



Adjustment Costs and the Theory of Supply

Robert E. Lucas, Jr.

The Journal of Political Economy, Volume 75, Issue 4, Part 1 (Aug., 1967), 321-334.

Stable URL:

<http://links.jstor.org/sici?sici=0022-3808%28196708%2975%3A4%3C321%3AACATTO%3E2.0.CO%3B2-3>

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The Journal of Political Economy is published by The University of Chicago Press. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/ucpress.html>.

The Journal of Political Economy
©1967 The University of Chicago Press

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR

THE JOURNAL OF POLITICAL ECONOMY

Volume 75

AUGUST 1967, PART I

Number 4

ADJUSTMENT COSTS AND THE THEORY OF SUPPLY

ROBERT E. LUCAS, JR.*
Carnegie Institute of Technology

I. INTRODUCTION

THIS paper contains a re-examination from a dynamic point of view of some familiar problems in the theory of supply. In particular, it is concerned with the response of firms in a competitive industry to once-and-for-all shifts in the industry demand curve, to changes in the *rate of shift* in this demand curve, and to changes in the prices of factors of production. In most respects, our discussion of these issues will proceed along the lines of the traditional Marshall-Viner theory of supply. The novelty of our treatment will lie in its attempt to introduce the relative "fixity" of capital explicitly into the formulation of the firm's maximum problem and to use this formulation to obtain a precise definition of the industry's "short-run" and "long-run" equilibrium posi-

tions and an analysis of the passage from one to the other.¹

Aside from its theoretical interest, such a revision of the theory of a competitive industry appears to be necessitated by recent empirical work in industrial organization. Traditional supply theory affords a full explanation of firm and industry behavior only for the case in which each firm's long-run average cost curve is "U-shaped." In this case, the theory predicts that profit-maximizing behavior on the part of both existing and potential firms will lead firms to operate at the minimum point of this U-shaped curve. This prediction is certainly at variance with the casual observation that in most

¹ The principal result concerning the relation of the long run to the short run is the famous "envelope theorem" relating long- and short-run average cost curves, each of which is obtained from an appropriate cost minimization problem (see Samuelson, 1947, p. 34). Aside from the fact that this theorem sheds no light on the way in which fixed inputs become "unfixed" through time, there is no reason to believe the equilibrium indicated by the solution to a "long-run cost minimization problem" will coincide with the equilibrium which arises from a succession of a "short-run" decisions. This point has been well developed in a different context in Phelps (1961, pp. 638-43) and Pearce (1962, pp. 1088-97).

* Some of the material in this paper is taken from the author's "Constant Returns to Scale and Measurement of Technological Change," given at the March, 1966, conference of the Inter-University Research Program on the Micro-Economics of Technological Change and Economic Growth. I would like to thank Sidney Winter for his helpful comments on the earlier paper.

industries the bulk of demand increases is met by existing firms rather than by new entrants. More systematic investigation is similarly unfavorable to the notion of an optimal firm size. The studies of Stigler, Simon and Bonini, Mansfield, and others,² while differing widely in conception and execution, all indicate that, with the exception of very small firms, the percentage rate of firm growth is roughly independent of its size. All of these authors have interpreted these findings as evidence in favor of constant returns to scale over a wide range of firm sizes.³

This evidence cannot be regarded as refuting traditional supply theory, since this theory nominally treats all conceivable sorts of average cost curves, but it must certainly call into question the empirical usefulness of this theory. With constant or increasing returns to scale, one can draw short- and long-run cost curves and discuss, in loose terms, long-run *tendencies* for the output of individual firms, but a unique long-run equilibrium cannot be exhibited on this diagram nor can it be obtained analytically by solving a "long-run maximum problem." Thus except for the case of "U-shaped cost curves," the traditional theory yields no unambiguous testable implications. One may also "reconcile" the industrial organizational evidence with U-shaped cost curves by supposing that these curves differ across firms in an industry and that they shift unpredictably over time, but such a reconciliation

² See Simon and Bonini (1958); Stigler (1958); Mansfield (1962). The cost curve evidence surveyed and augmented in Johnson (1960), while more difficult to interpret, also appears to support the constant-returns hypothesis for all but very small firms.

³ Stephen Hymer and Peter Pashigian (1962) interpreted data similar to Mansfield's as evidence for *increasing* returns to scale. Their conclusion is disputed in Simon (1964).

would only rob the hypothesis of whatever content it now has.

II. PRODUCTION POSSIBILITIES WITH "FIXED" INPUTS

In order to develop a supply theory in which adjustments to demand shifts are staggered over time, it is first necessary to define the terms "fixed" and "variable" inputs and to determine the sorts of production possibilities or market opportunities under which a particular input will fall into one or the other of these two categories. In this section, after a discussion of notational matters, we shall examine a fairly recent view of the firm in which *all* inputs are variable and then discuss a modification of this theory which introduces "fixity" of one of the inputs.

Consider a firm whose net receipts at time t are given by sales less labor costs less investment goods purchases, or:

$$R(t) = p(t)Q(t) - w(t)L(t) - v(t)I(t),$$

where $Q(t)$, $L(t)$, $I(t)$ denote output, labor input, and gross investment; and $p(t)$, $w(t)$, and $v(t)$ are the respective prices of these goods. If the firm finances entirely by borrowing at a fixed rate r , its present value at time 0 is:

$$V(0) = \int_0^{\infty} e^{-rt} R(t) dt.$$

We shall assume that the firm selects $Q(t)$, $L(t)$, and $I(t)$, for $t \geq 0$, so as to maximize $V(0)$ subject to a *production possibility constraint* relating these three flows, or time paths, and given suitable assumptions on *market opportunities* indicating how the three prices will vary over time. For the latter, it is assumed throughout this paper that in drawing up a plan at time 0, the firm regards these prices as beyond its control, and, further, it behaves as though the present

price levels will be maintained forever.⁴ It remains then only to characterize the firm's production possibilities.

