ESSAYS IN ECONOMIC THEORY

by

MICHAEL ROTHSCHILD
B.A. Reed College
(1963)
M.A. Yale University
(1965)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF
PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June, 1969.

Signature of Author. . . . .
Department of Economics, May 16, 1969

Certified by. . . . .
Thesis Supervisor

Accepted by. . . . .
Chairman, Departmental Committee
on Graduate Students.

ABSTRACT

ESSAYS IN ECONOMIC THEORY

Michael Rothschild

Submitted to the Department of Economics on May 16, 1969 in partial fulfillment of the requirement for the degree of Doctor of Philosophy.

This thesis consists of three self-contained but related essays. The first examines the effect which a cost of changing its capital stock has on the investment behavior of the firm. Cost of adjustment functions have been used to justify the use of distributed lags in econometric studies of investment. It is sometimes forgotten that this justification depends on the convexity of the cost of adjustment function. We argue that convexity is not a compelling assumption and show that non-convex cost of adjustment functions have implications different from, but no less interesting than, convex ones. In particular they provide a justification for the familiar distinction between fixed and variable factors of production.

In the second essay we attempt to answer the question: when is a random variable Y riskier than a random variable X? We show that two seemingly different approaches to this question, formalizing the intuitive concept that the density of X has less weight in its tails than that of Y and tracing the implications of the fact that risk averters prefer X to Y, lead to the same criterion for determining when Y is riskier than X. This criterion is not the same as that of mean-variance analysis. That it is more appealing seems to us obvious; that it is tractable, that is, that it can be used to give answers to questions of economic interest, is demonstrated.

In the final essay a model of firm behavior under conditions of changing demand is presented. The distinction between fixed and variable factors of production is used to examine the effects of changing demand on firm actions. The difficulties of defining increased variability in demand are discussed. For a special case, where these difficulties are absent, it is shown that profits decrease as the variability of demand increases. The effects of increasing variability of output on costs of production and choice of technique are examined. Econometric implications of the model are discussed briefly.

Thesis supervisor: Franklin M. Fisher

Title: Professor of Economics

ACKNOWLEDGMENTS

I have picked many brains in the course of writing this thesis. Joseph Stiglitz co-authored Chapter Two and criticized drafts of Chapter Three. David Ragozin helped me through many mathematical difficulties. Peter Diamond, Avinash Dixit, Martin Feldstein, and Robert Solow helped me to clarify my thoughts during early stages of this work. Duncan Foley, Franklin Fisher, and David Wallace made helpful comments on later drafts. To Stanley Fischer and Stephen Kennedy I owe a special debt, for they have listened to and pointed out the errors in countless false proofs. Christian von Weizsäcker agreed to serve on my thesis committee on very short notice.

Mrs. Inez Crandall typed the final version of a difficult manuscript. The figures were drawn by Katherine LaPerche.

Finally, I would like to take this opportunity to thank James Tobin for the elegant and exciting economic theory course I took at Yale in 1964-1965. It was responsible for my becoming an economist.

TABLE OF CONTENTS

Page No.

Chapter 1

ON THE COST OF ADJUSTMENT. 1

Chapter 2

THE RISK ANALYSIS OF CHOICES INVOLVING UTILITY. 26

Chapter 3

CHANGING DEMAND; ITS COSTS AND CONSEQUENCES. . 57

Chapter 1

ON THE COST OF ADJUSTMENT

Much of the current interest in cost of adjustment functions stems from their ability to provide a rigorous theoretical justification for the use of distributed lags in econometric studies of investment behavior.[1] It is sometimes forgotten how critically this justification depends on the assumption that the cost of adjustment function is convex. In this essay we argue that other forms of the cost of adjustment function are plausible and examine the dependence of the firm's investment program on the nature of the cost of adjustment function. In particular, we show that while convex cost of adjustment functions lead to distributed lags, other forms will cause the firm to operate with an unchanging capital stock in changing market conditions -- market conditions which would, were there no cost to adjustment, lead the firm to alter its capital stock. The firm adapts to changing market conditions by adjusting its labor force. Thus, non-convex cost of adjustment functions provide an explanation for the familiar distinction between fixed and variable factors of production.

The organization of this essay is as follows: In Section I we develop a simple general model of the profit-maximising firm and demonstrate, under rather restrictive assumptions, the existence of an optimal investment program. Section II examines the relation between the cost of adjustment function and the firm's investment program in the context of static expectations. Section III summarizes the results and suggests how they may

---

[1] See e.g., [1], [2], [4] and [5].

be extended. Most of these results may be found in Eisner and Strotz [1]; our model is somewhat different from theirs, our derivations somewhat more rigorous, and our emphasis and interpretations very different.

I

THE MODEL

(a) Production

The firm produces output with a neoclassical production function $Q = Q(K,L)$ which is twice continuously differentiable, homogeneous of degree one and concave.

(b) Revenue and Sales

The firm is a monopolist, facing market conditions at time t, described by a revenue function $R(Q,t)$, which we assume to be bounded above and strictly concave in Q for all t. Since the model allows no inventories, sales are necessarily equal to or less than production. We could allow the firm to sell less than it produces -- free disposal -- but since it always costs the firm something to produce output, nothing would be gained by doing so.[1]

---

[1]This is not strictly true; if it is possible to produce output without labor, the firm may find itself with an inherited capital stock -- which it does not want to change because of adjustment costs -- which will produce by itself more output than will maximize revenue. It is trivial to take care of this case by allowing the firm to sell less than it produces. This is a rare case and the gain of taking account of it is slight; the cost in terms of complexity of notation is great and we ignore it.

(c) Labor

At time t the firm may hire as much labor, L, as it desires at the wage rate w(t). This wage represents the only cost of using labor.

(d) Quasi-rents

Let

$$G(K,t) = \underset{L \geq 0}{\text{Max}} \; R(Q(K,L),t) - w(t)L.$$

Clearly G(K,t) represents the quasi-rents which the firm gains from having a capital stock K in time t. We now prove

Lemma 1: G(K,t) is a strictly concave function of K.

Proof: Suppressing the argument t, we must show that if

$$\hat{K} = \lambda K_1 + (1-\lambda)K_2 \text{ for } 0 < \lambda < 1 \text{ and } K_1 \neq K_2 \text{ then}$$

$$G(\tilde{K}) > \lambda G(K_1) + (1-\lambda)G(K_2) \; .$$

Let L_i satisfy $R(Q(K,L_i)) = G(K_i)$ and $T_i = Q(K,L_i)$ for 1=1,2, and define $\tilde{L}$ by $\hat{L} = \lambda L_1 + (1-\lambda)L_2$. The first order conditions for maximization require that the wage equal the marginal revenue product of labor, or that

$$R'(T_i)Q_2(K_i,L_i) = w.$$

Thus, if $K_1 \neq K_2$, $T_1 \neq T_2$. Since Q is concave and R strictly concave, we have

$$Q(\check{K},\check{L}) \geq \lambda Q(K_1,L_1) + (1-\lambda)Q(K_2,L_2) = \tilde{T}.$$

and

$$R(\widetilde{T}) > \lambda R(T_1) + (1-\lambda)R(T_2).$$

However,

$$G(\hat{K}) \geq R(\tilde{T}) - w\tilde{L}$$

$$> \lambda(R(T_1) - wL_1) + (1-\lambda)(R(T_2) - wL_2)$$

$$= \lambda G(K_1) + (1-\lambda)G(K_2)$$

as was to be shown.

(e) Capital

We assume that the firm rents capital, paying a charge of $\rho(t)K$ for the use of K units of capital in period t.

(f) Changes in the Capital Stock, the Cost of Adjustment

For simplicity we assume that there is no depreciation. It will be apparent that dropping this assumption will complicate our results without altering their spirit.[1] Changes in the capital stock come about through investment or disinvestment, that is, capital grows or declines according to the rule.

(1) $$K_t = K_{t-1} + I_t.$$

The firm pays a cost every time it changes its capital stock, that is, there is a cost of adjustment function C(I) with the following properties:

(2.i) $C(0) = 0$

(2.ii) $C(I) > 0$ for $I \neq 0$

(2.iii) Sign $C'(I)$ = Sign I for $I \neq 0$.

C is zero if and only if I is zero, and is an increasing function of the absolute value of I. (2.iii) also implies that C'(I) exists everywhere

[1] We could introduce depreciation into the model without altering our results at all, simply by defining the argument of the cost of adjustment function, see (2) below, to be net rather than gross investment.

except at 0. Since C has a minimum at 0, if $C'(0)$ exists, it is necessarily equal to 0.

If the cost of adjustment function is convex, that is, if as the scale of investment, or disinvestment, increases in any one period, the unit cost of investment, or disinvestment, increases, the firm will distribute its response to a change in the optimal capital stock through several periods, rather than concentrate it in one period. We prove this below rigorously. Here we merely note that this result justifies use of distributed lags in econometric work. For this reason, most of the interest in cost of adjustment functions has centered on convex cost of adjustment functions. The arguments given as to why the cost of adjustment function should be convex are quite weak. Eisner and Strotz [1, p.77] give two: The first is that as the firm increases its demand for investment goods in a single period pressure will be put on the supply of investment goods which will lead to an increase in the price of investment goods. Our model is that of a firm which is a price taker in factor markets; for it such considerations are clearly inappropriate. The second argument is that there are "increasing costs associated with integrating new equipment into a going concern: reorganizing production lines, training workers, etc." This is simply an assertion, and hardly a compelling one. Decreasing costs are just as plausible as increasing costs. Eisner and Strotz' examples will do as well as any to demonstrate this point. No reason is seen why training necessarily entails increasing costs. Training involves the use of information (once one has decided how to train one worker, one has in effect decided how to train any number of them), which is a classic cause of decreasing costs. Furthermore, the process is subject to some indivisibilities. It requires at least one teacher to train one worker. Presumably no more teachers are required to train two or three workers. Although there may be practical limits on teacher-trainee ratios which, coupled with the expenses involved in administering large training programs, could cause decreasing returns to set in eventually, it seems likely that at low levels of activity, increasing rather than decreasing returns would be characteristic of training programs. Similarly, reorganizing production lines involves both the use of information as a factor of production -- once one has decided how to reorganize one production line one has figured out how to reorganize two, three or n -- and indivisibilities -- one may not be able to reorganize only half or a tenth of a productive line.

Indivisibilities and the use of information as a factor of production would seem to be plausible components of much of what we would want to include under the rubric of adjustment costs. It is possible that there are some fixed costs to adjustment which are incurred whenever the capital stock is changed, regardless of how much it is changed. These fixed costs could represent such things as shut-down time -- it is necessary to stop the plant for a while to install new equipment -- or a break in period necessary to get the plant running again, or a portion of planning cost incurred simply because there were changes in the capital stock but which is independent of the size of the change. This is sufficient to establish that there is no compelling a priori reason to restrict our interest in adjustment costs to convex cost of adjustment functions. Below we examine the implications of other cost of adjustment functions.

(g) Profits and the Optimal Capital Stock

Let

(3) $$V(K,t) = G(K,t) - \rho(t)K.$$

Then $V(K,t)$ represents the profits the firm can make at time t if it has capital stock K. $V(K,t)$ is a strictly concave function of K (as $G(K,t)$ is). Since $G(K,t)$ is bounded above, $V(\infty,t) = -\infty$ and $V(K,t)$ attains its maximum on $[0,\infty]$. Let K_t^* be the unique maximizing K. K^* is a function of market conditions,

$$K^* = K^*(w,R(Q),\rho).$$

(h) Maximization

The firm discounts cash flows received in period t by the discount factor $r(t)$. We assume that the firm's investment policy is designed to maximize discounted cash flows. That is, the firm chooses the sequence of investment (I) to maximize

(4) $$\mathcal{L}\{I\} = \sum_{t=1}^{\infty} (V(K_t,t) - C(I_t))r(t)$$

subject to

$$K_t = K_{t-1} + I_t$$

which, of course, implies that $\mathcal{L}\{I\}$ is a function of K_0. We shall have occasion to stress this fact by writing $\mathcal{L}(\{I\},K_0)$ instead of $\mathcal{L}\{I\}$.

Before we can proceed, we must examine the conditions under which the series on the RHS of (4) converges. Obviously, for all $\{I\}$,

$$\mathcal{L}\{I\} \leq \sum_{t=1}^{\infty} [V(K_t^*,t)]r(t)$$

so that

(5) $$\sum_{t=1}^{\infty} [V(K_t^*,t)]r(t) < \infty$$

is a sufficient condition for $\mathcal{L}\{I\}$ to be bounded above. We shall assume that (5) holds in what follows. It is a condition with some bite; if maximum one period profits are growing at rate g (i.e., $V(K_t^*,t) = V(K_1^*,1)(1+g)^t$) and the discount rate is constant at r $(r(t) = (1+r)^{-t})$, then (5) will hold only if the discount rate is greater than the growth rate. If $V(K_t^*,t)$ is constant or bounded then (5) is equivalent to

(6) $$\sum_{t=1}^{\infty} r(t) < \infty,$$

a necessary condition for almost any economically sensible action. (If (6) failed, the value of a consol paying a mill a millennium would be infinite.)

As we wish to maximize $\mathcal{L}\{I\}$, we are not concerned that $\mathcal{L}\{I\}$ be bounded below for all $\{I\}$ only that there exist some $\{I\}$ such that $\mathcal{L}\{I\} > -\infty$. If there are no $\{I\}$ such that $\mathcal{L}\{I\} \geq 0$, the firm will not stay in business.

(i) <u>Existence of the Optimum Program</u>

We prove in this section that our problem is well defined, that is, that an optimum program exists. This may seem obvious and trivial, but the method of proof is somewhat novel and may be useful in other contexts. We shall make somewhat stronger assumptions than are strictly necessary in order to make the outlines of the argument as clear as possible. It will be clear that the proof can be modified so as to be appropriate in a considerably more general setting. Specifically, we assume that the maximum one period profit is bounded, or

(7) $$\bar{V} = \sup_t V(K^*(t),t) < \infty;$$

that the optimal capital stock is bounded,

(8) $$\bar{K} = \sup_t K^*(t) < \infty;$$

and that the rental price of capital is bounded,

(9) $$\bar{\rho} = \sup_t \rho(t) < \infty.$$

It is clear that (7) and (8) are sufficient to guarantee (5) -- that $\mathcal{L}\{I\}$ is bounded above.

It seems clear that if K^*_t is bounded, it will never pay to invest so much as to have a larger capital stock than K^*_t. This is the content of

Lemma II: Let $\{\hat{I}\}$ be an investment program if $\hat{K}_t > M = \text{Max}(K_0,\bar{K})$ for any t, there is an $\{I'\}$ such that,

$$\mathcal{L}\{I'\} > \mathcal{L}\{\hat{I}\}.$$

Proof: Let t_1 be the first t such that $\hat{K}_{t_1} > M$. Then either

(10) $$\hat{K}_{t_2} \le M \text{ for some } t_2 > t_1$$

or

(11) $$\hat{K}_t > M \text{ for all } t > t_1 .$$

We define $\{I'\}$ as follows. For $t < t_1$, let $I'_t = \hat{I}_t$, let $I'_{t_1} = M - K'_{t_1-1} = M - \hat{K}_{t_1-1} < \hat{I}_{t_1}$. If (10) holds, let $I'_t = 0$ for $t_1 < t < t_2$ and $I'_{t_2} = \hat{K}_{t_2} - M < \hat{I}_{t_2}$ (since $\hat{K}_{t_2-1} > M$), and for $t > t_2$, $I'_t = \hat{I}_t$. If (11) holds let $I'_t = 0$ for $t > t_1$; in either case $V(K'_t,t) \ge V(\hat{K}_t,t)$ for all t while $C(I'_t) \le C(\hat{I}_t)$ for all t and $C(I'_{t_1}) < C(\hat{I}_{t_2})$ so that $\mathcal{L}\{I'\} > \mathcal{L}\{\hat{I}\}$.

This means that if there is an optimal program $\{\tilde{I}\}$, then $\tilde{K}_t < M$ which implies both that

(12) $$\tilde{I}_t \leq M$$

and

(13) $$V(\tilde{K}(t),t) \geq -\bar{\rho}M.$$

We can thus restrict our search for an optimum to programs satisfying (12) and (13). This fact and some notation, is all we need to prove that an optimum exists. We use the symbol # to denote Cartesian product. If E^1 is the real line, by $H = \underset{t}{\#}E^1$, we denote the space of real sequences. In what follows we consider H as a topological space endowed with the product topology. Let $B = [-M,M]$. If $H_B = \underset{i}{\#}B$, then a way of stating (12) is to say that we need only look for a maximum to $\mathcal{L}\{I\}$ on H_B. Note that H_B -- the Cartesian product of compact sets -- is itself compact.[1] To show an optimum exists we then merely need show that $\mathcal{L}$ is a continuous mapping and the result follows from the fact that a continuous real valued function on a compact set achieves its maximum.

We cannot however be so direct; instead we show that there is a continuous function $\mathcal{M}$ on H_B which has the property that if $\{\tilde{I}\}$ maximizes $\mathcal{M}$, $\{\tilde{I}\}$ must maximize $\mathcal{L}$. Define $U(K(t),t)$ as $\mathrm{Max}(V(K(t),t), -\bar{\rho}M)$, and let $\mathcal{M}\{I\} = \sum_{t=1}^{\infty} [U(K(t),t) - C(I_t)]r(t)$. It is clear from (13) that if $\{\tilde{I}\}$ maximizes $\mathcal{M}$ it maximizes $\mathcal{L}$ as well. Thus, it remains to show that $\mathcal{M}$ is continuous. It is hardly surprising that this crucial step of the argument depends on (6).

---

[1] This is the well-known Tychnoff theorem. See any topology book, [3, p.143].

We show that $\mathcal{M}$ is continuous by showing that for any $\{\hat{I}\} \in H_B$ and any $\varepsilon > 0$ there is an open set S of H_B containing $\{\hat{I}\}$ such that $\{I'\} \in S$ implies

$$\left| \mathcal{M}\{I'\} - \mathcal{M}\{\hat{I}\} \right| < \varepsilon .$$

Let $\mathcal{M}_a^b\{I\} = \sum_{t=a}^{t=b} (U(K_t,t) - C(I_t))r(t)$. By convention $\mathcal{M}^b = \mathcal{M}_1^b$ and $\mathcal{M}_a = \mathcal{M}_a^\infty$. If $R = \bar{V} + |\rho M| + C(M) + C(-M)$ then $|U(K(t),t) + C(I_t)| < R$ for all t when $\{I\} \in H_B$.

By (6) there is $T(\varepsilon)$ such that $\left| \mathcal{M}_{T(\varepsilon)}\{I\} \right| \leq R \sum_{t=T(\varepsilon)}^{\infty} r(t) < \varepsilon/4$ for all $\{I\} \in H_B$. Also since U(K,t) and C(I) are continuous, there is a $\delta > 0$ such that $|I_t' - \hat{I}_t| < 2\delta$ for all $t \leq T(\varepsilon)$ implies

$$\left| \mathcal{M}^{T(\varepsilon)}\{I'\} - \mathcal{M}^{T(\varepsilon)}(\hat{I}) \right| < \varepsilon/2.$$

Now define S as $S = \#_t A_t$ where $A_t = (\hat{I}_t - \delta, \hat{I}_t + \delta)$ for $t \leq T(\varepsilon)$ and $A_t = B$ for $t > T(\varepsilon)$. S is an open set in H_B. If $\{I'\} \in S$ then,

$$\begin{aligned} \left| \mathcal{M}\{\hat{I}\} - \mathcal{M}(I') \right. &= \left| \mathcal{M}^{T(\varepsilon)}\{\hat{I}\} + \mathcal{M}_{T(\varepsilon)+1}\{\hat{I}\} \right. \\ &\left. - \mathcal{M}^{T(\varepsilon)}\{I'\} - \mathcal{M}_{T(\varepsilon)+1}\{I'\} \right| \leq \\ &\left| \mathcal{M}^{T(\varepsilon)}\{\hat{I}\} - \mathcal{M}^{T(\varepsilon)}\{I'\} \right| + \\ &\left| \mathcal{M}_{T(\varepsilon)+1}\{\hat{I}\} \right| + \left| \mathcal{M}_{T(\varepsilon)+1}\{I'\} \right| \\ &\leq \varepsilon/2 + \varepsilon/4 + \varepsilon/4 = \varepsilon . \end{aligned}$$

This completes the proof of

Theorem III Conditions (6) through (9) guarantee the existence of an optimum program.

