IRREVERSIBILITY, UNCERTAINTY, AND CYCLICAL INVESTMENT*

BEN S. BERNANKE

This paper builds on the theory of irreversible choice under uncertainty to give an explanation of cyclical investment fluctuations. The key observation is that, when individual projects are irreversible, agents must make investment timing decisions that trade off the extra returns from early commitment against the benefits of increased information gained by waiting. In an environment in which the underlying stochastic structure is itself subject to random change, events whose long-run implications are uncertain can create an investment cycle by temporarily increasing the returns to waiting for information.

> But I suggest that the essential character of the trade cycle . . . is mainly due to the way in which the marginal efficiency of capital fluctuates . . . . (T)he marginal efficiency of capital depends, not only on the existing abundance or scarcity of capital-goods and the current cost of production of capital-goods, but also on current expectations as to the future yield of capital-goods . . . . But, as we have seen, the basis for such expectations is very precarious. Being based on shifting and unreliable evidence, they are subject to sudden and violent changes.
>
> Keynes, *The General Theory*, Ch. 22.

INTRODUCTION

Sharp swings in the production of durable goods are yet, as in Keynes's day, an important feature of the business cycle. We are not without explanations for these fluctuations: The most prominent derive from various "accelerator" models of investment.[1] Accelerator models are based on the observation that relatively small percentage changes in desired capacity *levels* may equal large percentage changes in desired capacity *first differences*, which drive investment. In an

* I am indebted to Stanley Fischer for helpful conversations.

1. Basic references on the accelerator include Chenery [1952], Koyck [1954], and Hall and Jorgenson [1967]. Samuelson [1939] first discussed the relation of the accelerator to the business cycle.

The Quarterly Journal of Economics, February 1983 CCC 0033-5533/83/010085-22$03.20

accelerator model with adjustment costs or delivery lags, unanticipated shifts in desired capacity can generate cycle-like swings in durable goods spending.[2]

There is little question that the accelerator has a role in explaining durable goods cycles; but it is also true that this model does not square with some common observations about recession. Businessmen cutting investment spending during a recession, for example, do not seem to be aware of being on a smooth transition path to a lower level of capital intensity.[3] They do not typically have a firm expectation of when or in what form capital spending will revive, nor do they move up future projects to take advantage of lower costs of capital and shorter delivery lags in the trough. Not a purposeful adjustment of stocks to a lower level, investor behavior in recession is better described as a cautious probing, an avoidance of commitment until the longer run status of both the national economy and the investor's own fortunes are better known. There is a general perception of low points in the cycle as times of low "confidence," a fact that the accelerator is inadequate to explain.

This paper offers an alternative but potentially complementary analysis of cyclical fluctuations in durables, one emphasizing uncertainty and the flow of information. This analysis is based on two assumptions: First, that individual investment projects are economically *irreversible;* once constructed, they cannot be "undone" or made into a radically different type of project without high costs. (The existence of used-capital markets can be accommodated; see note 16.) Second, that new information relevant to assessing long-run project returns arrives over time; so that, by waiting, the potential investor can improve his chances of making a correct decision.

Under these assumptions, the optimizing investor must decide not only which projects to undertake but also which point in time is the best for making a commitment. Postponing commitment will be desirable if improved information is more valuable to the investor than short-run return. Thus, the dynamics of investment become very sensitive to expectations about the rate of information arrival. We shall show that the interactions of investor learning and the optimal timing of investments may plausibly give rise to sharp fluctuations in the demand for capital goods, thereby helping to explain the

2. The assumption of adjustment costs also serves to eliminate the implication of infinitely fast investment response.

3. Indeed, until recently the rate of investment spending during recovery has nearly always matched or exceeded the pre-recession rate.

business cycle. This application of irreversible investment theory to business cycle analysis is unique, as far as we know.[4]

The paper has three principal sections. Section I reviews the problem of choosing among irreversible investments under uncertainty. The "option value" associated with irreversibility (a concept drawn from the economics of environmental preservation literature) is developed for the multiple investments case. The relationship of the option value to investor caution is examined.

Section II, which looks at irreversible investment dynamics, contains the more novel aspects of the paper. A useful specification of the stochastic economic environment, based on Howard's [1964] dynamic inference model, is introduced. In that context, an extended example is used to show how the occurrence of a single event with uncertain implications can cause a cyclic movement in durables production as those implications work themselves out. Finally, Section III examines the link between the microeconomic analysis of investment and aggregate phenomena.

I. Choosing Among Irreversible Alternatives

A. The Basic Problem

To prepare the ground for our macroeconomic application, we consider the problem of choosing among irreversible alternatives when information pertinent to the choice is arriving over time. This section notes some selected earlier work in this area and develops an investor decision rule.

Since the insightful contribution of Marschak [1949], economists have studied the irreversibility (or "illiquidity") problem in a number of contexts. A particularly productive line of research is to be found in the economics of environmental preservation. The key question addressed there is, "When can an effectively irreversible action, such as the cutting down of a sequoia forest, be economically justified?" Weisbrod [1964], Arrow and Fisher [1974], Henry [1974a, 1974b], and others have developed the point that, under uncertainty, there is an

4. Since submitting this paper, I have been informed by a referee of recent work by Cukierman [1980], which takes a similar approach to mine in explaining investment variations. Both of our papers stress the idea that, when projects are irreversible, uncertainty can depress current investment by making waiting for information more attractive. The present paper is distinguished from that of Cukierman principally by its development, in this context, of the option value concept; by its explicit analysis of the investment dynamics that can result when agents make optimal timing decisions; by its use of the dynamic inference model; and by its extended discussion of irreversibility as a source of aggregate, not just micro-level, investment fluctuations.

option value associated with avoiding irreversible actions. The continued existence of sequoias, for example, is valuable not only for the direct recreational benefits; it is valuable also because it means that society, by not taking the irreversible step of clearing the forest, has retained the option of choosing between the sequoias and alternatives (such as logging) in the future. In contrast, cutting the trees down today eliminates the option of preservation, even if new information should reveal that preservation is society's truly preferred alternative. Because of this option value (these papers show), irreversible acts will not be justified in some cases where their reversible counterparts should be undertaken.

