Entry and Exit Decisions under Uncertainty

Avinash Dixit

The Journal of Political Economy, Vol. 97, No. 3 (Jun., 1989), 620-638.

Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28198906%2997%3A3%3C620%3AEAEDUU%3E2.0.CO%3B2-J

Entry and Exit Decisions under Uncertainty

Avinash Dixit

Princeton University

A firm's entry and exit decisions when the output price follows a random walk are examined. An idle firm and an active firm are viewed as assets that are call options on each other. The solution is a pair of trigger prices for entry and exit. The entry trigger exceeds the variable cost plus the interest on the entry cost, and the exit trigger is less than the variable cost minus the interest on the exit cost. These gaps produce "hysteresis." Numerical solutions are obtained for several parameter values; hysteresis is found to be significant even with small sunk costs.

I. Introduction

Many investment decisions are made in an uncertain environment and are costly to reverse later. A prominent recent instance is the entry and exit of firms in foreign markets in response to real exchange rate fluctuations. Other examples abound. As the prices of oil or gas fluctuate, producers of these fuels must decide how to expand or contract their operations, and users must decide whether to switch from one fuel to another; both decisions entail sunk costs. Similar choices arise in labor markets when future demand or productivity is uncertain, for firms facing specific training costs and for workers with moving or search costs. Closer to home, as student demand for different subjects fluctuates, university administrators have to decide

I am grateful to Sanford Grossman and Paul Krugman for valuable discussions; to Alan Blinder, Gene Grossman, Marcus Miller, Sherwin Rosen, Carl Shapiro, and two referees for comments on previous drafts; and the National Science Foundation (grant SES-8509536) and the International Finance Section, Princeton University, for financial support.

[*Journal of Political Economy*, 1989, vol. 97, no. 3]

whether and how to reallocate their faculty positions across departments.

In this paper I aim to elucidate the general nature and significance of such decisions. I do so by following a recent line of research that exploits an analogy between real and financial investment decisions. An opportunity to make a real investment is a call option on a stock that consists of the capital in place. Making the investment is like exercising the option, and the cost of the investment is the strike price of the option. Standard techniques of financial economics give us the price of the option (the value to the firm of the opportunity) and the rule that tells the firm when to exercise it optimally (the investment criterion).

I begin by stating the relation of the paper to some recent literature on this theme. McDonald and Siegel (1986) analyze a discrete project, and Pindyck (1987) and Bertola (1987) one with a variable scale, assuming that the operating profits never become negative. In McDonald and Siegel (1985), operating losses are possible, but investment once made is permanently available for use: operation can be suspended when operating profits are negative and resumed at no additional cost if they turn positive again.[1] But many investments decay or rust rapidly when they are not used. If a foreign firm withdraws from the U.S. market as the dollar falls, then its distribution network and brand recognition will disintegrate quite rapidly and must be rebuilt should it decide to reenter when the dollar appreciates again. A mine is subject to cave-ins and flooding when not used, and the costs of switching a furnace from oil to gas must be incurred over again should one decide to switch back. A firm that fires a trained worker cannot rely on hiring the same person, and must expect to train a new one should it decide to expand again. A worker must incur search cost for each job change.[2]

Such rusting adds an interesting new twist to the option pricing view. The asset that is acquired by exercising the option to invest includes another option, namely, to abandon the investment and revert to the original situation. We have two interlinked option pricing problems, which must be solved simultaneously, and the prices of both must be obtained in terms of the underlying uncertainty in exchange rates, demand, and so forth.

Brennan and Schwartz (1985) have done this for the decisions to open and close a mine, and Constantinides (1986) for consumption

[1] Mossin (1968) has a model in which investment can be mothballed at a cost, but the analysis is confined to a stochastic stationary state with zero discounting.

[2] To capture the effect most simply and clearly in my theoretical model, I shall assume that rusting occurs immediately on abandonment. In reality it is more gradual, and the rate differs across industries.

and portfolio choices with two assets. The model of investment is the same as the one here, but in each case it is mingled with other complications, forcing immediate recourse to numerical simulations. My simpler setting yields some important analytical results and brings out the generality of the idea and the variety of its applications.

The paper also relates to an older literature on the effects of sunk costs. For example, Becker (1962, pp. 22–23) and Oi (1962) considered specific training costs and discussed a firm's decision to lay off workers, and a worker's decision to quit, when there is an unexpected temporary decrease in demand. My model allows a better integrated treatment of such problems. It shows how the optimal ex post decisions are affected by the status quo and how the optimal choice of sunk investments is in turn affected by rational expectations concerning future fluctuations.

II. Hysteresis and Its Causes

The most important feature of entry and exit decisions in an environment of ongoing uncertainty is "hysteresis." This is defined as the failure of an effect to reverse itself as its underlying cause is reversed. For example, the foreign firms that entered the U.S. market when the dollar appreciated did not exit when the dollar fell back to its original levels. It would be useful to explain briefly the reasons for hysteresis before constructing a model that lets us assess its magnitude.

Consider a single discrete project with sunk investment cost k, no physical depreciation (but immediate rusting if unused), and avoidable operating cost w per unit of time. I shall call the entity that makes decisions about the project the firm, although in other applications this may be a worker (job search), a consumer (fuel use), or a university dean (faculty hiring). Let ρ be the rate of interest. Define the output flow of the project as a unit, so the revenue from the project is simply the output price P. Let us consider the investment decision under alternative assumptions about P. In all cases, the optimal decision rule consists of two triggers, P_H and P_L, with $P_H > P_L$, such that the investment should be made if P rises above P_H and should be abandoned if P falls below P_L.