If C is convex -- a contingency which includes the linear case $C(I) = \alpha|I|$ -- then $\mathcal{J}$ is a strictly concave function of $\{I\}$ and the optimal program is unique.

II

OPTIMAL INVESTMENT POLICY UNDER STATIC EXPECTATIONS

(a) Static Expectations

It is always simpler to analyze dynamic problems when assumptions are made which remove the dynamics from the problem. This case is no exception. We begin our analysis of optimal programs with the case where nothing is expected to change. That is, we assume that the firm expects present market conditions; as represented by $\rho(t)$, $w(t)$, and $R(Q,t)$ to persist forever. This allows us to write

(14) $$V(K,t) = V(K).$$

We further assume that the discount rate is constant, or that

(15) $$r(t) = (1+r)^{-t}.$$

These assumptions are sometimes referred to as static expectations. It is not assumed that static expectations are correct. It is a paradox whose familiarity is one of the signs of the low level of development of the principles of dynamic analysis in economics that the assumption that the firm expects no future changes is used to analyze its response to present, unanticipated changes. The firm is assumed to be initially in equilibrium, which in our problem means that its capital stock is at -- or near, see subsection (f) below -- K*, when market conditions change, changing with them

V(K) and K*. Static expectations, the assumption that the firm expects these changes to be permanent, simplifies considerably the problem of analyzing what the firm does when such changes occur. It is, of course, an unreasonable assumption and we would like to know how crucially our results depend on it. In Section III we suggest that, at least for concave and linear cost of adjustment functions, the analysis of the case of static expectations affords a good deal of insight into firm behavior under more general and more reasonable assumptions.

We make the further assumption that

$$K_0 < K^*. \tag{16}$$

Although this assumption is innocent, it is not without content. It is clear that if the firm stays in business, it really does not matter if we analyze investment or disinvestment. However, it might be that no matter what the firm did, it could not make a profit. In such a case the firm's best policy would be simply to go out of business. Its capital stock would not follow the rule set out in equation (1). The firm would not incur the cost of changing the capital stock described in (2). In short, our model would not apply to such a firm. If $K^* > K_0$, this cannot be the case. For since V(K) is strictly concave, $V(K^*) > V(K_0) \geq V(0) = 0$. The firm could make a profit were it simply to retain K_0 and do no investing.

For $K^* < K_0$ we have no such guarantee that the firm would not do best by going out of business. (Consider, for example, the case where $K^* = 0$ and $V(K) < 0$ for all $K > 0$.) If the firm chooses to stay in business, the analysis we are about to give applies to the firm's disinvestment policy with only a few obvious sign changes. If the firm chooses to go out of business, its behavior is very simple to analyze, and not very interesting. It is, for example, clear that there is an optimal policy -- that of declaring immediate bankruptcy -- which yields a total profit of 0. Thus the assumption that $K^* > K_0$ has no essential consequence beyond the agreement to limit our analysis to non-trivial cases.

(b) Some Useful Lemmas and Conventions

This subsection lists some tools and conventions we will need in our investigation of the optimal program and its dependence on the cost of adjustment function.

The argument of the proof of Lemma II, can be trivially adapted to prove,

Lemma IV: If (14) to (16) hold, and I' is an optimal program, then

(17) $$I'_t \geq 0 \text{ for all } t$$

and

(18) $$K'_t \leq K^* \text{ for all } t.$$

If in any period it makes sense to do no investing, the problem will, because of (14) and (15), look no different in the next period, and it will not pay to invest then, or in the following period and so on. This is the content -- and except for some technical details -- the proof of

Lemma V: If there are static expectations and $\{I'\}$ is an optimal program for which $I'_t = 0$, then there is an optimal program $\{\hat{I}\}$ such that $\hat{I}_{t+j} = 0$ for j any non-negative integer.

Proof: Clearly it will suffice to prove the assertion for the case j = 1. Define $\{I'^t\}$ by $I_s'^t = I'_{t+s}$. If $\{I'\}$ maximizes $\mathcal{L}(\{I\}, K_0)$, then $\{I'^{t-1}\}$ maximizes $\mathcal{L}(\{I\}, K'_{t-1})$ and $\{I'^t\}$ maximizes $\mathcal{L}(\{I\}, K'_t) = \mathcal{L}(\{I\}, K'_{t-1})$ since $I_t = 0$. The preceding equality holds for any $\{I\}$. Thus $\{I'^{t-1}\}$ maximizes $\mathcal{L}(\{I\}, K'_t)$. Define $\{\hat{I}\}$ by

$$\hat{I}_s = I'_s \text{ for } s \leq t$$

$$\hat{I}_{s+1} = K'_s \text{ for } s > t$$

Then $\{\hat{I}\}$ maximizes $\mathcal{L}(\{I\}, K_0)$ and $\hat{I}_{t+1} = I'_t = 0$ which is what was to be proved.

We observed above that if C'(0) exists it must be equal to 0. In the proofs to follow this is an inconvenience, but an avoidable one. Consider a function $\tilde{C}$ such that for $I \geq 0$, $\tilde{C}(I) = C(I)$ which also has the property that $\tilde{C}'(0)$ exists and is equal to $C'^{+}(0)$, the right hand derivative of C at 0.[1] We have shown that the optimal program must satisfy (17). Thus, it is clear that if $\{I\}$ maximizes $\tilde{\mathcal{L}}\{I\} = \sum_{t=1}^{\infty} [V(K_t) - \tilde{C}(I_t)]\,(1+r)^{-t}$ subject to (1) and (17) it also maximizes $\mathcal{L}\{I\}$ subject to (1). In what follows, we shall adopt (17) as a restriction and assume that C '(0) exists and is not necessarily equal to 0.

(c) Convex Cost of Adjustment Functions and Distributed Lags

In subsection I(f) we asserted that if adjustment costs were convex, then the response to change in market conditions would not be concentrated in one period, but would be distributed over several. In the context of static expectations a stronger theorem can be proved.

---

[1] $C'^{+}(I)$ so exists since C is piecewise monotonic.

<u>Theorem VI</u>: If C is twice differentiable and strictly convex, ($C''(I) > 0$ for all I), then, if $\{I'\}$ is optimal, $I_1' > 0$ implies $I_t' > 0$ for all t. This theorem says that if a firm responds at all[1] to changes in market conditions, its response will be distributed over all time, that is, investment will never stop. It is, of course, the justification for the use of distributed lags in econometric work.

<u>Proof</u>: Assume $I_1' > 0$ and there is a t such that $I_t' = 0$. We may, without loss of generality, assume that $t = 2$. (If $\{I'\}$ maximizes $\mathcal{L}(\{I\},K_0)$, $\{I'^{\,t-2}\}$ maximizes $\mathcal{L}(\{I\},K_{t-2}')$.) By Lemma V and the uniqueness of the optimal program $I_{2+j}' = 0$ for all $j \geq 0$. Thus, I_1' and I_2' must be the solution to the following problem:

$$(P)\quad \max_{I_1,I_2 \geq 0} H(I_1,I_2) = V(K_0 + I_1) - C(I) + \sum_{t=1}^{\infty} V(K_0 + I_1 + I_2)(1+r)^{-t} - C(I_2)(1+r)^{-1}$$

$$= V(K_0 + I_1) - C(I_1) + V(K_0 + I_1 + I_2)r^{-1} + C(I_2)(1+r)^{-1}.$$

We prove the theorem by showing that (P) cannot possibly have a solution with $I_1 > 0$ and $I_2 = 0$. The maximizing I_i' satisfy $H_i(I_1',I_2') \leq 0$ and $I_i'\, H_i(I_1',I_2) = 0$ for $i=1,2$. We must have then, $H_1(I_1',I_2') = 0$, or

$$V'(K_0 + I_1') + \frac{V'(K_0 + I_1' + I_2')}{r} = C'(I_1).$$

---

[1] This is a necessary condition. If the change in market conditions is slight, the firm may do no investing at all. See subsection II(f) below.

Since $I_2^1 = 0$, this becomes

$$V'(K_0 + I_1^1) \frac{(1+r)}{r} = C'(I_1^1).$$

From $H_2(I_1^1, I_2^1) \leq 0$, we have

$$\frac{V'(K_0 + I_1^1)}{r} \leq \frac{C'(0)}{1+r} .$$

Rearranging and combining, we have:

$$C'(I_1^1) \leq C'(0),$$

which is impossible since $I_1^1 > 0$ and $C''(I) > 0$. This contradiction completes the proof.

(d) Concave and Linear Adjustment Cost Functions

If the cost of adjustment function is concave or linear it would seem that the best policy for the firm would be to take advantage of decreasing costs of investment and concentrate its response to changes in market conditions in a single period. We demonstrate this in

Theorem VII: If there are static expectations and if C is strictly concave or linear,[1] then if $\{I'\}$ is an optimal program there is at most one t such that $I_t' > 0$.

Proof: We prove this by assuming that $I_1^1 > 0$, and $I_2^1 > 0$ and proving a contradiction. It will be apparent how the same method may be applied to the case when I_t' and I_{t+k}' are greater than zero.

---

[1]Concave or linear for $I \geq 0$, that is.

We begin by isolating the contribution of I_2' to $\mathcal{L}\{I'\}$. Let $D_t = V(K'_{t+1}) - V(K'_{t+1} - I_2)$. Since $\{I'\}$ is optimal, $K'_t < K^*$ and $D_t > 0$. As V is concave and $I'_t > 0$, D_t is monotone decreasing, in particular

$$D_t \geq D_{t+1}.$$

The contribution of I_2' to $\mathcal{L}\{I'\}$ can then be written as

$$\Delta\mathcal{L}(I_2';\{I'\}) = [\sum_{t=1}^{\infty} D_t(1+r)^{(1-t)} - C(I_2)](1+r)^{-2} .$$

Since $\{I'\}$ is optimal, $\Delta\mathcal{L}(I_2',\{I'\}) \geq 0$. Consider the program $\{\hat{I}\}$ defined by

$$\hat{I}_1 = I_1' + I_2'; \quad \hat{I}_2 = 0; \quad \hat{I}_t = I'_t \text{ for } t > 2.$$

We wish to isolate once again the contribution of the additional investment of I_2' in period 1 to $\mathcal{L}\{\hat{I}\}$ -- which we shall write as $\Delta\mathcal{L}(I_2',\{\hat{I}\})$ -- and compare it to $\Delta\mathcal{L}(I_2')$. The cost of the investment is

(19) $$C(I_1' + I_2') - C(I_1') \leq C(I_2')$$

as C is concave and increasing. The addition to cash flow in each period is $\hat{D}_t$ where $\hat{D}_1 = D_1$ and $\hat{D}_t = D_{t-1}$ for $t > 1$. Since D_t is monotone decreasing, $\hat{D}_t \geq D_t$ and

$$\begin{aligned}\Delta\mathcal{L}(I_2',\{\hat{I}\}) &= [\sum_{t=1}^{\infty} \hat{D}_t(1+r)^{1-t} - (C(I_2+I_1) + C(I_1))](1+r)^{-1} \\ &\geq (\sum_{t=1}^{\infty} D_t(1+r)^{1-t} - C(I_2))(1+r)^{-1} \\ &= \Delta\mathcal{L}(I_2',\{I'\})(1+r) \geq \Delta\mathcal{L}(I_2',\{I'\}).\end{aligned}$$

Therefore, $\mathcal{L}\{\hat{I}\} \geq \mathcal{L}\{I'\}$. If C is strictly concave, the inequality in (19) will be strict, as $I_1' > 0$, and $\mathcal{L}\{\hat{I}\} > \mathcal{L}\{I'\}$ so that $\{I'\}$ is not optimal. If C is linear, the optimal program is unique and $\{I'\}$ cannot be optimal if $\mathcal{L}\{\hat{I}\} = \mathcal{L}\{I'\}$. This completes the proof.

Thus, investment takes place in one period only. If it yields positive returns, that is if

$$\Delta\mathcal{L}(I) = (\sum_{t=0}^{\infty} V(K_0,I)(1+r)^{-t} - C(I))(1+r)^{-1} > 0$$

where I is the optimal one period investment, then it will be more profitable to invest in the first period than in any later period. For the returns from investing in the k^{th} period are $\Delta\mathcal{L}(I)(1+r)^{-k}$ which is at a maximum if k = 1 and if $\Delta\mathcal{L}(I) > 0$. If $\Delta\mathcal{L}(I) = 0$, it does not matter in what period investment takes place, or indeed, whether it takes place at all. In this case the optimal investment program is not unique.

(e) <u>More General Cost of Adjustment Functions</u>

There is no reason to believe that cost of adjustment functions are necessarily either convex or concave. Fortunately, there is no difficulty in extending our results to more general functions, to those which are piecewise convex or concave. This class includes functions which are initially concave -- reflecting indivisibilities in the adjustment process -- and later convex -- reflecting the increasing costs of the disruption caused by large-scale hurried changes which we feel on *a priori* grounds are likely to be prevalent.

Piecewise concave and convex adjustment functions lead to investment programs which are a combination of the infinitely distributed programs of convex adjustment functions and the one shot programs of concave adjustment functions. It is easy to construct cost of adjustment functions of this general type which will call for investing in two, three, four or any finite number of periods so that the consideration of these more general cost of adjustment functions destroys the simple dichotomy between one shot and infinitely distributed responses to market changes. Nonetheless, the pattern of investment is dominated by the shape of the cost of adjustment functions near zero in the following sense: only if C is convex near zero can investment continue forever. The reason is that for any optimal program, $\Sigma I_t < \infty$ so that for any I there is a T such that t $> T$ implies $I_t < I$. Thus after a finite number of periods, only the first part (the first convex portion or the first concave portion) of the cost of adjustment functions is relevant. Only if this part is convex can investment continue forever. Initial convexity is not sufficient to guarantee that investment will take place forever for its effects can, in some cases, be washed out by later concavity. Consider a function which is initially linear and later strongly concave, see Figure 1.

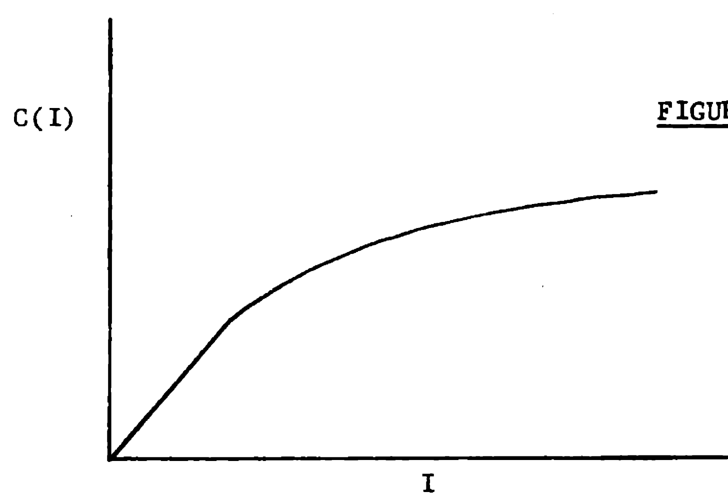


FIGURE 1

It leads to a one-period investment program. If the function is changed just slightly so that its first segment is now convex and if the discrepancy between K* and K_0 is large enough that the concave part of the function is relevant, a one-period investment program will still be in order. If the discrepancy between K* and K_0 is small, then only the convex part of the function will be relevant and an infinite investment program may be optimal. In this example, the character of the investment program depends not only on the nature of the cost of adjustment function, but also, and crucially, on the initial conditions of the problem, on the relation between K_0 and K*. We examine another and more significant case of this dependence in the following subsection.

(f) Sensitivity

Suppose adjustment costs are linear, that $C(I) = \alpha|I|$. Then investment is bunched in the first period and the optimal investment program is found by solving the following problem.

(Q) $$\max_{I>0} H(I) = V(K_0+I) \sum_{t=1}^{\infty} (1+r)^{-t} - C(I)(1+r)^{-1}$$

$$= (1+r)^{-1} [V(K_0+I)\frac{(1+r)}{r} - C(I)].$$

If I' is the optimal investment, then,

$$H'(I) \leq 0 \text{ and } I'H'(I') = 0,$$

or if $I' > 0$, then I' must be a solution to $H'(I) = 0$, so that,

(20) $$V'(K_0+I) = (\frac{r}{1+r}) C'(I).$$

In the linear case, this becomes

(20') $$V'(K_0+I) = \alpha(\frac{r}{1+r}) .$$

It is entirely possible that (20') will have no positive solution.

Let $\bar{K}$ be the unique solution to $V'(K) = (\frac{r}{1+r})\alpha$. Then if $K_0 > \bar{K}$, the optimal investment program will be one of no investment at all. If $K_0 < \bar{K}$ then $I' = \bar{K} - K_0$ is the optimal investment program.

This result easily generalizes to cases of convex and concave cost of adjustment functions. In the convex case, if there is investment in the first period there will be investment in every period. But for there to be any investment at all, it must be that investment in the first period yields a profit, or that the problem (Q) has a positive solution. But (Q) has a positive solution only if (20) has a solution. Since C is convex, the RHS of (20) is an increasing function of I bounded below by $C'(0)\frac{r}{1+r}$. The LHS of (20) is a decreasing function of I bounded above by $V'(K_0) > V'(K^*)=0$. A solution is, therefore, possible only if

$$V'(K_0) > C'(0)\frac{r}{1+r} . \tag{21}$$

Since K_0 can be arbitrarily close to K^*, (21) will be satisfied for arbitrary K_0, only if $C'(0) = 0$. Thus, unless $C'(0) = 0$ small changes in market conditions which will lead to no investment whatsoever.

The same argument and the same conclusions can be made for concave adjustment cost functions. The condition that all changes in market conditions lead to changes in the capital stock rather than simply changes in the optimal capital stock is that $C'(0) = 0$. This, because of (2), is only possible if C is convex around or in a neighborhood of zero.

III

SUMMARY AND EXTENSIONS

We may summarize our results for the case of static expectations as follows:

If adjustment costs are concave or linear, the entire response to a change in market conditions is concentrated in the first period following the change. If adjustment costs are strictly convex, if the firm responds at all to the change it will plan to distribute its response over all time. The nature of the firm's response to changes in market conditions depends critically on initial conditions, on the distance between the firm's inherited capital stock, K_0, and the new desired capital stock, K^*. Unless $C'(0) = 0$, for which convexity near 0, is a necessary but not a sufficient condition, the firm will not in general respond to small changes in market conditions. For general C, the nature of the response may depend on the size of the change.

We now turn to the question of how these results may be extended beyond static expectations to more general, and more realistic, circumstances. We have little insight to offer in the strictly convex case. For the concave case, it is clear that the analysis is not much affected. Suppose the firm faces a sequence of market conditions, represented by a $V(K,t)$ for each t. Suppose further, that for some periods market conditions do not change, that $V(K,t) = V(K,t+j)$ for $j=1,\ldots,k$. Then the argument of Theorem VII can be adapted to prove that the firm will hold a single capital stock in periods t to t+k, that, if the firm responds at all to the change in market conditions from those prevailing in period t-1 to those of period t to t+k, it will concentrate its response in the first period of the new conditions. Furthermore, if instead of the $V(K,j)$ being identical for j=t,...,t+k, they are merely very similar, then the argument of the last subsection can be adapted to show that the firm may not choose to respond to all the minor changes in market conditions occurring between t and t+k and decides instead to save on adjustment costs by making do with a single capital stock and adjusting to changes in market conditions by varying the amount of labor it employs.

This phenomena, the firm adjusting to changes in market conditions, by changing one factor of production and not the other, is the basis of the classical distinction between fixed and variable factors of production. We have shown that while strictly convex adjustment costs may justify the use of distributed lags in econometric work, concave and linear adjustment costs provide a rigorous theoretical justification for the distinction between fixed and variable factors of production. We show in Chapter 3 below that this distinction is not without econometric implications.

