

NBER WORKING PAPER SERIES

INTEREST RATES AND BACKWARD-BENDING INVESTMENT

Raj Chetty

Working Paper 10354

<http://www.nber.org/papers/w10354>

NATIONAL BUREAU OF ECONOMIC RESEARCH

1050 Massachusetts Avenue

Cambridge, MA 02138

March 2004

I thank Alan Auerbach, Gary Chamberlain, Ben Friedman, Boyan Jovanovic, Anil Kashyap, Robert Pindyck, Michael Schwarz, Arnold Zellner, seminar participants, two anonymous referees, and especially George Akerlof, Martin Feldstein, and Jerry Green for helpful discussions and suggestions. Funding from the National Bureau of Economic Research and National Science Foundation is gratefully acknowledged. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

©2004 by Raj Chetty. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Interest Rates and Backward-Bending Investment

Raj Chetty

NBER Working Paper No. 10354

March 2004

JEL No. D83, D92, E22, E52, H3

ABSTRACT

This paper studies the effect of interest rates on investment in an environment where firms make irreversible investments and learn over time. In this setting, changes in the interest rate affect both the cost of capital and the cost of delaying investment. These two forces combine to generate an aggregate investment demand curve that is always a backward-bending function of the interest rate. At low rates, increasing the interest rate stimulates investment by raising the cost of delay. Existing evidence supports the hypothesis that firms change the time at which they invest in response to changes in interest rates. The model also generates a rich set of additional predictions that can be tested empirically.

Raj Chetty

Department of Economics

UC- Berkeley

521 Evans Hall #3880

Berkeley, CA 94720

and NBER

chetty@econ.berkeley.edu

Macroeconomic policies to stimulate investment are frequently motivated by the downward sloping relationship between investment demand and interest rates derived from neoclassical models of investment. The intuition underlying this relationship is straightforward: Lowering the cost of capital via monetary or tax policies stimulates investment by enlarging the set of projects that are sufficiently profitable to warrant investment.¹

This paper shows that this canonical result breaks down when firms making irreversible investment decisions can learn over time, as in standard real options or "time to build" models.² To see the intuition, consider a pharmaceutical company deciding how quickly to proceed with investments in operations to produce new drugs. The firm is uncertain about which drugs will be successful, and can acquire further information by delaying investment via R&D. The cost of delaying investment is that the firm cannot retire its outstanding debt as quickly, raising its interest expenses.³ Now consider how an increase in the interest rate will affect the firm's behavior. A higher interest rate reduces the set of drugs that surpass the hurdle rate for investment, creating the standard cost of capital effect that acts to reduce the scale of investment. But a higher interest rate also makes the firm more eager to retire its debt quickly by investing immediately so that it has a chance to earn profits sooner. This second "timing effect" acts to *raise* current investment. These two forces combine to generate a non-monotonic investment demand curve.

To formalize this intuition, I first analyze a simple dynamic model where a continuum of profit-maximizing firms make binary investment decisions and can observe a noisy signal about the parameters that control payoffs by postponing investment. In this model, expected profits grow at a rate $g > 0$ when firms delay investment because they acquire more information and increase the probability of investing only in successful ventures. Profits earned in subsequent

¹Haavelmo (1960) pioneered the neoclassical theory of investment and Jorgenson (1963) derived equations to estimate the effect of the user cost of capital on investment.

²Models of the timing of investment were first analyzed by Marglin (1967, 1970) in the context of government investment. Pindyck (1991) and Dixit and Pindyck (1994) provide extensive reviews of the real options literature.

³This example ignores the additional cost of delay due to loss of rents in a competitive industry. This important issue is addressed below.

periods are discounted at the interest rate, r . Therefore, firms invest immediately only if the expected profit from investment is positive *and* the expected growth in profits from delaying (g) is less than the interest rate. The backward-bending shape of the aggregate investment demand curve, $I(r)$, arises directly from this firm-level optimality condition. If r is low, g is likely to exceed r , compelling many firms to delay investment rather than investing in period 1. On the other hand, when r is high, the expected return to immediate investment is negative for many firms, making investment in period 1 suboptimal for them. Consequently, aggregate investment is maximized at an intermediate $r^* > 0$, and $I(r)$ is upward-sloping from 0 to r^* and downward-sloping above r^* .⁴

The backward-bending property of the investment demand curve is robust to several generalizations of the basic model. First, permitting choices about the scale of investment does not affect the result. The aggregate economy in the basic extensive-marginal model is isomorphic to a single firm making scale choices, so the main intuition still goes through. Second, I study the effects of competition in a model where prices and profit rates are determined endogenously in equilibrium to equate supply and demand. If firms can earn sufficiently high quasi-rents (producer surplus) from investment in the short-run, the equilibrium level of investment remains a backward-bending function of r . Intuitively, as long as the marginal firm values the option to delay in equilibrium -- which will be true if entry by identical competitors does not occur immediately -- the interest rate continues to affect both the cost of delay and the cost of capital, thereby generating two opposing forces on investment demand in equilibrium. Third, in a model where firms have additional margins of choice beyond scale, other behavioral responses such as changes in the composition of investment reinforce the backward-bending shape that arises from learning effects. For instance, if construction is cheaper when firms take a longer time to build (as in Alchian, 1959), they have an incentive to

⁴In recent work independent of this study, Jovanovic and Rousseau (2001, 2004) and Capozza and Li (2001) point out that interest rate changes can have non-monotonic effects on IPOs and real estate development decisions. Their models differ from the present analysis in several respects, which are discussed along with their empirical results in section 4.

switch to slower building technologies when interest rates are low, reducing aggregate investment for reasons independent of learning.