The most familiar way to complete this model is to define the firm's capital stock by the declining balance formula:⁵

$$\frac{dK(t)}{dt} = I(t) - \delta K(t), \quad (1)$$

and an initial stock $K(0)$, and to assume that output is produced according to a production function:

$$Q(t) = F[L(t), K(t)], \quad (2)$$

which relates output to current labor and capital inputs. Under these assumptions, it is known that if a rental price or user cost of capital is defined by $v(r + \delta)$, then present value is maximized by the policy which maximizes:

$$pF[L(t), K(t)] - wL(t) - v(r + \delta)K(t)$$

at each point in time, $t > 0$.⁶ With constant prices, this rule determines an optimal stock K^* which will be held for all $t > 0$, and this stock will equal the historically given $K(0)$ only by coincidence. If we define a variable input as one whose optimal level at $t > 0$ is inde-

⁴ An alternative to this assumption of "static expectations" which is more attractive from some points of view is the hypothesis that firms' price expectations are rational (as defined by John F. Muth, 1961) or that future prices are correctly anticipated. I have developed some of the implications of this alternative hypothesis in "Optimal Investment with Rational Expectations," unpublished Carnegie Tech research memorandum, 1966. If expectations are taken to be rational, however, one cannot in general obtain the supply and factor demand functions for the individual firms in an industry (although one *can* characterize the behavior of industry output and price through time) so that some of the parallels between traditional, static theory and the theory developed here cannot be exhibited.

⁵ If capital is regarded as an observable magnitude, equation (1) is an assumption rather than a definition.

⁶ This result is given in Haavelmo (1960).

pendent of its level at $t = 0$, it is thus clear that capital and labor are, in this theory, both variable inputs, regardless of the institutional fact that one is owned by the firm and the other is hired. There is then *no* distinction between short- the long-run supply behavior: following a once-and-for-all change in prices, the firm proceeds *immediately* to its new long-run position.

If one wishes to build the "fixity" of capital explicitly into the formulation, one must introduce costs per unit of gross investment which *rise* with the investment rate, either by postulating monopsonistic capital goods markets (letting v be a rising function of I) or by introducing internal costs of investment in the form of output foregone. In this paper, capital fixity will be treated as arising from such internal costs,⁷ with the production function (2) replaced by:

$$Q(t) = F[L(t), K(t), I(t)]. \quad (3)$$

The investment rate is assumed to enter F with a negative and decreasing marginal productivity. For fixed $I(t)$, capital and labor have positive and diminishing marginal productivities.

The inclusion of the gross investment

⁷ There are a number of alternative ways to introduce capital fixity into the firm's maximum problem, ranging from inequality constraints on the rate of investment to various types of "smooth" adjustment cost hypotheses. The formulation used here is adopted from Eisner and Strotz (1963). One can also motivate the distinction between short- and long-run firm behavior by reference to more specific aspects of the production process, as Armen Alchian (1959) does with his distinction between the rate of output (of a particular process within a firm) and the scheduled volume of output. The virtue of an approach such as Alchian's is that by being specific about the source of the short-run-long-run distinction one is able to get some "feel" for the shapes of the functions involved. The difficulty with this approach, as opposed to the one used here, is that the resulting theory cannot be applied to data at the balance sheet and income statement aggregation level or at higher levels of aggregation.

rate in the production function may be motivated by some examples. First, consider a firm composed of a production division and a planning division, each of which uses the same inputs, $L(t)$ and $K(t)$. The production division produces an output $Q(t)$ according to an ordinary production function, while the activity of the planning division is proportional to the firm's investment rate (that is, a unit of investment requires a fixed amount of planning). Suppose that both divisions operate under decreasing returns to scale. Then an increase in investment for *fixed* levels of L and K requires a withdrawal of labor and capital from the production division, reducing output $Q(t)$. Because of decreasing returns, additional increases in investment require successively larger output sacrifices. (If the firm "hires" planning services from the outside, at rising unit costs, one would observe a similar effect.)

As a second example, suppose that the introduction of new capital goods introduces new production methods and that new capital becomes fully effective only after a learning period. In particular, suppose that current investment goods yield a fraction k , $0 < k < 1$, of the productive services they will ultimately provide. Then from (1), measured capital stock at time t is

$$K(t) = K(0)e^{-\delta t} + \int_0^t e^{-\delta(t-s)} I(s) ds,$$

while "effective" capital is $K^e(t) = K(t) - (1 - k)I(t)$. Inserting $K^e(t)$ in place of $K(t)$ in the static production function (2) yields a special instance of (3).

In either of these examples, one might argue that only *net* investment should enter into F , since replacement involves neither planning nor learning. This argument has some force if one thinks of re-

placement as typified, say, by changing a light bulb. It has less appeal if one thinks of replacing a computer with a newer model. In any case, the choice between gross and net investment is not crucial to any of the arguments which follow.

One could proceed immediately to obtain necessary conditions for the maximization of $V(0)$ under production possibilities given by (1) and (3), but without further restrictions on the production function F , it will be difficult to proceed beyond an empty statement of marginal equalities. In addition, the conditions so derived will not in general be consistent with the independence of firm growth and size. Accordingly, we shall add the assumption that $F(L, K, I)$ is *homogeneous of degree one* in its three arguments. It will be noted that this restriction differs from constant returns to scale in the usual sense, since a doubling of capital and labor *and* the investment rate will double output, but a doubling of capital and labor with a fixed investment rate will yield more than a doubled output.