REFERENCES

[1] Eisner R, and Strotz R., "Determinants of Business Investment," Research Study Two in Impacts of Monetary Policy, prepared for the Commission on Money and Credit, (Englewood Cliffs, N. J., 1963).

[2] Gould, J. P., "Adjustment Costs in the Theory of Investment of the Firm," Review of Economic Studies, XXXV, (January, 1968).

[3] Kelley, J. L., General Topology (New York, 1955).

[4] Lucas R., "Optimal Investment Policy and the Flexible Accelerator," International Economic Review, VIII, (February, 1967).

[5] Treadway, A., "Rational Entrepreneurial Behavior and the Dynamics of Investment," Ph.D. Dissertation, University of Chicago, 1967.

Chapter 2

THE RISK ANALYSIS OF CHOICES INVOLVING UTILITY

In discussing economic behavior under uncertainty, we often argue that economic agents make decisions so as to maximize the expected value of some function; that is, if X is a random variable with cumulative distribution function G and α is a control parameter, we hypothesize that the economic agent will choose α so as to maximize

$$\int U(x,\alpha)\,dG(x) \quad .$$

The first order condition for maximization is

$$\int \frac{\partial U(x,\alpha)}{\partial \alpha}\,dG(x) = 0.$$

We often are interested in the response which economic agents make to an increase or decrease in the variability or uncertainty of X. For instance we would like to know the effect on savings decisions of an increase in the riskiness of investment opportunities. In this paper we give a simple general method for answering such questions. To do this, it is essential to define what is meant by an increase in the riskiness of a random variable.

There seems to be three possible approaches to this question. We could try and formalize what seems to be the intuitive meaning of an increase in the variability of the distribution of X. Second, we could seek to characterize the class of changes implied by the commonly accepted definitions of risk aversion. Suppose, for example, X represents wealth, we could agree to call all those changes in the distribution of X which would make a risk averter worse off (and a risk lover better off) as increqses in the riskiness of X and then try to discover what class of changes in the distribution of X had this property. A third approach entails characterizing the distribution of X with a few parameters and then describing increases in the variability of X in terms of changes in these parameters. The familiar analysis of distributions in mean-variance terms is an example of this approach. Since we are dealing here with matters of definition as much as of analysis, there is no strict sense in which one approach could be said to be more correct than any other. This fact, however, is not a license for agnosticism or for the suspension of judgment. Definitions are chosen for their consistency. For this reason, it is of some interest to note that the first two questions lead to the same answer while the third does not. To us, this is yet another reason for abandoning mean-variance analysis in the theory of behavior under uncertainty and seeking more general approaches, like those to be presented below.[1]

I. A DEFINITION OF INCREASING RISK

(a) Mean Preserving Spreads

Intuitively, what we mean by a random variable increasing in variability is that the tails of the distribution gain weight at the expense of

---

[1]The problem is not a new one, nor is our approach completely novel; our result is, we think, new. Our interest in this topic was whetted by Peter Diamond [3]. R. M. Solow used a device similar to our Mean Preserving Spread to compare lag structures in [10]. The problem of "stochastic dominance" is a standard one in the (statistics) operations research literature. For other approaches to the problem see, for instance, [2], which takes a tack similar to our Section 2.2. Sinee this paper was written, papers by Hadar and Russel [5] and Hanoch and Levy [7] have come to our attention which contain results similar to ours.

the center. It is easy to make this notion precise. Suppose that f is a density function defined on a closed interval $[\underline{x}, \bar{x}]$. (Figure 1). Consider a step function s(x) on this interval satisfying the following conditions:

(1) $s(x) = \alpha \geq 0$ for $a < x \leq a + t$; $b < x \leq b + t$

$s(x) = -\alpha$ for $a + d < x \leq a + d + t$; $b - d < x \leq b - d + t$

$s(x) = 0$ otherwise

where

(2) $a + d < b - d$

Such a function is pictured in Figure 2. If

$$\alpha \leq f(x) \text{ for } a + d < x < a + d + t$$
$$\text{and } b - d < x \leq b - d + t$$

then $g(x) = f(x) + s(x) \geq 0$ for all x. Furthermore, (1) implies that $\int s(x)dx = 0$, so $\int g(x)dx = \int f(x)dx$ and if f is a density function so is g.[1] As Figure (3) shows, g differs from f in that g has more probability weight in its tails and less in its center.

Straightforward calculation will show that $\int xs(x)dx = 0$ so that $\int xf(x)dx = \int xg(x)dx$. For this reason we call such a function as s(x) a Mean Preserving Spread (MPS). It seems reasonable to agree that g is riskier than f if it can be obtained from f by a MPS. Whatever definition of greater riskiness we adopt, we would certainly require that it be

---

[1]When not otherwise specified, the limits of integration are the endpoints of the closed interval over which the functions in question are defined. We shall also write $\int S(x)dx$ as $\int S(x)$.

Figure 1

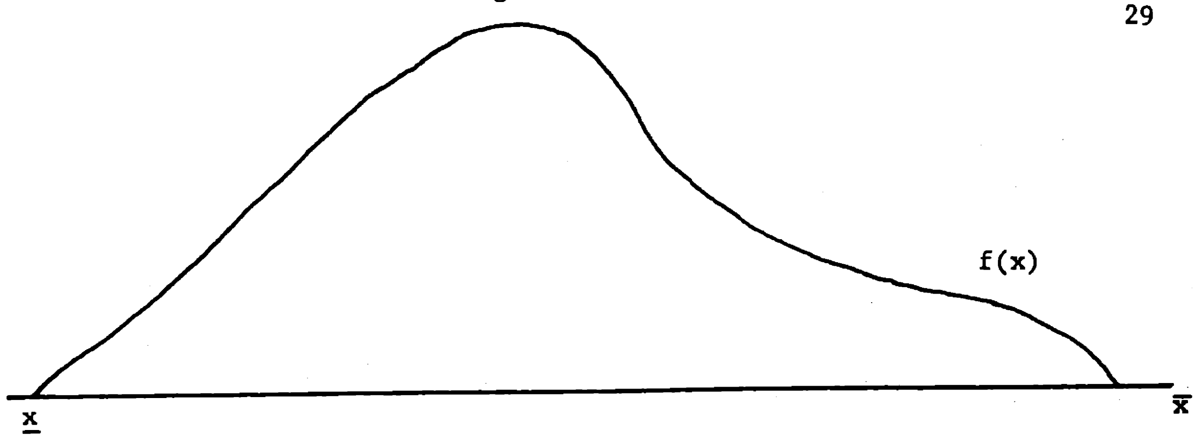


Figure 2

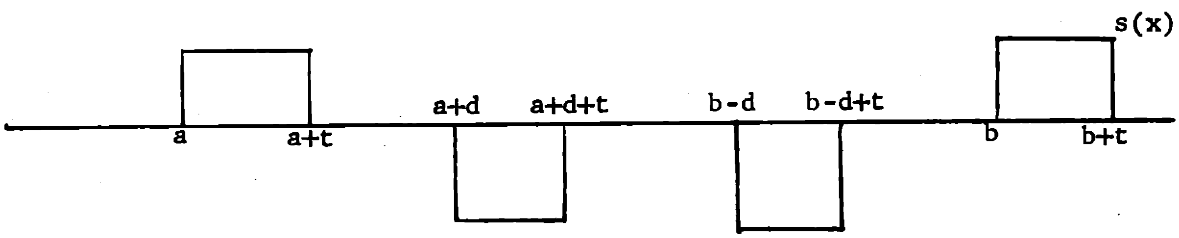
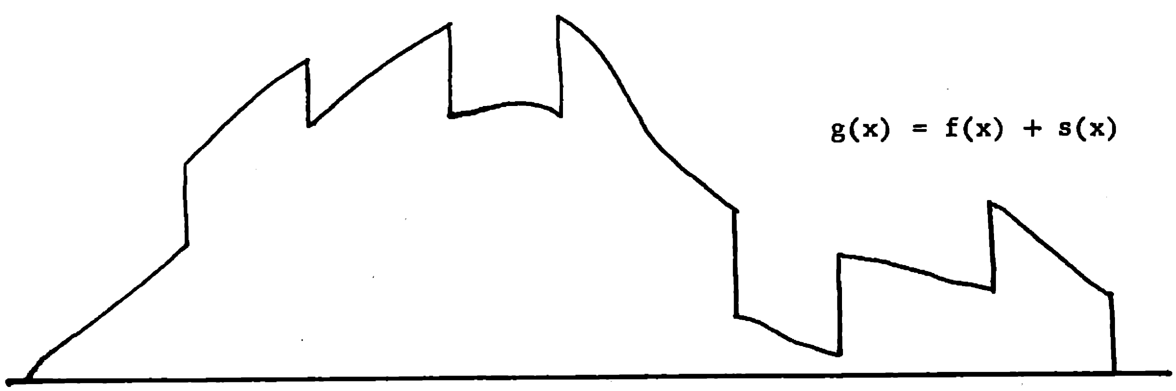


Figure 3



transitive. That is if g is riskier than h which is in turn riskier than f, then we would like to say that g is riskier than f. We then could tentatively define a relationship in the set of distribution functions by saying that g is more variable than f if there exist MPS's s_i such that

$$g = f + \Sigma s_i \tag{3}$$

There is no conceptual reason to require the sum in (3) to have a finite number of terms. However, since the space of density functions is not complete, we would have to be exceptionally careful if we were to allow infinite sums on the RHS of (3). Considerations of mathematical convenience of this sort make it desirable to phrase a definition of greater variability in terms of distribution functions rather than density functions. This is easy todo. All that is required is the integration of (1), and by implication, (3). Let $F(x) = \int_{\underline{x}}^{x} f(y)dy$ and define G(x) and S(x) similarly. We say that two random variables with cumulative distribution functions F and G differ from each other by a MPS if G = F + S where S is the integral of a MPS. (See Figure 4, the integral of Figure 2.) We would say then that G is riskier than F if $S \equiv G - F$ is the sum of integrals of MPS's. The question is, is there any simple criterion for determining whether the difference between two distributions could be accounted for by a sequence of MPS's? The answer is yes, but before we find the criterion, let us consider a second intuitive approach to the problem of defining riskiness.

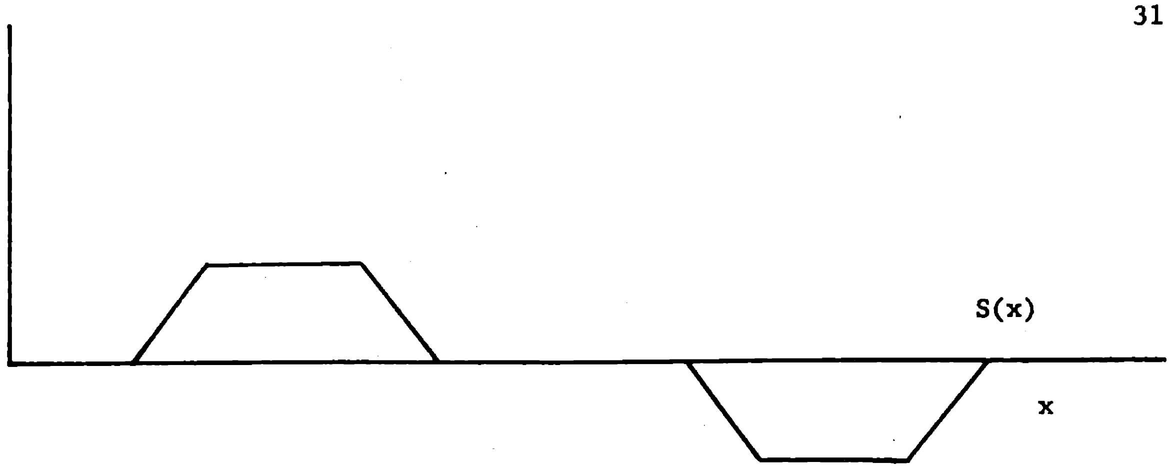


FIGURE 4

(b) Distribution Functions

There is another approach that we could take. In Figure 5 we have drawn the distribution function for X, a sure thing with mean M, and the distribution function for a random variable Y with same mean as X. Finally, we have drawn a third distribution Z which is a convex combination (mixture) of the two distributions, and therefore has the same mean. It is reasonable to say that the distribution which is a mixture of a safe and a risky distribution is "safer" than the risky distribution. There is no reason, however, to restrict ourselves to considering only distributions which intersect at their mean. More generally, we could say if two distributions have the same mean and intersect only once, the one which is initially below the other (and therefore has less weight in the tail) is safer than the other. Now let us consider three distributions, F, G, and H, all of which have the same mean. There is a unique value of x, x_1, for which F = G, and $F \lessgtr G$ as $x \lessgtr x_1$, and there is a unique value of x, $x_2 \neq x_1$, for which G = H, and $G \lessgtr H$ as $x \lessgtr x_2$. F is safer than G, and G is safer than H.

Once again we would like our definition to imply transitivity, that is, we would like to say that F is safer than H. Since F and H intersect twice, it is necessary to examine another characterization of the relationship between F and G (and G and H) to explore the implications of transitivity.

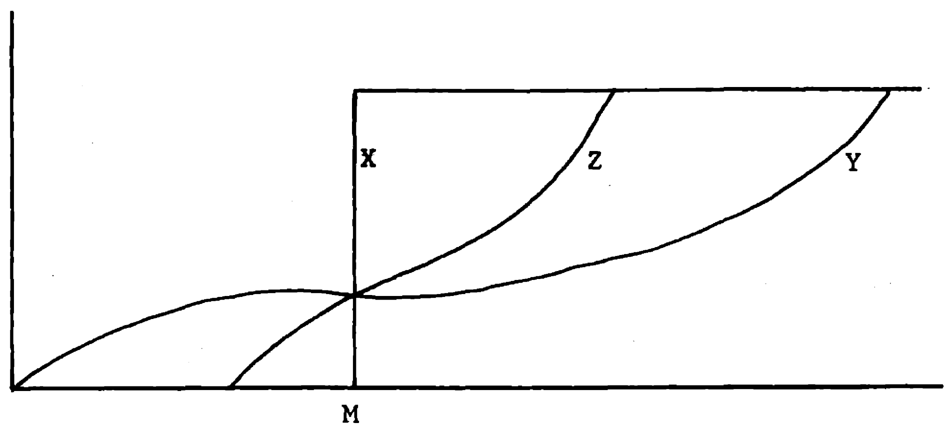


FIGURE 5

Let

(4) $$S(x) \equiv G(x) - F(x).$$

Then

(5) $$T(y) = \int_{\underline{x}}^{y} S(x)dx \geq 0 \text{ for } \underline{x} \leq y \leq \bar{x},$$

(6) $$T(\bar{x}) = 0,$$

and there exists a $\underline{x} \leq \hat{x} \leq \bar{x}$, such that

(7) $$S(x) \gtrless 0 \text{ as } x \gtrless \hat{x}$$

We may use integration by parts to show that (6) is equivalent to the statement that F and G have the same mean:

$$\int_{\underline{x}}^{\overline{x}} x dF - \int_{\underline{x}}^{\overline{x}} x dG = x(F - G) \Big]_{\underline{x}}^{\overline{x}} - \int_{\underline{x}}^{\overline{x}} (F - G) dx = 0 - T(\overline{x}) = 0$$

Condition (5) follows immediately from (6) and (7).

If we say that one distribution is safer than another if (5) and (6) hold (but not necessarily (7)), then if F is safer than G and G is safer than H, clearly F is safer than H. More generally, if M and N are distribution functions and if there exists a sequence of functions, $S_1, \ldots, S_j \ldots$ satisfying (5) and (6), such that,

$$M = N + \Sigma S_i$$

then under this definition M is riskier than N, since the difference between M and N

$$M - N = \Sigma S_i = S$$

satisfies (5) and (6). We now ask whether given two arbitrary distributions, M and N, such that the difference S satisfies (5) and (6), can S be written as the sum of sequence of functions satisfying (5), (6) <u>and</u> (7)? This question is answered affirmatively by

<u>Theorem 1</u>. Let S(x) be a function of bounded variation[1] on $[\underline{x}, \overline{x}]$ with $S(\underline{x}) = S(\overline{x}) = 0$. Then the following are equivalent:

(A) S(x) satisfies (5) and (6); and

---

[1]The restriction to functions of bounded variation is no restriction at all. We are concerned only with characterizing functions which could be the difference between two distribution functions. Since the latter are monotonic, their difference is perforce of bounded variation. The restriction (6) can be viewed as a normalization.

(B) For every $\varepsilon > 0$, there exist functions $S_i(x)$, $i = 1,\ldots,M$ satisfying (5), (6) and (7) such that

$$\int \left| \sum_{i=1}^{M} S_i(x) - S(x) \right| dx < \varepsilon$$

Proof: We have already observed that (B) implies (A). We demonstrate the reverse implication with a constructive argument which will be familiar to those who recall the development of the indefinite integral from elementary calculus. We approximate S(x) from above by a step function, each of whose steps consists of an integral number of the same size boxes. By pairing boxes above and below the axis we show that the step function can be closely approximated by a sum of functions each of which satisfies (5), (6), and (7). Let (P) = $p_0, \ldots\ldots p_n$ be a partition of $[\underline{x}, \bar{x}]$ with $\Delta p_j \equiv p_j - p_{j-1} = (\bar{x} - \underline{x})/n$ for all j. Since S is of bounded variation, it can be written as the difference between two monotone increasing functions:

$$S(x) = U(x) - V(x) \ .$$

Let

$$U_j = U(p_j) \ , \quad V_j = V(p_{j-1})$$

Then

$$0 \leq \Sigma\, U_j \Delta p_j - \int_{\underline{x}}^{\bar{x}} U(x)\,dx \leq \frac{[U(\bar{x}) - U(\underline{x})](\bar{x} - \underline{x})}{n} = \frac{D}{n}$$

$$0 \geq \Sigma\, V_j \Delta p_j - \int_{\underline{x}}^{\bar{x}} V(x)\,dx \geq \frac{[V(\bar{x}) - V(\underline{x})](\bar{x} - \underline{x})}{n} = \frac{C}{n} \ .$$

Define

$$R_j = U_j - V_j, \ R_j \geq S(x) \quad \text{for } x \in [p_{j-1}, p_j].$$

Then

$$\Sigma R_j \Delta p_j = \Sigma U_j \Delta p_j - \Sigma V_j \Delta p_j \geq \int_{\underline{x}}^{\bar{x}} U(x)\,dx - \int_{\underline{x}}^{\bar{x}} V(x)\,dx = \int_{\underline{x}}^{\bar{x}} S(x)\,dx = 0$$

Thus

$$0 \leq \Sigma R_j \Delta p_j - \int_{\underline{x}}^{\bar{x}} S(x)\,dx \leq \frac{C + D}{n}$$

We now define Q_j as a rational number such that

$$0 \leq Q_j - R_j < 1/n$$

so

$$\Sigma Q_j \Delta p_j \geq \Sigma R_j \Delta p_j \geq \int_{\underline{x}}^{\bar{x}} S(x)\,dx.$$

Hence

$$0 \leq \Sigma Q_j \Delta p_j - \int_{\underline{x}}^{\bar{x}} S(x)\,dx = \Sigma(Q_j - R_j)\,\Delta p_j + \Sigma R_j \Delta p_j - \int_{\underline{x}}^{\bar{x}} S(x)\,dx$$

$$\leq \frac{1}{n}\,\frac{(\bar{x} - \underline{x})}{n} + \Sigma R_j \Delta p_j - \int_{\underline{x}}^{\bar{x}} S(x)\,dx$$

$$\leq \frac{\bar{x} - \underline{x} + C + D}{n} = \frac{E}{n}$$

Because Q_j is rational, we can write $Q_j = h_j/k_j$, where h_j and k_j are integers, and $k_j > 0$. Define $k = \Pi k_j$, and $M_j = h_j \prod_{i \neq j} k_i$. Then

$$Q_j = M_j/k$$

Consider the step function Q(x) defined by $Q(x) = Q_j$ for $x \in [p_j - p_{j-1})$, $j = 1,\ldots,n$. We know that $Q(x) \geq S(x)$ for all x so that

$$\int_{\underline{x}}^{y} Q(x)dx \geq \int_{\underline{x}}^{y} S(x)d(x) \geq 0 \ .$$

The j^{th} step of Q(x) consists of $|M_j|$ boxes of size $B = \frac{1}{k}(\frac{\bar{x} - \underline{x}}{n})$. The boxes are stacked below the axis if $M_j < 0$ and above the axis if $M_j > 0$. Let us number all the boxes above the axis, starting at the first interval of the partition P for which $M_j > 0$, and working up and then to the right. See Figure 6. This figure displays two properties which we shall exploit. First, there are more boxes above the axis than below. This must be the case for if M^+ $(= \sum_{M_j > 0} M_j)$ is the number of boxes above the axis amd M^- $(= \sum_{M_j < 0} |M_j|)$ the number of boxes below the axis, then

$$0 = \int_{\underline{x}}^{\bar{x}} S(x) \leq \int_{\underline{x}}^{\bar{x}} Q(x) = (M^+ - M^-)B$$

so that $M^+ > M^-$. For future reference we note that $(M^+ - M^-)B \leq \int_{\underline{x}}^{\bar{x}} |Q(x) - S(x)|dx \leq \frac{E}{n}$.