First, suppose that the firm does not have an investment in place and that it believes that P will persist unchanged forever. It will make the investment if $P > w + \rho k$; the right-hand side is the annualized full cost of making and operating the investment. Conversely, suppose that a firm has such an investment in place and that the price falls unexpectedly to a new level P, where the firm believes it will persist forever. The firm will abandon the project if $P < w$. Thus the full cost serves as the entry trigger P_H and the variable cost as the exit trigger P_L. This is the standard Marshallian theory of the long run versus the short run, based on the gap between the full cost and the variable cost. It is also the basic model of Oi (1962, pp. 541–42).

Hysteresis could be explained along these lines: the price was initially between w and $w + \rho k$. Then it rose to a level above $w + \rho k$, the investment was made, and the price fell to its original level, but that was insufficient to induce abandonment. This story is unsatisfactory in two ways. First, the firm's expectations are irrational; the price changes are a succession of surprises. Second, we shall see later that other causes are quantitatively more important. Of course, if there were literally no sunk costs, there could be no hysteresis, but given some sunk costs, other forces play bigger roles.

As the second case, suppose that the usual value P^* of P is in the range $(w, w + \rho k)$. Suppose that P has risen to a higher level and is expected to revert to P^*. Now a price of $w + \rho k$ will not suffice to induce investment. The entry trigger P_H will be higher, so above-normal returns for a while can compensate for below-normal returns later. Similarly, if a firm has the investment in place and P falls temporarily to P_L, the firm will not abandon the project unless P_L is sufficiently far below w. When Oi (1962, p. 542) discusses uncertainty, he means a risk of such mean reversion. This case is also discussed in Baldwin (1986) and is formally similar to Baldwin and Krugman (1986), where the exchange rate fluctuates in a Markov fashion between a high value and a low one.

The third case has ongoing uncertainty, even without the need for any reversion to the mean. Suppose that the current price is $w + \rho k$, and from here on at each point in time it will take equal steps up or down with equal probabilities. If the firm invests right away and continues active forever after, its expected present value net of the investment cost is zero. But it can do better by waiting. Suppose that it waits one period. If at the end of this period the price has gone up, it can invest and get positive expected present value. If the price has gone down, it need not invest, so the expected present value is zero. If we weight these by the probabilities of ½ each and add, the expected present value of waiting one period is positive. If the current price exceeds $w + \rho k$, the current sacrifice of operating profits becomes more important, and eventually for some higher P_H it will be optimal to invest at once. This is the option value feature. At the price $w + \rho k$, the investment opportunity is an option that is only just in the money. It is not optimal to exercise it unless it goes deeper in the money. Similarly, the price P_L that triggers abandonment will be less than w.

This case, and the determination of P_H and P_L in terms of the underlying parameters of the problem, will be the focus of this paper. I shall argue that the option value aspect is quantitatively quite important; the gap between P_H and $w + \rho k$, and the one between P_L and w, are quite large even for modest sunk costs that make the gap between w and $w + \rho k$ quite small, and even for modest degrees of fluctuations in the price.

III. The Basic Model

I shall develop the ideas in the simplest possible context. A firm is defined by its access to a particular production technology. It can become active by investing a lump sum k. Then it can produce a unit flow of output at the variable cost w. It can decide to suspend operations but must pay a lump-sum exit cost l to do so.[3] It must incur the entry cost k again should it decide to reenter at some future time. The cost of capital, or the firm's discount rate, is ρ. The magnitudes w, k, l, and ρ are constant and nonstochastic. The uncertainty arises from the market price. The firm is risk neutral and maximizes its expected net present value. Many of these special assumptions will be relaxed in later sections.

The market price P evolves exogenously over time as a Brownian motion, which is the continuous-time formulation of the random walk. This is the standard setting in option pricing theory and also a good first approximation for real exchange rates and some natural resource prices.[4] Specifically,

$$\frac{dP}{P} = \mu dt + \sigma dz, \tag{1}$$

where dz is the increment of a standard Wiener process, uncorrelated across time, and at any one instant satisfying

$$E(dz) = 0, \qquad E(dz^2) = dt. \tag{2}$$

I shall consider some alternative processes in Section VI.

Start at $t = 0$, with the initial price $P_0 = P$, and consider the random price P_t at a later date t. By the standard theory of Brownian motion, we know that $\ln P_t$ is normally distributed with mean $\ln P_0 + (\mu - ½\sigma^2)t$ and variance $\sigma^2 t$. Then, from standard properties of the lognormal distribution, we have $E(P_t|P_0) = \exp(\mu t)$. Thus μ is the trend rate of growth of the market price. For convergence we need $\mu < \rho$.

[3] In the labor demand context, this may be a statutory firing cost. In the formal theory, l can be negative but less than k in magnitude; this covers the case in which a part of the entry cost can be recovered on exit.

[4] In the job search context, P stands for the earnings from, and w for opportunity cost of, a particular job. The uncertainty can occur in either or both; what ultimately matters is the behavior of $P - w$.

The firm's decision problem has two state variables, the current price P and a discrete variable that indicates whether the firm is active (1) or not (0). In state $(P, 0)$, the firm decides whether to continue being idle or to enter. Likewise, in state $(P, 1)$, it decides whether to continue being active or to exit. Let $V_0(P)$ be the expected net present value of starting with a price P in the idle state and following optimal policies.[5] Similarly define $V_1(P)$ for the active state. The solution consists of these functions and the rules for optimally switching between the states 1 and 0. This is a problem in stochastic dynamic programming. I give the formal theory in appendix A (available from the author); here I develop the solution using the option pricing analogy.[6]

Over the range of prices P where it is optimal for an idle firm to continue in this state, the asset of the investment opportunity must be willingly held. There being no operating profit, the only return to this asset is the expected capital gain, $EdV_0(P)/dt$, as the value $V_0(P)$ changes with P. This must then equal the normal return $\rho V_0(P)$.