The unconventional relationship between I and r derived here is of interest for two reasons. First, the model in this paper is representative of the extensive literature on irreversible investment under uncertainty, a concept that Caballero (1999) emphasizes is "at the center of modern theories." From a normative perspective, it is useful to understand how interest rates should affect investment in what is increasingly viewed as the leading theory of investment behavior. Second, the non-monotonic relationship is interesting from an empirical perspective because several econometric studies have searched for a negative relationship between exogenous changes in the cost of capital and aggregate investment demand without success.⁵ This paper proposes a model that could explain the lack of a clear, monotonic relationship between I and r at least in certain high-risk sectors of the economy.

A natural question in this regard is whether the timing effects that generate the non-monotonic investment curve are empirically important. Existing microeconomic studies, reviewed in section 4, find that firms change the time at which they invest in response to changes in interest rates in several industries, ranging from mining to real estate development. In addition, studies show that many firms in high-risk sectors explicitly or implicitly use real-options approaches to make investment decisions, and that the timing of investment appears to be a real choice variable. Hence, it is clear that at least some firms follow decision rules that generate a non-monotonic relationship between investment and the interest rate.

While this evidence suggests that timing effects are empirically relevant, it of course does not directly indicate the importance of these issues for aggregate investment demand. Sharper

⁵Several studies have found that interest rates have little or no effect on the level of aggregate investment; see e.g., the early studies of Eisner and Nadiri (1968) and Feldstein and Flemming (1971), or Chirinko (1993a,b) for a review. A modern literature that exploits cross-sectional variation in the user cost finds a larger role for the cost of capital in the long run in some sectors (see e.g., Cummins, Hassett, and Hubbard 1994, who exploit tax reforms as natural experiments). Caballero (1999) and Hassett and Hubbard (2002) review this literature. In interpreting these results, note that in the model proposed here, tax changes can affect investment differently than changes in r and interest elasticities are more negative in the long run than the short run. These points are discussed in greater detail in section 5.

tests of the relationship between interest rates are required. The learning structure of the model yields many predictions that could be tested in future work. The most important is perhaps that an increase in the interest rate is more likely to increase investment in sectors or times when the potential to learn is greater, i.e. when signals about future payoffs are more informative and the variance of payoffs is large. Examples that satisfy these conditions include startups or small businesses, especially in high-tech fields. Intuitively, the effect of changes in r on the cost of learning become amplified when the potential to learn is large. The learning effect thus dominates the cost of capital effect for a larger range of r , raising the investment-maximizing r^* .

Several other empirical implications are derived as well. First, a permanent increase in r is more likely to raise investment in the short-run than the long run, because the benefits to additional learning diminish over time. Second, a change in r affects not only the size but also the quality of investment non-monotonically. Raising r when $r < r^*$ expands the pool of current investors and drives down average observed profits by bringing in less successful ventures; raising r when $r > r^*$ raises the average observed profit rate. Finally, temporary changes in interest rates create incentives to substitute investment intertemporally, in different directions depending on whether the changes are anticipated or not. The model predicts when the yield curve becomes steeper, current investment should rise relative to subsequent investment.

In interpreting these results, it is important to keep in mind that the interest rate is taken as exogenous throughout this paper. The determinants of the supply of capital are therefore left unspecified.⁶ This partial-equilibrium approach is appropriate in analyzing policy questions such as the effects of exogenous changes in the user cost of capital via tax or monetary policies. This question has been the primary focus of the empirical literature, which has at least attempted to generate exogenous variation in r using instrumental-variable and other econometric techniques (see Chirinko 1993b). The relationship between investment and interest rates derived

⁶The supply of capital could itself be a non-monotone function of r for several reasons, including countervailing price and wealth effects and asymmetric information (Stiglitz and Weiss 1981). In addition, if consumers purchasing durables can learn more about their properties by delaying purchase, the same non-monotonicities that arise from learning effects on the investment demand side could also affect the schedule of the supply of funds.

here should not be expected to hold when r is endogenously determined in general equilibrium, as changes in r could occur because of autonomous shocks to investment demand (a simultaneity problem).