Second, if we define $j^+(i)$ as the interval in the partition in which the i^{th} box above the axis occurs $(i \leq M^+)$ and define $j^-(i)$ as the interval in which the i^{th} box below the line appears, then $j^+(i) < j^-(i)$. For, suppose $j^-(i) < j^+(i)$. (Equality is not an alternative.) Then clearly

$$\int_{\underline{x}}^{p_{j^-(i)+1}} Q(x) \leq -B < 0.$$

But, by hypothesis, $\int_{\underline{x}}^{y} Q(x) \geq \int_{\underline{x}}^{y} S(x) \geq 0$ for all y.

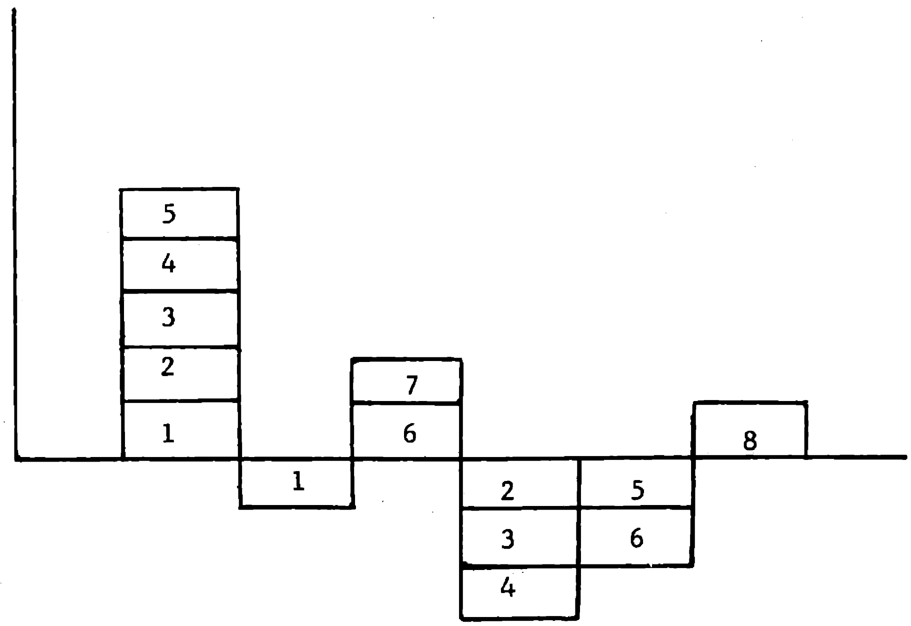


FIGURE 6

For $1 \leq i \leq M^-$, define $S_i(x)$ as follows:

$$S_i(x) = \begin{cases} \frac{1}{k} \text{ for } x \in (p_{j^+(i)}, p_{j^+(i)+1}) \\ -\frac{1}{k} \text{ for } x \in (p_{j^-(i)}, p_{j^-(i)+1}) \\ 0 \text{ otherwise.} \end{cases}$$

Clearly $S_i(x)$ satisfies (5), (6), and (7) with $\hat{x}_i = p_{j^+(i)+1}$.

By construction

$$Q(x) = \sum_{i=1}^{M^+} S_i(x) + (M^+ - M^-)B.$$

Therefore

$$\int \left| \Sigma s_i(x) - S(x) \right|$$

$$\leq \int \left| \Sigma S_i(x) - Q(x) \right. + \int \left| Q(x) - S(x) \right|$$

$$\leq \frac{E}{n} + \frac{E}{n} = \frac{2E}{n} < \varepsilon \text{ for n sufficiently large}$$

This completes the proof.

It is clear that integrals of MPS's satisfy (5), (6), and (7). The functions $S_i(x)$ constructed in the proof of Theorem II are not, however, integrals of MPS's -- as the latter are shaped like trapezoids and the former like rectangles. They are, however, limits of integrals of MPS's, and therefore we may use the proof of Theorem I to establish

Theorem II. Let S(x) be a function of bounded variation defined on some closed interval $[\underline{x}, \overline{x}]$ satisfy $S(\underline{x}) = S(\overline{x}) = 0$. Then the following are equivalent:

(A) S satisfies (5) and (6).

(B) For any $\varepsilon > 0$, there exist functions $S_i(x)$, $i = 1,\ldots,M$ such that each S_i is an integral of an MPS and

$$\int \left| \sum_{i=1}^{M} S_i(x) - S(x) \right| dx < \varepsilon$$

(c) Definition of $\underset{V}{\leq}$.

These theorems allow us to give the following simple and easy to apply definition of greater or lesser riskiness of random variables, which is the most extensive generalization of the rather clumsy notion of MPS's.

Definition 1. If F and G are two distribution functions, defined on $[\underline{x}, \overline{x}]$, then $F \underset{V}{\leq} G$ (read G is more variable than F) if and only if $S(x) = G(x) - F(x)$ satisfies (5) and (6) above.

<u>Theorem III.</u> $\underset{V}{\leq}$ is a partial order on the set of all distribution functions defined on $[\underline{x}, \overline{x}]$, considered as a subset of $L_1(\underline{x}, \overline{x})$. By this we mean that we consider $F = G$ if and only if $F = G$ a.e., or $\int |F(x) - G(x)| dx = 0$.[1]

<u>Proof</u>: It is immediate that $\underset{V}{\leq}$ is transitive and reflexive. We need only show that $F \underset{V}{\leq} G$ and $G \underset{V}{\leq} F$ implies $F = G$ a.e. If $F \underset{V}{\leq} G$ and $G \underset{V}{\leq} F$, then there are S_1 and S_2 satisfying (5) and (6) such that $F = G + S_1$ and $G = F + S_2$, so $S_1 + S_2 = 0$. Let $T_i(y) = \int_{\underline{x}}^{y} S_i(x)dx$. Then

$$0 = \int_{\underline{x}}^{y} [S_1(x) + S_2(x)] = T_1(y) + T_2(y)$$

but since $T_i(y) \geq 0$, $T_1(y) \equiv 0$. We prove that this implies $S_i(x) = 0$ a.e. Since $S_i(x)$ is of bounded variation its discontinuities form a set of measure zero. Let us call this set N. Define

$$\begin{aligned} \tilde{S}_1(x) &= 0 \text{ for } x \in N \\ &= S_1(x) \text{ otherwise.} \end{aligned}$$

Then

$$\int_{\underline{x}}^{y} \tilde{S}_1(x) = \int_{\underline{x}}^{y} S_1(x) = T_1(y) .$$

Suppose there is an $\hat{x}$ such that $\tilde{S}_1(\hat{x}) \neq 0$, say $\tilde{S}_1(\hat{x}) > 0$, then $\tilde{S}_1(x) > 0$ for $x \in (\hat{x} - \varepsilon, \hat{x} + \varepsilon)$ for some $\varepsilon > 0$ (since $\tilde{S}_1(x)$ is continuous at $\hat{x}$). Therefore $T_1(\hat{x} - \varepsilon) < T_1(\hat{x} + \varepsilon)$, a contradiction.

---

[1]It should also be observed that if $F \underset{V}{\geq} G$, and H is any other distribution defined over the interval $[\underline{x}, \overline{x}]$, $\lambda F + (1 - \lambda)H \underset{V}{\geq} G + (1 - \lambda)H$ for $0 \leq \lambda \leq 1$.

It is important to emphasize that $\underset{V}{\leq}$ is a partial rather than a complete ordering. It is almost as interesting to know which distributions are not comparable with respect to riskiness as which are. We return to this point later.

(d) Risk aversion

This definition is compatible with the standard definition of risk aversion. A risk averter is a person with a concave Von-Neuman Morgenstern utility function. We should expect, if our definition were a good one, that an increase in risk would leave such a person worse off -- decrease his expected utility -- while an increase in risk would raise his expected utility. This is the content of

Theorem IV. If $F \underset{V}{\leq} G$ and $U(x)$ is concave, then $\int_{\underline{x}}^{\bar{x}} U(x)dF(x) \geq \int_{\underline{x}}^{\bar{x}} U(x)dG(x)$.

Proof: Let $S = G - F$ and $T(y) = \int_{\underline{x}}^{y} S(x)dx$. Thus if U'' exists it is non-positive and we may prove the theorem simply by integrating by parts twice:

$$\int_{\underline{x}}^{\bar{x}} U(x)dF(x) - \int_{\underline{x}}^{\bar{x}} U(x)dG(x) = -\int_{\underline{x}}^{\bar{x}} U(x)dS(x) = \int_{\underline{x}}^{\bar{x}} U'(x)S(x)dx$$

$$= -\int_{\underline{x}}^{\bar{x}} T(x)U''(x)dx \geq 0 \ .$$

More generally, since U is concave it is absolutely continuous (which the reader may demonstrate by showing that $U(y) - U(\underline{x}) = \int_{\underline{x}}^{y} U'_-(x)dx = \int_{\underline{x}}^{y} U'_+(x)dx$

where U'_- and U'_+ are the left and right hand derivatives of U). We may therefore write

$$-\int_{\underline{x}}^{\overline{x}} U(x)\,dS(x) = \int_{\underline{x}}^{\overline{x}} U'(x)S(x)\,dx = \int_{\underline{x}}^{\overline{x}} U'(x)\,dT(x)$$

where U'(x) is the Radon-Nikodym derivitive of the measure induced by U. (See Feller [4], pp.136-138.) T is continuous and U' monotonic so the second mean value theorem of the integral calculus guarantees the existence of an $\hat{x}$ such that

$$\int_{\underline{x}}^{\overline{x}} U'(x)\,dT(x) = U'(\underline{x})\int_{\underline{x}}^{\hat{x}} dT(x) + U'(\overline{x})\int_{\hat{x}}^{\overline{x}} dT(x) = [U'(\underline{x}) - U'(\overline{x})]T(\hat{x}) \geq 0.$$

We now show that our definition of increasing risk is the widest definition which is consistent with the standard definition of risk aversion.

<u>Theorem V</u>. If

$$\int_{\underline{x}}^{\overline{x}} U(x)\,dG(x) \geq \int_{\underline{x}}^{\overline{x}} U(x)\,dG(x)$$

for every convex funtion U, then $F \underset{V}{\leq} G$.

<u>Proof</u>: Let S = G-F and $T(y) = \int_{\underline{x}}^{y} S(x)\,dy$. Since U(x) = x and V(x) = -x are both convex

$$\int_{\underline{x}}^{\overline{x}} x\,dS(x) \geq 0 \text{ and } -\int_{\underline{x}}^{\overline{x}} x\,dS(x) \geq 0,$$

so $\int_{\underline{x}}^{\bar{x}} x dS(x) = 0$ or $T(\bar{x}) = 0$. It remains to show that $T(y) \geq 0$. For fixed y let $b_y(x) = \text{Max}(y-x,0)$. Then $b_y(x)$ is convex and by hypothesis

$$0 \leq \int_{\underline{x}}^{\bar{x}} b_y(x) dS(x) = \int_{\underline{x}}^{y} (y-x) dS(x)$$

$$= yS(y) - \int_{\underline{x}}^{y} x dS(x)$$

Integrating by parts we see that

$$-\int_{\underline{x}}^{y} x dS(x) = -xS(x)\Big]_{\underline{x}}^{y} + \int_{\underline{x}}^{y} S(x) dx$$

$$= -yS(y) + T(y).$$

Therefore

$$T(y) = \int_{\underline{x}}^{\bar{x}} b_y(x) dS(x) \geq 0$$

as was to be proved.[1]

---

[1] We are indebted to David Wallace for the present form of this proof.

If F and G have the same mean, but $\int_{\underline{x}}^{y} [F(x) - G(x)]dx = T(y)$ changes sign, F and G cannot be ordered. Then there always exists two concave functions, U_1 and U_2, such that $\int U_1 dF(x) > \int U_1 dG(x)$ while $\int U_2 dF(x) < \int U_2 dG(x)$; i.e. there is some risk averse individual who prefers F to G, and another who prefers G to F. Similarly, given any differentiable function, U, which, over the interval $[\underline{x}, \bar{x}]$ is neither concave nor convex, then there exists distribution functions, F, G, and H, $F \underset{V}{\leq} G \underset{V}{\leq} H$, such that $\int U(x)dF < \int U(x)dG$ and $\int U(x)dH < \int U(x)dG$.

In short, $\underset{V}{\leq}$ defines the set of all concave functions; i.e. a function U is concave if and only if $X \underset{V}{\leq} Y$ implies $EU(X) \geq EU(Y)$.

(e) Mean variance analysis

Consider the following ordering on the set of all distribution functions with the same mean; $X \underset{q}{\geq} Y$ if $EX^2 \geq EY^2$. By arguments closely analogous to those used earlier, it can be shown that a function U is quadratic if and only if $X \underset{q}{\leq} Y$ implies $EU(X) \leq EU(Y)$. An immediate consequence of this is that if U(x) is any non-quadratic concave function, then there exists random variables X_i, i = 1,2,3, all with the same mean such that $EX_1^2 < EX_2^2$ and $EX_3^2 < EX_2^2$ but $E(U(X_1)) < EU(X_2) < EU(X_3)$, i.e. the ranking by variances and the ranking by expected utility are different. Some further properties of the ordering $\underset{q}{\leq}$ should be noted. First, $\underset{q}{\leq}$ is a complete ordering of the set of distributions with the same mean, i.e. if F and G have the same mean, either $F \underset{q}{\leq} G$ or $G \underset{q}{\geq} F$. Secondly, if $F \underset{V}{\leq} G$ then $F \underset{q}{\leq} G$. This follows as an immediate consequence of Theorem IV, since the variance is a convex function. However, even if $F \underset{q}{\geq} G$, F may not be comparable to G, under the partial ordering $\underset{V}{\leq}$, i.e. neither $F \underset{V}{\geq} G$, nor $G \underset{V}{\geq} F$.

Tobin has conjectured that mean-variance analysis may be appropriate if the class of distributions -- and thus the class of changes in distributions -- is restricted. This is true but the restrictions required are as far as is presently known, very severe. Tobin's proof is (as he implicitly recognizes in [13, p.20-21]) valid only for distributions which differ only by "location parameter," (See Feller [4, p.134] for a discussion of this classical concept.) That is, Tobin is only willing to consider changes in distributions from F to G if there exist a and b ($a > 0$) such that $F(x) = G(ax + b)$. Such changes amount only to a change in the centering of the distribution and a uniform shrinking or stretching of the distribution -- equivalent to a change in units.

There has been some needless confusion along these lines about the concept of a two-parameter family of distribution functions. It is undeniable that all distributions which differ only by location parameters form a two-parameter family. In general, what is meant by a "two-parameter family"? To us a two-parameter family of distributions would seem to be any set of distributions such that one member of the set would be picked out by selecting two parameters. As Tobin has put it, it is "one such that it is necessary to know just two numbers in order to describe the whole distribution." Technically that is, a two-parameter family is a mapping from E^2 into the space of distribution functions.[1] It is clear that for this broad definition of two-parameter family, Tobin's conjecture cannot possibly hold, for nothing restricts the range of this mapping.

---

[1] Or from some subset of E^2; we might restrict one or both of our parameters to be non-negative.

Other definitions of two-parameter family are of course possible. They involve essentially restrictions to "nice" mappings from E^2 into the space of distribution functions, e.g. a family of distributions with an explicit algebraic form containing only two parameters which can vary. It is easy, however, to construct examples where if the variance, σ^2, changes with the mean, μ, held constant, $\frac{\partial T(y)}{\partial \sigma^2}$ changes sign in the interval $(\underline{x}, \bar{x}]$ where $T(y, \sigma^2, \mu) = \int_{\underline{x}}^{y} F(x; \sigma^2, \mu)$; i.e. there exist individuals with concave utility functions who are better off with an increase in variance.[1]

[1] Consider, for instance, the family of distributions defined as follows: $(a, c > 0)$

$$F(x; a,c) \begin{cases} 0 & x \leq 1 - .25/a \\ ax + .25 - a & 1 - .25/a \leq x \leq 1 + 2c - .5/c - a \\ cx + .75 - 3c & 1 + 2c - .5/c - a \leq x \leq 3 + .25/c \\ 1 & x > 3 + .25/c \end{cases}$$

Two members of the family with the same mean but different variances are depicted in Figure 7(a). They clearly do not satisfy (5). The density functions are illustrated in Figure 7(b).

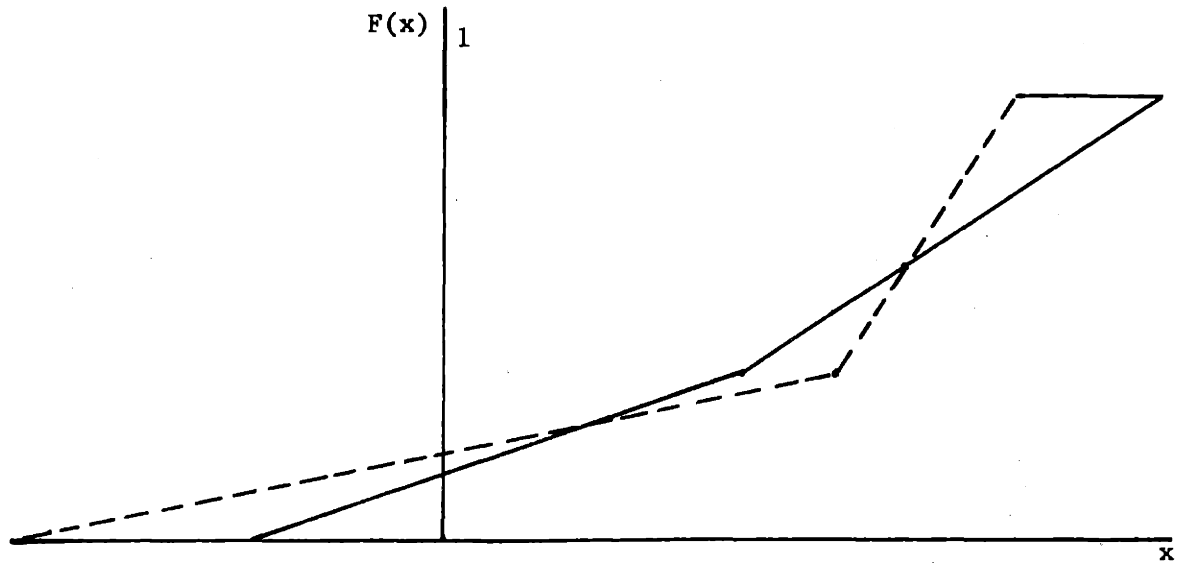


FIGURE 7(a)

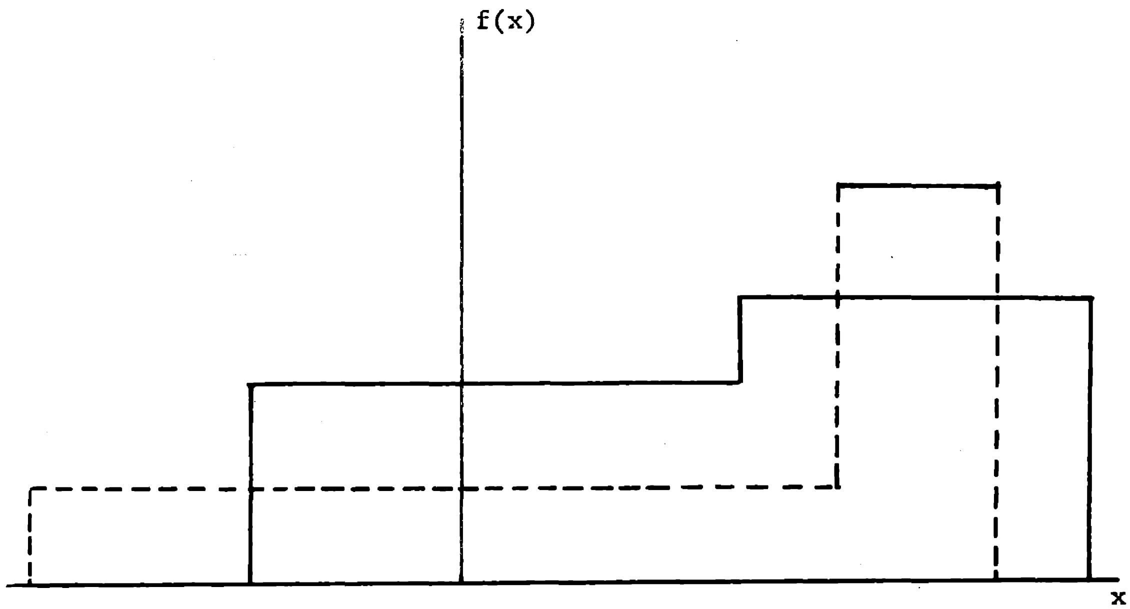


FIGURE 7(b)

II. THE EFFECTS OF INCREASING RISK

Our original problem was to examine the effect of an increase in risk on some control parameter α, where α was chosen to maximize

$$\int U(X, \alpha)\, dF(x) \tag{8}$$

The first order condition for utility maximization is

$$\int \frac{\partial U(X, \alpha)}{\partial \alpha} dF(x) = 0 \tag{9}$$

Assume there is a unique solution to (9), α * and, that, in the neighborhood of α *, U_α is monotone decreasing in α.[1] Then if $U_\alpha(X,\alpha)$ is a convex (concave) function of X, an increase in riskiness will lead to an increase (decrease) in $EU_\alpha(X,\alpha)$. But since $EU_\alpha(X,\alpha^*) = 0$, before the change, now $EU_\alpha(X,\alpha^*) > 0$ (< 0). Hence α * is increased (decreased). In any particular problem, the question becomes one of ascertaining the conditions under which $U_\alpha(X,\alpha)$ is concave (convex) in X. We examine below four problems of some economic interest.

