L that can be purchased by the firm. The equation of the isocost line is found by solving the cost equation for K:

$$K = \frac{\bar{C}}{r} - \frac{w}{r}L$$

By assigning various values to L we can find all the points of the isocost line.

Case I: maximisation of output subject to a cost constraint (financial constraint)

We assume: (*a*) A given production function

$$X = f(L, K, v, \gamma)$$

and (*b*) given factor prices, w, r, for labour and capital respectively.

The firm is in equilibrium when it maximizes its output given its total cost outlay and the prices of the factors, w and r.

In figure 3.34 we see that the maximum level of output the firm can produce, given the cost constraint, is X_2 defined by the tangency of the isocost line, and the highest isoquant. The optimal combination of factors of production is K_2 and L_2, for prices w and r. Higher levels of output (to the right of e) are desirable but not attainable due

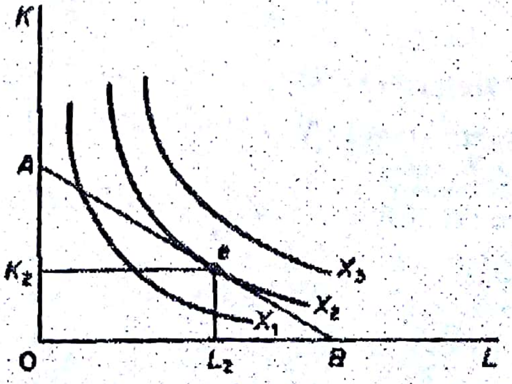


Figure 3.34

to the cost constraint. Other points on AB or below it lie on a lower isoquant than X_2. Hence X_2 is the maximum output possible under the above assumptions (of given cost outlay, given production function, and given factor prices). At the point of tangency (e) the slope of the isocost line (w/r) is equal to the slope of the isoquant (MP_L/MP_K). This constitutes the first condition for equilibrium. The second condition is that the isoquants be convex to the origin. In summary: the conditions for equilibrium of the firm are:

(a) Slope of isoquant = Slope of isocost

or
$$\frac{w}{r} = \frac{MP_L}{MP_K} = \frac{\partial X/\partial L}{\partial X/\partial K} = MRS_{L,K}$$

(b) The isoquants must be convex to the origin. If the isoquant is concave the point of tangency of the isocost and the isoquant curves does not define an equilibrium position (figure 3.35).

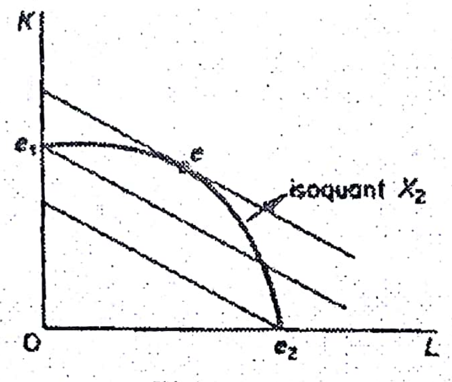


Figure 3.35

Output X_2 (depicted by the concave isoquant) can be produced with lower cost at e_2 which lies on a lower isocost curve than e. (With a concave isoquant we have a 'corne solution.)

Formal derivation of the equilibrium conditions

The equilibrium conditions may be obtained by applying calculus and solving a 'constrain maximum' problem which may be stated as follows. The rational entrepreneur seeks the ma misation of his output, given his total-cost outlay and the prices of factors. Formally:

Maximise $\quad X = f(L, K)$

subject to $\quad \bar{C} = wL + rK \quad$ (cost constraint)

This is a problem of constrained maximum and the above conditions for the equilibrium of t firm may be obtained from its solution.