The option value comes into play with equal force when there are many irreversible alternatives rather than one. Because the option value and its determinants are central to the subsequent discussion, we shall develop this fact in a simple model. Only a modest generalization in itself, this model is intended to help fix ideas and provide a basis for the subsequent analysis.

Consider an investor (or a central planner) who may either (1) choose at most one of k irreversible investment projects,[5] or (2) defer commitment (forfeiting current investment returns) in order to gather more information. It is important that the investment projects be mutually excluding *alternatives*. This may arise because of limits on resources or time; or, think of an investor choosing among putty-clay technologies for a proposed new plant. The indivisibility of the various projects is an inessential restriction.[6]

For simplicity, take the decision period to be discrete and the horizon finite; the current period is t and time ends after period T. The set of possible information-states in period s is I_s, $s = t, t+1, \ldots, T$. Also letting the number of information-states be finite, take I_s to be an $n_s \times 1$ vector, with I_s^j a representative element. The information-state vectors are assumed to have a (subjective) joint probability distribution, so that conditional probability matrices of the form,

$$p(I_r | I_s) \qquad r > s$$

are well-defined. The matrix $p(I_r | I_s)$ is dimensioned $n_s \times n_r$.

5. Since a "project" is to be defined only as a set of date-state-returns combinations, our model handles the case of randomly arriving or departing opportunities without loss of generality. Thus, although we shall speak as if the investor faces a fixed menu of projects, some of the k projects represent opportunities that cannot currently be acted on. "Choosing" such a project at time t means forswearing all alternatives, even if the opportunity never actually materializes. That such precommitment is generally inferior to remaining flexible is an obvious corollary of the decision rules derived below.

6. The case of divisible projects was handled in an earlier version of this paper.

Once undertaken, each of the k projects pays a return in each period that depends on the contemporaneous information-state. This return (which may be positive or negative) is thus known for the decision period and is a random variable for future periods. Let the set of possible returns to be paid by the ith investment in period s be designated by the $n_s \times 1$ vector $r_{i,s}$. Then $r^j_{i,s}$ is the return corresponding to information-state I^j_s.[7]

We assume present-value maximization (risk neutrality). The purpose of this assumption (which can easily be relaxed) is to underscore the independence of "investor caution" analyzed below and usual concepts of risk aversion. We define recursively the $n_s \times 1$ vector of *total expected returns* to investment i, $R_{i,s}$, by

$$\begin{aligned} R_{i,s} &= r_{i,s} + \beta p(I_{s+1}|I_s)R_{i,s+1} \\ &= r_{i,s} + \beta E_s R_{i,s,+1} \qquad s = t,t+1,\ldots,T, \end{aligned}$$

where the vector of scrap values $R_{i,T+1}$ is given exogenously, and β is a scalar discount factor that may be time- and state-dependent. For project i, the element $R^j_{i,s}$ is the total expected return in s corresponding to information-state I^j_s.

In each period t the investor can either lock into some project i (and receive the realizations of $r_{i,s}$ for $s = t,\ldots,T$) or defer his decision. Deferral means that the investor receives no current return but may reconsider the projects after observing the information-state drawn in $t + 1$.

Clearly no investment i will be made, given I^j_t, unless both $r^j_{i,t} \geq 0$ and $R^j_{i,t} = \max(R^j_{1,t},\ldots,R^j_{k,t})$. Yet with irreversibility these conditions are not sufficient. The Appendix proves

PROPOSITION 1. The value-maximizing strategy for this problem is as follows: Invest in project i, given I^j_t, if and only if

$$R^j_{it} \geq \max(R^j_{1t},\ldots,R^j_{kt},V^j_t), \tag{1}$$

where

$$V_s = \beta E_s\{\overline{\max}(R_{1,s+1},\ldots,R_{k,s+1},V_{s+1})\} \qquad (n_s \times 1) \tag{2}$$

and

$$V_T = (0,0,\ldots,0)^{\text{transpose}} \qquad (n_T \times 1). \tag{3}$$

If no project i satisfies (1), defer commitment.

7. These returns may be thought of as including changes in the initial cost of investment, amortization, and depreciation costs. The assumptions that the initial return is known and that the gestation period is zero are easily relaxed; see Bernanke [1979].

The operator $\overline{\max}(A,B)$ forms a new vector whose jth element is $\max(A^j,B^j)$. V_t^j is the jth element of V_t, corresponding to I_t^j.

Proposition 1 tells us that not only must the expected return to an acceptable investment be the largest of the k alternatives, but in addition it must exceed a reservation level V_t^j. V_t^j may be identified with the value function of this problem; it is the "expected value of deferring commitment." Note that V_t^j is always nonnegative, and that (1) therefore leads to the possibility that no investment will be made even if positive returns are available.

The relation of Proposition 1 to the option value concept can be seen clearly if we rewrite it as

PROPOSITION 1′ (EQUIVALENT TO PROPOSITION 1). To maximize, invest in project i, given I_t^j, if and only if

$$R_{it}^j = \max(R_{1t}^j, \ldots, R_{kt}^j) \tag{4}$$

and

$$r_{it}^j \geq Z_{it}^j = j\text{th element of } \beta E_t \times \{\overline{\max}(R_{1,t+1} - R_{i,t+1}, \ldots, R_{k,t+1} - R_{i,t+1}, V_{t+1} - R_{i,t+1})\}. \tag{5}$$

An acceptable investment must have both the highest total return and a current return in excess of the nonnegative quantity Z_{it}^j. Z_{it}^j is the (discounted) weighted average of what the investor would be willing to pay, in each possible state in $t + 1$, for the right to "undo" a previous commitment to project i. Z_{it}^j is thus the Weisbrod-Henry option value, defined relative to the most profitable of multiple alternative investments. The investment rule can be read, "Accept the most profitable irreversible investment if and only if its current return exceeds the value of the options thus forfeited."

Proposition 1′ contains the basic results of irreversible investment theory:

a. Irreversible investments should not be made using myopic decision rules (Arrow [1968]; see also Sargent [1980]). Calculation of V or Z requires a deep look into the future.

b. Irreversible investments are less frequently justified than their reversible counterparts [Arrow-Fisher, 1974; Henry, 1974a]. Suppose, for example, that there is only one project. If the project is reversible, investment will take place if $r_t^j \geq 0$. With irreversibility, $r_t^j \geq Z_t^j$ is required, where Z_t^j is strictly positive except when there is no state in $t + 1$ in which current commitment would be regretted.

c. The option value is positive whenever "disinvestment" may be desired in the future (Arrow [1968] for the deterministic case;

Henry [1974b] for the stochastic case). Z^j_{it} equals zero when project i would be freely undertaken again in every state in $t + 1$.