(a) <u>Savings and uncertainty</u>[2]

There are at least two stories of how uncertainty about the rate of return on savings affects the savings rate. (i) A risk averse individual, in order to ensure his "minimum standard of living" saves more in the face of uncertainty. (ii) A risk averse individual is discouraged from saving

---

[1]We assume $U_{\alpha\alpha} < 0$ for all (α, x) in the relevant region. In the examples below we shall have occasion to demonstrate the uniqueness of *.

[2]For a fuller discussion of this and related problems, see [6,12].

by the uncertainty of the return -- "a bird in the hand is worth two in the bush." We shall show that whether the savings rate increases or decreases (in our simplified model) depends on whether relative risk aversion [1,9] is less than or greater than unity.

We consider a risk averse individual who has a given wealth, W_o, which he wishes to allocate between consumption today and consumption tomorrow. What he does not consume today, he invests; at the end of the period his investment yields the random return e per dollar invested. He wishes to allocate W_o between the two periods to maximize two period expected utility,

$$E\{U(C_1) + (1-\delta (U(C_2)\} = E[U((1-s)W_o) + (1-\delta)U(sW_o e)]$$

where s is the savings rate and δ the pure rate of time preference. The necessary and sufficient condition for utility maximization is that

$$U'((1-s)W_o) = E\ [U'(sW_o e)\ (1-\delta)e$$

Whether s decreases or increases as risk increases depends on whether $U'(C_2)e$ is concave or convex in e, which can be shown after some manipulation to depend on whether

$$U''(C_2)(1-R(C_2)) - U'(C_2)(R'(C_2)) \lesseqgtr 0$$

where $R(C) = -U''(C)C/U'(C)$ is relative risk aversion and $C_2 = sW_o e$ is consumption in the second period. If relative risk aversion is constant, savings is unaffected if relative risk aversion is unity (the Bernoulli utility function), decreased if it is less than unity, increased if it is greater. If relative risk aversion is increasing, but less than or equal to one throughout the relevant range, then savings are increased.

(b) A Multi-stage planning problem[1]

Consider a simple economy in which the final consumption good is produced by labor and an intermediate commodity y,

$$Q = F(L_2, y)$$

while y is produced by labor alone:

$$y = C(L_1)$$

The economy faces an overall labor constraint L, so

$$L_1 + L_2 = L,$$

In the absence of uncertainty, maximization of Q simply requires

$$F_1 = F_2 G'.$$

Assume that there is uncertainty associated with the production of y:

$$y = G(L_1) + e$$

where e has mean zero. We wish to maximize EQ; we require

$$E[F_1 - F_2 G'] = 0$$

If e becomes more variable, what happens to L_1 (and L_2?). This depends on the sign of

$$F_{122} - G' F_{222}$$

Assume F is a constant elasticity of substitution production function:

$$F(L_2, y) = (\delta L_2^{-\rho} + (1-\delta) y^{-\rho})^{-1/\rho}$$

---

[1]This problem was posed to us by M. Weitzman.

where σ, the elasticity of substitution is given by, $\sigma = \frac{1}{1+\rho}$. Since F is concave and homogeneous of degree one, $F_{22} < 0$ and

$$F_{222} = -(F_{122}L_2 + F_{22})/y \ .$$

Thus if $F_{122} < 0$, then $F_{222} > 0$ and $F_{122} - G'F_{222} < 0$ and an increase in uncertainty will call for an increase in L_2. It can be shown that if $Z = F/y$ and $H = \delta(1+\rho)/y > 0$, then $F_{22} = H \cdot (\delta Z^{2\rho+1} - Z^{\rho+1})$. Since $F_{22} < 0$, $b(2) = (\delta Z^{2\rho+1} - Z^{\rho+1}) < 0$. Furthermore,

$$F_{122} = F_{221} = H\, b'(Z) \frac{dz}{dL_2} \ .$$

Since $\frac{dz}{dL_2} > 0$, the sign of $b'(Z)$ determines the sign of F_{221}. Observe that sign $b'(Z)$ = sign $Zb'(Z)$, while $Z'b(Z) = b(Z) + (2\rho\delta Z^{2\rho+1} - \rho z^{2\rho+1})$. If $\sigma > 1$, then $0 < \rho < 1$ and $2\rho\delta Z^{\rho+1} < \rho\sigma Z^{2\rho+1}$ from which it follows that $Zb'(Z) < (1+\rho)b(Z) < 0$. Thus if $\sigma > 1$, increases in uncertainty will lead to the allocation of more labor to the earlier stage of production.

Consider the other extreme case, where Q is produced by a fixed coefficients production function $Q = \min(L_2, y)$. Then if Ψ is the distribution function of e

$$E(Q) = \int_{-\infty}^{L_2-G(L_1)} [G(L_1) + e]\,d\Psi(e) + L_2(1 - \Psi(L_2 - G(L_1))$$

$$= \int_{-\infty}^{\bar{L}-L_1(G(L_1)} [G(L_1) + e]\,d\Psi(e) + (\bar{L} - L_1)(1 - \Psi(\bar{L} - L_1 - G(L_1))$$

so that maximization of EQ requires

$$[G'(L_1) + 1]\Psi(\bar{L} - L_1 - G(L_1)) = 1 \ .$$

The second order conditions are satisfied, since $G''\psi - \phi(G' - 1)^2 < 0$, where ϕ is the density function corresponding to ψ; hence there is a unique maximum. Whether L_1 increases or decreases depends solely on whether $\psi'(\bar{L} - L_1 - G(L_1))$ increases or decreases; either is clearly possible. Note that if y is also produced by a constant returns to scale production function

$$y = L_1$$

then the optimal value of L_1 is simply given by

$$F(\bar{L} - 2L_1) = \frac{1}{2}$$

so what happens to L_1 depends completely on whether the median of e increases or decreases.

(c) A portfolio problem

An individual with initial wealth of W_o, wishes to allocate a fixed amount of wealth, W_o, between money, which yields a zero rate of return, and a risky asset which yields a random rate of return e, so as to maximize the expected utility of his terminal wealth:

$$EU(W) = E\,(U(W_o(1 + ae)))$$

where a is the fraction of his wealth invested in the risky asset. U is assumed to be concave. A necessary and sufficient condition for utility maximization is

$$EU'e = 0$$

What happens to a if e becomes riskier depends on whether U'e is concave or convex, i.e. whether

$$U''(1 - R + W_oA) + U'(W_oA' - R') \gtrless 0 ,$$

where, as before, $R = -U''W/U'$, the Arrow-Pratt measure of relative risk aversion, and $A = -U''/U'$, the measure of absolute risk aversion. A sufficient condition for an increase in uncertainty leading to an increased allocation to the safe asset is that relative risk aversion be less than or equal to unity, and that absolute risk aversion be non-increasing and relative risk aversion be non-decreasing. The Bernoulli utility function clearly satisfies these conditions.

Observe that taxation of earnings from investments amounts to a particular kind of change in the distribution of the payoff from an investment. The results we obtain here are much weaker than the corresponding results for the effect of an income tax with full loss offset, but they are identical to those obtained in [11] for an income tax with no loss offset. Such a tax can be viewed as a mean preserving reduction in risk plus a reduction in mean (see Figure 8), by shifting the distribution to the left. The latter will lead to an increase in the demand for the safe asset if there is decreasing absolute risk aversion, a condition already included in the conditions for a mean preserving reduction in risk leading to an increase in the demand for the safe asset.

(d) Choice of output level for a competitive firm

In the examples considered so far, the conditions we have obtained under which unambiguous statements about the effects of increased variability have been essentially identical to those obtained earlier in comparisons between safe and risky situations. There are however problems in which the latter comparisons can be made under weaker conditions than the former.

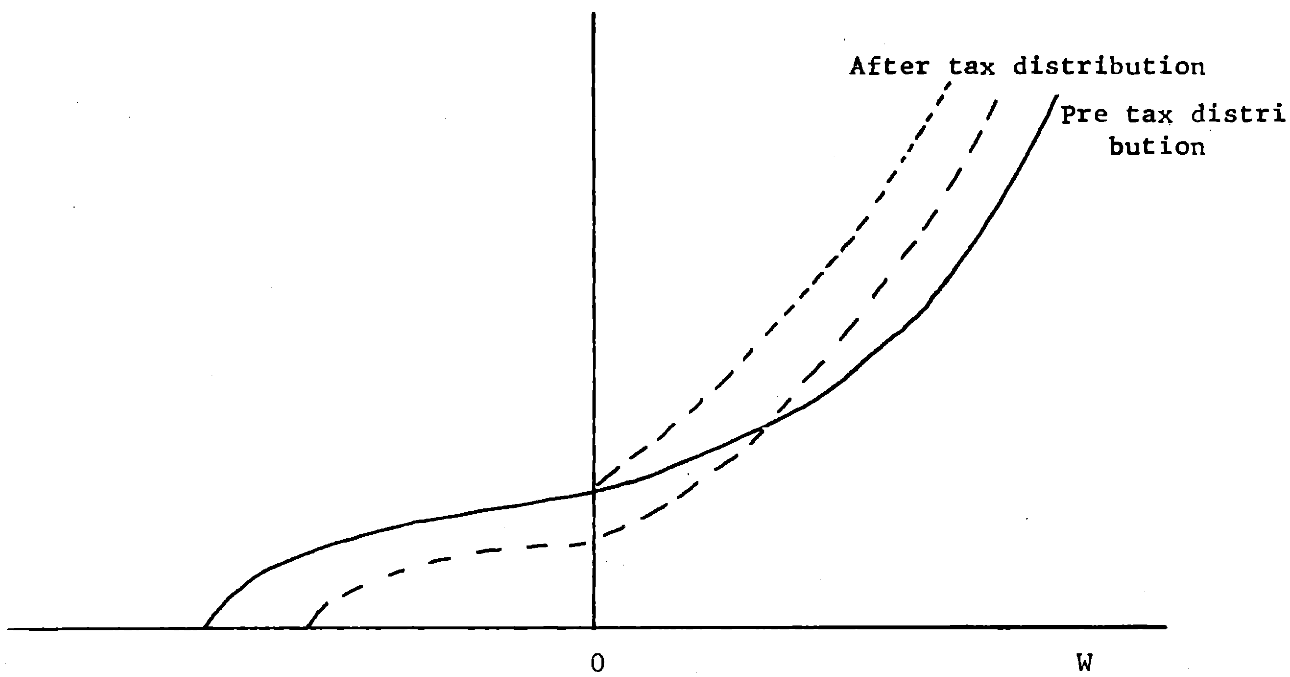


FIGURE 8

In the notation of Section 3.1, in a certain situation, we choose x so that

$$U_{\alpha}(x, \hat{\alpha}) = 0$$

Whether $\hat{\alpha} \gtrless \alpha^*$, where α^* is the solution to (9), depends simply on whether

$$EU_{\alpha}(x, \hat{\alpha}) \gtrless U(Ex, \hat{\alpha})$$

Jensen's inequality allows us to make unambiguous statements whenever U is concave or convex in x; but this is the same condition under which we have been able to make unambiguous statements for a wider class of problems. In the following problem,however, we can make unambiguous statements even when the first order condition is neither concave nor convex.

Consider a competitive firm which must decide today on the level of output tomorrow, although the price, p of output Q is uncertain. It wishes to maximize expected utility of profits, $U(\pi)$, where U is concave[1] and where

$$\pi = pQ - C(Q)$$

where C(Q) is the cost function, and is convex. A necessary and sufficient condition for an optimum is that

$$\frac{EU'p}{EU'} = C'(Q^*)$$

If the producer is risk neutral or if there is no variability in p, profit maximization requires price equal marginal cost,

$$Ep = C'(\hat{Q}).$$

[1]For a discussion of the case of constant absolute risk aversion, see [8].

$Q^* \lesseqgtr \hat{Q}$ as $\frac{EU'p}{EU'} \gtreqless Ep$, i.e., as $E[(U' - EU')(p - E(p))] \gtreqless 0$. But since $U'' < 0$, $U'(p) \gtreqless U'(E(p))$ as $p \lesseqgtr E(p)$, so $E[(U' - EU')(p - E(p))] = E[(U' - U'(Ep))((p) - E(p))] < 0$. Hence, there is always less output under uncertainty than under certainty.

REFERENCES

[1] Arrow, K. J., "Aspects of the Theory of Risk Bearing," (Helsinki, 1965).

[2] Bessler, S. A. and A. F. Veinott, Jr. "Optimal Policy for a Dynamic Multi-Echelon Inventory Model," Naval Research Logistics Quarterly, XIII (Dec., 1966).

[3] Diamond, P. A., "Savings Decisions Under Uncertainty," Working Paper No. 71, Institute of Business Economic Research, (University of California, Berkeley, June, 1965).

[4] Feller, William, An Introduction to Probability Theory and Its Applications, Vol. II (New York, 1966).

[5] Hadar, J. and Russel, W. R., "Rules for Ordering Uncertain Prospects," American Economic Review (forthcoming).

[6] Hahn, F. H., "Savings Under Uncertainty" LSE mimeo (May, 1968).

[7] Hanoch, G. and H. Levy, "The Efficiency Analysis of Choices Involving Risk," Review of Economic Studies (forthcoming).

[8] McCall, J. J., "Competitive Production for Constant Risk Utility Function," Review of Economic Studies, XXXIV, (October, 1964).

[9] Solow, R. M., "A Note on Dynamic Multipliers" Econometrica XIX, (July, 1961).

[10] Stiglitz, J. "The Effects of Wealth, Income and Capital Gains Taxation on Risk Taking," Quarterly Journal of Economics (forthcoming).

[11] ___________, "Liquidity Preference and the Term Structure of Interest Rates," Cowles Foundation Discussion Paper, 252, (August, 1968).

[12] Tobin, James, "The Theory of Portfolio Selection," in Hahn, F. and F. Brechling, eds., The Theory of Interest Rates, (London, 1965).

Chapter 3

CHANGING DEMAND: ITS COSTS AND CONSEQUENCES

This paper uses the familiar distinction between fixed and variable factors of production to examine the effects of changing demand on firm behavior. This distinction, central to the classical theory of the firm makes little sense if attention focuses only on the short run or, if equivalently, the firm is conceived of as operating in a static certain world.

The first section of this paper shows how varying market conditions justify distinguishing between the two classes of productive factors. The source of this variation may be either changes through time or random fluctuations. The formal equivalence of a dynamic model of firm behavior with a model of firm behavior under uncertainty is demonstrated and the relationship of changes over time with random fluctuations is discussed. The equivalence of the time and uncertainty models is useful for it makes it possible to use the results of chapter two on the definitions and consequences of increased variation in both models.

In section II random demand functions are discussed; the difficulties of defining increased variability in demand are examined. The effects of increased variability on profits are explored in a special case, which avoids these difficulties. In section III the consequences of increased variability in output on costs of production and factor proportions are discussed.

I

FIXED AND VARIABLE FACTORS OF PRODUCTION

The standard theory of the firm funs roughly as follows. A firm may produce an output Q by combining variable factors L which it buys with fixed factors K which it already has on hand. The amount of variable factors required to produce Q depends on K; that is, Q is a function of K and L, $Q = Q(K,L)$. For any fixed K this function can be inverted to give L as a function of $Q, L = L(Q,K)$. If $C = C(L)$ is the cost of hiring variable factors in amount L then the variable cost to a firm with fixed factors in amount K of producing an output Q is $M(Q;K) = C[L(Q;K)]$. A revenue curve $R(Q)$ gives the revenue the firm receives as a function of the amount it sells. The firm chooses output so as to maximize its cash flow $G(Q,K) = R(Q) - M(Q;K)$. If cash flow is maximized at Q^*, profits will be $C(Q^*,K)$ less the cost of the fixed factor of production, $T(K)$. Clearly Q^* is a function of K and thus so are profits; $\pi(K) = G(Q,K) - T(K)$. If the firm could choose K, it would do so in such a way as to maximize $\pi(K)$. But if the firm is allowed to choose K, the distinction between fixed and variable factors of production is embarrassingly bereft of content. On the other hand, if K is simply assumed given, the theory does not say much about what effects the demand for fixed as opposed to variable factor of production.

To restore meaning to the distinction between fixed and variable factors of production is straightforward. We need only recognize that the short run is but a part of the long run and lift the firm out of a static certain world and puts it in a richer, dynamic or uncertain world.[1] We assume that instead of operating in only one epoch, the firm operates in different epochs with different market conditions. The fixed factor of production is distinguished from the variable factor in that the firm is constrained to operate with one level of the fixed factor through these different epochs, while it may adjust the amount of the variable factor it uses to different market conditions. We may conceive of these different epochs either as different periods in time or as different states of the world in an uncertainty context. The firm is presumed to maximize a weighted sum of its profits in the various epochs in which it operates -- the weights being discount factors in the time interpretation and probability weights in the uncertainty interpretation. We shall continue to use K to refer to fixed factors of production and L to variable factors. We assume that the firm has a neo-classical production function. That is Q(K,L) is homogeneous of degree one (constant returns), concave (diminishing returns), and has as many continuous derivatives as the argument requires.

(a) Time interpretation

The firm is to operate from time $\underline{\theta}$ to time $\overline{\theta}(\underline{\theta} < \overline{\theta})$ with a fixed capital stock K. It can adjust $L(\theta)$, the amount of labor it hires to market conditions at time θ, defined by a revenue function $R(Q,\theta)$ and a wage rate $w(\theta)$. We showed in chapter one that the existence of cost of adjustment functions may lead the firm to behave in just this way. We analyze now the implications of the firm having to operate from $\underline{\theta}$ to $\overline{\theta}$ with a single

[1] Our approach is similar to that taken by Stigler [4] and Hart [2].

capital stock. We assume that it can adjust $L(\theta)$, the amount of labor it hires, at each moment of time to market conditions.