By Itô's lemma, we have

$$dV_0 = V_0'(P)dP + \tfrac{1}{2}V_0''(P)\sigma^2P^2dt.$$

Therefore,

$$E(dV_0) = [V_0'(P)\mu P + \tfrac{1}{2}V_0''(P)\sigma^2P^2]dt.$$

The asset equilibrium condition becomes the differential equation

$$\tfrac{1}{2}\sigma^2P^2V_0''(P) + \mu PV_0'(P) - \rho V_0(P) = 0. \qquad (3)$$

The return on the asset constituting an active project can be calculated similarly. The only difference is that there is a dividend, namely the flow of operating profit, in addition to the expected capital gain. Therefore, over the interval of prices at which it is optimal for an active firm to continue being active, the value of the asset $V_1(P)$ must satisfy the differential equation

$$\tfrac{1}{2}\sigma^2P^2V_1''(P) + \mu PV_1'(P) - \rho V_1(P) = w - P. \qquad (4)$$

The general solutions of (3) and (4) are easy to obtain. Both are linear and have the same homogeneous part. Therefore, we can find the complementary functions together. Try a solution of the form P^{ξ}. Substitution yields

$$\tfrac{1}{2}\sigma^2\xi(\xi - 1) + \mu\xi - \rho = 0$$

[5] When such a model of the firm is used in construction of the industry's equilibrium, there must be restricted access to the technology, e.g., a given distribution of firms across values of (w, k, l), to allow $V_0(P) \neq 0$. If there is free entry, the price process must adjust endogenously to ensure $V_0(P) = 0$ for all relevant P. I leave this step for future research.

[6] Of course the formal theory of option pricing is itself an application of stochastic dynamic programming.

or

$$\phi(\xi) \equiv \xi^2 - (1 - m)\xi - r = 0, \tag{5}$$

where I have defined $m \equiv 2\mu/\sigma^2$ and $r \equiv 2\rho/\sigma^2$. The convergence condition is now $r > m$. Therefore, $\phi(0) = -r < 0$ and $\phi(1) = -(r - m) < 0$. Since $\phi''(\xi) \equiv 2 > 0$, one root must be greater than one (call it β), and the other must be less than zero (call it $-\alpha$). Written out explicitly,

$$\beta = \frac{(1 - m) + [(1 - m)^2 + 4r]^{1/2}}{2},$$

$$-\alpha = \frac{(1 - m) - [(1 - m)^2 + 4r]^{1/2}}{2}.$$

Next consider a particular solution for the nonhomogeneous equation (4). Trying a linear form and solving for the coefficients, we get $[P/(\rho - \mu)] - (w/\rho)$. Thus we can write the general solution of (3) as

$$V_0(P) = A_0 P^{-\alpha} + B_0 P^{\beta}$$

and that of (4) as

$$V_1(P) = A_1 P^{-\alpha} + B_1 P^{\beta} + \left(\frac{P}{\rho - \mu} - \frac{w}{\rho}\right),$$

where A_0, B_0, A_1, and B_1 are constants to be determined.

The terms in parentheses in the expression for $V_1(P)$ have a very useful interpretation. Since the expected value of P rises at the trend rate μ, we have

$$\frac{P}{\rho - \mu} - \frac{w}{\rho} = E\left[\int_0^{\infty} (P_t - w)\exp(-\rho t)dt\right].$$

This is just the expected present value that can be obtained by keeping the project active forever and, come what may, starting from an initial price P. Therefore, the remaining part of the solution, the complementary function, must be the value of the option to shut down optimally. Similarly, since the idle firm has no current operating profit, the whole of the solution for $V_0(P)$ must be the value of the option to become active at the optimal time.

This interpretation also gives us some natural endpoint conditions. If P is very small, the event of its rising to P_H in any given finite amount of time has a very small probability. Therefore, the option of activating should be nearly worthless. For this, we need $A_0 = 0$. Similarly, considering large P, we have $B_1 = 0$. Then we can omit the subscripts on the remaining coefficients and write the solutions as

$$V_0(P) = BP^{\beta} \tag{6}$$

and

$$V_1(P) = AP^{-\alpha} + \frac{P}{\rho - \mu} - \frac{w}{\rho}. \tag{7}$$

Note that an economically meaningful solution for V_0 must be nonnegative, so in (6) we must have $B \geq 0$. Similarly, an active firm has the expected value of $[P/(\rho - \mu)] - (w/\rho)$ from the feasible strategy of never shutting down; therefore, the optimal strategy must yield no less, and in (7) we must have $A \geq 0$.

It remains to determine the two parameters A and B. For this we must link the two regimes and consider the optimal transitions from the idle to the active state and vice versa.