III. THE PRESENT VALUE-MAXIMIZATION PROBLEM FOR A SINGLE FIRM

In the preceding section, we defined production possibilities of a particular type (equations [1] and [3]) and asserted that a single firm with these production possibilities, operating in competitive product and factor markets, would regard capital as a fixed input in the sense that its optimal capital stock for at least some positive t would depend on its initial stock $K(0)$. In this section, we shall work through in detail the present value-maximization problem for such a firm. In the course of this discussion, the assertion that capital is (partially) fixed under (1) and (3) will be proved, and the dependence of the firm's output supply,

labor demand, and investment demand on prices, the interest rate, and the initial capital stock will be examined. In other words, we shall derive the short- and long-run supply and factor demand functions for the firm.

The maximum problem of the firm has been stated in piecemeal fashion in the preceding section. It will be helpful here to collect these assumptions. The firm's objective at time 0 is to choose continuous investment and labor force plans, $I(t)$ and $L(t)$, $t \geq 0$, so as to maximize:

$$V(0) = \int_0^\infty e^{-rt} \{ pF[L(t), K(t), I(t)] - wL(t) - vI(t) \} dt,$$

where r , p , w , and v are current ($t = 0$) prices and where F is homogeneous of degree one with continuous first and second derivatives satisfying:

$$\begin{aligned} F_L > 0, \quad F_K > 0, \quad F_I < 0, \\ F_{LL} < 0, \quad F_{KK} < 0, \quad F_{II} < 0. \end{aligned} \tag{4}$$

To simplify the analysis somewhat, it will be assumed that F can be written as a sum of an ordinary production function, with arguments L and K , and an internal "adjustment cost function," with arguments I and K , or that:

$$F_{LI} = 0. \tag{5}$$

Next, F is required to be concave in the following sense: let (L_0, K_0, I_0) and (L_1, K_1, I_1) be any two input combinations, let $L_\theta = \theta L_0 + (1 - \theta)L_1$, and define K_θ and I_θ similarly. Assume that:

$$\begin{aligned} F(L_\theta, K_\theta, I_\theta) &\geq \theta F(L_0, K_0, I_0) \\ &+ (1 - \theta)F(L_1, K_1, I_1), \end{aligned} \tag{6}$$

with equality if and only if (L_0, K_0, I_0) and (L_1, K_1, I_1) are proportional.

Finally, the derivatives of F are required to satisfy:

$$F_L(I, K) \rightarrow 0 \text{ as } L/K \rightarrow \infty;$$

$$F_L(L, K) \rightarrow \infty \text{ as } L/K \rightarrow 0; \tag{7}$$

$$F_I(K, I) \rightarrow -\infty \text{ as } I/K \rightarrow r + \delta - \alpha,$$

where α is some positive number.

Under these conditions, it is shown in the Appendix to this paper that a plan $[L(t), I(t)]$ with $L(t)$ and $I(t)$ strictly positive and $0 \leq I(t)/K(t) < r + \delta - \alpha$, for all t , is optimal if and only if it satisfies:

$$pF_L(L, K) = w, \tag{8}$$

$$pF_K(L, K, I) = [v - pF_I(K, I)](r + \delta), \tag{9}$$

in addition to the depreciation formula (1).

Equation (8) is simply the usual marginal productivity condition for labor. The left side of (9) is the value of the marginal product of capital. The right side of (9) may be termed the "marginal user cost" of capital and may be interpreted by analogy to the expression $v(r + \delta)$ for user cost to which pF_K is equated when capital is variable. In either case, the term $r + \delta$ converts a stock price to a flow price. The term $v - pF_I$ is the marginal cost of accumulating capital, with v measuring cash outlay per unit of investment goods, and $-pF_I$ measuring the value of output foregone with each unit of investment.

Next we derive from (1), (8), and (9) the firm's short-run labor demand function, its short-run supply function, its investment demand function, and its long-run growth plan. Since F is homogeneous of degree one, its first derivatives can be treated as functions of the ratios L/K and I/K . From (4) and (7), (8) can be solved for a positive value of L/K for any price ratio w/p . The short-run demand curve for labor is then:

$$L(t) = K(t)D_1(w/p), \quad D_1'(w/p) < 0. \tag{10}$$

A change in w/p thus induces an instantaneous movement of L along this curve; as capital is adjusted, this demand curve will shift proportionally.

To determine the investment demand function, re-arrange (9), apply Euler's Theorem to F , and use (10) to eliminate L/K to obtain:

$$p \left\{ F \left[1, D_1 \left(\frac{w}{p} \right), \frac{I}{K} \right] - D_1 \left(\frac{w}{p} \right) \frac{w}{p} + \left(r + \delta - \frac{I}{K} \right) F_I \left(\frac{I}{K} \right) \right\} = v(r + \delta). \quad (11)$$

where $\partial D_2 / \partial (w/p) < 0$, $\partial D_2 / \partial (v/p) < 0$, and $\partial D_2 / \partial r = \partial D_2 / \partial \delta < 0$.

