COMMUNICATIONS

On the Theory of Expansion of the Firm[1]

Economists who have spent time observing the operations of business enterprises come away impressed with the extent of management's occupation with growth. Expansion is a theme which (with some variations) is dinned into the ears of stockholders, is constantly reported in the financial pages and in the journals devoted to business affairs. Indeed, in talking to business executives one may easily come to believe that growth of the firm is the main preoccupation of top management. A stationary optimum would doubtless be abhorrent to the captains of industry, whose main concern is surely not at what size their enterprises should finally settle down (except where sheer size endangers their standing with the administrators of the antitrust laws) but rather, how rapidly to grow.[2]

Although the static theory of the firm is a helpful snapshot description of a system in motion,[3] it is useful also to have an alternative construction of the kind which is described in this paper—another equilibrium analysis in which the *rate of growth* of output, rather than its *level,* is the variable whose value is determined by optimality considerations.

I. *A Simple Growth Equilibrium Model*

For simplicity the first model is confined to a case in which input and output prices are fixed (pure competition), and where the production function is linear and homogeneous. Thus I am either dealing only with the period of time before the firm grows so large that the prices become variables which are

[1] This paper owes much to the growing literature of the dynamics of the firm. Particularly, I am indebted to Robin Marris for permitting me to read his unpublished manuscript [5] and to Herbert Frazer who wrote his doctoral dissertation on the subject. Highly relevant and stimulating is Edith Penrose [7]. In addition I owe much to the work of Richard E. Quandt [8, esp. pp. 156-66]. I am also very grateful for their comments to A. Heertje, Fritz Machlup, Burton Malkiel, Richard Quandt, Harold Shapiro, and John Williamson. Finally, I must express my great appreciation to the National Science Foundation whose grant helped materially in the completion of this manuscript.

[2] This view is not unrelated to one of Kaldor's well-known arguments. See N. Kaldor [3]. Here the author reminded us that equilibrium of the competitive firm requires some sort of increasing costs to make it unprofitable for the company to expand indefinitely. But under pure competition there seems to be no obvious source of diminishing returns, and hence little reason for *any* scale of operations of the competitive firm to constitute a long-run stationary equilibrium situation.

[3] Thus I am emphatically *not* proposing that the conventional theory of the firm be relegated to the garbage heap or the museum of curious antiquities. Static analysis of a nonstationary phenomenon can be immensely illuminating, and the received theory of the firm contains many very helpful results, both from the point of view of the understanding of the workings of the economy and the applied work of the operations researcher. It would be folly to deny ourselves the use of this body of analysis just because its domain of applicability is somewhat limited.

subject to the influence of the firm, or we must assume that all firms grow together and that in this process no one of them outgrows the others sufficiently to constitute it a significant force in the market. This premise permits me to evade the problem of demand for the expanding outputs of the firm. So long as it operates under conditions of pure competition its demand curves will be perfectly elastic and no marketing problems will affect its plans.[4]

It is assumed that management considers only a very simple growth pattern—a fixed percentage rate of growth, to be continued into the indefinite future. This heroic assumption is adopted to permit a simple characterization of the optimal growth path by means of a single variable, the permanent percentage rate of growth, g.[5]

Finally, it is assumed, at least for the moment, that the company's objective (which determines the optimal rate of growth of its output) is conventional profit maximization.

It is posited that costs can be divided into two categories: ordinary production and operating costs, and costs which arise only as a result of the expansion process. That is, any costs which would be associated with a given level of output if the output rate were not changing may be classed under output costs; any additional outlays above and beyond the output costs are called expansion costs. Output costs will only be taken into account implicitly, in the net revenue figures. That is, in discussing revenues, net revenue figures, from which output costs have already been deducted, will be employed.