Suppose that P_H is the price that triggers entry.[7] The firm pays k to exercise that option and gets an asset of value $V_1(P)$. Therefore, P_H must satisfy the standard pair of conditions for optimal exercise, the value-matching condition

$$V_0(P_H) = V_1(P_H) - k \tag{8}$$

and the high-order contact or smooth pasting condition[8]

$$V_0'(P_H) = V_1'(P_H). \tag{9}$$

Similarly, the price P_L that triggers exit satisfies the value-matching condition

$$V_1(P_L) = V_0(P_L) - l \tag{10}$$

and the smooth pasting condition

$$V_1'(P_L) = V_0'(P_L). \tag{11}$$

Written out in terms of the functional forms of the candidate solutions, the value-matching conditions (8) and (10) become

$$AP_L^{-\alpha} + \frac{P_L}{\rho - \mu} - \frac{w}{\rho} = BP_L^{\beta} - l \tag{12}$$

and

$$AP_H^{-\alpha} + \frac{P_H}{\rho - \mu} - \frac{w}{\rho} = BP_H^{\beta} + k, \tag{13}$$

[7] To be rigorous, we should prove that there is a unique solution with such a property. See app. A for this.

[8] For a discussion of this in the option pricing problem, see Merton (1973, p. 171, n. 60). A heuristic derivation based on stochastic dynamic programming is in app. A. The intuition is that if the slopes of these two value functions differ, then one can use the kink formed by them to depart from the supposedly optimal policy and improve the payoff.

while the smooth pasting conditions (9) and (11) become

$$-A\alpha P_L^{-\alpha-1} + \frac{1}{\rho - \mu} = B\beta P_L^{\beta-1} \tag{14}$$

and

$$-A\alpha P_H^{-\alpha-1} + \frac{1}{\rho - \mu} = B\beta P_H^{\beta-1}. \tag{15}$$

The four equations (12)–(15) determine A, B, P_H, and P_L, completing the solution. The equations are nonlinear in P_H and P_L. They yield some important analytical results, but numerical simulations are needed to get a better and more quantitative idea of the properties of the solution. I shall consider the two in turn.

IV. Analytical Results

The most important general result concerns the nature of hysteresis discussed in Section II. To obtain it, define

$$G(P) = V_1(P) - V_0(P). \tag{16}$$

Using the solutions (6) and (7), we have

$$G(P) = AP^{-\alpha} - BP^{\beta} + \frac{P}{\rho - \mu} - \frac{w}{\rho}. \tag{17}$$

The value-matching and smooth pasting conditions can be written in terms of G as

$$G(P_L) = -l,\ G(P_H) = k,\ G'(P_L) = 0,\ G'(P_H) = 0. \tag{18}$$

The general shape of $G(P)$ is as shown in figure 1. Then to solve the problem, we must adjust A and B until $G(P)$ becomes tangent to the horizontal lines at $-l$ and k, and the respective points of tangency define P_L and P_H. Note that

$$G''(P_L) > 0, \qquad G''(P_H) < 0. \tag{19}$$

Now subtract (3) from (4) to see that $G(P)$ satisfies the differential equation

$$\tfrac{1}{2}\sigma^2 P^2 G''(P) + \mu G'(P) - \rho G(P) = w - P. \tag{20}$$

Evaluate this at P_H and use (18) and (19) to get

$$\begin{aligned} w - P_H &= \tfrac{1}{2}\sigma^2 P_H^2 G''(P_H) + \mu G'(P_H) - \rho G(P_H) \\ &< -\rho k \end{aligned}$$

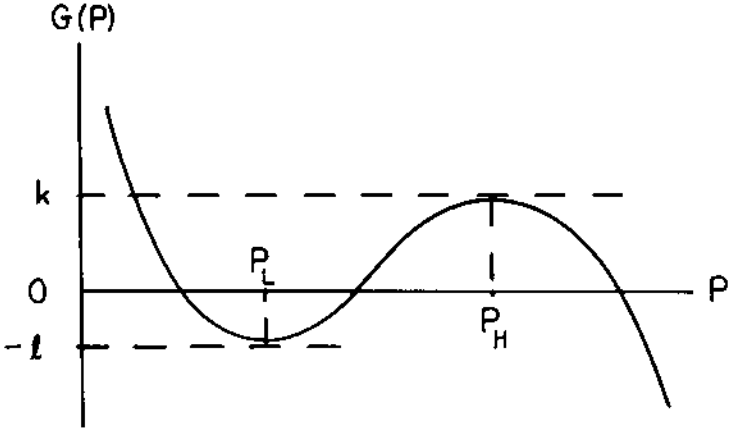

FIG. 1.—Determination of P_H and P_L

or

$$P_H > w + \rho k \equiv W_H. \quad (21)$$

Similarly,

$$P_L < w - \rho l \equiv W_L. \quad (22)$$

This is the effect of uncertainty discussed in Section II. The Marshallian trigger prices for investment and abandonment are, respectively, W_H and W_L; the former is the usual full cost, but the latter differs from the variable cost w of Section II because we now have a lump-sum exit cost. At a price between these limits, an idle firm does not invest and an active firm does not exit. Now (21) and (22) show that uncertainty widens this Marshallian range of inaction. In the next section I shall examine the quantitative significance of this widening for a range of parameter values.

Next consider some limiting results. These are so obvious that I shall merely state them and omit the derivations. As both k and l tend to zero, both P_H and P_L tend to the common limit w; thus sunk costs are essential for hysteresis. If only one of k and l tends to zero while the other stays positive, however, both inequalities (21) and (22) remain strict. For example, even if there is no exit cost as such ($l = 0$), the exit trigger P_L remains below w. The firm knows that by remaining active it can avoid incurring k for reentry should future developments turn favorable; therefore, it is willing to incur some current loss to preserve this option.