Inserting these solutions for L/K and I/K into the production function gives the short-run supply function of the firm:

$$Q(t) = K(t) F \left[1, D_1 \left(\frac{w}{p} \right), D_2 \left(\frac{w}{p}, \frac{v}{p}, r, \delta \right) \right]. \quad (13)$$

Output will be a rising function of r, δ ,

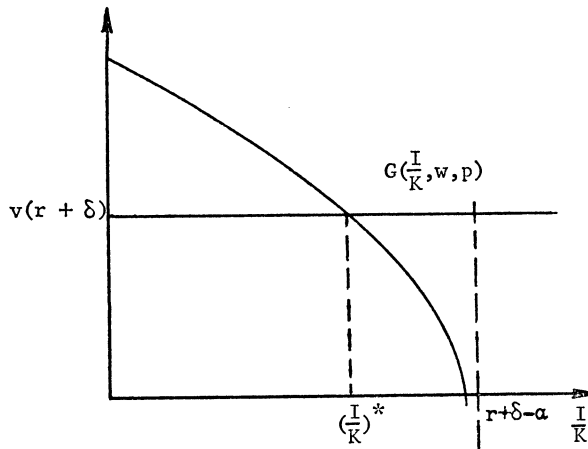


FIG. 1

The left side of (11), which we shall call $G(I/K, w, p)$ is drawn in Figure 1. The downward slope of G , as a function of I/K , follows from (4). The fact that G must cross the I/K axis to the left of the point $r + \delta - \alpha$ follows from (7). It may also be verified that $\partial G / \partial w < 0$ and $\partial G / \partial p > 0$. For any fixed w , there will be some critical price p_c (say), such that if $p > p_c$, (11) yields a positive solution for I/K . As indicated earlier, this analysis will consider *only* prices above this critical level. Then if $p > p_c$, (11) may be solved for the investment demand function:

$$I(t) = K(t) D_2 \left(\frac{w}{p}, \frac{v}{p}, r, \delta \right), \quad (12)$$

and q/p , but the effect on $Q(t)$ of changes in w/p is indeterminate. This leads to the odd possibility that the slope of the short-run supply curve (the derivative of Q with respect to p) may have either sign. This result—which is of course at variance with the implication of conventional theory that the short-run supply curve must slope upward—arises from the presence of investment in the production function and from the role that price plays as a perfect indicator of future price. The initial effect of a price increase is to induce the firm to expand its labor force to meet current demand and *also* to increase the investment rate so as to be prepared to meet future de-

mand. These two responses affect current output in opposite directions. While one might strain for examples in which the negative effect outweighs the positive one, thus making a downward sloping short-run supply curve "reasonable," I think it is probably best simply to rule this case out by assumption. In what follows, then it is assumed that $\partial Q/\partial p > 0$.

Equations (10), (12), and (13) give the firm's response to any once-and-for-all price or interest rate change for any given capital stock. If the price change occurs at $t = 0$, $K(0)$ is historically fixed, and these equations give the initial or short-run response. To obtain the long-run response, assuming the new prices to be maintained over time, we need only to know the time path of $K(t)$. Combining (1) and (12):

$$\frac{\dot{K}(t)}{K(t)} = D_2 \left(r, \delta, \frac{w}{p}, \frac{q}{p} \right) - \delta. \quad (14)$$

Thus for constant prices, capital stock will grow (or decline) at a constant percentage rate, and this rate will be independent of size (as measured by assets). From (10), (12), and (13) it is evident that gross investment, labor force, and output will all grow at the rate $D_2 - \delta$ as well.

IV. THE ORGANIZATION OF A COMPETITIVE INDUSTRY

In the last section, we determined the optimal growth plan for a single firm with a particular type of production possibilities. In this section, we examine an industry composed of a large number of such firms, each with identical production functions as given by (3) and each faced with the same set of prices p , w , q , and r , but with possibly different capital stocks.

Before deriving the supply curve of the industry, it is important to inquire

whether the composition of the industry will remain the same in the face of shifting demand or, in particular, to raise the issues of merger and dissolution of firms and of entry and departure. A natural way to deal with both these questions is to examine the way in which present value under an optimal policy will vary with initial capital stock. Retaining the assumption of constant prices, (7) and (8) imply that the optimal ratios of labor to capital and gross investment to capital remain constant over time. Then if $R^*(t)$ and $K^*(t)$ denote receipts and capital stock at t under the optimal policy, the ratio R^*/K^* will remain constant, with both R^* and K^* expanding at the percentage rate $D_2 - \delta$, as given in (14). Then present value under the optimal policy is:

$$\begin{aligned} V^* [K(0)] \\ = (r + \delta - D_2)^{-1} \frac{R^*(t)}{K^*(t)} K(0). \end{aligned} \quad (15)$$

Thus if each firm is selecting inputs optimally, the present value of *all* of the firms in the industry will be independent of how the industry's capital stock is distributed across firms.⁸ There is, therefore, no incentive for firms to merge or dissolve, regardless of the initial asset distribution.

From (15) it is evident that the value of a firm with *no* capital is zero, so that mere membership in the industry has no value. From this, one is tempted to conclude that there will be no incentive for new firms to enter the industry and begin accumulating capital "from scratch." Unfortunately, the entry problem cannot be handled quite so neatly for rea-

⁸ The proportionality of present value to initial capital stock as in (15) is not generally true with adjustment costs. In fact, with depreciation as assumed in (1), it will hold if and only if the function F is homogeneous of degree one. This is shown in the paper cited in note (*).

sons which will be clearer after industry growth in the absence of entry is discussed.