B. The Option Value and the Importance of "Bad News"

Some insight into the option value can be gained by further analysis of Proposition 1′. Let us define $x_{i,t+1}$ to be a random variable that equals, for each state in $t + 1$, the maximum value attainable other than by investing in i, less the value of investing in i.[8] Thus, positive values of $x_{i,t+1}$ attach to states in $t + 1$ in which project i would not be undertaken by choice. From the point of view of period t, a positive realization of $x_{i,t+1}$ is a "bad news" outcome for i, with the implication that an investment in i made in period t is "regretted" (by an amount equal to the realization of $x_{i,t+1}$) in $t + 1$.

Noting that $x_{i,t+1}$ has a probability density induced by the probability of the associated elements of I_{t+1}, we can write the option value criterion for investing in project i in t (equation (5)) as

$$(6) \qquad r^j_{i,t} \geq j\text{th element of } \beta E_t(x_{i,t+1} | x_{i,t+1} \geq 0) = Z^j_{it}.$$

The option value equals the expectation of $x_{i,t+1}$, conditional on $x_{i,t+1} \geq 0$.

This restatement of the basic investment rule has an implication that is a bit surprising. This condition says that, given the current return of the most profitable investment i, the willingness to invest in t depends only on the average expected severity of *bad* news for i that may arrive in the next period. Potential good news for the investment does not matter at all. Thus, let the investor in period t be indifferent just between committing to project i and deferring his decision. Transform his beliefs about information to arrive in $t + 1$ as follows: In one state in $t + 1$ for which $x_{i,t+1} > 0$, make $x_{i,t+1}$ slightly more positive (i.e., make a piece of potential bad news a bit worse). Simultaneously, make the realization of $x_{i,t+1}$ in a state where $x_{i,t+1} < 0$ much more negative (creating the possibility that extraordinarily good news will arrive about project i). This can be done so as to make $R^j_{i,t}$ as large as desired. Yet this transformation unambiguously increases the right side of (6), causing no investment to take place.

This "bad news principle of irreversible investments"—that of possible future outcomes, only the unfavorable ones have a bearing on the current propensity to undertake a given project—is easily explained once we return to the basic option value idea. The investor

8. Let $X_{i,t+1} = \overline{\max}(R_{1,t+1}, \ldots R_{i-1,t+1}, R_{i+1,t+1}, \ldots, R_{k,t+1}, V_{t+1}) - R_{i,t+1}$. Then the random variable $x_{i,t+1}$ takes on the value of the jth element of $X_{i,t+1}$ when state j is drawn in $t + 1$.

who declines to invest in project i today (but retains the right to do so tomorrow) gives up short-run returns.[9] In exchange for this sacrifice, he enters period $t + 1$ with an "option" that entitles him to invest in some project other than i (or to wait longer) if he chooses. This option is valueless in states where investing in i is the best alternative. In deciding whether to "buy" this option (by declining to make a commitment in t), the investor therefore considers only possible "bad news" states in $t + 1$, in which an early attachment to i would be regretted.

Although equivalent to the original option value concept, the bad news principle is useful as an aid to understanding. Among other things, it helps to clarify the relationship of the option value approach to some other parts of the literature. There are a number of papers in related areas, such as financial illiquidity (see, for example, Baldwin and Meyer [1979]) or sequential research-and-development planning (see Roberts and Weitzman [1980]), which do not explicitly use option concepts; yet their derived decision rules end up depending only on one tail of the distribution of outcomes. At least in the two papers mentioned, the authors give no intuitive explanation of the asymmetrical aspect of their results. It turns out that the approaches of both papers are isomorphic to the option value model, an observation that allows an easy explanation of the decision rule forms via the bad news principle. One-tail decision rules are also common in search theory, a fact that brings out the close formal similarities between that field and the option value analysis of irreversible investment.

Recognizing that the expected value of bad news is what determines the option value also provides information about the comparative statistics of irreversible investment. Consider the problem of the effect of increased uncertainty on investment. A natural specification of increased uncertainty is a "spread" of the outcomes $x_{i,t+1}$.[10] Indeed, it is easily shown that any spread (whether mean-preserving or not) of the density of $x_{i,t+1}$ around $x_{i,t+1} = 0$ will increase the option value and depress current investment.[11] Superficially, this confirms the deterrent effect of uncertainty on investment. However, as the

9. Sometimes an investment must be made today or not at all, e.g., in a competitive R and D situation where the first firm to invest gets the patent. To fit this type of project into our formal framework, we must define $r_{i,t} = R_{i,t}$ = total expected payoff to the project, $r_{i,s} = 0$ for $s > t$. The option value still exists, and for a given $r_{i,t}$, the bad news principle still applies. See Baldwin and Meyer [1978].

10. Uncertainty is used (in the sense of Hart [1942]) to mean the reducible component of ignorance about outcomes; as distinguished from risk, the irreducible part. The distribution of $x_{i,t+1}$ measures uncertainty, not risk, since the realization of $x_{i,t+1}$ can be observed simply by waiting.

11. Since $E_t(x_{i,t+1} | x_{i,t+1} \geq 0)$ must increase.

bad news principle tells us, the spread can affect the option value only through the bad news side of the distribution ($x_{i,t+1} \geq 0$); what happens in the region $x_{i,t+1} < 0$ does not matter. We see then that what irreversible investment is sensitive to is "downside" uncertainty. Here is the source of the "caution" of investors and consumer durables buyers described in the introduction; a small increase in the probability of disaster cannot be offset by any potential good news in its effect on current purchases.

Note that the impact of downside uncertainty on investment has nothing to do with risk preferences; we have, in fact, assumed risk neutrality in our analysis. The negative effect of uncertainty is instead closely related to the search theory result that a greater dispersion of outcomes, by increasing the value of information, lengthens the optimal search time.