The remainder of the paper is organized as follows. The next section develops a stylized model of investment by learning firms. It solves for optimal investment behavior, and aggregates the model to derive an investment demand curve. The main backward-bending investment result is derived for this model in section 2. Section 3 generalizes the result to more realistic environments. Section 4 describes existing empirical evidence that supports the main timing intuition, and section 5 derives additional empirical implications of the model. The final section offers concluding remarks. All proofs are given in the appendix.

1 A Stylized Model of Investment by Learning Firms

I analyze the effect of interest rates on investment in a standard discrete-time learning model where firms are Bayesian updaters. Firms are assumed to be residual claimants in all states of the world and make one irreversible investment decision with the objective of maximizing profits. Two simplifying assumptions are made in the basic case: Firms only decide whether to invest or not (the scale of investment at the firm level is not flexible), and competitive forces are ignored by taking profit rates as exogenous. The basic model thus best describes a firm that has already obtained a patent on an idea (e.g. a chemical compound) and is deciding when to market its innovation (e.g. a new drug) by building a factory. I first analyze the investment decision of a single firm of this type, and then aggregate over firms with heterogeneous expectations to characterize total investment in the economy.

1.1 Structure and Assumptions

Suppose a firm is deciding whether to invest in a new plant that can be built at cost C . The economy is stationary in the sense that nominal revenues from the project and the cost of investment are constant over time. In a world of perfect information, the manager's decision rule is simple: invest if the rate of return from the project is the higher than that of his best alternative. But the manager is uncertain about how strong demand for the firm's product will be. To model his uncertainty, assume that there are two possible distributions that govern the characteristics of demand z for the product in each period. Labeling the two distributions $f(z)$ and $g(z)$, index the true distribution by $\mu \in \{0, 1\}$, where

$$\mu = 0 \Rightarrow z \sim f(z) \text{ and } \mu = 1 \Rightarrow z \sim g(z)$$

The value of μ determines the stream of revenues that the firm gets from investment. Let R_μ denote the manager's expectation of total revenue from the project in state μ and let $\mu = 1$ denote the good state, i.e. assume $R_1 > R_0$. To make the problem nontrivial, assume that investment is unprofitable in the bad state $\forall r > 0$, i.e. $R_0 < C$. Note that the two-state assumption simplifies the exposition but is not essential; the results hold with a continuous state space.

Investing in the plant allows the firm to start production in the next period, so revenue starts accruing one period after the investment is made. The plant generates revenue via sales of the product for a fixed number of periods, after which it is worthless. The decision to invest is irreversible -- once the plant is built, it cannot be sold at any price.⁷ This assumption, which is equivalent to assuming a large non-convex adjustment cost for the capital stock, underlies most recent investment models (Caballero, 1999). It is motivated by evidence that investment is a very lumpy process in practice. For instance, Doms and Dunne (1993) document that nearly 40% of the median firm's investment over a 17 year span takes place within the span of one year in the U.S. More recently, Goolsbee and Gross (1997) find strong evidence of non-convex adjustment costs in data on investment decisions of US airlines.

⁷Full irreversibility is not essential. If there were a cost to investing and then liquidating the plant, the firm would still be reluctant to plunge resources into a venture of uncertain value. But if the investment decision were fully reversible and all money put in could be recovered, there would be no reason not to invest immediately.

To model learning, assume that the manager can gain further information about the probability distribution governing demand by delaying his investment decision and observing a signal z . These observations can be used to update prior beliefs about the project's payoff, allowing the firm to make a more informed decision. Let $\lambda_0 = P(\mu = 1)$ denote the manager's prior belief that the project will succeed. By postponing his decision to the next period, he updates his estimate of the probability of success to $\lambda_1 = P(\mu = 1|z)$ after observing a realization of z .

The cost of this reduction in uncertainty is that a delayed investment yields revenues one period later, which have lower present value. To simplify the analysis, I abstract from additional costs of delay that the firm may incur, such as the cost of performing the research needed to obtain the signal or the permanent loss of one period of profits. Section 3.2 shows that the key results hold as long as the additional interest-invariant cost of delaying is small relative to the potential benefits of delay.

Having outlined the basic features of the model, we can define the firm's action space and profit functions formally. Let i denote the decision to invest immediately, l the decision to delay, and r the real interest rate. Assume that the investment opportunity is available for T periods; after T periods, the opportunity disappears, perhaps because the patent expires and all rents are bid away.⁸ In the terminal period T , the firm therefore must decide either to invest immediately or reject the project. The profit function, $\pi_t(\mu)$, identifies the expected payoff (in period 1 dollars), to investing in period t when the true state is μ :

$$\pi_t(\mu) = \frac{1}{(1+r)^{t-1}} \left\{ \frac{R_\mu}{1+r} - C \right\} \quad (1)$$

To simplify the discussion below, I restrict attention to the case in which the manager must decide whether to invest or not within $T = 2$ periods: here, delaying investment more than once

⁸Though the phrasing below refers to finite T , the results apply to $T = \infty$ as well.

is not possible. However, all the results for the basic model are proved in the appendix for general T , including the limiting case of an infinite decision horizon.