Assume an industry-wide demand function of the form:

$$Q(t)e^{-\lambda t} = z(t) = h[p(t)], \quad (16)$$

where h is required to satisfy:

$$h'(p) < 0, \quad h(0) = \infty, \quad (17)$$

$$\text{and } h(\infty) = 0.$$

intersection of the short-run supply function (13) and the demand curve (16).⁹

The behavior of price and industry output over time can now be handled geometrically with a slightly modified supply-and-demand diagram (Fig. 2).¹⁰ The downward sloping curve in Figure 2 is obtained from (16) and (17): it is the relative (to the "size of the market") demand curve facing the industry, with price on the vertical axis and $z = Qe^{-\lambda t}$

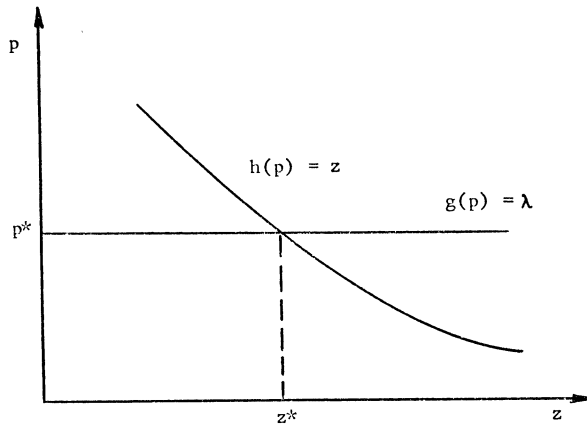


FIG. 2

If one thinks of GNP growth as the principal factor shifting the demand curve, the growth of demand λ can be thought of as the product of the growth rate of income times the income elasticity of demand for the industry's product.

The growth rate of output supplied is given by the right side of (14) in the previous section. For fixed levels of r , w , q , and δ , this rate may be written $D_2(r, w/p, q/p, \delta) = g(p)$, where $g'(p) > 0$ (see Fig. 1). Then assuming the product market to be cleared at each point in time, the industry supply function is:

$$\frac{\dot{z}(t)}{z(t)} = g[p(t)] - \lambda. \quad (18)$$

The initial value $z(0)$ is given by the

on the horizontal axis. The horizontal line on Figure 2 gives that value of p which satisfies $g(p) = \lambda$, for fixed $\lambda \geq 0$. To draw this line above the z axis requires the assumption that a positive price is needed to induce firms to grow at a non-negative rate. Since $g'(p) > 0$, an increase in λ will shift this line upward. Under these assumptions, the

⁹ Under the assumptions used here, there must be at least one such intersection, but the possibility of multiple short-run equilibria is not ruled out.

¹⁰ With rational rather than static expectations (see n. 4), Figure 2 remains an accurate description of price-quantity movements in the industry. In general, the initial (short-run) response is different in the two cases, and the speed of adjustment along $z = h(p)$ will differ. In the context of this model, the equilibrium (p^*, z^*) is stable under either expectations hypothesis.

curve $z = h(p)$ will intersect the line $g(p) = \lambda$ at exactly one value of z , say z^* . At this point, (18) implies $\dot{z} = 0$. For values of z to the left of z^* , product price (which is read off the relative demand curve) will exceed the price p^* which satisfies $g(p) = \lambda$, so that $g(p)$ will exceed λ . From (18) this implies that $\dot{z} > 0$. Similarly, if $z < z^*$, \dot{z} will be negative. Thus z^* is a position of stable equilibrium for the industry, and the corresponding price p^* is the equilibrium price.

until it reaches the new equilibrium z_1^* , with the speed of approach governed by (18). Thus the analysis of a once-and-for-all shift in relative demand proceeds exactly as in the traditional treatment of a constant cost industry.

The analysis of a change in the rate of demand growth, λ , is simpler, since in this case there is no short-run change in z or p . A rise in λ will raise the line $g(p) = \lambda$ but leave $h(z)$ unaffected. The industry will move left along the relative demand curve until the new equilibrium is

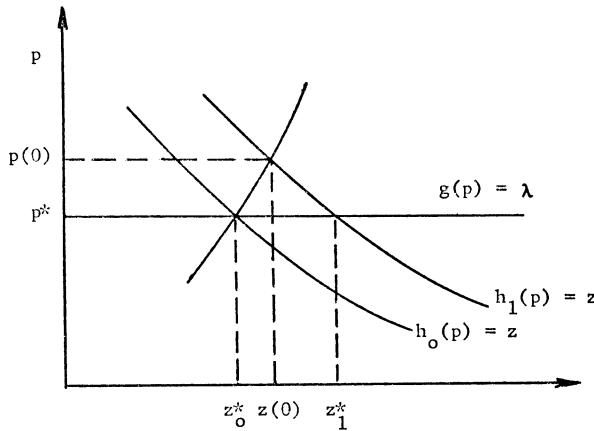


FIG. 3

Using Figure 2, we can analyze the effects of either a shift in the relative demand curve or of a change in the demand growth rate λ . The first of these problems is depicted (for an upward demand shift) in Figure 3. Consider an industry initially in a long-run equilibrium position (such as z^* in Fig. 2). Suppose that demand shifts as shown from $h_0(p)$ to $h_1(p)$, the rate λ remaining the same. The industry will *immediately* move to a new short-run equilibrium position $z(0)$, $p(0)$ which, if the short-run supply curve is upward sloping, will be to the right of the former equilibrium z_0^* . The industry will then move to the right along $h_1(p)$

reached, at a higher equilibrium price and lower z than before. This latter case helps to illustrate both the similarities and the differences between the line $g(p) = \lambda$ and the traditional flat supply curve of a constant-cost industry. For λ fixed at zero, the price p which satisfies $g(p) = 0$ is a supply price in the traditional sense: it is the price at which firms will be willing to supply an output to the market on a continuing basis, and as in the usual constant-cost case, this price is the same for all outputs. Unlike the usual constant-cost case, however, this supply price cannot be derived from a static, long-run cost minimization problem. In

addition, if one considers positive values of λ , it is seen that supply price is a rising function of the growth rate of the industry. The height of the curve $g(p) = \lambda$, unlike the height of the usual average cost curve, depends on the rate of shift of demand as well as on cost considerations.