We can solve this problem by using Lagrangian multipliers. The solution involves the followi steps:

Rewrite the constraint in the form

$$\bar{C} - wL - rK = 0$$

Multiply the constraint by a constant λ which is the Lagrangian multiplier:

$$\lambda(\bar{C} - wL - rK) = 0$$

The Lagrangian multipliers are undefined constants which are used for solving constrain maxima or minima. Their value is determined simultaneously with the values of the oth

unknowns (L and K in our example). There will be as many Lagrangian multipliers as there are constraints in the problem.
Form the 'composite' function

$$\phi = X + \lambda(\bar{C} - wL - rK)$$

It can be shown that maximisation of the ϕ function implies maximisation of the output.

The first condition for the maximisation of a function is that its partial derivatives be equal to zero. The partial derivatives of the above function with respect to L, K and λ are:

$$\frac{\partial \phi}{\partial L} = \frac{\partial X}{\partial L} + \lambda(-w) = 0 \tag{3.1}$$

$$\frac{\partial \phi}{\partial K} = \frac{\partial X}{\partial K} + \lambda(-r) = 0 \tag{3.2}$$

$$\frac{\partial \phi}{\partial \lambda} = \bar{C} - wL - rK = 0 \tag{3.3}$$

Solving the first two equations for λ we obtain

$$\frac{\partial X}{\partial L} = \lambda w \quad \text{or} \quad \lambda = \frac{\partial X/\partial L}{w} = \frac{MP_L}{w}$$

$$\frac{\partial X}{\partial K} = \lambda r \quad \text{or} \quad \lambda = \frac{\partial X/\partial K}{r} = \frac{MP_K}{r}$$

The two expressions must be equal; thus

$$\frac{\partial X/\partial L}{w} = \frac{\partial X/\partial K}{r} \quad \text{or} \quad \frac{MP_L}{MP_K} = \frac{\partial X/\partial L}{\partial X/\partial K} = \frac{w}{r}$$

This firm is in equilibrium when it equates the ratio of the marginal productivities of factors to the ratio of their prices.

It can be shown[1] that the second-order conditions for equilibrium of the firm require that the marginal product curves of the two factors have a negative slope.

The slope of the marginal product curve of labour is the second derivative of the production function:

$$\text{slope of } MP_L \text{ curve} = \frac{\partial^2 X}{\partial L^2}$$

Similarly for capital:

$$\text{slope of } MP_K \text{ curve} = \frac{\partial^2 X}{\partial K^2}$$

The second-order conditions are

$$\frac{\partial^2 X}{\partial L^2} < 0 \quad \text{and} \quad \frac{\partial^2 X}{\partial K^2} < 0$$

and

$$\left(\frac{\partial^2 X}{\partial L^2}\right)\left(\frac{\partial^2 X}{\partial K^2}\right) > \left(\frac{\partial^2 X}{\partial L\, \partial K}\right)^2$$

These conditions are sufficient for establishing the convexity of the isoquants.

[1] See Henderson and Quandt, *Microeconomic Theory* (McGraw-Hill, 1958) pp. 49–54.

Case 2: minimisation of cost for a given level of output

The conditions for equilibrium of the firm are formally the same as in Case 1. That is, there must be tangency of the (given) isoquant and the lowest possible isocost curve, and the isoquant must be convex. However, the problem is conceptually different in the case of cost minimisation. The entrepreneur wants to produce a given output (for example, a bridge, a building, or $\bar{X}$ tons of a commodity) with the minimum cost outlay.

In this case we have a single isoquant (figure 3.36) which denotes the desired level of output, but we have a set of isocost curves (figure 3.37). Curves closer to the origin show a lower total-cost outlay. The isocost lines are parallel because they are drawn on the assumption of constant prices of factors: since w and r do not change, all the isocost curves have the same slope w/r.