1.2 Optimal Investment Rule

The optimal action in each period can be computed by solving the firm's dynamic programming problem using backwards induction. To reduce notation, assume that the signal z is a scalar, and that the likelihood ratio of the two densities, $\frac{g(z)}{f(z)}$, is monotonically and continuously increasing in z . This monotonic likelihood ratio property holds for many distributions, including all one parameter Natural Exponential Families. The motivation for these assumptions will become clear shortly. Let $V(i)$ denote the expected value of investing in period 1 and $V(l)$ the expected value of delay.

Lemma 1 In period 2, the firm invests iff $z > z^*$ where z^* satisfies

$$\frac{g(z^*)}{f(z^*)} = \frac{1 - \lambda_0}{\lambda_0} \frac{C - R_0/(1+r)}{R_1/(1+r) - C} \quad (2)$$

In period 1, the firm invests iff

$$\begin{aligned} V(i) &= \lambda_0 \left(\frac{R_1}{1+r} - C \right) + (1 - \lambda_0) \left(\frac{R_0}{1+r} - C \right) \\ &> V(l) = \frac{1}{1+r} \left\{ \lambda_0 \beta(z^*) \left(\frac{R_1}{1+r} - C \right) + (1 - \lambda_0) \alpha(z^*) \left(\frac{R_0}{1+r} - C \right) \right\} \end{aligned} \quad (3)$$

where $\beta(z^*) \equiv \int_{z^*}^{\infty} g(z) dz$ and $\alpha(z^*) \equiv \int_{z^*}^{\infty} f(z) dz$.

When making his period 2 decision, the manager needs to determine the relative likelihood that the observed signal z came from the distributions corresponding to $\mu = 0, 1$. He refines his estimate of $P(\mu = 1)$ using Bayes Rule, and compares the expected payoffs to investing and not

investing in period 2 given his updated belief. If the likelihood that the observed demand z came from the good distribution g is high -- that is, if $\frac{g(z)}{f(z)}$ exceeds some threshold value -- he invests. Hence, the firm's second period decision rule is formally identical to a likelihood ratio hypothesis test. The test here has power $\beta(z^*)$ and type one error rate $\alpha(z^*)$, where z^* is chosen via profit maximization in period 2. In the limiting case of noiseless signals, $\beta(x) = 1$ and $\alpha(x) = 0 \forall x$. The probability that the manager will invest in period 2 when $\mu = 1$ is β , while the probability that he will invest when $\mu = 0$ is α .

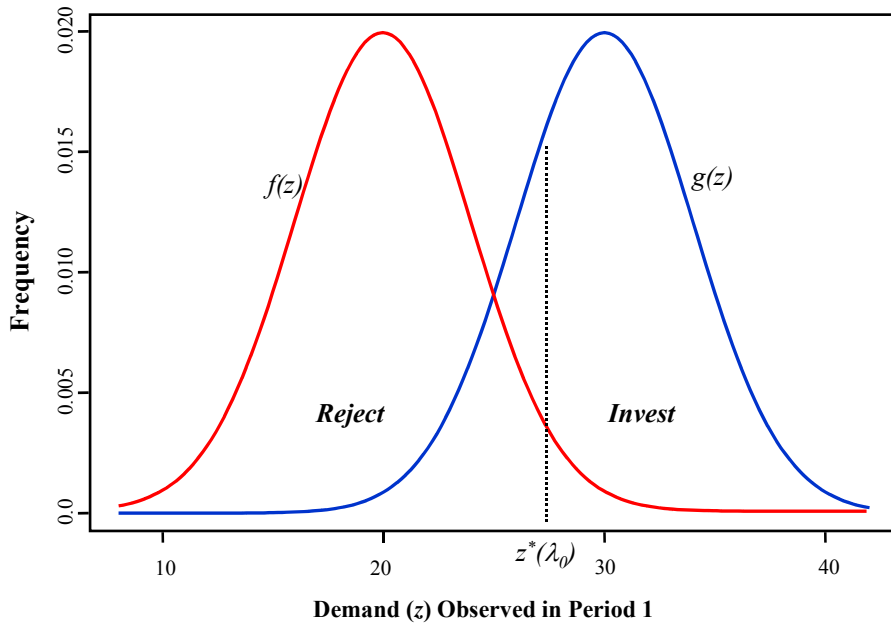


FIGURE 1. PERIOD 2 INVESTMENT DECISION AFTER OBSERVING DEMAND SIGNAL

If $\frac{g(z)}{f(z)}$ is monotonic, this likelihood ratio test translates into a cutoff value for investment in the second period (z^*) determined by the manager's prior odds and the profit-loss ratio. The cutoff z^* is computed as in (2) so that the expected profit from investing in period 2 conditional on observing a signal of exactly z^* is zero. Intuitively, at the optimal threshold, the manager should be indifferent between investing and not investing in period 2; if he were not, there would

either be a region of the state space where he is investing and earning negative expected profits or one where he is not investing when he could have earned positive expected profits.