Let R represent the initial net revenue of our firm, g be the rate of growth (which is to be determined), and i be the rate of interest relevant in discounting future revenues. Then, because of the constancy of the prices of all of the firm's inputs and outputs and the linear homogeneity of the production function, net revenues will grow precisely in the same proportion as inputs. Thus, in t periods, the firm's net revenue will be $R(1 + g)^t$, and the discounted present value of that net revenue will be $R[(1 + g)/(1 + i)]^t$. The present value of the expected stream of revenues will therefore be:

$$P = \sum_{t=0}^{\infty} R\left(\frac{1+g}{1+i}\right)^t = R\frac{1}{1-\dfrac{1+g}{1+i}} = R\frac{1+i}{i-g}, \tag{1}$$

[4] However, if all firms expand simultaneously in this way they may encounter secularly declining prices and problems of Keynesian excess supply. This macroeconomic problem is not discussed here since it merits being considered by itself in some considerable detail. I have elsewhere taken the optimistic position that if all firms expand rapidly enough they will usually create sufficient purchasing power to constitute a market for their products. No doubt many readers will question this hypothesis which appears to be a distant relative of the Say's Law family.

[5] If this premise is not employed and the optimal rate of growth at every future moment of time is left to be determined, we are forced in to the morass of the theory of functionals and we cannot escape without at least some recourse to the calculus of variations.

provided only[6] that $g < i$ so that $(1+g)/(1+i) < 1$ as is required for convergence of the geometric series (1).

It is perfectly obvious in this situation that we have

$$\frac{\partial P}{\partial g} > 0, \tag{2}$$

that is, the present value of the net revenue stream will grow indefinitely with the rate of expansion g. In fact, P will grow at an increasing rate with g, and its value will exceed any preassigned number as g approaches i, as shown in the net revenue curve, RR' in Figure 1. There is clearly nothing here to place a limit on the rate of expansion of the firm.

The firm will only be constrained from accelerating its activities without limit by its expansion costs, the present value of which we designate as $C(g)$. The literature is replete with discussions of the administrative costs of growth and there is no point in recapitulating these materials here. It is enough to point out that growth is what strains the firm's entrepreneurial resources and adds to the company's risks, and it may be expected that after some point the resulting increases in costs will catch up with the marginal revenues derived from more rapid expansion.[7] That is, it may be assumed that the slope of the cost curve CC', which is the graph of the function $C(g)$, will normally be less than that of RR' near the horizontal axis, but that eventually the slope of the former will catch up with and finally exceed that of the latter. (It may be, however, that in some cases the slope of the cost curve will ex-

[6] The problems caused for such a model if the rate of growth exceeds the rate of interest are well known. Specifically, the geometric series (1) will then not converge and the present value of the firm's profit stream will no longer be finite. See, e.g., David Durand [2]. However, as Miller and Modigliani have shown, the case $g > i$ is not a serious possibility. They write [6, fn. 14]: "Although the case of (perpetual) growth rates greater than the discount factor is the much-discussed 'growth stock paradox' . . . it has no real economic significance. . . . This will be apparent when one recalls that the discount rate, . . . though treated as a constant in partial equilibrium (relative price) analysis of the kind presented here, is actually a variable from the standpoint of the system as a whole. That is, if the assumption of finite value for all shares did not hold, because for some shares g was (perpetually) greater than i, then i would necessarily rise until an over-all equilibrium in the capital markets had been restored." (The notation has been changed from the original to that employed in this paper.)

An alternative way of avoiding this problem is to drop the (unrealistic) premise that the horizon is infinite. However, a finite horizon (say one involving 5 periods) will yield an expression for total revenue which is somewhat more messy than (1). Though it will be a fifth-degree polynomial, it will have only positive coefficients and so any equilibrium will still be unique. Indeed, the results of the infinite horizon model all seem to continue to hold in the finite horizon case.

[7] This view of the shape of the cost function can also be defended with the aid of the usual (somewhat shaky) appeal to the second-order maximum conditions. For, given the shape of our revenue function, the cost curve must behave in the manner shown in Figure 1 or there would be no profit-maximizing growth rate.

Note also that C is likely to be a function of other variables in addition to g, i.e., it is apt to depend on the initial absolute level of output—a small firm is likely to find it less costly to expand 10 per cent than does a large company. However, since $C(g)$ is the present value of all expected future costs taken together, initial cost differences may not play a very important role.