Perhaps an even more instructive implication of the bad news principle is this: Events that threaten to reverse the profitability rankings of irreversible projects—even though they may increase the absolute returns to all projects—tend to reduce the current propensity to invest. If we use (6), the option value for project i is increased if, for some state in $t + 1$, project i is demoted from first to a lower rank. This increase is not avoided even if (1) the absolute return to i is increased in the demotion state, and (2) i's returns in states where it retains the first rank are increased an arbitrary amount, so that it remains first overall in total return.

An illustration of this is as follows. Suppose that Congress is currently debating an increase in the investment tax credit. With the goal of avoiding postponements of projects during the debate period, the bill specifies that the credit (if passed) will be retroactive. This provision avoids the postponement problem if capital is putty-like and reversible, but not if alternative projects are irreversible and have distinct characteristics. Consider a firm that needs to replace an old factory. A more labor-intensive plant, Project A, is preferred when the tax credit is at the historical level; with the proposed credit increase, a capital-intensive Project B is more profitable. If no credit increase were being considered, A would be built immediately. Discussion of the tax bill, however, creates an option value; there is now the possibility that hasty commitment to A will be regretted. The firm's dominant strategy may well become deferral of investment until after the legislative vote is taken. Thus, a change from the status quo that is favorable to capital—the introduction of the tax credit bill—depresses current investment and destabilizes the investment path by bunching new projects in a future period.

The assumption of irreversibility can have some perverse results, from the point of view of traditional investment theory. This is seen even more clearly in the discussion of dynamics.

II. The Dynamics of Irreversible Investment

In this section of the paper we are concerned with the short-run dynamics of irreversible investment flows. For the certainty case these dynamics are well understood.[12] However, for the case of imperfect foresight (with which we are concerned), this is essentially untraveled ground.[13]

In accord with the goals of this paper, our discussion of irreversible investment paths will not aim to present the most general characterization possible. Instead, we set the more modest task of trying to establish the plausibility of our claim that irreversibility and uncertainty can help to explain observed cyclical fluctuations in durables. We shall do this by the vehicle of an extended example, which incorporates the essential elements of our argument in the simplest possible manner.

Before specializing the discussion to that example, however, it will be useful to ask in more general terms what sorts of restrictions on the economic environment are required to give us the results we want. This problem in economic modeling is addressed in the next part of the paper.

A. The Dynamic Inference Specification

If we wish to explain short-term fluctuations in irreversible investment, we must find a source of similar variations in the option value. The problem is to locate a plausible specification of the stochastic environment that can generate such variations.

It is clearly of no use for this purpose to take the underlying stochastic structure of returns to be time-independent with known parameters, as in Baldwin and Meyer [1979] or Sargent [1980]. With a fixed menu of projects, this will imply that the option value (or "hurdle rate") is constant over time, except possibly for the uninteresting reason (in finite-time models) that the horizon is drawing nearer.

A somewhat better approach is to assume that, while the sto-

12. Arrow [1968]; Arrow and Kurz [1970].

13. Nickell's analysis [1974] does not really depart from the certainty model. A paper by Sargent [1980] derives a (long-run) steady state for irreversible capital but differs greatly from this work both in object and in many important particulars.

chastic structure of returns is unchanging, its parameters are unknown and must be inferred by investors. In earlier versions of this paper,[14] an example using the Dirichlet distribution was given to show that more diffuse investor priors led to higher option values. Thus, in early stages of the learning process, waiting for information is a valuable activity, and investment tends to be deferred. As knowledge accumulates and priors concentrate, the propensity to invest increases. Here, at least, the option value can vary—but only monotonically downward over time as uncertainty diminishes.

This suggests that, for an interesting analysis, we must permit uncertainty to be periodically renewed. One way to model this is to assume that an unknown underlying distribution is replaced at random intervals by a new one drawn from a meta-distribution. It is probably most realistic to assume that agents cannot observe when distributions change. The problem of predicting future returns in this environment was dubbed the dynamic inference problem by Howard [1964]. Howard showed how correct inferences could be made by repeated applications of Bayes' Law.[15]

The dynamic inference model is a good representation of many important situations. For a macroeconomic application, make the realistic assumption that the returns in certain industries (e.g., homebuilding) depend on the Federal Reserve's monetary rule. The rule cannot be observed directly but must be inferred from the money growth rates it randomly generates. Although the probability of it occurring in a given period is small, it is known that the Fed may change its rule in a manner imperfectly correlated with its announcements. How are the homebuilders to estimate the rule and predict the money supply?

The novelty of this problem is the necessity of inferring when a distribution (monetary rule) change has taken place, given only periodic drawings (money growth rates) as clues. This calculation depends on two factors: the homebuilders' beliefs about the frequency and serial dependence of changes in the underlying distribution, and the degree to which observations conform to prior expectations. "Unusual" recent observations tend to increase the subjective probability that the distribution has shifted. For example, a sequence of unusually low money growth rates is not necessarily inconsistent with Fed maintenance of the status quo; but observing such a sequence

14. Or see Bernanke [1979].
15. Our description of this procedure is brief and heuristic. For a formal statement of the dynamic inference methodology, see Howard [1964] or the discussion in Bernanke [1979]. A simple example of dynamic inference is worked in the next section.

increases the probability that a new rule has been adopted, especially if expectations are that any shift in the rule will be toward greater restrictiveness.

Our claim is that new information, when filtered through a dynamic inference model, can plausibly be the source of short-run cycle-like swings in investment, as follows: Unusual drawings, which occur periodically, make investors less sure about what distribution is generating the returns to projects. If distribution changes are thought to be relatively infrequent, then it will pay to defer commitments in order to gather information; investment declines. Over time the nature of the new distribution (or the fact that no change has occurred) is revealed. Investment revives as waiting for information becomes less desirable. Thus, a complete short-run cycle is created. This cycle has the characteristic that its low points correspond to periods of high investor uncertainty and caution.

The next portion of the paper expands and attempts to justify this claim through the presentation of an extended example.

B. An Example: Investment and the Formation of an Energy Cartel

We introduce a simple model of investment and output in an energy-importing economy after the unanticipated formation of an energy-exporters' cartel. It will be shown that uncertainty can create an investment cycle, even though capital goods appear to dominate the alternative asset in each period.