In period 1, the firm again chooses the action that maximizes its expected payoff, where possible actions are now to invest or learn by delaying. The payoff to investing is the expected profit in period 1, where the weight in the expectation is given by the prior belief, λ_0 . The payoff to learning, $V(l)$, is also a weighted average of profits in each state, but there are two changes in the formula. First, the relevant payoff outcomes are π_2 instead of π_1 -- revenue is discounted more steeply because it is earned one period later. Second, the weights in the profit expression are multiplied by the factors $\beta(z^*)$ and $\alpha(z^*)$. The term corresponding to the good state, $\pi_2(1)$, decreases by the weight $\beta(z^*) < 1$ because of the chance of rejecting the project when it is profitable. The test's benefit is that $\alpha(z^*) < 1$, placing less weight on the negative term corresponding to the bad state. The sole benefit of delaying investment is to reduce the probability of undertaking an unprofitable venture.

The period 1 investment rule is closely linked to the results of existing real options models. To see this, let g denote the expected growth in profits by delaying, which is the undiscounted expected profit in period 2 divided by the expected profit in period 1 (minus 1):

$$g = \frac{\{\lambda_0\beta(z^*)(\frac{R_1}{1+r} - C) + (1 - \lambda_0)\alpha(z^*)(\frac{R_0}{1+r} - C)\}}{\lambda_0(\frac{R_1}{1+r} - C) + (1 - \lambda_0)(\frac{R_0}{1+r} - C)} - 1 \quad (4)$$

We can rewrite the period 1 optimality condition for investment given in (0) as

$$V(i) > 0 \text{ and } r > g \quad (5)$$

As we will see shortly, this expression is the key condition that drives the backward-bending result. The intuition underlying this condition is that it is optimal to invest if (a) the expected profit from investment is positive *and* (b) the growth rate of profits from delaying, g , is smaller than the interest rate, r . If the second condition is not satisfied, the firm will delay since doing so yields a higher expected rate of return than the market interest rate.

Importantly, the intuition in (5), and therefore the subsequent results, apply much more generally than in the simple stylized model considered here. The same condition characterizes investment behavior in a large set of models in the optimal "tree-cutting" literature, pioneered in the analysis of public investment by Marglin (1967, 1970), and developed further in the real options literature reviewed by Dixit and Pindyck (1994).⁹ A basic insight of these models is that the optimal time to cut a growing tree, once it has already been planted, is precisely when the rate of return on the best alternative (r) begins to exceed the rate at which the tree grows (g). The present model gives a learning interpretation to the "growth" of the tree, which generates several additional predictions that can be used to test the model and refine understanding of investment behavior more generally.

The two parts of equation (5) drive the two effects of interest rate changes in this model. The second part shows that a reduction in r causes investors to cut trees later (postpone investment), because it is more likely that $g > r$. The first part shows that a reduction in r also makes more individuals plant trees (increasing the scale of investment), because more projects have positive expected value. These two effects are hard to see at the firm level because there are discontinuous jumps in investment as r changes in this extensive-margin model. To analyze how changes in r affect investment more intuitively, I now aggregate the model over heterogeneous firms and derive a smooth aggregate investment demand curve.

1.3 Aggregation

Consider an economy populated by a continuum of firms with heterogeneous prior probabilities of success (λ_0 s). Assume that the density of λ_0 , denoted by $d\eta(\lambda_0)$, is continuous and places non-zero weight on all $\lambda_0 \in [0, 1]$. Revenues from investment in each state and the learning technology are identical across firms. In addition, assume for now that each firm's profit

⁹See, for example, Cukierman (1980), Bernanke (1983), McDonald and Siegel (1986), Pindyck (1988), Demers (1991), Leahy (1993), and Bertola and Caballero (1994).

realization is independent of other firms' outcomes, so firms can ignore the behavior of other firms when making investment decisions.

Under these assumptions, it follows that each firm follows Lemma 1 in determining its investment rule. The period 1 investment decisions of each firm can be identified by computing the action d that maximizes $V(d; \lambda_0)$ for each λ_0 .¹⁰ This allows us to characterize the decisions of all firms in the economy by a single threshold value λ_0^* that determines who invests in period 1 and who does not, as shown in Figure 2. The next lemma establishes this result formally.

Lemma 2 There is a unique λ_0^* at which the value of investing equals that of postponing. In period 1, firms with $\lambda_0 < \lambda_0^*$ delay their investment decision. Firms with $\lambda_0 \geq \lambda_0^*$ invest in period 1.