Revenue is a function of time and sales, $R = R(S,\theta)$. The firm is assumed to sell all it produces, i.e., $S = Q(K,L)$, so we shall generally make no distinction between sales and output and write $R = R(Q,\theta)$. For all θ, $R(Q,\theta)$ is a strictly concave function of $Q \geq 0$, that is, marginal revenue is a decreasing function of sales. It is convenient as well as sensible to assume that zero revenue is gained by selling nothing, that is, $R(0,\theta) = 0$. We shall also assume that $R(Q,\theta)$ is bounded above for all θ.

If $w(\theta)$ is the cost of hiring one unit of labor in period θ, a firm with capital stock K will choose L so as to gain a cash flow of

$$(1) \qquad G(K,\theta) = \max_{L > 0} \left\{ R(Q(K,L),\theta) - w(\theta)L \right\} .$$

This maximum clearly exists and is finite since $R(Q,\theta)$ is bounded. The firm discounts cash flow received in period θ by $\delta(\theta) > 0$. If $\rho(> 0)$ is the cost, discounted to $\underline{\theta}$, of renting one unit of capital from $\underline{\theta}$ to $\overline{\theta}$ (we assume no physical depreciation during the period) the profits of a firm with capital stock K are given by

$$(2) \qquad V(K) = \int_{\underline{\theta}}^{\overline{\theta}} G(K,\theta)\, \delta(\theta)d\theta - \rho K.$$

Since we shall take δ to be bounded -- it is reasonable to assume $\delta \leq 1$ -- $V(K)$ is also bounded above. We have shown (Lemma I, Chapter 1) that $G(K,\theta)$ is strictly concave. From this it follows immediately that $V(K)$ is a strictly concave function of K. This guarantees that V is continuous, differentiable almost everywhere and that first order conditions for a maximum are both necessary and sufficient.

(b) Uncertainty Interpretation

The same model can be used to analyze the behavior of a firm which is uncertain about market conditions and must choose the amount of capital which it will use before it finds out either the cost of labor or the demand function it faces. Once these are known, the firm does the best it can given the decision it has already made about the fixed factor of production. Let $R(Q,\theta)$ be a family of revenue functions indexed by the random variable θ. Different values of θ represent different states of the world. The cost of hiring a unit of labor is a function of the state of the world. The cost of hiring a unit of labor is a function of the state of the world, $w(\theta)$. It is convenient to assume that θ belongs to some closed interval $[\underline{\theta},\overline{\theta}]$ and has a distribution function, $F(\theta)$; $F(\underline{\theta}) = 0$, $F(\overline{\theta}) = 1$. Assume for the moment $F(\theta)$ is differentiable so that $f(\theta) = F'(\theta)$ is a density function. We assume R and w satisfy the conditions set out above and define again $G(K,\theta) = \max_{L \geq 0} \{R(Q,\theta) - w(\theta)L\}$. If r is the cost hiring one unit of capital -- which is independent of θ --

$$w(K,\theta) = G(K,\theta) - rK$$

gives the maximum profits which the firm can make given that it has decided to hire K units of capital and that the state of the world turned out to be θ. If the firm is an expected profit maximizer, it will try to choose K so as to maximize $w(K) = E[w(K,\theta)] = \int_{\underline{\theta}}^{\overline{\theta}} G(K,\theta)f(\theta)d\theta - rK$. We will call this maximum $\pi^*(K)$. That is, $\pi^*(K) = \max_K \pi(K)$. The argument used above to prove V(K) concave and bounded applies to $\pi(K)$. The only mathematical difference between the time and uncertainty interpretation is that the function used to weight cash flow in the time interpretation did not integrate to unity while the function weighting G(K,θ) in the uncertainty interpretation is a probability density function and perforce its integral is one. This is not an essential difference for we can reformulate the time problem to make it equivalent to the uncertainty problem. Let

$$\gamma = \int_{\underline{\theta}}^{\overline{\theta}} \delta(\theta)d\theta.$$

Then if K maximizes V(K), it also maximizes

$$\pi_\gamma(K) = 1/\gamma\ V(K) = \int_{\underline{\theta}}^{\overline{\theta}} G(K,\theta)f(\theta)d\theta - rk \text{ where } f(\theta)$$

$$= \frac{\delta(\theta)}{\gamma} \text{ and } r = \rho/\gamma.$$[1]

---

[1]Nothing about the uncertainty problem requires that the distribution function F(θ) be differentiable. If θ has a discrete distribution, we may define V(K) by

$$\pi(K) = \sum_{i=0}^{\infty} G(K,\theta_i)P_i - rK,$$

or using Stieltjes integrals, combine both cases by writing

$$\pi(K) = \int_{\underline{\theta}}^{\overline{\theta}} G(K,\theta)dG(\theta) - rK.$$

The interpretation of these equations for the time problem is straight forward. A discounting function which gives weight to only a countable number of time periods is a discounting function appropriate for use in discrete time problems.

The equivalence of the time and uncertainty formulations is convenient, for it suggests that the tools of probability theory may be used to analyze both models; in particular chapter two's definition of increasing variability may be applied to variation in time. We discuss in the next section how this definition may be applied to demand functions. Before doing so it will be well to examine somewhat more closely the relation of random variables to functions which vary over time. After exploring this connection we shall for the most part use the language of probability theory to discuss the model.

Suppose f(t) is a function on (0,1). Then we may define a distribution function on the range of f as follows:

(3) $$G(x) = \mu\{t : f(t) \leq x\}$$

where $\mu\{A\}$ denotes the measure of the set A. Clearly G is an increasing non-negative function with $G(-\infty) = 0$ and $G(\infty) = 1$ so that G is a distribution function. If f is defined on (0,T) we can similarly define a distribution function by $G(x) = \frac{1}{T}\,\mu\{t : f(t) \leq x\}$. Every function f(t) on (0,T) corresponds to a distribution function and thus a random variable. If f(t) and $\hat{f}(t)$ correspond to G and $\hat{G}$ then we shall say that f is more variable than $\hat{f}$ if and only if $\hat{G} \underset{v}{\leq} G$ where $\underset{v}{\leq}$ is the partial ordering defined in chapter two. This may seem somewhat at odds with common usage for it pays no attention to the order of increase or decrease in f(t). It would, for example, allow us to say that a smoothly rising monotonic function was more variable than one that jumped about all the time.

In Figure 1, for example, f(t) could be more variable than g(t).

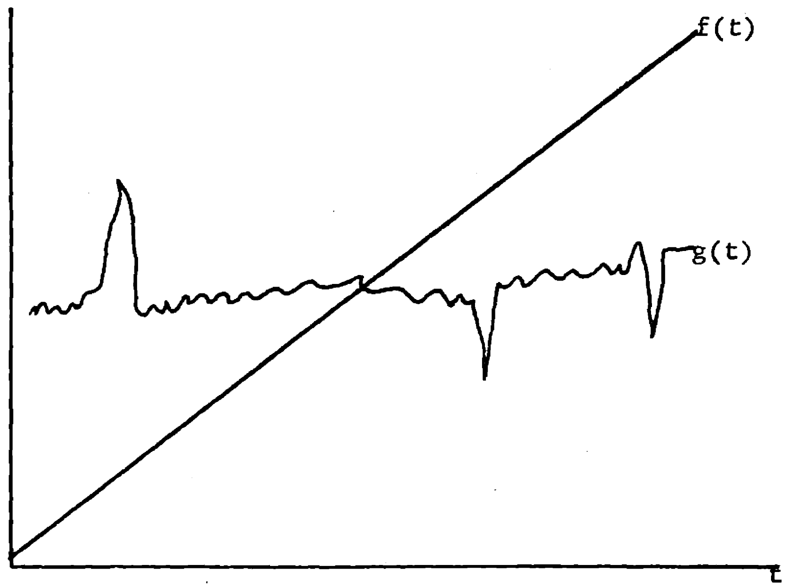


Figure 1

Nonetheless, even though this convention may appear an unnatural wrench of the language, it is appropriate for our model. In it the cost of adjustment depends on neither speed of adjustment nor on the number of turning points. What counts is how much the firm has to adjust, how long <u>in toto</u> the firm spends in a particular position, not how often it moves in or out of that position.

Discounting complicates matters, but not seriously. If f(t) is defined on [0,T], but what happens at t is discounted back to time 0 by b(t)(>0), it would still be possible to define a distribution function $G(x) = \frac{1}{T}\mu\{t: f(t) \leq x\}$ but it would be misleading. For our purposes we need a new measure, one which reflects the weighting provided by the discount function. It is easy to form a distribution function which reflects the fact that what happens at t is more important than what happens at t + t' Consider first the case where f(t) is a step function. Let $[\alpha_0, \ldots, \alpha_n]$ be a partition of [0,T] and define f(t) by

$$f(t) = f_i \text{ for } \alpha_{i-1} < t \leq \alpha_i; \qquad i = 1, \ldots, n$$

Further suppose that what happens between α_{i-1} and α_i is worth b_i. We are going to stretch and/or contract the interval [0,T] in a way which reflects this discounting.

$$\text{Let, } \beta_i = \sum_{j=0}^{i} b_j(\alpha_j - \alpha_{j-1}); \qquad i = 1,\ldots,n, \text{ and } \beta_0 = \alpha_0 = 0.$$

Consider a new function $f_b(t)$ defined on $(0,\beta_n)$ by

$$f_b(t) = f_i; \quad \text{for } \beta_{i-1} < t \leq \beta_i$$

Then $G(x) = \frac{1}{\beta_n} \mu[t: f_b(t)\ x]$ is a distribution function which reflects the fact that different weights are given to what happens at different times.

The same trick works for continuous functions. Let $b(t) > 0$ denote the weight accorded f(t) at time t. Define $\beta(t) = \int_0^t b(t)dt$. If f(t) is defined on [0,T], we consider $f_b(t)$ defined on [0, (T)] by

$$f_b(t) = f[\beta^{-1}(t)].$$

Since $b(t) > 0$, $\beta(t)$ is monotone increasing and thus invertible. The distribution function corresponding to f(t) discounted by b(t) is given by $G(x) = [\beta(T)]^{-1} \mu[t:f_b(t) \leq x]$.

II

INCREASING VARIABILITY OF DEMAND

How is the firm's behavior changed by its inability to vary both factors of production? What are the effects on profit, output, and employment of factors? It is questions like these which this model provokes and we shall now try to answer them. To begin, it is useful to note that if demand and wages are constant, it makes no difference at all if some factors of production are fixed and some variable. If $R(Q,\theta) = R(Q)$ and $w(\theta) = w$ for all θ[1] then the firm will obviously in each period of time choose to operate with the same labor force. The freedom to change the labor force is worth nothing. Similarly, if the firm could costlessly change its capital stock it would not choose to do so. The constraint that the firm may not vary its capital stock is no constraint at all if demand is constant. However, if demand is not constant, the firm probably will choose to change its labor force and would, if it could, change its capital stock to adjust to different market conditions. Only if demand varies does it matter that the firm cannot adjust all factors of production costlessly and instantaneously. One suspects that, in some sense, the more demand varies the more it matters, the more would, for example, the profits a firm could make if it could vary both factors of production exceed those it could make if capital were fixed. Unfortunately, this conjecture is so vaguely and loosely phrased that we cannot decide whether it is true or false. To do so, it is necessary to give rather precise mathematical content to the notion of increasing variability of demand.

---

[1] We henceforth assume that wages are constant in order to focus on the effects of changing demand.

This is very difficult as the demands we are considering are not capable of summary by a few -- or even a countable number of -- parameters. Although the partial ordering, $\underset{v}{\leq}$, defined in chapter two does a reasonable job of classifying random variables as more or less variable, it cannot in general be applied to demand functions indexed by a random variable. The structure of demand can be described by a selection and a weighting of some revenue functions from among all possible revenue functions. (Technically, one could describe the structure of demand by a measure on the space of all concave functions.) It is not clear what one would in general mean by a change in the structure of demand which increased the variability of demand, but we imagine that one requirement would be, that if output were produced at constant marginal cost, and profits maximized by setting production at the point at which marginal revenue equalled marginal cost, then an increase in the variability of demand would result in an increase, or at the very least no decrease, in the variability, of output. It is, however, easy to give an example of a change in demand which with one cost structure leads to an increase in the variability of output and which with another leads to a decrease in the variability of output.

In Figure 2 are drawn four marginal revenue curves, (Q is represented on the horizontal axis). The configuration A,B is the same as the configuration C,D. We may define a structure of demand by assigning weights to these curves. One such demand structure D might involve weighting each of the curves equally. In an obvious notation

$$D = (P_A = 1/4;\ P_B = 1/4;\ P_C = 1/4;\ P_D = 1/4).$$

Another possible structure D' involves weighting only curves A and D

$$D' = (P_A = 1/2;\ P_B = 0;\ P_C = 0;\ P_D = 1/2).$$

Is the move from D to D' an increase or decrease in the variability of demand? If the (constant) marginal cost is S, output is unchanged by the shift in demand, if it is greater than S, the shift in demand leads to an increase in the variability of output. But if cost is less than S, the shift in demand leads to a decrease in the variability of output. Clearly our question has no answer.

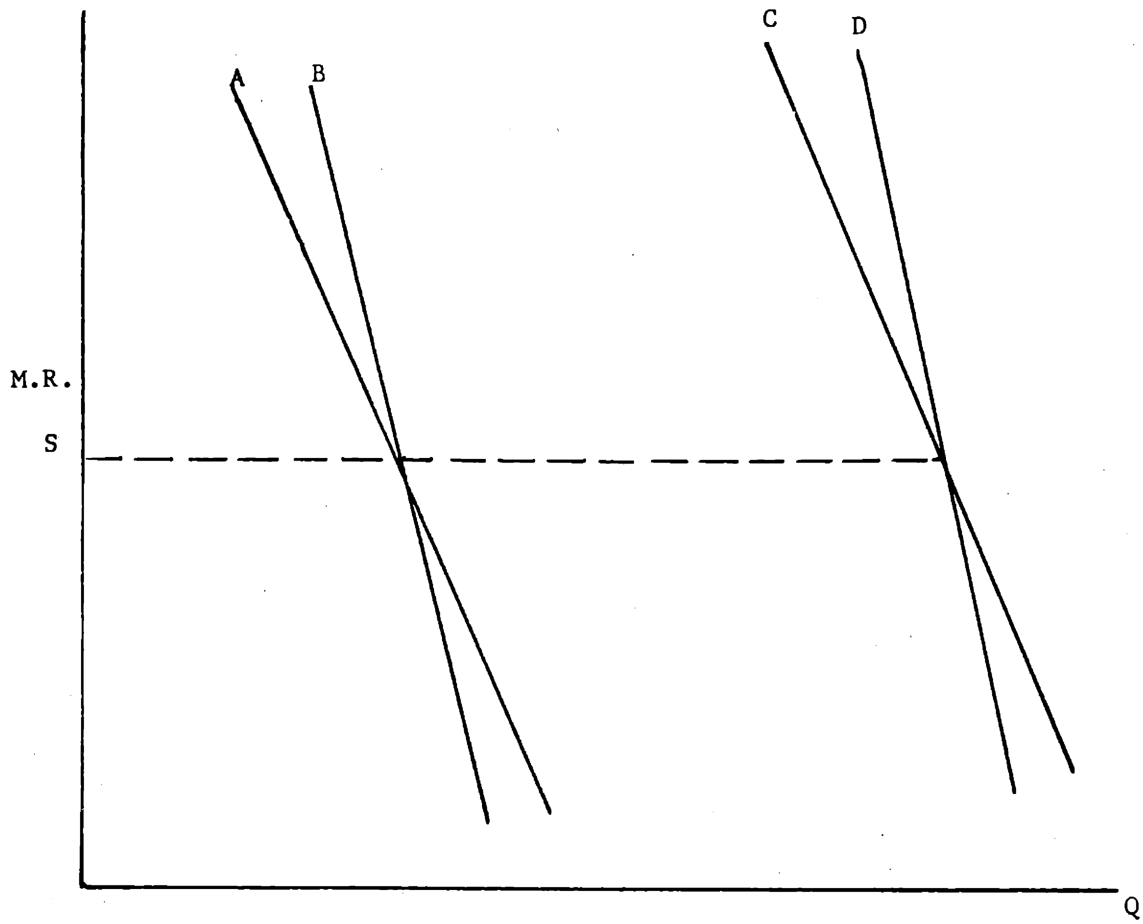
This is a dramatic demonstration of the rather obvious fact that there exists no reasonable complete ordering of random demand functions with respect to variability. In the rest of this section we shall consider a special case where the random demand functions depend simply on a single random variable and may be reasonably ordered by $\frac{\sigma}{v}$. This is the case of multiplicative uncertainty which has appeared, with slightly different interpretations, in other models of firm behavior under uncertainty. (See, e.g., Zabel [6]).

Suppose that

$$R(Q,\theta) = P(Q/x(\theta))Q \tag{3}$$

where x is some positive function of θ -- in the uncertainty interpretation of the model $x(\theta)$ is simply a random variable -- and $P(z)$ is an ordinary demand function giving price as a function of its argument. We assume that $P(z)$ is twice continuously differentiable, the requirement that $R(Q,\theta)$ be concave in Q for all θ is equivalent to

$$P''(z)z + 2P'(z) < 0 \text{ for all } z > 0. \tag{4}$$


M.R.
S
A
B
C
D
Q

This specification of the way in which demand changes has three notable properties: First, we can consider $x(\theta)$ as a measure of the strength of demand at θ. If $x(\theta)$ doubles, the entire demand curve is shifted out (horizontally in Figure 3) by a factor of two. Twice as much can be sold at the same price. In the time interpretation of the model x could be some function of θ and represent seasonal ebbs and flows of demand. In the context of a growth model $x(\theta) = e^{g\theta}$ could represent exponentially growing demand or $x(\theta) = b\theta + c$ linearly growing demand. The growth interpretation is more plausible than the seasonal interpretation; we could consider $x(\theta)$ to represent the number of people at time θ in the market for Q, as time passes, the market grows but individuals' tastes[1] -- as represented by their demand curves -- remain constant. Thus the growth of the market is accounted for by aggregating ever more of the same shaped demand curves. Such growth is described by (3). Not all changes in demand are. For example, the difference in the summer and winter demand for air conditioners is more likely to be caught by Figure 4 than Figure 3. The shift from the higher and relatively inelastic $P(Q,\theta_s)$ to the lower and relatively elastic $P(Q,\theta_w)$ cannot be described by (3).

Second, equation (3) represents the multiplicative uncertainty on which Edward Zabel [6] has based his model of monopoly behavior under uncertainty. In our model it is arbitrary whether we write $R(Q,\theta)$ or $R(P,\theta)$,

---

[1] It is unnecessary to assume that all individuals have the same tastes; we require only that people of different types enter -- and leave -- the market in the same proportions.