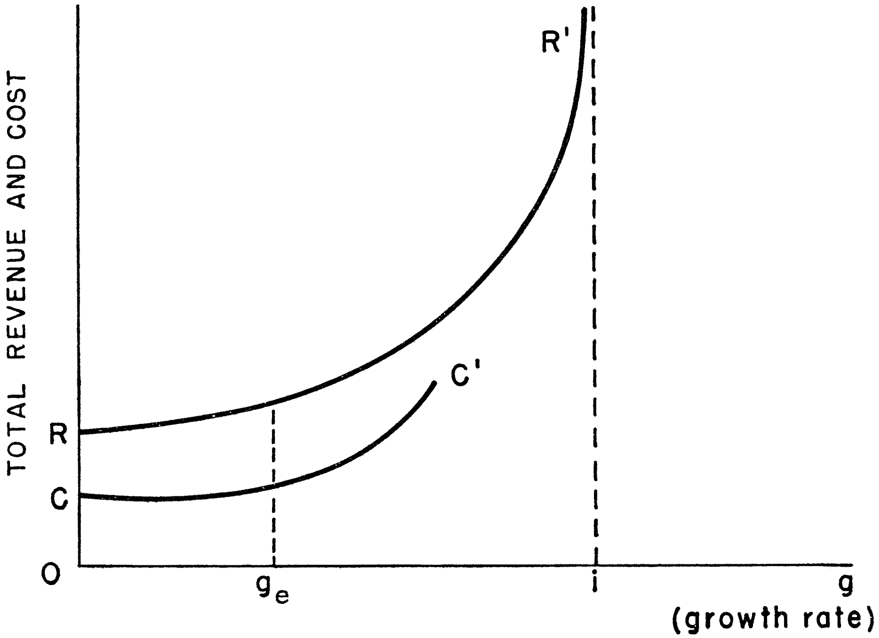

FIGURE 1

ceed that of the revenue curve throughout the positive quadrant so that the optimal growth rate will be zero negative.)

Specifically, we obtain the growth-profit function:

$$\Pi = P - C(g) = R\frac{1+i}{i-g} - C(g). \tag{3}$$

The profit-maximizing conditions are then (using the notation Π_g for $\partial\Pi/\partial_g$, etc.),

$$\Pi_g = P_g' - C'(g) = R\frac{1+i}{(i-g)^2} - C'(g) = 0, \tag{4}$$

(the first-order marginal revenue equals marginal cost condition), and

$$\Pi_{gg} = 2R\frac{1+i}{(i-g)^3} - C''(g) < 0, \tag{5}$$

the second-order condition.

Graphically, the equilibrium rate of growth is given by Og_e in Figure 1, the value of g at which the slope of the expansion cost curve CC' and that of RR' are equal.[8]

[8] We might even envision a long-run zero-profit competitive growth equilibrium in which entry has caused shifting of the curves and produced a zero-profit tangency position at which growth level has settled. There is some question in my mind whether, in a growth model such as this, much relevance can be ascribed to that type of long-run adjustment.

II. *Comparative Statics in the Analysis of the Model*

This simple growth model can easily be made to yield some results in terms of comparative statics. While some of these are not particularly surprising, they may offer some reassurance that the model does not possess particularly perverse properties, and that it can serve as an instrument of analysis much like the standard stationary equilibrium model.

First, a rise in the interest rate will reduce the present value of the stream of expected revenues, for we have[9] by (1)

$$P_i = R\frac{(i-g)-(1+i)}{(i-g)^2} = -R\frac{1+g}{(i-g)^2} < 0. \tag{6}$$

Moreover, a rise in the interest rate will reduce the *marginal* revenue yield of increased economic growth, P_g, for we have, differentiating (6) partially with respect to g,

$$P_{ig} = P_{gi} = -R\frac{(i-g)^2+2(i-g)(1+g)}{(i-g)^4} < 0 \tag{7}$$

by our basic assumption $g < i$.

It is now rather simple to prove that (at least in a perfect capital market where some market rate of interest determines the relevant discount factor) a rise in the interest rate will reduce the equilibrium rate of growth of the firm. For differentiating the first-order maximum condition (4) totally and setting $d\Pi_g = 0$ (so that the equilibrium condition continues to hold) we obtain:

$$d\Pi_g = P_{gi}di + \Pi_{gg}dg = 0$$

or

$$\frac{dg}{di} = -\frac{P_{gi}}{\Pi_{gg}} < 0 \tag{8}$$

by (5) and (7).

Geometrically, this obvious result is a consequence of the fact that a rise in i reduces the slope of the RR' curve in Figure 1 throughout its length, as indicated by (7), so that the equilibrium growth level, Og_e, must move to the left.