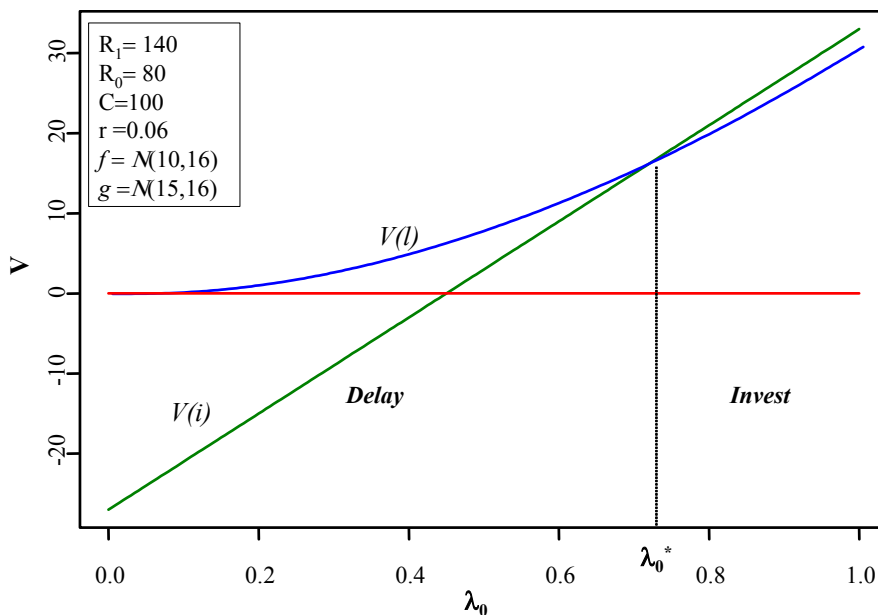


FIGURE 2. EXPECTED PAYOFFS AND INVESTMENT BEHAVIOR IN THE ECONOMY

Investment behavior in the economy follows a simple pattern: Confident firms (λ_0 high) do not want to forego profits by delaying and invest immediately. The remaining firms, who are

¹⁰Unless otherwise noted, all subsequent figures use the parameter values given in this figure.

less certain about whether they have a profitable project, choose to wait and decide what to do in the next period based on the information they observe. The threshold λ_0^* thus determines the scale of investment in the economy.

It follows from Lemma 2 that aggregate period 1 investment is

$$I = \int_{\lambda_0^*}^1 C d\eta(\lambda_0) \quad (6)$$

Note that "period 1" investment is always equal to "current" investment; in other words, the economy is always currently in period 1. The reason is that the only state variable in any firm's dynamic programming problem, irrespective of when it started learning, is its current belief, which we call λ_0 .¹¹ Firms that existed prior to the current period and already acquired information about their projects simply have a different value of λ_0 . Hence, we use the terms "period 1 investment" and "current investment" interchangeably below.

2 Interest Rates and Investment Demand

In the model above, the level of current (period 1) aggregate investment is *always* a backward-bending function of the interest rate. Irrespective of the underlying parameters, $I(r)$ has an upward-sloping segment from $r = 0$ to $r = r^* > 0$ followed by a downward sloping segment thereafter.

Proposition 1 Investment demand is a backward-bending function of the interest rate.

- (i) $I(r = 0) = 0$ and $\lim_{r \rightarrow 0} \frac{\partial I}{\partial r}(r) = +\infty$
- (ii) $r^* \equiv \operatorname{argmax}_r I(r) > 0$ and $r < r^* \Rightarrow \frac{\partial I}{\partial r} > 0$ ¹²

¹¹With a finite decision horizon, the current period t is also a state variable, but with appropriate redefinition of T the current belief remains a sufficient statistic to compute investment behavior.

¹²More precisely, $\frac{\partial I}{\partial r} = 0$ for $r > \frac{R_1}{C} - 1$, the uninteresting case in which the interest rate is so high that investing is suboptimal even in the good state.

To see the intuition for this result, first note that if $r = 0$, no one invests in the first period. Firms certain of success ($\lambda_0 = 1$) are indifferent between postponing and investing today, and all firms with lower priors must therefore strictly prefer delay (Lemma 2). Hence, $I(r = 0) = 0$. In this stylized model, there is no reason to forego the free information one gets by waiting and learning if $r = 0$. Increasing r from $r = 0$ raises the cost of learning by delaying and *increases* aggregate investment by making the most confident firms invest immediately. At the other extreme, if $r > \frac{R_1}{C} - 1$, projects are unprofitable in both states for all firms, and hence no one invests. Since few firms invest when r is low or high, it follows that $I(r)$ is non-monotonic.

A natural concern with this result is that the prediction that investment falls to zero at low interest rates is empirically implausible. However, when certain unrealistic assumptions of the stylized model are relaxed, this prediction disappears, while the backward-bending shape of $I(r)$ remains intact. Two differences between the stylized model and the real world are important in this respect. First, some types of investment, such as replacement of depreciating machines, involve virtually no learning. This component of investment has a conventional downward-sloping relationship with r . In a more general model that allows for both non-learning and learning investment, total investment is positive at $r = 0$. Nonetheless, $I(r)$ remains upward-sloping at low r because $\lim_{r \rightarrow 0} \frac{\partial I}{\partial r}(r) = +\infty$ for the learning component. Second, even within the learning component, there are other non-interest costs to delay such as research expenditures and loss of profits due to competition that are ignored in the model above. Incorporating these other costs eliminates the prediction that $I(r = 0) = 0$, because the most confident investors will not want to incur these additional costs at any interest rate. In section 3.2, I show that $I(r)$ remains backward-bending provided that these costs are not too large. I proceed here with the stylized model that abstracts from non-learning investment and other waiting costs since the main intuitions are most transparent in this setting.