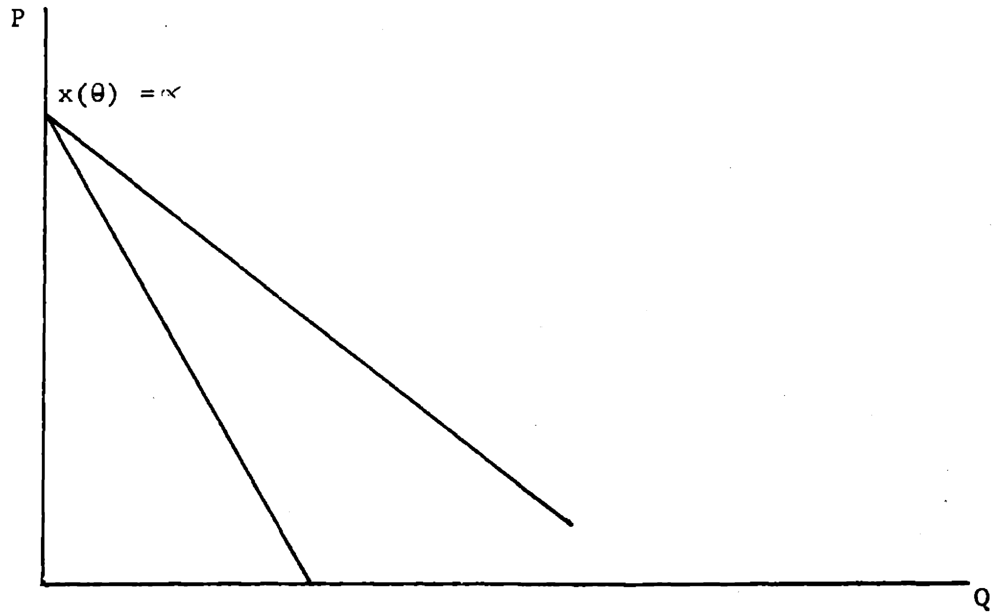


FIGURE 3

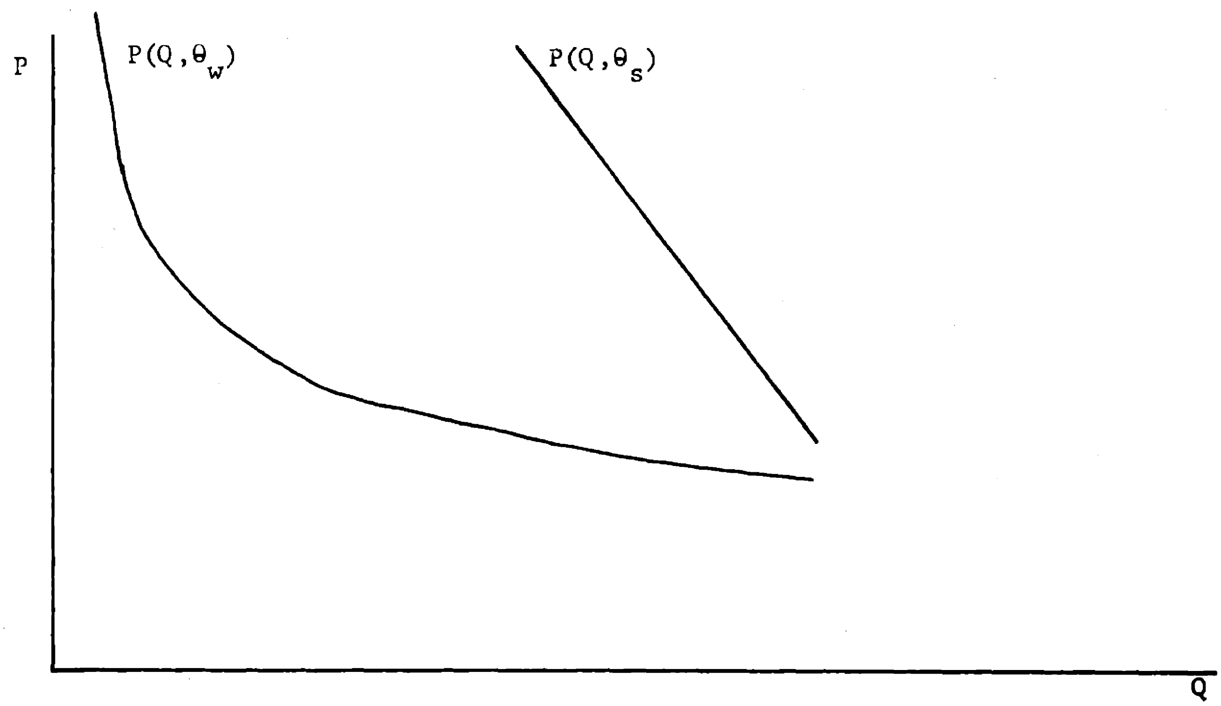


FIGURE 4

for, once θ is known, the firm is assumed to know its entire demand curve. On any single demand curve, P is an invertible function of Q; knowledge of one implies knowledge of the other. Some models of firm behavior under uncertainty -- e.g., Mills [3], Zabel [6] -- assume that the demand curve itself is unknown, that the firm sets a price P and that the quantity demanded in a random variable which is a function of θ. For his model Zabel chooses the specification

$$Q(P,\theta) = x(\theta)\ \Psi(P) \tag{5}$$

where $\Psi(P)$ is an ordinary non-stochastic demand function. Mills chose to analyze the case where uncertainty entered additively

$$Q(P,\theta) = \Psi(P) + x(\theta).$$

It is easy to see that (3) is equivalent to (5), although it should be emphasized that the way in which Zabel's model works is quite different from the way in which our model works. Since P(z) in (3) is an ordinary demand function, we may invert it and write $Q/x(\theta) = \Psi(P)$, from which (5) follows immediately.

The third useful fact about the specification (3) is that $R(Q,\theta) = P(Q/x(\theta))Q$ is homogeneous of degree one in Q and x.

If equation (3) describes demand, it is natural to define an increase in the variability of demand as an increase in the variability of $x(\theta)$. As we have argued that $x(\theta)$ scales the intensity of demand or the number of people in the market at θ, this definition is sensible. We adopt it. We are now in a position to prove the conjecture with which we began Section II, that as the variability of demand increases, the difference between the profits[1] the firm could make if all factors were freely variable and the profits the firm could make if some factors of production were fixed -- in other words, the cost of not being able to vary some factors of production -- increases as the variability of demand increases. We begin by examining what happens to profits if all factors are variable. In this case, for any r and w factor proportions are chosen for all θ so as to minimize unit costs of production.

If we define this minimum by

$$(6) \qquad c = \min_{\substack{L \geq 0 \\ K \geq 0 \\ Q(K,L) = 1}} w \cdot L + r \cdot K,$$

the firm's cost function is linear; for all θ, the cost of producing Q is cQ. The firm maximizes profits by choosing $Q(\theta)$ so that marginal revenue equals marginal cost, or

$$(7) \qquad P'(Q(\theta)/x(\theta))(Q(\theta)/x(\theta)) + P(Q(\theta)/x(\theta)) = c.$$

The right-hand side of (7) is a function of -- and by (4) a decreasing function -- of the ratio $Q(\theta)/x(\theta)$. Thus if (7) has a solution[2] it is unique and of the form

---

[1]In this section, "profits" refer to average or expected profits in the sense of Section I, unless the context clearly implies the contrary.

[2]If (7) has no solution, it is never possible to sell any Q at profit. In Section I we required that there exist a z such that $P'(z)z < c$; thus (7) can fail to have a solution only if $\sup_z P'(z)z < c$ or if marginal -- and in this case average -- cost exceeds marginal revenue at all levels of sales. The question of the existence of a solution to (7) depends only on the function P and not on $x(\theta)$.

(8) $$Q(\theta)/x(\theta) = \alpha$$

for some constant α. Thus the price at which output is sold is a constant, say $\bar{P} = P(\alpha)$. From (5) we see that revenue is a linear function of $x(\theta)$ $R(Q(\theta),\theta) = x(\theta)\ \bar{P}\ \Psi(\bar{P})$, and from (8) we see that $cQ(\theta) = c\alpha x(\theta)$, thus profits at θ are given by

$$\pi(\theta) = x(\theta)\ (\bar{P}\ \Psi(\bar{P}) - c\alpha) = x(\theta)B$$

where B is some positive constant, and expected profits by

$$\pi^* = E[\pi(\theta)] = E[x(\theta)B] = \mu B.$$

π^* is only a function of the mean of $x(\theta)$. We have proved

Theorem I. If demand is given by (3), π^* (maximum expected profits when all factors of production are variable) is unaffected by increases in the variability of demand.

Next, we show that $\pi^*(K)$ (maximum expected profits when K is a fixed factor) decreases as the variability of demand increases. It is necessary first to prove

Lemma II. If $g(x,y)$ is homogeneous of degree one, twice continuously differentiable, and $g_{11} < 0$, then $g_{22} < 0$ and g is concave.

Proof: From Euler's theorem we have

$$\frac{x}{y}g_{11} = - g_{12} = \frac{y}{x}g_{22},$$

so that $g_{11} = (\frac{y}{x})^2\ g_{22}$ and g_{11} and g_{22} are necessarily of the same sign.

Concavity is equivalent to the matrix

$$H = \begin{pmatrix} g_{11} & g_{12} \\ g_{12} & g_{22} \end{pmatrix}$$

being negative semi-definite. Since $g_{11} < 0$, by the well-known theorem on principal minors, H will be negative semi-definite if det $H \geq 0$. But det $H = g_{11}g_{22} - g_{12}^2 = g_{11}^2(\frac{y}{x})^2 - g_{11}^2(\frac{y}{x})^2 = 0$.

This completes the proof.

We now prove

Theorem III: Increases in the variability of demand decrease $\pi^*(K)$ when demand is given by (3).

Proof:

$$\pi^*(K) = \underset{K}{\text{Max}}\ E[\underset{L}{\text{Max}}\ P(\frac{Q(K,L)}{x(\theta)})\ Q(K,L) - wL] - rK$$

or, in an adaptation of the notation of Section I,

$$\pi^*(K) = \underset{K}{\text{Max}}\ E[G(K,x(\theta))] - rK.$$

If we can show that $G(K,x)$[1] is a concave function of x for all K, then an increase in the variability of x will decrease $E[G(K,x)]$ for all K and $\pi^*(K)$ will necessarily decrease.

For fixed K, let $\tilde{x} = \lambda x_1 + (1-\lambda)x_2$ and define L_i by $G(K,x_i) = P(\frac{Q(K,L_i)}{x_i})Q(K,L_i) - wL_i$ and Q_i by $Q_i = Q(K,L_i)$. Define $\hat{L}$ by $Q(K,\hat{L}) = \lambda Q_1 + (1-\lambda)Q_2$. From the concavity of the production function Q, it follows that $\hat{L} \leq \tilde{L} = \lambda L_1 + (1-\lambda)L_2$. By the definition of G,

---

[1]We drop the argument θ when this can be done without confusion.

$$G(K,\tilde{x}) \geq P(\frac{Q(K,\hat{L})}{\tilde{x}})\, Q(K,\hat{L}) - w\hat{L}$$

$$\geq P(\frac{Q(K,\hat{L})}{\tilde{x}})Q(K,\hat{L}) - w\tilde{L}$$

$$\geq \lambda[P(Q_1/x_1)Q_1 - wL_1] + (1-\lambda)[P(Q_2/x_2)Q_2 - wL_2]$$

$$= \lambda G(K,x_1) + (1-\lambda)G(K,x_2).$$

The first step follows from the fact that $\hat{L} \leq \tilde{L}$, the second from the concavity of $P(Q/x)Q$ in Q and x (Lemma II). This completes the proof.

Consider

$$\pi^*(K,L) = \max_Q E[\pi(Q,x(\theta))]$$

That $\pi^*(K,L)$ (maximum expected profits when all factors of production are fixed) decreases as the variability of $x(\theta)$ increases is an immediate -- and hardly astonishing -- consequence of Lemma II which tells us that $\pi(Q,x(\theta))$ is concave in x for all Q.

It is clear that $\pi^*(K,L) \leq \pi^*(K) \leq \pi^*$ since any strategy available to the firm when one (both) factor(s) of production is (are) fixed, is available to the firm when no (one) factor(s) of production are (is) fixed.

We illustrate these concepts with an example. Consider the firm with a Cobb-Douglas production function

$$Q = K^{\beta}L^{1-\beta}\ ;\ 0 \leq \beta \leq 1$$

which faces a constant elasticity demand curve,

(9) $$P(Q,\theta) = (Q/\theta)^{-\alpha};\ R(Q,\theta) = QP(Q,\theta) = \theta^{\alpha}Q^{1-\alpha}$$

where α^{-1} is the elasticity of demand. To assure concavity of the revenue function we must have the elasticity of demand greater than unity or $0 < \alpha < 1$. If r and w are the rental prices of capital and labor, then

$$c = (w/(1-\beta))^{1-\beta} (r/\beta)^{\beta}$$

is the cost of producing one unit of output at best factor proportions. We now examine the behavior of the firm in the three cases distinguished above. In case 2 the firm is allowed to vary both factors of production <u>ex post</u>, i.e., after demand is known; in case 1, it may change only its labor force <u>ex post</u>, in case 0 both factors of production and the level of output and sales are chosen <u>ex ante</u>, before demand is known. In case 2 the firm will always choose to operate at optimal factor proportions and the unit cost of output will be c. Thus, the firm will choose $Q(\theta)$ so as to maximize $R(Q(\theta),\theta) - cQ$ for each θ. Thus, we must have that

$$\frac{\partial R(Q(\theta),\theta)}{\partial Q} = c$$

or

$$Q(\theta) = c^{-1/\alpha} (1-\alpha)^{1/\alpha} \theta$$

We find price as function of output by substituting in the demand function (9): $P_2(\theta) = c/1-\alpha$ which is independent of θ. Since price is constant for all θ,

$$P_2 = E(P_2(\theta)) = c/(1-\alpha) \; .$$

Profits on the other hand change as θ does

$$\pi_2(\theta) = \theta\alpha((1-\alpha)/c)^{((1-\alpha)/\alpha)}$$

so that

$$\Pi_2 = E(\Pi_2(\theta)) = \mu\alpha((1-\alpha)/c)^{((1-\alpha)/\alpha)}$$

where $\mu = E(\theta)$.

In case 0 when the firm chooses both factors of production (as well as the level of output) ex ante it will also produce at minimum average cost. Thus, it will choose Q so as to maximize

$$E(R(Q,\theta)) - cQ = \mu_\alpha Q^{1-\alpha} - cQ$$

where

(10) $$\mu_\alpha = E(\theta^\alpha).$$

Thus,

$$Q_0 = \left(\frac{(1-\alpha)\mu_\alpha}{c}\right)^{1/\alpha}$$

while

$$P_0(\theta) = \frac{\theta^\alpha}{\mu_\alpha}\,\frac{(1-\alpha)}{c}$$

and,

$$P_0 = E(P_0(\theta)) = (1-\alpha)/c$$

so that

$$\Pi_0(\theta) = c^{-(\frac{1-\alpha}{\alpha})}\left[\theta^\alpha(1-\alpha)^{\frac{1-\alpha}{\alpha}}\mu_\alpha^{\frac{1-\alpha}{\alpha}} - \mu_\alpha^{1/\alpha}(1-\alpha)^{1/\alpha}\right]$$

and

$$\Pi_0 = E(\Pi_0) = \mu_\alpha^{1/\alpha}\,\alpha((1-\alpha)/c)^{\frac{1-\alpha}{\alpha}}.$$

Case 1, when the firm chooses capital ex ante and labor ex post, is more tedious to analyze. We begin by inverting the production function to get labor as a function of Q and K. Then cash flow for a given choice of K,Q, and θ is given by

$$\theta^{\alpha} Q^{1-\alpha} - wQ^{1/1-\beta} K^{-\beta/1-\beta} \tag{11}$$

Q(K,θ) is chosen to maximize (11). Substituting this maximizing Q into (11) we find after considerable manipulation that, in the notation of Section I,

$$G(K,\theta) = \gamma^{\frac{\gamma}{1-\gamma}} (1-\gamma) w^{-\frac{\gamma}{1-\gamma}} \theta^{\frac{\alpha}{1-\gamma}} K^{\frac{(1-\alpha)\beta}{1-\gamma}}$$

where

$$\gamma = (1-\alpha)(1-\beta) \tag{12}$$

Choosing K* to maximize E(G(K,θ)) - rK, we find that

$$K^* = \frac{(1-\alpha)^{1/\alpha} (1-\beta)^{\gamma/\alpha} \beta^{\frac{1-\gamma}{\alpha}}}{r^{\frac{1-\gamma}{\alpha}} w^{\frac{\gamma}{\alpha}}} \left[\mu_{\left(\frac{\alpha}{1-\gamma}\right)}\right]^{\frac{1-\gamma}{\alpha}} .$$

Substituting K* into the results already obtained we find $P_1(\theta)$, $Q_1(\theta)$ and $\pi_1(\theta)$ and their expected values, P_1, Q_1, and π_1. All these results are summarized in Table 1, which gives P(θ), Q(θ) and π(θ) and their expected values for all three cases.

The most interesting aspect of Table 1 is line 6 which states that expected profits are a constant times a function of the form $\mu_{a_i}^{1/a_i}$ where $a_0 = \alpha$, $a_1 = \frac{\alpha}{1-\gamma}$, and $a_2 = 1$. Since we know that $\pi_2 \geq \pi_1 \geq \pi_0$ and since a_1 and a_2 are arbitrary except that they satisfy $0 \leq a_0 \leq a_1 \leq 1$, this constitutes a proof of the fact that

$$\phi(a) = \mu_a^{1/a} = [E(|X|^a)]^{1/a}$$

is a non-decreasing function of a for $0 < a \leq 1$. From this we can prove

<u>Lemma IV</u>. $\phi(a)$ is non-decreasing for all positive a.

<u>Proof</u>: Suppose $a' > a > 0$. Then $a/a' < 1$. Let $Y = |X|^{a'}$. Then, $[E(Y^{a/a'})]^{a'/a} \leq E(Y)$. Raising each side of this inequality to the $1/a'$<u>th</u> power, we have

$$[E((Y)^{a/a'})]^{1/a} \leq [E(Y)]^{1/a'}$$

or,

$$[E(|X|^{a'})^{a/a'}]^{1/a} = [E(|X|^{a})]^{1/a} \leq [E(|X|^{a'})]^{1/a'}.$$

We will find this well-known inequality -- which can also be obtained from Holder's inequality -- useful in what follows.

We use it now to show that π_1, expected profits when one factor is variable and the other factor is fixed, increase as the output elasticity of the variable factor increases. π_1 is a constant times $\mu_\delta^{1/\delta}$ where $\delta = \alpha/(\alpha+\beta(1-\alpha))$ is an increasing function of $(1-\beta)$ the output elasticity of labor. This is a reasonable result, particularly considering the extreme values of π_1; when $\beta = 0$, $\delta = 1$ and $\pi_1 = \pi_2$; and if $\beta = 1$, $\delta = \alpha$, and $\pi_1 = \pi_0$. If output is produced by labor alone, as $\beta = 0$ implies; then the constraint that capital must be chosen before demand is known as no constraint; if output is produced by capital alone, as $\beta = 1$ implies, the freedom to hire labor, but not capital, <u>ex post</u> is worth nothing.

This is not a <u>ceteris paribus</u> result. The constant which multiplies $\mu_\delta^{1/\delta}$ depends on c, minimum average costs, which is a function of r, w, <u>and</u> β.

Thus, we cannot be sure that π_1 increases as β decreases unless r and w change so as to hold c constant, or unless $r/w > \beta/1-\beta$. This last condition stems from writing $\pi_1 = f(c)\mu_\delta^{1/\delta}$ and observing that $f'(c) < 0$ while $\frac{dc}{d\beta} > 0$ if and only if $r/w > \beta(1-\beta)$, as can be seen by examining $\frac{d \log c}{d\beta}$.

It is interesting to note that average price is lowest in case 1. $P_0 = P_2$ and

$$P_1 = (\mu_{\frac{\alpha}{1-\gamma}})^{-\beta}(\mu_{\frac{\alpha\beta}{1-\gamma}}) = [E(Y)]^{-\beta}(E(Y^\beta))P_0$$

where $Y = \theta^{\frac{\alpha}{1-\gamma}}$. But, $E(Y^\beta)^{1/\beta} \leq E(Y)$ since $\beta < 1$. Take each side to the β $\underline{th}$ power and multiply by $(E(Y))^{-\beta}$ to get $[E(Y)]^{-\beta}(E(Y^\beta)) \leq 1$. Therefore, $P_1 \leq P_2 = P_0$.