If $l > w/\rho$, the project is never abandoned. However, P_H does not go to infinity; there is a finite price that will attract a firm to a project that is impossible to get out of. In equations (12)–(15), A goes to zero because the option of exiting is now worthless. Then we can solve (13) and (15) in closed form for B and P_H. We have[9]

$$P_H = \frac{\rho - \mu}{\rho} \frac{\beta}{\beta - 1} W_H, \tag{23}$$

where W_H is as defined above in (21).

Conversely, if k goes to infinity, the entry option becomes worthless and B goes to zero. Then we can solve (12) and (14) for A and P_L. We have

$$P_L = \frac{\rho - \mu}{\rho} \frac{\alpha}{\alpha + 1} W_L. \tag{24}$$

This is how bad things must get before a project will be abandoned, if one knows that one can never reinvest in it later.

If $\sigma \rightarrow 0$, $P_H \rightarrow W_H$ and $P_L \rightarrow W_L$. Thus we verify that in the absence of uncertainty, only the Marshallian zone of inaction remains.

Consider next the comparative statics of P_H and P_L with respect to w, k, and l. As usual, we totally differentiate (12)–(15) and solve. This is simple but tedious, and I relegate the details to appendix B (available from the author). Most of the results are obvious. As w increases, both P_L and P_H increase. However, when k increases, P_L decreases and P_H increases; that is, the hysteresis effect becomes more pronounced. Similar effects stem from l. Of course, P_L and P_H are homogeneous of degree one in (w, k, l) jointly.

A limiting comparative static result is of considerable importance. If we keep σ fixed at a positive level and let $k \rightarrow 0$, we have $dP_H/dk \rightarrow \infty$ and $dP_L/dk \rightarrow -\infty$. In other words, when there is some uncertainty, hysteresis emerges very rapidly even for very small sunk costs; similarly when $l \rightarrow 0$. I shall show this more vividly in numerical calculations in the next section.

V. Numerical Results

I shall establish a set of central values for the parameters and examine a wide range of variations around this. Begin with the relation between variable and sunk costs. One might think of these as labor and capital costs, respectively, and try a ratio of $w:\rho k = 2:1$. However, some capital costs arise from depreciation and are more properly

[9] This satisfies $P_H > w + \rho k$ if $\rho/\mu > \beta$; that is true since (5) gives $\phi(\rho/\mu) > 0$. The result is not quite that of McDonald and Siegel (1986) because their operating profit follows a different process.

thought of as recurrent.[10] Other capital costs are recoverable on exit, while a significant portion of labor costs is sunk. In all, a ratio of 10 : 1 seems reasonable for the central case. With this in mind, I chose $w = 1$ (merely a normalization), $k = 4$, and $\rho = 2.5$ percent, so $\rho k = 0.1$. I let $l = 0$; exit costs are probably more important in Europe than in the United States.

As for uncertainty, I take $\sigma = 0.1$ for the central case. This means that ln P has a variance of 1 percent per year. Since the standard deviation goes as the square root of time, its value is 10 percent over 1 year or 20 percent over 4 years. This is reasonable for real exchange rate fluctuations (see Frankel and Meese 1987) and low for the prices of many natural resources (see Brennan and Schwartz 1985). Finally, I set $\mu = 0$.

In the central case of parameter values, I find $P_H = 1.4667$, which is 33 percent above the full cost $W_H = 1.1$, and $P_L = 0.7657$, which is 24 percent below the variable cost $W_L = w = 1$. These gaps are much bigger than the Marshallian gap between full and variable costs. In other words, hysteresis is very significant, and a major part of the full gap between P_H and P_L arises from the uncertainty and option values.

In international trade, hysteresis bears on dumping. Most economists think of dumping as international price discrimination, which occurs under conditions of imperfect competition. But the trade laws of the United States and most other countries deem dumping to occur also if the price of an imported good is below its cost of production abroad plus delivery to the home market, irrespective of the market structure. The cost is usually taken to be average cost. My model shows that even small firms can practice dumping in this sense. For example, if P is only slightly above P_L for the numerical values above, the firm will remain in the market and go on selling at a price that is more than 30 percent below full cost, and even 24 percent below variable cost.[11] But it is important to note that there is nothing irrational or strategically manipulative about this. Whatever the merits of antidumping tariffs when the dumping is an act of strategic price discrimination, such tariffs have no economic merit when price falls below cost for a while as uncertainty unfolds but atomistic firms rationally remain in the market. In the same way, discrepancies between the marginal product of a worker and the wage need not be the result

[10] Exponential depreciation is easily handled by thinking of it as a maintenance expenditure needed to keep the project alive. Let δ be the rate of depreciation. Then we merely replace w by $w + \delta k$.

[11] See Ethier (1982) for a different model of nonstrategic dumping under uncertainty.

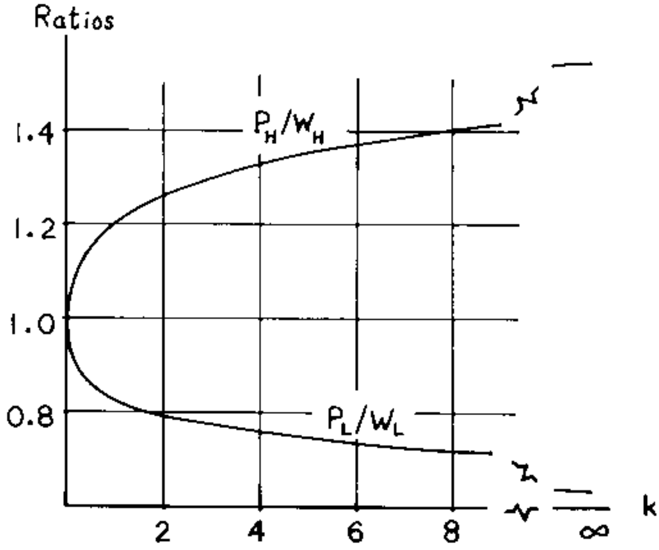

FIG. 2.—Effects of changes in k

of monopsony power, but could result from rational decisions by atomistic firms to hire and fire workers (see Becker 1962, p. 22).