The identical, risk-neutral agents in our model energy-importing economy have three perfectly durable assets with which to form portfolios. These are (1) energy-using capital (K^e), (2) energy-saving capital (K^s), and (3) investible resources (W). The two forms of capital can be used to produce a homogeneous consumption good according to a relation to be specified below. Investible resources are the real storable goods that are used up in the investment process: land, materials, energy, and (to the extent that leisure and consumption can be intertemporally substituted) labor. These resources are assumed to have no direct productivity and to pay no return; they are valuable, however, because they can be converted costlessly into units of K^e or K^s. Conversion of resources into a specific form of capital is irreversible. For simplicity, we shall assume that the investment resources of the society are augmented at a constant rate ΔW per period. All that is required for our results, however, is that there be increasing real marginal costs to cumulative investment; hence, a less-thaninfinite elasticity of supply for W would be a sufficient assumption.

At time t_0 an energy-exporters' cartel is assumed to have formed. The state of nature in each period depends on the status of this cartel. Let s_t, the state of nature in t, be defined by

$$s_t = \begin{cases} 1, & \text{if the cartel exists in period } t \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

The duration of the cartel is uncertain. We first give a description of investor beliefs about the cartel's continued existence, then offer a justifying derivation. Let common subjective probabilities be described by

$$Pr(s_{t+1} = 1 | s_t = 1) = p_t \tag{8}$$

$$Pr(s_{t+1} = 1 | s_t = 0) = 0 \tag{9}$$

$$dp_t | dt > 0, \lim_{t \to \infty} p_t = 1. \tag{10}$$

Thus, if the cartel fails, everyone assumes that it will be gone forever. The longer the cartel lasts, the greater is the common subjective probability of its survival through the next period. If the cartel survives long enough, it is assumed to be permanent.

The assumption that any collapse of the cartel is believed to be once-and-for-all is made only for ease of calculation. What is essential to this example is that the subjective probability of cartel permanence increase monotonically as the cartel continues to persist. This necessary assumption is in fact an implication of virtually any reasonable dynamic inference model of the basic sources of the cartel. To be concrete, let us suppose that, upon observing the formation of the cartel, investors have two hypotheses about its duration. The null hypothesis (H_0) is that the underlying economic-political situation remains unchanged from the precartel era, so that there is a probability q of cartel collapse in each period. The alternative hypothesis (H_1) is that the economic-political situation has been irrevocably transformed into one which guarantees the existence of a cartel in every future period. Cartel formation by either cause is thought to be very unlikely a priori, so that we can neglect the possibilities of (1) a political-economic shift in any period other than t_0 and (2) a second cartel arising under the status quo following a collapse of the first.

Let the probability as of date t that H_1 is true be denoted by P_t. Bayes' Law implies that

$$(11)\quad P_t = P_{t-1}/(P_{t-1} + (1-q)(1-P_{t-1})) \qquad \text{if } s_\tau = 1, \text{ all } \tau \text{ such that } t_0 \le \tau \le t,$$

and $P_t = 0$ otherwise. P_{t_0} is given by prior beliefs. Note that P_t increases monotonically toward 1 as the cartel remains in existence.

We can now calculate p_t, the probability that the cartel will exist in $t + 1$ given that it has persisted up to t:

$$(12)\qquad p_t = P_t + (1-q)(1-P_t).$$

It is easily shown that p_t has the properties assumed in (10). Also, if the cartel fails, it is proof that the status quo prevails, so that the possibility of future cartel formation can be neglected, as assumed by (9).

The status of the cartel is important to investors in the energy-importing economy because it determines the net productivity of capital. Assume that each unit of energy-intensive capital (K^e) has a net per-period output of $r^{e,0}$ consumption goods when there is no cartel (i.e., $s_t = 0$) and a smaller output of $r^{e,1}$ when the cartel exists ($s_t = 1$). Energy-saving capital (K^s) produces an intermediate quantity of r^s per period per unit, independently of the state of nature. Net production is always nonnegative. In summary, the net production of the consumption good Y_t in each period is given by

$$(13)\qquad Y_t = \begin{cases} r^{e,1} K_t^e + r^s K_t^s, & \text{if } s_t = 1 \\ r^{e,0} K_t^e + r^s K_t^s, & \text{if } s_t = 0, \end{cases}$$

where $0 \le r^{e,1} < r^s < r^{e,0}$.

The agents' goal is to maximize the discounted consumption streams,

$$E_t \sum_{\tau=t}^{T} \beta^{\tau-t} Y_\tau,$$

over the set of possible portfolio strategies. This is a trivial problem if the cartel disappears; all resources would go to expanding the energy-using capital stock. We shall therefore consider the dynamics of aggregate investment (that is, the economy-wide rate of conversion of investible resources to capital) only in the case where the cartel stubbornly refuses to disappear.

First, we define some total returns. Let

$$(14)\qquad \begin{aligned} R^{e,0} &= \sum_{\tau=t}^{T} \beta^{\tau-t} r^{e,0} \\ R^s &= \sum_{\tau=t}^{T} \beta^{\tau-t} r^s \\ R^{e,1} &= \sum_{\tau=t}^{T} \beta^{\tau-t} r^{e,1} \end{aligned}$$

$$\bar{R}^e(P) = PR^{e,1} + (1-P)$$
$$\times \left[r^{e,1} \sum_{\tau=t}^{T} \beta^{\tau-t}(1-q)^{\tau-t} + r^{e,0} \sum_{\tau=t}^{T} \beta^{\tau-t}\,(1-(1-q)^{\tau-t}) \right].$$

$R^{e,0}$ and $R^{e,1}$ are the total returns to a unit of K^e when (1) the cartel has vanished and (2) the cartel is known to be permanent, respectively. R^s is the total return to K^s. $\bar{R}^e(P)$ is the expected total return to a unit of K^e, given a current probability P that Hypothesis 1 (the permanent cartel) is true.