Having discussed why $I(r)$ is non-monotonic, I now explain why it has a *backward-bending* shape more precisely. If a given manager has λ_0 such that $V(i; \lambda_0) < V(l; \lambda_0)$ for all $r \geq 0$, his behavior is unresponsive to changes in r and he does not affect the aggregate investment demand curve. Therefore, to understand how the shape of $I(r)$ emerges from microeconomic decisions, restrict attention to firms that do invest for some value of r .¹³ To analyze the firm's behavior, let us examine how the two payoff functions, $V(i; \lambda_0)$ and $V(l; \lambda_0)$ change with respect to r . Decomposing $\frac{\partial\{V(i)-V(l)\}}{\partial r}$ into the NPV ($\frac{\partial V(i)}{\partial r}$) and learning ($-\frac{\partial V(l)}{\partial r}$) effects gives:

$$\frac{\partial\{V(i) - V(l)\}}{\partial r} = NPV + L \quad (7)$$

where

$$NPV = \frac{-1}{(1+r)^2} \{\lambda_0 R_1 + (1-\lambda_0)R_0\} < 0 \quad (8)$$

$$L = \lambda_0 \beta(z^*) \left\{ \frac{2R_1}{(1+r)^3} - \frac{C}{(1+r)^2} \right\} + (1-\lambda_0) \alpha(z^*) \left\{ \frac{2R_0}{(1+r)^3} - \frac{C}{(1+r)^2} \right\} > 0$$

The NPV effect makes an increase in r reduce the value of immediate investment, as in static investment models. The learning effect arises because the value of delaying is also affected by changes in the interest rate. The two terms of L reflect the fact that an increase in r causes the proceeds of investment at $t = 2$ to be discounted more steeply and also reduces the value of the investment at $t = 2$ in period 2 dollars. Via the L effect, a higher r reduces $V(l)$, creating a force that counteracts the conventional effect by making immediate investment *more* attractive.

These expressions show that the magnitude of $L(r)$ diminishes relative to the magnitude of $NPV(r)$ as r gets larger. Hence, for any given λ_0 , there is exactly *one* value r at which $NPV(r) = -L(r)$. This implies that for a given firm, $V(i; \lambda_0)$ and $V(l; \lambda_0)$ intersect for at most two values of r , say $r_L(\lambda_0)$ and $r_U(\lambda_0)$. The individual investment demand curves thus all have the same form: invest iff $r_L(\lambda_0) \leq r \leq r_U(\lambda_0)$, as shown in Figure 3. The source of the

¹³Such firms exist: For $\lambda_0 = 1$, $V(i; \lambda_0) > V(l; \lambda_0) \forall r > 0 \Rightarrow \exists \lambda'_0 < 1$ s.t. $V(i; \lambda'_0) > V(l; \lambda'_0)$ for some $r > 0$ by continuity.

non-monotonicity with respect to r is that a small increase in r causes $V(l)$ to fall *more than* $V(i)$ at $r_L(\lambda_0)$, increasing period 1 investment by firm λ_0 , but an increase in r causes $V(l)$ to fall *less than* $V(i)$ at $r_U(\lambda_0)$, reducing the level of investment by the same firm.