Consider next the effect of varying the parameters around the central case. Figures 2–4 show a sample of the calculations I have performed. In each, one of the parameters is allowed to vary while the others are held fixed at the base levels.

Figure 2 shows the effect of variation in the sunk cost k. The upper curve is the ratio of the entry price P_H to the full cost W_H, while the lower curve is the ratio of the exit price P_L to W_L. Remember that at $k = 0$, the curves emerge from the common limiting value of one with respective slopes of ∞ and $-\infty$. Therefore, hysteresis is quite strong even for small values of k. For example, at $k = 0.4$, meaning that the interest on the sunk cost, ρk, is only 1 percent of the variable cost $w = 1$, we find the exit price 13 percent below the variable cost and the entry price 15 percent above the full cost. At the other extreme, for very large k, P_L goes to 0.64, the value given by (24).

Figure 3 shows the effect of changes in σ, the standard deviation of the price process. Once again hysteresis remains significant even for quite small σ. When $\sigma = 0.025$, the exit price is still 11 percent below the variable cost and the entry price is 10 percent above the full cost. These gaps are as large as that between W_L and W_H. Even a little uncertainty matters a lot.

Figure 4 shows the effect of variation in μ, the trend of the price process. An increase in μ reduces both P_H and P_L. This accords with

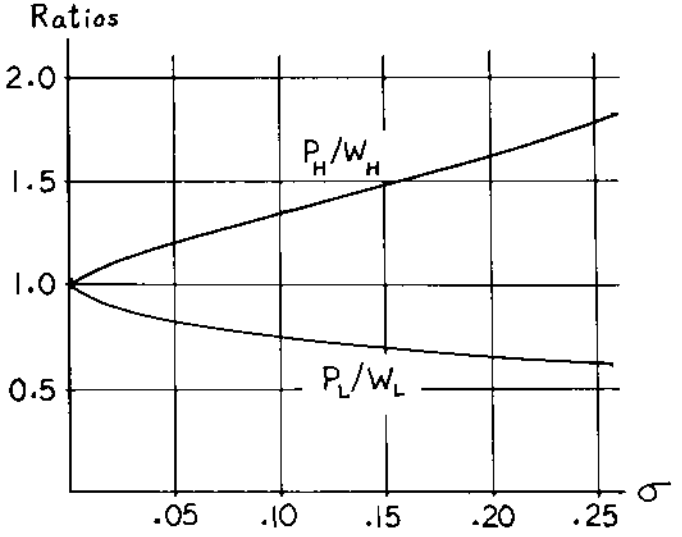

FIG. 3.—Effects of changes in σ

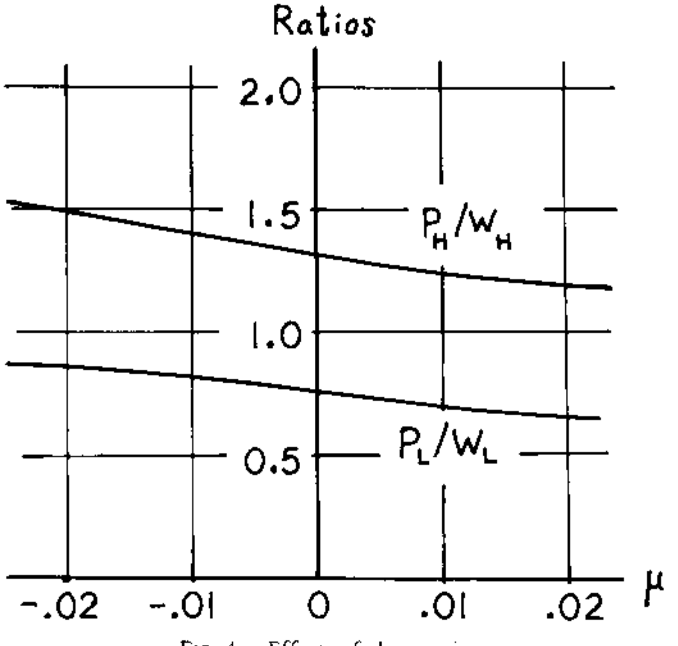

FIG. 4.—Effects of changes in μ

intuition: if the firm expects a favorable price trend, it enters the market at a lower threshold of current profitability and, once in, is more willing to hang on despite a temporarily adverse price. However, the quantitative significance of this effect seems relatively mild. It appears that the ratio of P_H to full cost goes to one as μ goes to infinity,[12] and that of P_L to w goes to one as μ goes to negative infinity.

Finally, as ρ increases, both triggers P_H and P_L rise, although quite slowly. To save space, I omit the details. Thus investment is more reluctantly made and more easily abandoned. This is just the overall decrease in investment as the interest rate rises.

VI. Some Extensions

Here I consider several ways in which the simple model above can be extended to handle other issues or make it more realistic. The general conclusion is that the method of analysis and the qualitative results carry over; many realistic modifications in fact add to the hysteresis.