Now we can make three observations about the path of aggregate investment in this economy as the cartel persists. The first two of these observations are not too surprising:

1. The commitment of all investible resources to energy-using capital may continue for some time after the formation of the cartel. Sufficient conditions for this to happen on the interval $[t_0,t_1]$ are

$$\bar{R}^e(P_t) > R^s \quad \text{and} \quad r^{e,1} > \beta p_t(R^{e,0} - \bar{R}^e(P_{t+1})), \qquad t_0 \le t \le t_1. \tag{15}$$

(See Appendix.) Equation (15) can always be made to hold by taking P_{t_0} and $(1-q)$ close enough to zero. That is, if investors do not take seriously at first the possibility of either an underlying shift or of cartel persistence under the status quo, they will not be willing to give up current returns in order to get more information. Of course, as time passes and the cartel continues to exist, beliefs will change, and (15) can no longer hold.

2. If the cartel continues to persist, there is some date t_2 after which all resources will be devoted to energy-saving capital. This follows directly from the assumptions that the cartel is eventually perceived as permanent and $R^s > R^{e,1}$.

The point of special interest about the investment path is that t_1 (the last period of investment is K^e) and t_2 (the first period for K^s) need not be adjacent.

3. There may exist an investment "pause," during which no investment is made, and "barren" investible resources simply accumulate. This pause may occur because it is possible for investors to prefer uncommitted investible resources to capital in their portfolios, despite the fact that capital is guaranteed a positive return in each period and uncommited resources pay no direct returns.

We shall show numerically that a pause can exist. This can be done simply if we first state

PROPOSITION 2. Let $s_t = 1$, all t (i.e., the cartel persists). Suppose that $\bar{R}^e(P_t) \ge R^s$ for $t \in [t_0,t_1]$, and that $R^s \ge \bar{R}^e(P_t)$ for $t \in [t_2]T$, where $t_2 = t_1 + 1$. If there is investment in every period (i.e., there is no pause), then the option value in period t_1 is given by

$$\beta p(t_1)(R^s - \overline{R}^e(P(t_2)))$$

and the option value in t_2 is given by

$$\beta(1 - p(t_2))\,(R^{e,0} - R^s).$$

(See Appendix.)

Now assign some convenient numerical values. Let $r^{e,0} = 10$, $r^s = 5$, $r^{e,1} = 1$, $\beta = 0.9$. Take T, the total number of periods, to be large. Then $R^{e,0} = 100$, $R^s = 50$, $R^{e,1} = 10$. Let P_{t_0} (the probability as of t_0 that there is a new regime) equal to 0.1, and let q (the probability of cartel failure in a period, given that the status quo regime exists) equal 0.2. Then by (11) and (14), we can show in tabular form the values for P_t and $\overline{R}^e(P_t)$:

Period	P_t	$\overline{R}^e(P_t) = 67.86 - 57.86\, P_t$
0	0.100	67.02
1	0.122	60.80
2	0.148	59.30
3	0.178	57.54
4	0.213	55.51
5	0.253	53.21
6	0.298	50.64
7	0.346	47.82

Notice how, as the cartel fails to disappear, the subjective probability of cartel permanence P_t grows; correspondingly, the expected total return to energy-intensive investment diminishes.

We now prove that, in this example, there must be at least one period when making no investment is the dominant strategy (that is, the most preferred portfolio consists entirely of uncommitted resources). Assume the contrary, that investment takes place in every period. Then investment in K^e must be occurring in periods 0 to 6, since in those periods $\overline{R}^e(P_t) > R^s$. Likewise, investments in K^s are only made in period 7 or later. Thus, $t_1 = 6$, $t_2 = 7$. Using Proposition 2, we can calculate the following option values $Z(t)$:

$$Z(6) = 0.9(0.860)(50 - 47.82) = 1.69$$

$$Z(7) = 0.9(1 - 0.869)(100 - 50) = 5.89,$$

where, by (12), $p(6) = 0.860$ and $p(7) = 0.869$. But by Proposition 1′,

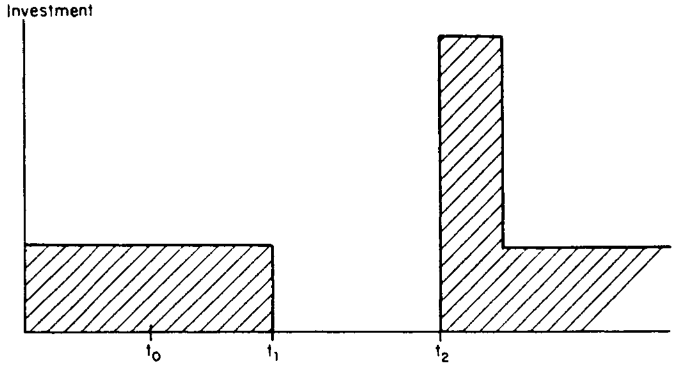

FIGURE I

if investment is to take place in periods 6 and 7, it must be true both that $r^{e,1} \geq Z(6)$ and that $r^s \geq Z(7)$. The failure of either of these conditions is sufficient to create a contradiction; in fact, since $r^{e,1} = 1$ and $r^s = 5$, both fail. We conclude that an investment path that is nonzero in every period cannot be optimal in this case; there must be a "pause."

How can it be that investible resources, which pay no return, can dominate both sorts of capital in portfolios? The answer is that these resources, by representing a generalized and uncommitted power to invest, do pay an implicit return; this return is equal to the option value. When the option value is large, investors are willing to sacrifice current returns in order to receive more information. The bad news principle is hard at work here: High uncertainty and the possibility that new information will change project rankings depress current investment. Note that, in the above example, the returns ($r^{e,0}$, r^s, and $r^{e,1}$) could be multiplied by any factor—putting the "average value of capital" beyond any arbitrary number—and still no investment will take place in certain periods.

Drawing together observations 1 to 3, we see that a possible investment path for this economy as the cartel persists is given in Figure I. The cartel forms in t_0. From t_0 until t_1, insufficient credence is given to the permanence of the new situation for investment in energy-using capital to stop. By t_1, however, enough uncertainty has been created for an investment pause to take place. Agents wait for new information until t_2, when the permanence of the cartel seems sure enough for resources to be devoted to energy-saving capital. Notice the investment spurt at t_2 as cumulated resources are converted into K^s.