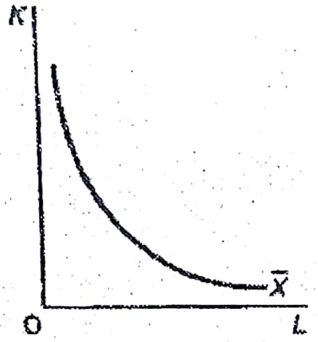


Figure 3.36

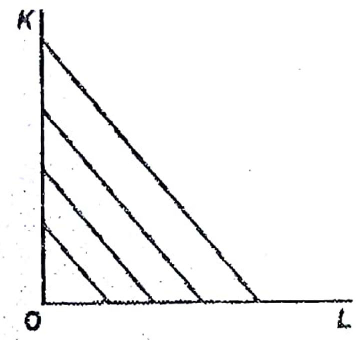


Figure 3.37

The firm minimises its costs by employing the combination of K and L determined by the point of tangency of the $\bar{X}$ isoquant with the lowest isocost line (figure 3.38). Points below e are desirable because they show lower cost but are not attainable for output $\bar{X}$. Points above e show higher costs. Hence point e is the least-cost point, the point denoting the least-cost combination of the factors K and L for producing $\bar{X}$.

Clearly the conditions for equilibrium (least cost) are the same as in Case 1, that is, equality of the slopes of the isoquant and the isocost curves, and convexity of the isoquant.

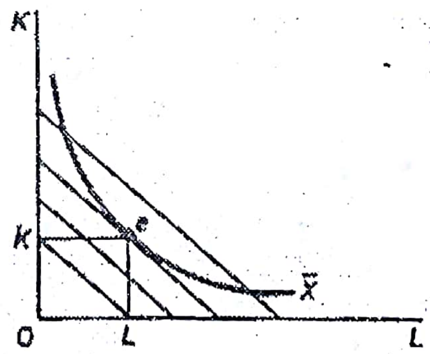


Figure 3.38

Formally:

Minimise $$C = f(X) = wL + rK$$

subject to $$\bar{X} = f(L, K)$$

Rewrite the constraint in the form

$$\bar{X} - f(L, K) = 0$$

Premultiply the constraint by the Lagrangian multiplier λ

$$\lambda(\bar{X} - f(L, K)) = 0$$

Form the 'composite' function

$$\phi = C - \lambda[\bar{X} - f(L, K)]$$

or

$$\phi = (wL + rK) - \lambda[\bar{X} - f(L, K)]$$

Take the partial derivatives of ϕ with respect to L, K and λ and equate to zero:

$$\frac{\partial \phi}{\partial L} = w - \lambda \frac{\partial f(L, K)}{\partial L} = 0 = w - \lambda \frac{\partial X}{\partial L}$$

$$\frac{\partial \phi}{K} = r - \lambda \frac{\partial f(L, K)}{\partial K} = 0 = r - \lambda \frac{\partial X}{\partial K}$$

$$\frac{\partial \phi}{\partial \lambda} = -[\bar{X} - f(L, K)] = 0$$

From the first two expressions we obtain

$$w = \lambda \frac{\partial X}{\partial L}$$

$$r = \lambda \frac{\partial X}{\partial K}$$

Dividing through these expressions we find

$$\frac{w}{r} = \frac{\partial X/\partial L}{\partial X/\partial K} = MRS_{L,K}$$

This condition is the same as in Case 1 above. The second (sufficient) condition, concerning the convexity of the isoquant, is fulfilled by the assumption of negative slopes of the marginal product of factors as in Case 1, that is

$$\frac{\partial^2 X}{\partial L^2} < 0, \quad \frac{\partial^2 X}{\partial K^2} < 0 \quad \text{and} \quad \left(\frac{\partial^2 X}{\partial L^2}\right)\left(\frac{\partial^2 X}{\partial K^2}\right) > \left(\frac{\partial^2 X}{\partial L\, \partial K}\right)^2$$

B. CHOICE OF OPTIMAL EXPANSION PATH

We distinguish two cases: expansion of output with all factors variable (the long run), and expansion of output with some factor(s) constant (the short run).

Optimal expansion path in the long run

In the long run all factors of production are variable. There is no limitation (technical or financial) to the expansion of output. The firm's objective is the choice of the optimal way of expanding its output, so as to maximise its profits. With given factor prices (w, r) and given production function, the optimal expansion path is determined by the points of tangency of successive isocost lines and successive isoquants.