TABLE 1

THE EFFECT OF VARIABILITY OF FACTORS ON PRICE OUTPUT AND PROFITS*

		Constant*	Case 0	Case 1	Case 2
1.	$P(\theta)$	$(1-\alpha)/c$	$\theta^{\alpha}/\mu_{\alpha}^{-1}$	$(\frac{\mu_{\alpha}}{1-\gamma})^{-\beta}\theta^{\frac{\alpha\beta}{1-\gamma}}$	1
2.	P	$(1-\alpha)/c$	1	$(\frac{\mu_{\alpha}}{1-\gamma})^{-\beta}\ \frac{\mu_{\alpha}\beta}{1-\gamma}$	1
3.	$Q(\theta)$	$((1-\alpha)/c)^{1/\alpha}$	$\mu_{\alpha}^{1/\alpha}$	$\theta^{\frac{\alpha(1-\beta)}{1-\gamma}}\ \mu_{\alpha/1-\gamma}^{\beta/\alpha}$	θ
4.	Q	$((1-\alpha)/c)^{1/\alpha}$	$\mu_{\alpha}^{1/\alpha}$	$(\frac{\mu_{\alpha}(1-\beta)}{1-\gamma})\mu_{\alpha/1-\gamma}^{\beta/\alpha}$	μ
5.	$\pi(\theta)$	$((1-\alpha)/c)^{\frac{1-\alpha}{\alpha}}\alpha$	$\left[\theta^{\alpha}\mu_{\alpha}^{\frac{1-\alpha}{\alpha}}\alpha^{-1}-\mu_{\alpha}^{1/\alpha}(1-\alpha)\alpha^{-1}\right]$	$[\frac{\mu_{\alpha}}{1-\gamma}]^{\frac{1-\gamma}{\alpha}}(\theta^{\frac{\alpha}{1-\gamma}}\mu_{\frac{\alpha}{1-\gamma}}^{-1}$	
				$\frac{(1-\gamma)}{\alpha}-\frac{(1-\alpha)\beta}{\alpha})$	θ
6.	$\tilde{\pi}$	$((1-\alpha)/c)^{\frac{1-\alpha}{\alpha}}\alpha$	$\mu_{\alpha}^{1/\alpha}$	$\frac{\mu_{\alpha}}{1-\alpha}^{\frac{1-\gamma}{\alpha}}$	μ

*The constant multiplies all entries in the row.

III

INCREASING VARIABILITY OF OUTPUT

Embedded in each profit maximizing problem is a cost minimizing problem. In this section we investigate it; for the most part we shall use the language appropriate to the uncertainty interpretation of the model. In each state of the world the firm sells a quantity of output $S(\theta)$. Thus $S(\theta)$ is a random variable with a distribution function $F(S)$ defined by:

$$F(S) = \mu\{A\}^{-1}\mu\{\theta \in A : S(\theta) \leq S\}$$

where θ ranges over the set A. The firm which maximizes expected profits will choose its fixed and variable factors of production so as to minimize the expected cost of producing $S(\theta)$. If K and L represent fixed and variable factors of production and w is the price of L and r that of K then the variable cost of producing $S(\theta)$ given an <u>ex ante</u> choice of K is $\mathcal{B}(K,S(\theta))$, the solution to

(13) $$\min_{L \geq 0} w \cdot L \text{ subject to } F(K,L) \geq S(\theta).$$

Total costs for state θ are $C(K,S(\theta)) = \mathcal{B}(L,S(\theta)) + r \cdot K$ and expected costs are $C(K) = r \cdot K + E(\mathcal{B}(K,S(\theta)))$.[1] The firm which maximizes expected profits, must choose K so as to minimize expected costs of production.[2]

---

[1]The notation for C parallels that for w.

[2]Since $\pi^*(K) = E[R(S(\theta),Q)] - C(K)$, for a given random variable $S(\theta)$, $\pi^*(K)$ is maximized only if $C(K)$ is minimized.

$S(\theta)$ is a random variable; we can use the techniques of the previous section to analyze the effects of increases in the variability of sales on costs and the employment of factors of production. The relationship between changes in the distribution of demand and those of sales is not clear cut -- as a glance at the example of the last section makes appallingly clear. We would conjecture that in most cases increases in the variability of demand lead to decreases in average sales, and perhaps to increases in some, but not all, measures of dispersion which take account of this shift in expected value. We have not studied extensively the nature of the relationship between changes in the distribution of demand and of sales. We doubt that any general results are to be had. Nonetheless there is good reason (aside from the obvious one that at present this is what we can do), for wanting to examine the effects of increases in the variability of sales. This concept of sales appears quite naturally in at least two models of firm behavior. Consider a firm which must[1] meet a random demand. An electric utility supplying power to a city is a good example. The firm's customers are all plugged into the power utility; to use power they simply turn on appliances. The utility has no choice as to whether to sell power to its customers. It treats the total amount of power which it supplies at any instant as a random variable. We may easily imagine that the firm has fixed (e.g., a steam generator) and variable factors (e.g., coal), which it combines to produce

---

[1]Or pay a penalty, of which more later.

electricity. The firm's task is to pick the size of the fixed factor of production which minimizes expected costs. Another model of firm behavior of which sales are an important component is a slight variant of the model of section I. Suppose a monopolistic firm names a price, P, at which it will sell output and that orders, which we assume that the firm must meet, are a random variable which is a function of θ. We can reasonably ask in such a model what happens when $S(P,\theta)$ becomes more variable.

The reader may object that the last two models are unrealistic. While $S(\theta)$ has no natural bound, our models would have the firm produce whatever was ordered. If the elasticity of substitution is low, this may mean that our results, particularly those concerning the choice of K will be dominated by extremely high S's of low probability. For example, if the production function is of the fixed coefficients type, $Q = \min(K,L)$, then $K = \sup_{\theta \in A} S(\theta)$ is necessary if the firm is to meet all orders. This is clearly unreasonable. If $S(\theta) = 1000$ for $\theta \in A'$ and $S(\theta) = 1$ for $\theta \in A-A'$ if $\mu(A')/\mu(A) < .001$, no reasonable model would predict that the firm would hold a capital stock of 1000 to meet the slight (less than 1 in 1000) chance that it will need a capital stock greater than one. If the firm should, by chance, encounter orders for a 1000 units of output it will simply not meet them, paying some sort of penalty rather than incurring the cost of keeping around a large idle capital stock.

We can take account of this objection if we expand our model to include n factor production functions. That is, K, instead of being a single kind of machine is now a vector of j fixed factors of production and L is a vector of n-j variable factors. We still require that Q(K,L) be homogeneous of degree one and concave. Moving from 2 to n factors enriches our model considerably without altering most of our results.[1] Variable factors can be interpreted to include both rental markets for what are ordinarily fixed factors and penalty costs for failing to meet orders. The firm may choose to rent machines for short periods of exceptionally high demand. The model now also allows us to consider different types of labor some of which are fixed and others of which are variable. A penalty cost for not filling orders is simply a rental market for the variable factor of production, units of the final product. The rental price of this input is the penalty cost for not filling orders.[2]

(a) <u>Effects on Costs</u>

We expect that as $S(\theta)$ increases in variability expected costs increase. This is an immediate consequence of

<u>Lemma V</u>. $\beta(K,S)$ is a convex function of S.

<u>Proof</u>: For any fixed K let $\tilde{S} = \lambda S_1 + (1-\lambda)S_2$, $0 \leq \lambda \leq 1$. We must show that $\lambda\beta(K,S_1) + (1-\lambda)\beta(K,S_2) \geq \beta(K,\check{S})$. Let L_i satisfy $\beta(K,S_i) = w \cdot L_i$ and let $Q_i = Q(K,L) \geq S_i$ for i=1,2. Since Q is concave, $Q(K,\lambda L_1+(1-\lambda)L_2) = Q(K,\tilde{L}) \geq \lambda S_1 + (1-\lambda)S_2 = \check{S}$. Therefore $\beta(K,\tilde{S}) \geq w \cdot \check{L} = \lambda(w \cdot L_1)+(1-\lambda)(w \cdot L_2) = \lambda\beta(K,S_1) + (1-\lambda)\beta(K,S_2)$.

[1] The proofs of section II go through with only minor changes.

[2] This is formally indistinguishable from an electric utility supplying its customers by buying electricity from another utility. This model can incorporate convex as well as linear penalty costs.

This lemma implies that expected variable costs for <u>any</u> choice of K increase (or at least do not decrease) as the variability of demand increases. Thus expected costs when K is chosen to minimize expected costs must also increase. We have proved

<u>Theorem VI</u>. C*(K) increases as the variability of sales increases.

This theorem has an interesting implication. A reasonable measure of the relative efficiency of firms producing the same output and facing the same factor prices is costs of production. If we imagine two identical firms, one of which faces a more variable demand than the other, the first firm will produce at a higher average cost than the second, and thus appear to be less efficient than the first even though technically they are identical. If we compare two economies, each made up of the same kind of firms and one economy is subject to more fluctuations in demand -- both aggregate and in composition -- than the other then the first economy will appear to be technically behind the second.

(b) <u>Effects on the Demand for Factors</u>

Since increases in the variability of output increase costs, they must increase the employment of some inputs. It is not in general possible to say which inputs are increased by increases in the variability of sales. It can be shown that, in the two-factor case, if the production function is of the constant elasticity (CES) class, then, if the elasticity of substitution (σ) is less than, or equal to, unity, increases in the variability of output increase K. If $\sigma > 1$ their effect is not clear cut. In the polar case of perfect substitutibility ($\sigma = \infty$) it is possible to show that some increases in variability increase K and others decrease it.

A more interesting question to ask is: what is the effect of variable demand on factor proportions? If output were constant, then the firm would choose its labor-capital ratio ($\ell = \frac{L}{K}$) so as to minimize unit costs. First order conditions for cost minimization are that the marginal rate of substitution, ($m = \frac{Q_K}{Q_L} = - \left.\frac{dL}{dK}\right|_{dQ=0}$) be equal to $\frac{r}{w}$. Since m is a function of ℓ we may write

(14) $$\frac{r}{w} = m(\ell).$$

If output is random expected cost minimization implies different conditions. For any K, $\ell(K,Q) = wL(Q;K)$ where $L(Q;K)$ is the amount of labor necessary to produce Q with a capital stock of K. Expected costs are then,

$$C(K) = rK + w\int L(Q;K)\,dQ.$$

First order conditions forminimization are that $C'(K) = 0$ or that

$$\frac{r}{w} = -\int \left.\frac{d(L(Q;K))}{dK}\right|_{dQ=0} dQ.$$

But,

$$\left.\frac{d(L(Q;K))}{dK}\right|_{dQ=0} = -\frac{Q_K}{Q_L} = -m(\ell),$$

so that efficiency with random output requires that

(15) $$\frac{r}{w} = E(m(\ell))$$

or that r/w equal the expected marginal rate of substitution.

For the special case of the CES production function, $Q = (\delta K^{-\rho} + (1-\delta)L^{-\rho})^{-1/\rho}$, where $\sigma = 1/(1+\rho)$, and

(16) $$m(\ell) = \left(\frac{\delta}{1-\delta}\right) \ell^{1/\sigma} = \left(\frac{\delta}{1-\delta}\right) \left(\frac{L}{K}\right)^{1/\sigma}$$

so that (14) implies that $\bar{\ell}$, the labor-capital ratio when demand is constant, is given by

$$(17) \qquad \bar{\ell} = (\frac{r}{w})^{\sigma} (\frac{1-\delta}{\delta})^{\sigma} .$$

If demand is random then we may substitute (16) into (15) to obtain:

$$\frac{r}{w} = E[(\frac{\delta}{1-\delta}) (\frac{L(Q,K^*)}{K^*})^{1/\sigma}],$$

where K* is the capital stock chosen to minimize expected costs. This last equation implies that

$$\frac{r}{w} (\frac{1-\delta}{\delta}) = (\frac{1}{K^*})^{1/\sigma} E[(L(Q,K^*)^{1/\sigma}],$$

or

$$\bar{\ell} = (\frac{r}{w})^{\sigma} (\frac{1-\delta}{\delta})^{\sigma} = \frac{E(L^{1/\sigma})^{\sigma}}{K^*} .$$

Let $\hat{\ell} = E(\frac{L}{K^*}) = \frac{E(L)}{K^*}$ be the expected or average labor-capital ratio when output varies. Then, $\hat{\ell} \gtreqless \bar{\ell}$ as $E(L^{1/\sigma})^{\sigma} \lesseqgtr E(L)$. But, we know from Lemma IV that $E(L^{1/\sigma})^{\sigma} \lesseqgtr E(L)$ as $\sigma \gtreqless 1$. We have thus proved

<u>Theorem VII</u>. If $\sigma > 1$, the firm producing a varying output with a fixed capital stock according to a CES production function will, on the average, use more labor intensive techniques of production than it would if output were constant or both K and L were allowed to vary. If $\sigma < 1$, then the reverse holds. If $\sigma = 1$, the Cobb-Douglas case, $\bar{\ell} = \hat{\ell}$.

Consideration of polar cases shows that this is a reasonable result. If $\sigma = 0$, fixed coefficients, the firm will have to keep a capital stock large enough to produce the largest output which it is going to produce in any state of the world. In any state of the world when output is not at its maximum some capital will be idle and the labor-capital ratio will be less than that consistent with full employment of both factors. Now, consider the case of perfect substitutability, $\sigma = \infty$. We may choose units so as to write the production function as $Q = K + L$. Then, if $r < w$, no labor will be employed if sales are constant. If sales vary the firm may choose to meet high demand by employing some labor rather than keeping a large idle capital stock on hand. If the firm hires any labor at all it will raise the average labor-capital ratio above zero, the value of the optimal labor-capital ratio when demand is constant. If $r > w$, the firm will hire no capital whether demand is fixed or varying.

This result can explain the disparity between the wage and the marginal product of labor which Thurow [5] has observed. Thurow estimated a production function for the U.S. private economy from data on output, capital, and labor. He then used the estimated production function to calculate a marginal product of labor series and compared this series with observed data on employee compensation. He found that from 1929 to 1965 the wage was always much lower than the marginal product of labor. Thurow offered several explanations of his finding. Our model suggests another. If our model describes the U.S. economy -- an heroic assumption but surely one which strains the imagination only a little more than those implicit in Thurow's work (all that is required beyond the existence of an aggregate production function of the CES class is that there be a cost to changing the capital stock which is described by a function which is <u>not</u> convex near zero)-- then the disparity which Thurow observes is explained by a low elasticity of substitution. For, if $\sigma < 1$, then each firm will, according to Theorem VII, on the average use less labor intensive methods of production than factor price ratios would imply if demand were constant. Thus, the marginal product of labor will on the average be higher than what it would be if output were constant. But, if output were constant the marginal product of labor would equal the wage rate; if output varies, it must be greater than the wage.

We do not suggest that this is the correct explanation of Thurow's results. We only wish to point out that this model of firm behavior has rather different implications than models which assume that all factors of production are variable. We emphasize this point by considering the problem of estimating production functions. Because reliable data on capital are very hard to come by, many estimation procedures rest on the presumed identity of wages with the marginal product of labor. In our model the wage is not the marginal product of labor; assuming that it is will lead to biases in estimation. This is easily illustrated by examination of the method proposed by Arrow, Chenery, Minhas and Solow [1] for estimating the elasticity of substitution. If the production function is of the CES class then

$$\log \frac{Q}{L} = -\sigma \log (1-\delta) + \sigma \log \frac{\partial Q}{\partial L}$$

if $w = \frac{\partial Q}{\partial L}$ then we may estimate,

(18) $$\log \frac{Q}{L} = \beta_0 + \beta_1 \log w + \varepsilon ,$$

and take $\hat{\beta}_1$ as an estimate of σ. Suppose we estimate (18) for a cross section of firms which treat capital as a fixed factor of production and labor as a variable factor. It seems reasonable to assume on the average that firms with high output per man will be using less labor than factor price ratios would indicate and thus that wages would be lower than the marginal product of labor. For firms operating at low output per man, the reverse is likely to be the case. The observed relationship between $\log w$ and $\log \frac{Q}{L}$ will have a lower slope than that of the actual relationship between $\log \frac{\partial Q}{\partial L}$ and $\log \frac{Q}{L}$. (See Figure 5). $\hat{\beta}_1$ the estimate of will be biased downward.[1]

[1] Theorem VII implies that the estimate of the intercept term will be biased as well.

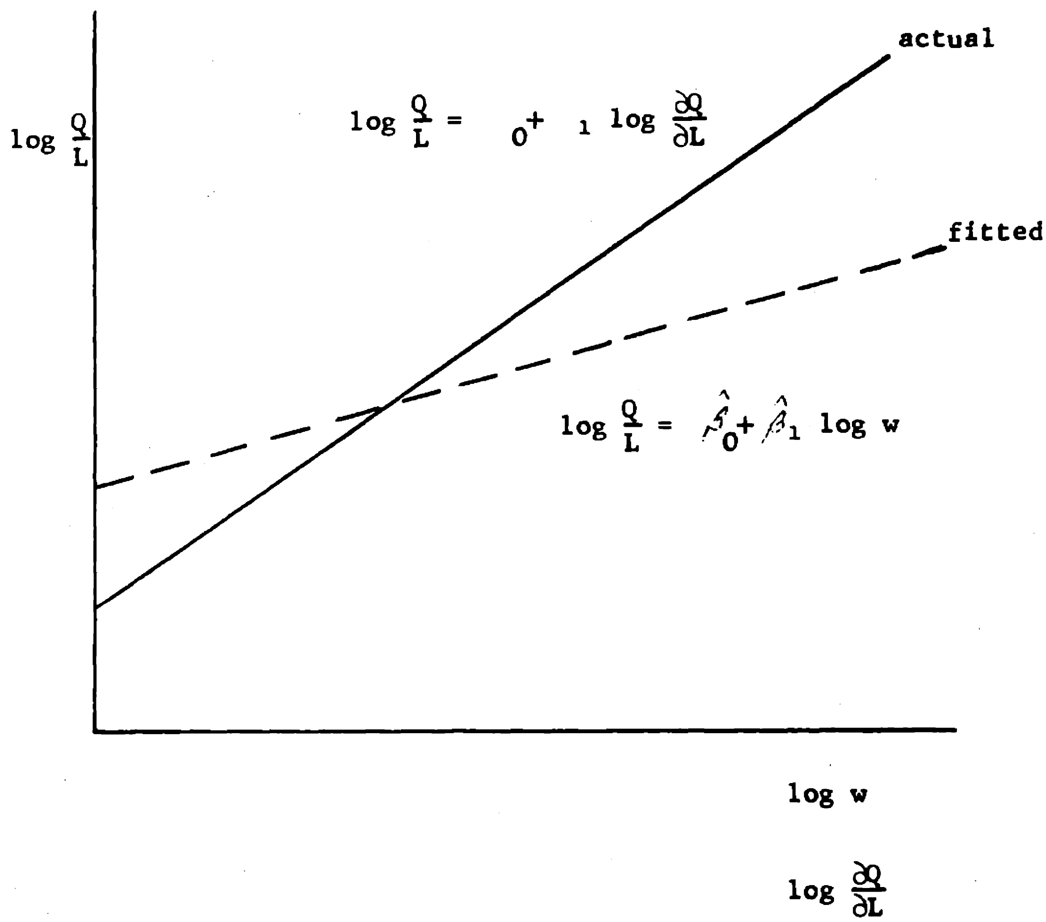

actual
log Q/L
log Q/L = 0+ 1 log ∂Q/∂L
fitted
log Q/L = β̂0 + β̂1 log w
log w
log ∂Q/∂L

FIGURE 5

REFERENCES

[1] Arrow, K. J., H. Chenery, B. S. Minhas, and R. M. Solow, "Capital-Labor Substitution and Economic Efficiency," Review of Economics and Statistics, XXXXIII, (1961).

[2] Hart, A. G. Anticipations, Uncertainty and Dynamic Planning, (New York, 1965)

[3] Mills. Price Output and Inventory Policy (N.Y., 1962).

[4] Stigler, G., "Production and Distribution in the Short Run," Journal of Political Economy, XXXXVII, (1939).

[5] Thurow, L. C., "Disequilibrium and the Marginal Productivity of Capital and Labor," Review of Economics and Statistics, L, (1968).

[6] Zabel, E., "Monopoly and Uncertainty," unpublished, (Feb., 1968).

BIOGRAPHICAL NOTE

Michael Rothschild was born in Chicago, Illinois, in 1942. He received a B.A. (in anthropology) from Reed College in 1963 and a M.A. from Yale University in International Relations in 1965. He is a member of Phi Beta Kappa. In September 1968 he was appointed Instructor of Economics at Boston College.