A somewhat more interesting application arises out of the recent proposals

[9] A complication is introduced by the fact that interest payments are among the output costs which have been subtracted from our net revenue figure, R, so that R should no longer be treated as a constant when differentiating with respect to i. This can be taken care of by noting that our assumptions of linear homogeneity and constant price imply that the quantity of money capital employed by the firm should be strictly proportionate with $R(1+g)^t$. Say it will equal $kR(1+g)^t$ and therefore incur an annual interest payment, $ikR(1+g)^t$. In that case we need merely write $R = R^*(1-ik)$ and make this substitution in our revenue function (1). It may then easily be verified by direct differentiation that the resulting expression for P_i will be slightly more complicated than (6) but that it will still be negative. A similar remark holds for (7) and (8).

to stimulate business growth by means of appropriate government subsidies.[10] Suppose one is considering two alternative subsidy plans for this purpose. The first plan involves payments (S_{1t}) proportionate with the percentage rate of increase of the firm's output growth,

$$S_{1t} = s_1 g_t = s_1 g.$$

The present value of all such expected future subsidy payments is

$$S_1 = s_1 g \sum_{t=0}^{\infty} \left(\frac{1}{1+i}\right)^t = s_1 g \left(1 + \frac{1}{i}\right). \tag{9}$$

The alternative plan proposes to offer a stream of subsidies (S_{2t}) proportionate to the absolute rate of increase of output:

$$S_{2t} = s_2[R(1+g)^t - R(1+g)^{t-1}] = s_2 g R(1+g)^{t-1}$$

whose capitalized present value is:

$$S_2 = \frac{s_2 g R}{1+g} \sum_{t=1}^{\infty} \left(\frac{1+g}{1+i}\right)^t = \frac{s_2 g R}{1+g} \left[\left\{\sum_{t=0}^{\infty} \left(\frac{1+g}{1+i}\right)^t\right\} - 1\right]$$

$$= \frac{s_2 g R}{1+g} \left[\frac{1+i}{i-g} - 1\right] = \frac{s_2 g R}{1+g} \cdot \frac{1+g}{i-g}$$

or

$$S_2 = \frac{s_2 g R}{i-g} \,. \tag{10}$$

Adding, in turn, the subsidy expressions (9) and (10) to our basic profit function (3) we obtain the two new profit expressions:

$$\Pi^* = R \frac{1+i}{i-g} - C(g) + s_1 g \left(1 + \frac{1}{i}\right) \tag{11}$$

and

$$\Pi^{**} = R \frac{1+i}{i-g} - C(g) + \frac{s_2 g R}{i-g} \,. \tag{12}$$

By the same procedure as was employed when the effect of an interest rate change was examined, we arrive at the respective comparative statics results:

$$\frac{dg}{ds_1} = -\frac{1 + 1/i}{\Pi^*_{gg}} > 0 \tag{13}$$

and

$$\frac{dg}{ds_2} = -\frac{iR}{\Pi^{**}_{gg}(i-g)^2} > 0 \tag{14}$$

[10] For example, in the Kennedy administration 1961 investment credit proposals. See also Klaus Knorr and W. J. Baumol [4], whose suggestion seems, in substance, to have been adopted by the Canadian Government (House of Commons, *Proceedings*, Apr. 10, 1962, pp. 2706-7).

if the second-order conditions $\Pi_{gg}^* < 0$ and $\Pi_{gg}^{**} < 0$ both hold. Thus both types of subsidy would, indeed, stimulate the growth of the profit-maximizing firm.

We can go beyond this somewhat uninteresting conclusion by asking which of these two types of subsidy will yield more growth per dollar of government outlay. For this purpose we must deal not with s_1 and s_2, the subsidy rates, but with the total subsidy outlays, S_1 and S_2, as given by (9) and (10). From these we obtain:

$$\frac{ds_1}{dS_1} = \frac{1}{g(1 + 1/i)} \tag{15}$$

and

$$\frac{ds_2}{dS_2} = \frac{i - g}{Rg}. \tag{16}$$

Multiplying (13) by (15) and (14) by (16) and writing out the expressions for Π_g^* and Π_{gg}^{**} we obtain:

$$\frac{dg}{dS_1} = -\frac{1}{g\Pi_{gg}^*} = -\frac{1}{g\left[2R\dfrac{1+i}{(i-g)^3} - C''(g)\right]} \tag{17}$$

and

$$\begin{aligned}\frac{dg}{dS_2} &= -\frac{i}{g(i-g)\Pi_{gg}^{**}} \\ &= -\frac{i}{g\left[2R\dfrac{1+i}{(i-g)^3} - C''(g) + \dfrac{2is_2R}{(i-g)^3}\right](i-g)}.\end{aligned} \tag{18}$$