This concludes the analysis of the response of industry output to demand changes, under the assumption that these shifts do not induce new firms to enter the industry. We now turn to the effects of entry on the conclusions of this analysis. From (15) and from Figure 2, one sees that at a position of industry equilibrium in the *absence* of entry, the curve relating the value of each firm under optimal management $\{V^*[K(0)]\}$ to its assets, $K(0)$, is a straight line passing through the origin whose slope is higher, the higher is the rate of demand shift λ . Thus the value of a unit of capital in an industry will depend upon how fast demand is growing in that industry *and* upon the ease with which this demand growth can be satisfied by existing firms. If capital goods were freely transferable, at constant unit cost, across industries (that is, if capital were a variable input) and if entry were free, the ratio $V^*[K(0)]/K(0)$ would be equated across industries. If capital is *not* a variable input, then one must treat each means of accumulating capital differently, so that internal investment by an established firm is associated with one set of adjustment costs, merger with another set of costs, and entry of new firms with still a third set of costs. We have treated formally only the first mode of capital accumulation. As remarked at the beginning of this section, firms will have no incentive, on our model, to merge even in the absence of costs associated with mergers, so that the introduction of such costs would imply that

firms will *never* merge. To deal with entry, some additional theory is needed.

We shall not attempt to develop a theory of entry from notions of the "optimal behavior of potential firms," but the outlines of what a reasonable entry hypothesis should look like are easily sketched. First, while one cannot assert that $V^*[K(0)]/K(0)$ should equalize across industries, the birth rate of firms should certainly be a rising function of this ratio. For fixed factor prices, this ratio in turn is a rising function of output price. Then if $b(p)$, $b'(p) > 0$ is the increment to output from new firms relative to current industry output, equation (18) is replaced by:

$$\frac{\dot{z}(t)}{z(t)} = g[p(t)] + b[p(t)] - \lambda. \quad (19)$$

The construction of Figure 2 and the subsequent analysis then proceeds as before. The effect of replacing (18) with (19) is to make the equilibrium product price less sensitive to changes in the growth rate λ . If, as one might expect, $b(p)$ becomes infinitely elastic at some price p_0 (whose value would of course depend on prices in other industries), then under (14) the horizontal line in Figure 1 would have to lie under the line $p = p_0$, regardless of the size of λ . In the limit, as costs of shifting capital between industries go to zero, the height of the horizontal line in Figure 1 would be fixed exactly by prices and costs in other industries, and changes in λ would affect *only* the entry rate.

V. INVESTMENT DEMAND IN A COMPETITIVE INDUSTRY

Once the behavior of industry output over time is determined, it is not difficult to obtain the industry factor demand functions. In view of the recent interest in the theoretical foundations of

econometric investment functions, most of which have been estimated from industry or even sectoral time series, it may be worthwhile to examine the implications for investment demand of the analysis of the preceding two sections. In so doing, we neglect the effects of entry of new firms.

From equation (12), which is readily aggregated over all firms, we have the gross investment function for the industry as a function of output price and other variables:

$$\frac{I(t)}{K(t)} = D_2 \left(r, \frac{w}{p}, \frac{q}{p} \right). \quad (20)$$

From (16) and (18), $p(t)$ is obtained as a function of $p(t)$ and λ . For econometric purposes, then, (20) and this first-order equation in $p(t)$ would be approximated by convenient linear or log-linear forms, and $p(t)$ eliminated between the two equations. The industry investment demand function so obtained will be a first-order differential (or difference) equation in $I(t)/K(t)$ with the levels and rates of changes of the factor prices w , v , and r as "forcing variables" and with the rate of demand shift, λ , as an additional variable. The properties of this investment function are easily developed from the preceding analysis and need not be discussed here.

Two important differences between the industry demand function for capital goods and the demand function for a single firm may be noted. First, the firm's demand function, equation (20), regarded as a differential equation in capital stock, is a first-order equation as is the familiar flexible accelerator. When one looks at adjustment from an industry-wide viewpoint, regarding product price as an "endogenous" variable, the order of the differential (or difference) equa-

tion in capital stock goes up by one. Second, when one aggregates from the firm to the industry, the rate of shift in the industry's demand curve, λ , becomes an argument in the investment function. The presence of λ in this function was, of course, a feature of the early formulations of the accelerator hypothesis.¹¹

VI. CONCLUSION

This paper has been an attempt to clarify some familiar notions of the theory of supply, in particular the concepts of short- and long-run adjustments to demand shifts, through the assumption of production possibilities which induce present value-maximizing firms to "stagger" their responses to price changes. Insofar as the effort has been successful, it seems to me to have relevance for empirical study of industrial organization and of the demand for investment goods.

First, the present value-maximization problem of the firm as we have formulated it yields an optimal investment plan characterized by the independence of firm size and the percentage growth rate of the firm. Such a model, if made stochastic in an appropriate way, has the empirical advantages of the "Gibrat's Law" theories of firm growth as well as the other well-known virtues of a hypothesis based on the assumption of maximizing behavior. In addition, it has been found that if firm growth is independent of size, the ratio of a firm's market value to the replacement value of its assets is also independent of size.