A. *Other Kinds of Uncertainty*

It can be argued that many prices will show a tendency toward some predictable long-run equilibrium levels or paths, even though they may fluctuate in response to various random short-run influences. As a simple example, consider the mean-reverting process

$$dP = \lambda(P^* - P)dt + \sigma P dz,$$

where P^* is the fixed long-run equilibrium or mean level. When this replaces (1), the differential equation (3) for V_0 is replaced by

$$\tfrac{1}{2}\sigma^2 P^2 V_0''(P) + \lambda(P^* - P)V_0'(P) - \rho V_0(P) = 0,$$

and similarly for V_1. These equations do not have closed-form solutions, but it is intuitively obvious that the result can be only a widening of the range of inaction, that is, an increase in P_H and a decrease in P_L. For example, when the current price is high, mean reversion makes the future outlook less favorable, and therefore the firm is more reluctant to enter.

If there are floors and ceilings on the price process, for example, because of government intervention, there should be a similar effect on the entry and exit trigger prices. Take the solution P_H of the unconstrained problem, but now suppose that the price process is constrained by a reflecting barrier at $\bar{P} > P_H$. Now at $P = P_H$, the prospects are less favorable than they were with the unconstrained

[12] This limit is a formalism since convergence requires $\mu < \rho$.

process; therefore, a still higher price is needed to trigger entry. Similarly, a lower reflecting barrier $\underline{P} < P_L$ should lower the exit trigger price. Sufficiently tight barriers may put an end to all entry and exit.

We can also consider Poisson jump processes, which are merely a probabilistic version of the mean reversion discussed in Section II.

B. *Variable Scale of Output*

Suppose that after the investment is in place, the output Q can be varied. If there is no fixed cost or any other source of economies of scale, then the investment need never be abandoned since it can be kept alive at an infinitesimal loss by choosing Q very small. But many investments must be operated at or above a minimum scale to prevent rusting. To model this, suppose that the production function in operation is Cobb-Douglas with diminishing returns but that there is a flow fixed cost. Then the operating profit is $P^\theta - f$, where the multiplicative constant is set at unity, $\theta > 1$, and f is the flow fixed cost. Define $\Pi = P^\theta$. By Itô's lemma,

$$d\Pi = \theta P^{\theta-1} dP + \tfrac{1}{2}\theta(\theta - 1)P^{\theta-2}(\sigma P)^2 dt$$

or

$$\begin{aligned}\frac{d\Pi}{\Pi} &= \theta\left[\mu + \frac{(\theta - 1)\sigma^2}{2}\right]dt + \theta\sigma dz \\ &= \hat{\mu}dt + \hat{\sigma}dz,\end{aligned}$$

for new constants $\hat{\mu}$ and $\hat{\sigma}$. This has exactly the same mathematical form as the process for P. Therefore, in the basic differential equations (3) and (4) we need only to replace P by Π and w by f and use $\hat{\mu}$ and $\hat{\sigma}$ instead of μ and σ. Then the equations yield solutions for V_0 and V_1 as functions of Π exactly as in Section III, and all the qualitative results of Sections IV and V remain valid. The added flexibility will make the firm quicker to invest and slower to abandon; that is, both P_H and P_L will be lower.

C. *Variable Scale of the Project*

We can parameterize the scale of the project by s and make the variable costs $w(s)$. The costs of entry and exit are more appropriately regarded as adjustment costs, that is, functions of changes in s. Then the problem is one of finding the optimal policies for s. It is intuitively clear that the trigger prices P_H and P_L will be replaced by functions $P_H(s)$ and $P_L(s)$, giving rise to a band of inaction. A similar model of industry equilibrium can be constructed. For that, one must move the uncertainty to a deeper level, for example demand or exchange rates, and endogenize P. Once again we have entry and exit functions relating the number of active firms to the underlying uncertainty, and a band of inaction between the two. The model of this paper proves a useful grounding for such extensions, but they need enough additional work to merit separate treatment. Lucas and Prescott (1974) develop the qualitative general theory using stochastic dynamic programming in the context of job search. Some very recent applications and numerical simulations using the option pricing approach are Bentolila and Bertola (1987) and Dixit (1989*a*, 1989*b*).

Finally, one might allow the firm to choose its optimal degree of flexibility, reducing k or making l more negative, by accepting a higher w along a specified schedule. It is intuitively clear that the higher σ, the greater the benefit from such a "rustproofing" strategy. But a formal treatment would need too much space to be attempted here.

D. *Risk Aversion*

The basic differential equations (3) and (4) for a risk-neutral firm were derived by equating the total rate of return in each of the idle and active states to the risk-free rate ρ. The simplest way to allow risk aversion would be to replace ρ by the appropriate risk-adjusted rate from a capital asset pricing model.

For example, the required rate of return for an idle firm, ρ_0, is given by

$$\rho_0 = \rho + [E(\rho_M) - \rho] \frac{\text{cov}(dV_0/V_0, \rho_M)}{\text{var}(\rho_M)},$$

where ρ_M is the rate of return on the market portfolio. With Itô's lemma, the covariance becomes

$$\sigma\left[\frac{PV_0'(P)}{V_0(P)}\right]\text{cov}(dz, \rho_M).$$

Then

$$\rho_0 = \rho + \eta\sigma\left[\frac{PV_0'(P)}{V_0(P)}\right], \tag{25}$$

where

$$\eta \equiv [E(\rho_M) - \rho] \frac{\text{cov}(dz, \rho_M)}{\text{var}(\rho_M)}. \tag{26}$$

Replacing ρ by ρ_0 in (3) and simplifying, we have

$$\tfrac{1}{2}\sigma^2 P^2 V_0''(P) + (\mu - \eta\sigma) P V_0'(P) - \rho V_0(P) = 0. \tag{27}$$

The calculation of ρ_1 for the active firm is similar, and its differential equation is

$$\tfrac{1}{2}\sigma^2 P^2 V_1''(P) + (\mu - \eta\sigma) P V_1'(P) - \rho V_1(P) + (P - w) = 0. \tag{28}$$

These equations are of the same form as (3) and (4), with μ changed to $\mu - \eta\sigma$. The solution can be completed in exactly the same way. If the risk in P is positively correlated with the market risk, then $\eta > 0$. The effect of risk aversion is to act as if μ is lower. From figure 4, this is to raise both P_H and P_L, which accords with intuition.