Hence subsidy S_2 will yield higher marginal returns than subsidy S_1 if and only if expression (18) exceeds expression (17), i.e., if and only if

$$-\frac{1}{g\Pi_{gg}^*} < -\frac{i}{g(i-g)\left[\Pi_{gg}^* + \dfrac{2is_2R}{(i-g)^3}\right]}.$$

This requires

$$(i-g)\left[\Pi_{gg}^* + \frac{2is_2R}{(i-g)^3}\right] > i\Pi_{gg}^*$$

or

$$-g\Pi_{gg}^* + (i-g)\frac{2is_2R}{(i-g)^3} > 0$$

and since (because $\Pi^*_{gg} < 0$ by the second-order condition) both terms in this last expression are positive, this requirement will always be satisfied. We conclude that in our model a subsidy of the second type will then always yield higher marginal growth returns than does a subsidy of the first type.

It is also noteworthy that a net investment tax credit of the sort originally proposed is essentially equivalent in our model to a growth subsidy proposal of type two. For the investment credit is a subsidy proportionate to the level of net investment. With our linear homogeneous production function, and with constant input prices, the capital-output ratio will be constant so that a subsidy proportionate to investment will automatically be proportionate to the absolute rate of increase in output.

Since so many other considerations must enter any decision among alternative growth stimulation methods there is no point in laboring this discussion further. The case serves, however, to illustrate how meaningful theorems can be derived from the growth equilibrium model of the firm.

III. *Profit versus Growth Maximization*

The discussion so far has been confined to the case of pure competition and has assumed that the firm's objective is to maximize profit. But larger *oligopolistic* firms may well have a different set of objectives.[11] Specifically, I have suggested that management's goal may well be to maximize "sales" (total revenue) subject to a profit constraint. Though I remain firmly convinced of the merit of the hypothesis as a static characterization of the current facts of oligopolistic business operation, in the present context—a growth equilibrium analysis—it is desirable to modify the hypothesis in two respects.

First, maximization of *rate of growth* of sales revenue seems a somewhat better approximation to the goals of many management groups in large firms than is maximization of the current *level* of sales. For example, most company publicity materials seem to emphasize the extent to which the firm has "progressed" rather than the sheer magnitude of its current operations. In my earlier static model I was forced to employ a sales-revenue-level objective as an approximation to a measure of the rate of growth of the firm's scale of operations. A growth equilibrium model now frees me from this necessity.

The second modification deals with the nature of the profit constraint, which in a static model may have seemed to be arbitrarily imposed from the outside—perhaps even a device to avoid explaining what had to be explained, very much like the fixed mark-up of doubtful origin which lies at the heart of the full-cost pricing discussions. A growth analysis enables me to give an explanation of the profit constraint which, I hope, is somewhat less superficial and rather more convincing.

From the point of view of a long-run growth (or sales) maximizer, profit no longer acts as a constraint. Rather, it is an intrumental variable—a means whereby management works towards its goals. Specifically, profits are a means for obtaining capital needed to finance expansion plans. Capital is raised both by direct retention of profits and by the payment of dividends

[11] See [1], esp. Ch. 6-8.

to induce outside investors to provide funds to the company. But, beyond some point, profits compete with sales. For the lower prices and higher marketing outlays which are necessary to promote sales also cut into net earnings. Hence, too high a level of profits will reduce the magnitude of the firm's current operations, while too low a profit level will prevent future growth. The optimal profit stream will be that intermediate stream which is consistent with the largest flow of output (or rate of growth of output) over the firm's lifetime.