From the point of view of empirical investment studies, the approach taken here has two advantages over an ap-

¹¹ See, for example, Clark (1917) and Samuelson (1939).

proach based on the theory of a single firm's behavior. First, the industrial organizational evidence is highly relevant to any investment theory (indeed, Gib-rat's Law is an investment theory), and restricting oneself to investment hypotheses which are consistent with this evidence provides a useful discipline in theoretical work. Second, in view of the aggregate nature of the time series used in most empirical investment studies, there are obvious advantages to develop-

ing an investment demand for an industry rather than for a single firm.

Finally, we have shown that in the presence of costs of adjustment the long-run supply price for a competitive industry cannot be interpreted as an average cost curve derived from a "long-run cost minimization problem." If the industry's long-run equilibrium position is not attained immediately, then the nature of this equilibrium depends on the costs of approaching and maintaining it.

APPENDIX

DERIVATION OF (8) AND (9)

It will be convenient to approach the maximum problem stated in Section III of the text in two steps: first, at each point in time, choose $L(t)$ so as to maximize $R(t)$ for arbitrary $K(t)$ and $I(t)$; second, use the solution to the first problem to eliminate $L(t)$ from the problem, and choose a path $I(t)$, $t \geq 0$, so as to maximize $V(0)$.

The first of these problems is:

$$\max_{L(t)} \{ pF[L(t), K(t), I(t)] - wL(t) \}.$$

Clearly, (8) is a necessary condition for this problem. From (4) and (7), (8) has exactly one solution $L_0[K(t), w/p]$, as given in (10), and this solution yields a maximum to receipts for given $K(t), I(t)$.

Let:

$$\begin{aligned} G[K(t), I(t)] &= pF \left\{ L_0 \left[K(t), \frac{w}{p} \right], K(t), I(t) \right\} \\ &\quad - wL_0 \left[K(t), \frac{w}{p} \right] - vI(t). \end{aligned}$$

Then using (10), it is easily verified that G is homogeneous of degree one, that:

$$\begin{aligned} G_K > 0, G_I < 0, G_{KK} < 0, G_{II} < 0, \\ \text{and } G_I \left(\frac{I}{K} \right) \rightarrow -\infty \text{ as } \frac{I}{K} &\rightarrow r + \delta - \alpha. \end{aligned} \tag{A1}$$

Further, if $0 < \theta < 1, I_\theta = \theta I_0 + (1 - \theta)I_1$, and $K_\theta = \theta K_0 + (1 - \theta)K_1$,

$$\begin{aligned} G(K_\theta, I_\theta) &\geq \theta G(K_0, I_0) \\ &\quad + (1 - \theta)G(K_1, I_1), \end{aligned} \tag{A2}$$

with equality if and only if (K_0, I_0) and (K_1, I_1) are proportional.

The problem is now reduced to that of selecting a continuous investment plan $I(t), t \geq 0$, satisfying $0 \leq I(t)/K(t) < r + \delta - \alpha$, so as to maximize:

$$W = \int_0^\infty e^{-rt} G[K(t), I(t)] dt,$$

subject to (1) and given $K(0)$. For any admissible plan $I(t)$, present value may be regarded as a functional $W[I(t)]$ of $I(t)$. In view of the bounds on $I(t)/K(t)$, $W(I)$ will have a finite value for all admissible plans. Let $I^*(t)$ be the solution to (11), satisfying $I^*(t) > 0$ for all t . (Recall that the existence of such a solution rests on an additional assumption that p is above some critical value p_c .) Let $I_1(t)$ be any admissible plan. Then if θ is sufficiently small, $I_\theta(t) = (1 - \theta)I^*(t) + \theta I_1(t)$ is also admissible, and $f(\theta) = W[I_\theta(t)]$ is well defined. We then show that $f'(0) = 0$ (that is, that $I^*(t)$ satisfies the Euler necessary condition and the transversality condition) and, further, that $W[I^*(t)] > W[I(t)]$ for all admissible $I(t)$ with $I(t) \neq I^*(t)$ for some t .

Form the Lagrangian expression:

$$\mathcal{L}(K, I, \lambda, t) = e^{-rt} [G(K, I) + \lambda (\dot{K} - I + \delta K)]$$

Then if a continuous plan $I(t), I(t) > 0$ for all t , maximizes $W(I)$, there must exist a multiplier function $\lambda(t)$ such that the following conditions are met:

$$0 = \frac{\partial \mathcal{L}}{\partial I} = e^{-rt} [G_I(K, I) - \lambda], \tag{A3}$$

$$0 = \frac{\partial \mathcal{L}}{\partial K} - \frac{d}{dt} (\partial \mathcal{L} / \partial \dot{K}) = e^{-rt} [G_K(K, I) + \lambda \delta] - \frac{d}{dt} e^{-rt} \lambda, \tag{A4}$$

$$0 = \frac{\partial \mathcal{L}}{\partial \lambda} = e^{-rt} (\dot{K} - I + \delta K), \tag{A5}$$

$$0 = \lim_{t \rightarrow \infty} (\partial \mathcal{L} / \partial \dot{K}) = \lim_{t \rightarrow \infty} e^{-rt} \lambda(t). \tag{A6}$$

Since G is homogeneous of degree one, the derivatives G_I and G_K are functions of I/K , and (A3) and (A4) involve only the variables I/K and λ . Setting $\dot{\lambda} = 0$ and eliminating λ between these two equations gives an equation in I/K . Using the definition of G in terms of F , this equation is found to be equivalent to (11). Hence if (11) has a solution $I^*(t)$, this solution also satisfies (A3) and (A4). Then the associated capital plan satisfies (A5), and since $I^*(t)$ is constant, (A6) is satisfied. This proves that $f'(0) = 0$, where $f(\theta) = W[I_\theta(t)]$.