The results of this example support the claim of the previous section. A single event, whose uncertain implications are resolved through the dynamic inference methodology, may create a short-run cycle in investment. This cycle may not begin until some time after the event, and it exhibits both a "trough" (as uncertain investors wait for new information) and a "boom" (as stored-up investment capacity is used). These basic features would remain in a model with risk aversion, heterogeneous investors, adjustment costs, and an endogeneous supply of investible resources. Indeed, these more realistic assumptions would soften the corner-solution aspects of this example, leading to a more sinusoidal investment path.

This analysis has been done without reference to prices and thus might be best thought of in a central planning context. But it must be emphasized that because (1) the economic outcomes described above are completely efficient and (2) agents are identical, there is no difficulty in constructing a sequence of relative prices for output, investible resources, and capital that would support the same investment path in a competitive, decentralized-decision-making environment. The investment pause still occurs, motivated by speculation in investible resource stocks and the realization by investors that choosing the wrong kind of capital will be penalized by low output and low second-hand values in the future. In particular, allowing agents to sell old capital does not change any result obtained by assuming agent-by-agent capital irreversibility.[16]

III. The Aggregation Problem

We have tried to show that the timing of irreversible investment projects can be quite sensitive to the arrival of new information. The essential assumptions are as follows:

1. Different investments are to some extent alternatives, so that expansion in some lines limits the capacity (of the individual, firm, or society) to pursue others.
2. Alternative investments are physically irreversible (although

16. In our example, in which agents are identical, the existence of a used capital market is irrelevant; prices in that market would always be such as to make agents indifferent between selling and holding their capital. In the more general case, the existence of second-hand markets makes the problem *look* different to the individual agent but does not materially affect general equilibrium results obtained by assuming irreversibility. The economy as a whole must still hold all irreversible investment, a fact that is reflected in second-hand prices. Second-hand markets are important only if beliefs or preferences are so heterogeneous that there is no agreement on what constitutes good or bad news for a particular investment.

ownership may be transferred) and have different patterns of returns across states of nature.

3. Information arriving over time is pertinent to the choice of investments.

Still, to claim that effects described above may operate at the level of a national economy or contribute to the aggregate business cycle requires an additional step. For example, we must take seriously the potential counterargument that, because of the law of large numbers, firm or industry fluctuations might wash out the aggregate level.

There are two separate (and to some extent alternative) arguments for aggregate investment instability arising from the interaction of irreversibility and uncertainty.

The first argument simply recognizes that macro-level factors are important in the microeconomic investment decisions. We can point to such phenomena as wars and other foreign policy shocks (e.g., OPEC); changes in monetary, fiscal, regulatory, or other policy regimes; variations in the conditions of international trade and competition; changes in the level of world supplies of basic industrial commodities; or the advent of a technology, such as computerization, with widespread implications. In addition, the U. S. economy is not so well diversified as to be immune in the aggregate from shocks to certain large industries, such as automobiles, housing, defense, and agriculture. If macro-uncertainty is important, then aggregate investment instability follows from our micro-analysis.[17]

Pursuing this line gives us a view of the business cycle reminiscent of the work of Schumpeter or Keynes (see Chapter 22 of *The General Theory*, the source of the introductory quote). Aggregate investment fluctuations are seen as driving the cycle. This fits equally well into the traditional IS-LM model (so that the cycle results from shocks to the IS curve) and the business cycle models of modern equilibrium theorists (in which variations in investment demand induce perfectly efficient intertemporal substitutions between labor and leisure). Indeed, the possibility of an aggregate investment pause solves an important problem of the equilibrium theorists: how to explain nontransitory deviations of aggregate production from trend in a manner consistent with rationality and efficiency.

This argument takes uncertainty to be completely exogenous to the economic system; the investment fluctuations created by this uncertainty provide (in the terminology of Ragnar Frisch) an "im-

17. The competing accelerator theory of cycles also requires that macro-level factors have great weight in investment decisions.

pulse" to cyclical swings. An alternative way to justify aggregation begins by pointing out that uncertainty can be generated within the system. Consider a model of imperfect current information in the tradition of R. E. Lucas, Jr. [1972]. We suppose that the (stochastic) demand for the product of each firm or the services of each worker is capable of both transitory and permanent variations (distribution shifts). The transitory component of individual demands is primarily associated, let us say, with random fluctuations in aggregate demand; the permanent component depends on tastes and technology. Let contemporaneous aggregate demand, or its relationship to individual demands, be imperfectly perceived. Then firms or workers experiencing a low current demand, being unsure of its source, will be uncertain about its permanence. Per the theory set forth above, they will delay irreversible investments (or, for the worker, purchases of illiquid durables) in order to learn the true state of affairs. But the reduced investments of these agents lower demands perceived by other agents in the economy. Thus, an intrinsically transitory variation in aggregate demand can be converted into a more persistent downward fluctuation. Irreversible investment is, here, the "propagating" mechanism of the cycle.

CONCLUSION

While this paper has set up a fairly general model of the irreversible investment decision and provided interpretations of the optimal decision rules, it was here neither our goal nor our achievement to present new results in the pure theory of investment. Rather, we have built on existing work on irreversible decisions to outline a new explanation of an important macroeconomic phenomenon. Taken in conjunction with traditional accelerator models, this explanation enhances our understanding of why recessions are felt so disproportionately in durable goods sectors. In future research we shall examine richer general equilibrium models of the relationship of uncertainty and irreversibility to the business cycle.

APPENDIX: MATHEMATICAL NOTES ON THE TEXT

Proposition 1 is straightforward. At each stage the investor has $k + 1$ mutually exclusive options: choose one of k investments, or wait. The expected value of committing to investment i, given I_t^j, is R_{it}^j, $i = 1,2, \ldots, k$. That (1) is value-maximizing follows immediately if we can show that V_t^j is the expected value of deferring commitment in t, given information I_t^j.

If we apply induction, no investment can be made after period T, so, clearly, $V_T^j = 0$, all j. Now assume that V_{s+1}^m is the expected value of making no investment in $s + 1$, given information I_{s+1}^m. The value of arriving in $s + 1$ without prior commitment, conditional on drawing I_{s+1}^m, is

$$\text{(i)} \qquad \max\,(R_{1,s+1}^m, \ldots, R_{k,s+1}^m, V_{s+1}^m).$$

From the point of view of an investor observing I_s^j in period s, the expected value of waiting is just the probability of each state in $s + 1$ times the (discounted) value of arriving in that state without previous commitment:

$$\text{(ii)} \qquad \beta \sum_{m=1}^{n_{s+1}} P(I_{s+1}^m | I_s^j) \max\,(R_{1,s+1}^m, \ldots, R_{k,s+1}^m, V_{s+1}^m)\,.$$

But the quantity in (ii) is just V_s^j, as expressed (in vector notation) by (2).