If the production function is homogeneous the expansion path will be a straight line through the origin, whose slope (which determines the optimal K/L ratio) depends on the ratio of the factor prices. In figure 3.39 the optimal expansion path will be $0A$,

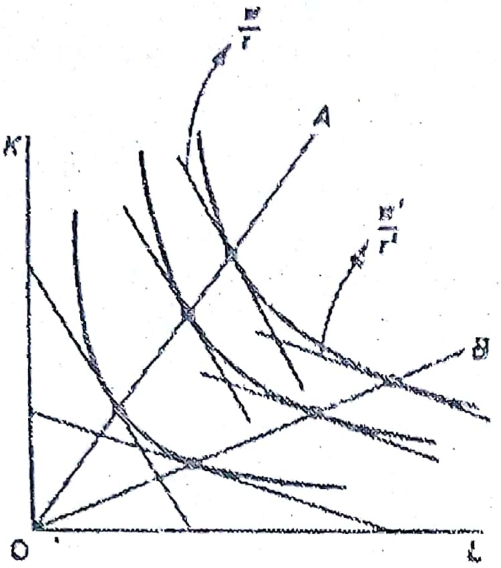


Figure 3.39

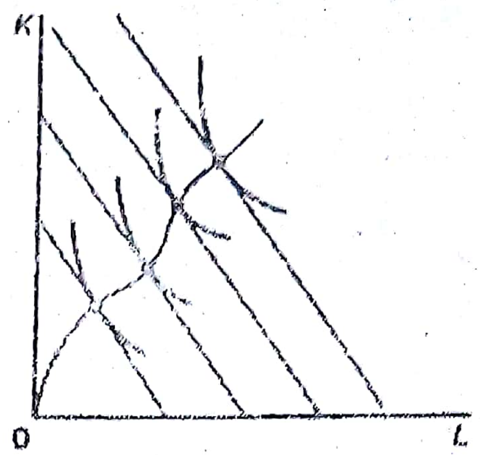


Figure 3.40

defined by the locus of points of tangency of the isoquants with successive parallel isocost lines with a slope of w/r. If the ratio of the prices increases the isocost lines become flatter (for example, with a slope of w'/r'), and the optimal expansion path will be the straight line $0B$. Of course, if the ratio of prices of factors was initially w/r and subsequently changes to w'/r', the expansion path changes: initially the firm moves along $0A$, but after the change in the factor prices it moves along $0B$.

If the production function is non-homogeneous the optimal expansion path will not be a straight line, even if the ratio of prices of factors remains constant. This is shown in figure 3.40. It is due to the fact that in equilibrium we must equate the (constant) w/r ratio with the $MRS_{L,K}$, which is the same on a curved isocline (see section II).

Optimal expansion path in the short run

In the short run, capital is constant and the firm is coerced to expand along a straight line parallel to the axis on which we measure the variable factor L. With the prices of factors constant the firm does not maximise its profits in the short run, due to the constraint of the given capital. This situation is shown in figure 3.41. The optimal expansion path would be $0A$ were it possible to increase K. Given the capital equipment, the firm can expand only along $\bar{K}\bar{K}$ in the short run.

The above discussion of the choice of optimal combination of the factors of production is schematically summarised on p. 94.

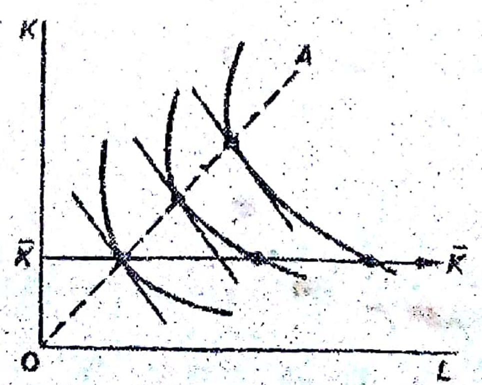


Figure 3.41