We next show that $W(I)$ is a strictly concave functional of $I(t)$. Let I_0 and I_1 be distinct (for some interval of time) admissible plans, and let $I_\theta = \theta I_0 + (1 - \theta) I_1$, where $0 < \theta < 1$. Let K_i be the capital plan implied by (1) and I_i ($i = 0, 1, \theta$). Since (1) is linear, $K_\theta = \theta K_0 + (1 - \theta) K_1$. Then:

$$\begin{aligned} &W(I_\theta) - \theta W(I_0) - (1 - \theta) W(I_1) \\ &= \int_0^\infty e^{-rt} [G(K_\theta, I_\theta) - \theta G(K_0, I_0) - (1 - \theta) G(K_1, I_1)] dt. \end{aligned} \tag{A7}$$

By (A2), the right side of (A7) is non-

negative and is strictly positive unless (K_0, I_0) and (K_1, I_1) are proportional. Let t_0 be the lower end point of the first interval on which $I_0 \neq I_1$, and let t_1 be the right end point of this interval. Suppose that $I_0(t) = a(t) I_1(t)$, $a(t) \neq 1$, on (t_0, t_1) . Then if (K_0, I_0) and (K_1, I_1) are proportional, we must have $K_0(t) = a(t) K_1(t)$ or:

$$\begin{aligned} e^{-\delta(t-t_0)} K(t_0) + \int_{t_0}^t e^{-\delta(t-s)} a(s) I_1(s) ds \\ = a(t) \left[e^{-\delta(t-t_0)} K(t_0) + \int_{t_0}^t e^{-\delta(t-s)} I_1(s) ds \right], \end{aligned}$$

for all $t_0 < t < t_1$, where $K(t_0)$ is the common value of K_0 and K_1 at t_0 . But this equality implies $a(t) \equiv 1$ on (t_0, t_1) . Hence (K_0, I_0) and (K_1, I_1) cannot be proportional for all t , which proves that the right side of (A7) is strictly positive, or that $W(I)$ is strictly concave.

Next, let $I^*(t)$ be the solution to (11) as before, and suppose that $I_1(t)$, with $I_1(t) \neq I^*(t)$ for some t , is another admissible plan with $W[I_1(t)] > W[I^*(t)]$. Let $I_\theta(t) = (1 - \theta) I^*(t) + \theta I_1(t)$, for $0 < \theta < 1$. Then:

$$\begin{aligned} &\frac{f(\theta) - f(0)}{\theta} \\ &= \theta^{-1} \{ W[I_\theta(t)] - W[I^*(t)] \} \\ &> \theta^{-1} \{ (1 - \theta) W[I^*(t)] + \theta W[I_1(t)] - W[I^*(t)] \} \\ &= W(I_1(t)) - W(I^*(t)) \\ &> 0. \end{aligned}$$

But since $f'(0) = 0$, this is a contradiction.

Alternatively, if $W(I_1[t]) = W(I^*[t])$ one can (by the strict concavity of W) construct a plan $I_\theta(t) = \theta I^*(t) + (1 - \theta) I_1(t)$ such that $W(I_\theta[t]) > W(I^*[t])$. Hence this case is also ruled out by the argument above.

This proves that the solution found in Section III gives the unique optimal policy.

REFERENCES

- Alchian, Armen. "Costs and Outputs," in Moses Abramovitz (ed.). *The Allocation of Economic Resources: Essays in Honor of B. F. Haley*. Stanford, Calif.: Stanford Univ. Press, 1959.
- Clark, J. M. "Business Acceleration and the Law of Demand: A Technical Factor in Economic Cycles," *J.P.E.*, Vol. XXV (March, 1917).
- Eisner, R., and Strotz, R. H., "The Determinants of Business Investment," in Commission on Money and Credit. *Impacts of Monetary Policy*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1963.
- Haavelmo, Trygve. *A Study in the Theory of Investment*. Chicago: Univ. of Chicago Press, 1960.
- Hymer, Stephen, and Pashigian, Peter. "Firm Size and the Rate of Growth," *J.P.E.*, Vol. LXX (December, 1962).
- Johnson, J. *Statistical Cost Analysis*. New York: McGraw-Hill Book Co., 1960.
- Mansfield, Edwin. "Entry Gibrat's Law, Innovation and the Growth of Firms," *A.E.R.*, Vol. LII (December, 1962).
- Muth, John F. "Rational Expectations and the Theory of Price Movements," *Econometrica*, Vol. XXIX (July, 1961).
- Pearce, I. F. "The End of the Golden Age in Solovia: A Further Fable for Growthmen Hoping To Be 'One Up' on Oiko," *A.E.R.*, Vol. LII (December, 1962).
- Phelps, Edmund. "The Golden Rule of Accumulation: A Fable For Growthmen," *A.E.R.*, Vol. LI, Part 1 (September, 1961).
- Samuelson, P.A. "Interactions between the Multiplier Analysis and the Principle of Acceleration," *Rev. Econ. and Statis.*, Vol. XXI (May, 1939).
- . *Foundations of Economic Analysis*. Cambridge, Mass.: Harvard Univ. Press, 1947.
- Simon, H. A. "Comment: Firm Size and the Rate of Growth," *J.P.E.*, Vol. LXXII (February, 1964).
- Simon, H. A., and Bonini, C. P. "The Size Distribution of Business Firms," *A.E.R.*, Vol. XLVII (September, 1958).
- Stigler, G. J. "Economies of Scale," *J. Law and Econ.*, Vol. I (October, 1958).