Specifically, this optimal profit rate can be described with the aid of a simple model such as the following:[12]

Let

g represent our firm's growth rate,

I be its level of investment as a per cent of the value of current capital assets (the percentage rate of growth of the firm's money capital),

Π be the profit rate as a per cent of present equity[13]

D be the dividend as a per cent of present equity, and

E be the retained earnings as a per cent of present equity per unit of time.

The objective then is to maximize:

$$g = f(I, \Pi)$$

subject to[14]

$$I = \phi(\Pi, D) + E$$
$$\Pi \equiv D + E.$$

The first of these equations, the objective function, expresses the competitive relationship between growth and profit rates, and states that the rate of growth of the firm's operations varies (directly) with investment, and (after a point) inversely with the profit rate (as indicated in Figure 1). The next equation, however, shows that the profit rate indirectly assists growth by providing capital through retained earnings, and by attracting funds from outside sources at a rate, ϕ (II, D), which depends both on the dividend rate and the company's profit rate. From this system we can then determine the optimal profit rate, Π, which from our long-run point of view enters into the

[12] For present purposes there is no need to take explicit account of such decision variables as prices, advertising outlay, etc.; but the model can easily be expanded to do so.

[13] In practice, of course, different profit rates may be optimal at different points in the company's history. But in the fixed-price constant-returns-to-scale model which is employed here there is no reason to depart from a single optimal profit level.

[14] Perhaps, in accord with the Miller-Modigliani view [6] that dividends do not matter, D should be omitted from the Φ function. Other possible variables that have been suggested for inclusion as elements which significantly affect the willingness of the public to supply funds to the firm are $d\Pi/dt$—the rate of growth of the firm's profit rate, and g, the rate of growth of its output.

It has been suggested that other, partly conventional, constraints are imposed by the capital market and should be incorporated in a more elaborate version of the model. These include restrictions on the debt-equity ratio, on the ratio between current assets and sales, and on the extent of reliance on noninternal financing.

constraints just as one of the variables in the system. Only in a static sales-maximization model, then, does profit appear as an independent datum arbitrarily given from the outside—a fixed minimal profit requirement which has somehow to be met by the firm.

Substantive theorems for a (sales) growth maximization model may be developed which contrast its consequences with those of profit maximization.[15] However, these propositions are completely analogous with those which I have already developed elsewhere for the case of sales maximization. For example, the growth maximizer's sales, advertising outlay, and (trivially) his growth rate will be larger than those of the profit maximizer and the pricing and output decisions of only the former may be expected to vary in response to changes in fixed costs. Since the logic of these results in our present analysis is exactly the same as it was in the sales maximization model there is no point in repeating the argument here.

I will only suggest what appears to be the most important point, that our discussion has shown the standard apparatus of marginal analysis and mathematical programming to be fully applicable to decision problems even when management's objective is not the venerable profit maximization of economic theory.

William J. Baumol*

References

1. W. J. Baumol, *Business Behavior, Value and Growth.* New York 1959.
2. D. Durand, "Growth Stocks and the Petersburgh Paradox," *Jour. Finance,* Sept. 1957, *12,* 348-63.
3. N. Kaldor, "The Equilibrium of the Firm," *Econ. Jour.,* Mar. 1934, *44,* 60-76; reprinted in N. Kaldor, *Essays in Value and Distribution.* Glencoe 1961.
4. K. Knorr and W. J. Baumol (ed.), *What Price Economic Growth?* Englewood Cliffs 1961.
5. R. Marris, *The Micro Economics of Managerial Capitalism.* Unpublished manuscript.
6. M. H. Miller and F. Modigliani, "Dividend Policy, Growth, and the Valuation of Shares," *Jour. Bus.,* Oct. 1961, *34,* 411-33.
7. Edith Penrose, *The Theory of the Growth of the Firm.* London 1959.
8. R. E. Quandt, "Effects of the Tax-Subsidy Plan," Appendix B in K. Knorr and W. J. Baumol [4].

[15] It is sometimes stated or implied that long-run growth and profit maximization must necessarily lead to identical decisions and results (see e.g., Penrose [7, p. 29]). But if, as in our model, it is sales rather than assets whose growth is being maximized, or if, even in the long run, investment in the firm can fall short of or exceed profit earnings, it is extremely easy to find counterexamples. In fact, only in the most unusual circumstances would sales (revenue) growth maximization be achieved by the maximization of profits.

* The author is professor of economics at Princeton University.