Proposition 1′ is equivalent to Proposition 1. Equation (1) is broken up into (4) and the condition,

$$\text{(iii)} \qquad R_{it}^j \geq V_t^j.$$

Subtracting $\{\beta E_t R_{i,t+1}\}^j$ from both sides of (iii) yields (5).

The cartel problem. First, we show that (15) is sufficient for energy-intensive investment to continue after the cartel has formed. By Proposition 1′, all we need to show is that $\beta p_t(R^{e,0} - \overline{R}^e(P_{t+1}))$ is greater than or equal to the option value. The option value is

$$\text{(iv)} \quad \beta p_t \max\,(R^s - \overline{R}^e(P_{t+1}), 0, V_{t+1}^{s=1} - \overline{R}^e(P_{t+1})) \\ + \beta(1 - p_t) \max(R^s - R^{e,0}, 0, V_{t+1}^{s=0} - R^{e,0}).$$

Given that $R^{e,0} > R^s$, sufficient conditions for $\beta p_t(R^{e,0} - \overline{R}^e(P_{t+1}))$ to exceed (iv) are $R^{e,0} > V_{t+1}^{s=1}$ and $R^{e,0} > V_{t+1}^{s=0}$. But these always hold, since the value of waiting can never exceed that of a guarantee of the maximum possible return. (This also follows from induction.)

Proposition 2 is proved using Proposition 1′. In t_1 we have $\overline{R}^e(P(t_1)) \geq R^s$, and the option value is as in (iv) above, with $t = t_1$. Since $R^{e,0} \geq V$ always, all we need is $R^s \geq V(t_2)$ when $s(t_2) = 1$. But this follows from the assumption that investing in K^s is preferable to waiting in t_2, when $s(t_2) = 1$.

In t_2 we have $R^s \geq \overline{R}^e(P(t_2))$, and the option value is given by

$$\text{(v)} \quad \beta p_t \max(\overline{R}^e(P(t + 1)) - R^s, 0, V_{t+1}^{s=1} - R^s) \\ + \beta(1 - p_t) \max(R^{e,0} - R^s, 0, V_{t+1}^{s=0} - R^s),$$

where $t = t_2$. Now $R^s > \overline{R}^e(P(t_2 + 1))$, since $\overline{R}^e(P(t_2)) > \overline{R}^e(P(t_2 + 1))$, and $R^s \geq V(t_2 + 1)$ by the assumption that investment in K^s takes place in $t_2 + 1$. Also, $R^{e,0} > V$, as before. Thus, (v) is equal to $\beta(1 - p_t)(R^{e,0} - R^s)$, as stated.

STANFORD UNIVERSITY

REFERENCES

Arrow, K. J., "Optimal Capital Policy and Irreversible Investment," in J. N. Wolfe, ed., *Value, Capital, and Growth* (Chicago: Aldine, 1968), pp. 1–20.

——, and A. C. Fisher, "Environmental Preservation, Uncertainty, and Irreversibility," this *Journal*, LXXXVIII (May 1974), 312–20.

——, and M. Kurz, "Optimal Growth with Irreversible Investment in a Ramsey Model," *Econometrica*, XXXVIII (March 1970), 331–44.

Baldwin, C. Y., and R. F Meyer, "Liquidity Preference Under Uncertainty: A Model of Dynamic Investment in Illiquid Opportunities," *Journal of Financial Economics*, VII (Dec. 1979), 347–74.

Bernanke, B. S., "Long-term Commitments, Dynamic Optimization and the Business Cycle," Ph.D. thesis, M.I.T., 1979.

Chenery, H. B., "Overcapacity and the Acceleration Principle," *Econometrica*, XX (Jan. 1952), 1–28.

Cukierman, A., "The Effects of Uncertainty on Investment under Risk Neutrality with Endogenous Information," *Journal of Political Economy*, LXXXVIII (June 1980), 462–75.

Hall, R. E., and D. W. Jorgenson, "Tax Policy and Investment Behavior," *American Economic Review*, LVII (June 1967), 391–414.

Hart, A. G., "Risk, Uncertainty, and the Unprofitability of Compounding Probabilities," in O. Lange *et al.*, eds., *Studies in Mathematical Economics and Econometrics* (Chicago: University of Chicago Press, 1942), pp. 110–18.

Henry, C., "Investment Decisions Under Uncertainty: the 'Irreversibility Effect,'" *American Economic Review*, LXIV (Dec. 1974a), 1006–12.

——, "Option Values in the Economics of Irreplaceable Assets," *Review of Economic Studies Symposium* (1974b), 89–104.

Howard, R. A., "Dynamic Inference," *Research in the Control of Complex Systems*, T. R. # 10, M.I.T., 1964.

Koyck, L. M., *Distributed Lags and Investment Analysis* (Amsterdam: North-Holland, 1954).

Lucas, R. E, "Expectations and the Neutrality of Money," *Journal of Economic Theory*, IV (1972), 103–24.

Marschak, J., "Role of Liquidity Under Complete and Incomplete Information," *American Economic Review*, XXXIX (May 1949), 183–95.

Nickell, S., "On the Role of Expectations in the Pure Theory of Investment," *Review of Economic Studies*, LXI (1974), 1–19.

Roberts, K. W. S., and M. L. Weitzman, "Funding Criteria for Research, Development, and Exploration Projects," unpublished working paper, M.I.T., July 1980.

Samuelson, P. A., "Synthesis of the Principle of Acceleration and the Multiplier," *Review of Economic Statistics*, XXI (May 1939), 75–78.

Sargent, T. J., "'Tobin's *q*' and the Rate of Investment in General Equilibrium," *Carnegie-Rochester Conference Series on Public Policy 12* (Amsterdam: North-Holland, 1980), pp. 107–54.

Weisbrod, B. A., "Collective-Consumption Services of Individual-Consumption Goods," this *Journal*, LXXVIII (Aug. 1964), 471–77.