Matters are even simpler if the output from the project is an asset, for example, foreign exchange, a natural resource, or a durable good. We know that the asset has an expected rate of capital gain μ; suppose that it pays a flow dividend at rate δ. The capital asset pricing model formula for it is

$$\begin{aligned}\delta + \mu &= \rho + [E(\rho_M) - \rho]\,\frac{\operatorname{cov}(\sigma dz, \rho_M)}{\operatorname{var}(\rho_M)} \\ &= \rho + \eta\sigma.\end{aligned} \tag{29}$$

Now we can replace $\mu - \eta\sigma$ by $\rho - \delta$ in (27) and (28). The advantage is that δ is more readily observable than η.

In financial theory, this argument is developed using a replicating portfolio. Consider being long the firm and short $V_0'(P)$ units of output. The value Z of this portfolio evolves according to

$$dZ = dV_0 - V_0'(P)dP = \tfrac{1}{2}\sigma^2 P^2 V_0''(P)dt.$$

Therefore, the portfolio is riskless. The capital gain dZ minus the dividend to be paid for being short on output $-\delta V_0'(P)dt$ should amount to the riskless return on the portfolio whose initial value is $Z = V_0(P) - PV_0'(P)$. Therefore,

$$\tfrac{1}{2}\sigma^2 P^2 V_0''(P) - \delta V_0'(P) = \rho[V_0(P) - PV_0'(P)].$$

In view of (29), this is the same as (27) above. The derivation of (28) is similar.

VII. Concluding Remarks

This paper has constructed a model of optimal inertia in investment decisions under uncertainty and suggested a wide variety of applications ranging from foreign trade to job search. The concept of option value in these contexts is frequently mentioned in the literature; the specific model makes the idea precise and demonstrates its considerable quantitative significance. Practitioners probably understand this at an intuitive level very well: witness the great reluctance of university deans to approve new faculty positions in departments that experience a surge of students. I hope that the theoretical treatment deepens economists' understanding of the issue and opens up the way for treating further problems of this kind.

References

Baldwin, Richard E. "Hysteresis in Trade." Working paper. Cambridge: Massachusetts Inst. Tech., July 1986.

Baldwin, Richard E., and Krugman, Paul R. "Persistent Trade Effects of Large Exchange Rate Shocks." Working paper. Cambridge: Massachusetts Inst. Tech., August 1986.

Becker, Gary S. "Investment in Human Capital: A Theoretical Analysis." *J.P.E.* 70, no. 5, pt. 2 (October 1962): 9–49.

Bentolila, Samuel, and Bertola, Giuseppe. "Firing Costs and Labor Demand: How Bad Is Eurosclerosis?" Working paper. Cambridge: Massachusetts Inst. Tech., December 1987.

Bertola, Giuseppe. "Irreversible Investment." Working paper. Cambridge: Massachusetts Inst. Tech., October 1987.

Brennan, Michael J., and Schwartz, Eduardo S. "Evaluating Natural Resource Investments." *J. Bus.* 58 (April 1985): 135–57.

Constantinides, George M. "Capital Market Equilibrium with Transaction Costs." *J.P.E.* 94 (August 1986): 842–62.

Dixit, Avinash. "Hysteresis, Import Penetration, and Exchange Rate Pass-through." *Q.J.E.* (1989), in press. (*a*)

———. "Intersectoral Capital Reallocation under Price Uncertainty." *J. Internat. Econ.* (1989), in press. (*b*)

Ethier, Wilfred J. "Dumping." *J.P.E.* 90 (June 1982): 487–506.

Frankel, Jeffrey A., and Meese, Richard. "Are Exchange Rates Excessively Variable?" In *NBER Macroeconomics Annual*, vol. 2. Cambridge, Mass.: MIT Press, 1987.

Lucas, Robert E., Jr., and Prescott, Edward C. "Equilibrium Search and Unemployment." *J. Econ. Theory* 7 (February 1974): 188–209.

McDonald, Robert L., and Siegel, Daniel R. "Investment and the Valuation of Firms When There Is an Option to Shut Down." *Internat. Econ. Rev.* 26 (June 1985): 331–49.

———. "The Value of Waiting to Invest." *Q.J.E.* 101 (November 1986): 707–27.

Merton, Robert C. "Theory of Rational Option Pricing." *Bell J. Econ. and Management Sci.* 4 (Spring 1973): 141–83.

Mossin, Jan. "An Optimal Policy for Lay-up Decisions." *Swedish J. Econ.* 70 (September 1968): 170–77.

Oi, Walter Y. "Labor as a Quasi-fixed Factor." *J.P.E.* 70 (December 1962): 538–55.

Pindyck, Robert S. "Irreversible Investment, Capacity Choice, and the Value of the Firm." Working paper. Cambridge: Massachusetts Inst. Tech., 1987.