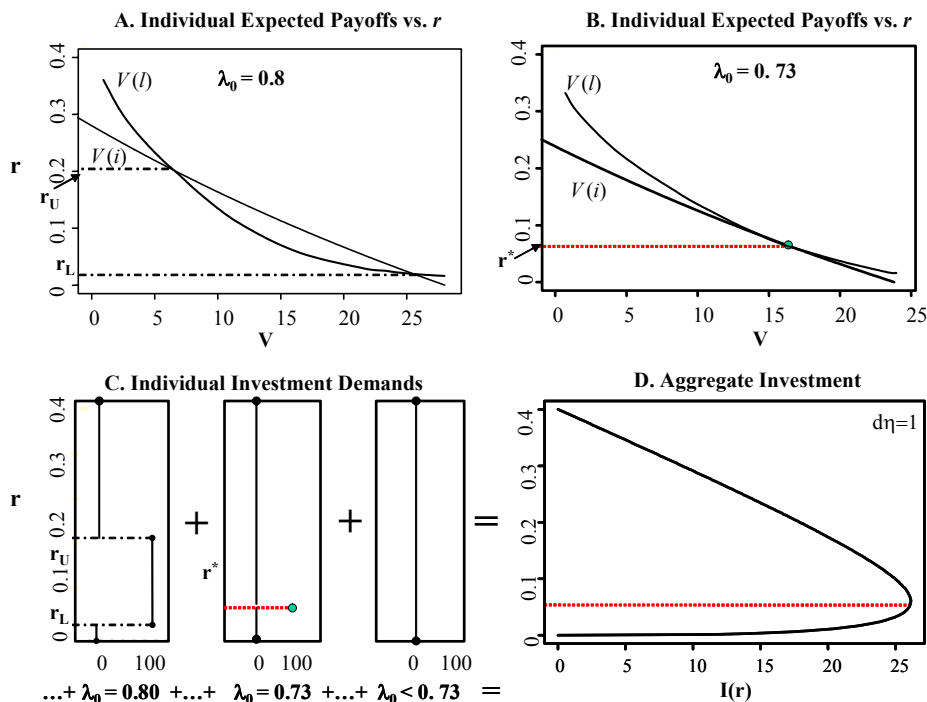


FIGURE 3. EFFECT OF INTEREST RATE ON PERIOD 1 INVESTMENT DEMAND
Notes: Firms compare $V(l)$ and $V(i)$ for each value of r (A,B) and compute their investment demands as functions of r (C). Summing these step functions horizontally yields $I(r)$ (D).

It can be shown that $r_L(\lambda_0)$ is decreasing in λ_0 and $r_U(\lambda_0)$ is increasing in λ_0 -- more confident firms have a larger range of interest rates for which they finding immediate investment optimal. At the extremes, investors with $\lambda_0 = 1$ strictly prefer i for any $r \in (0, \frac{R_1}{C} - 1)$, whereas investors with $\lambda_0 = 0$ prefer not to invest $\forall r > 0$. There is exactly one λ'_0 such that $r_L(\lambda'_0) = r_U(\lambda'_0)$. For this firm, $V(i; \lambda_0)$ and $V(l; \lambda_0)$ are tangent at $r^* = r_L(\lambda'_0) = r_U(\lambda'_0)$. The λ'_0 firm invests *only* if $r = r^*$. $I(r)$ is maximized at r^* because all firms who have $\lambda_0 > \lambda'_0$ also invest at r^* by Lemma 2.

Summing the individual non-monotonic step functions horizontally generates a smooth aggregate investment demand curve. Aggregate investment demand is a backward-bending

function of r because the firm-level investment demand curves are non-monotonic step functions that are strictly nested within each other as λ_0 falls, as shown in Figure 3. The slope of $I(r)$ approaches $+\infty$ as r tends to 0 because the most confident investors have little to gain by learning and immediately jump into the market when a small cost of delay is introduced.

Note that in contrast to investment demand, the value of the firm is a strictly downward sloping function of the interest rate, because both $V(i)$ and $V(l)$ rise when r falls. Lower interest rates essentially lead to more investment in information rather than physical capital such as equipment and structures, ultimately yielding higher profit rates. If the measure of "investment" is broadened to include the value of information, the conventional prediction that higher interest rates lower investment still holds. However, from a normative perspective, the distinction between investment in information and physical capital could be important. If a policy maker's goal is to stimulate job creation or aggregate demand, or if physical investment leads to spillovers that raise growth, the amount of investment in equipment and structures itself may matter.¹⁴ Therefore, while the results of this paper are in some sense empirical claims related to the measurement of investment, they also have real implications for economic welfare.

3 Extensions

3.1 Scale choice

To incorporate scale choice at the firm level, assume that each firm can set investment in periods 1 and 2, I_1 and I_2 , at any positive value. The restriction that investment must be positive captures irreversibility. There are two states of the world, which differ in the mean price at which the output good can be sold (p_μ). Investments generate a profit stream for T_P periods. An

¹⁴This is particularly clear when changes in r lead to compositional effects, as in section 3.3. If firms switch to slower construction methods because r is low, building permits fall. The fact that building permits are perceived as an indicator of the economy's strength suggests that this change in behavior could have real economic consequences.

investment of I_1 generates a profit of $p_\mu f(I_1)$ for T_P periods in state μ , starting in period 2 (one period after investment). A further incremental investment of I_2 in period 2 changes the profit stream to $p_\mu f(I_1 + I_2)$ in periods 3 to $T_P + 2$. Irreversibility of investment is a meaningful restriction only when $T_P > 2$, so assume that this condition holds below.

The scale problem has a solution only if $f(I)$ is concave, i.e., the marginal return to investment is diminishing. This concavity can arise from technological constraints or from a downward-sloping demand curve where price falls as supply rises. The information revelation structure of the model is the same as in section 2: A signal z is observed at the end of period 1 and beliefs are then updated. Let the ex-ante probability of state 1 be given by λ_0 . We can now generalize Proposition 1.

Proposition 2 $I_1(r)$ is backward-bending when firms choose scale.

To understand this result intuitively, recall that the key step in the proof in the extensive-margin case was to show that investment is zero at both low r and high r . This continues to hold here: Since there is no non-interest cost to delay in this simple model, there is no reason to invest immediately if $r = 0$. In other words, profit-maximizing firms will rationally choose a scale of 0 investment in period 1, implying $I_1(r = 0) = 0$. Similarly, if r is sufficiently high, investment is undesirable. Hence, investment-demand must be a non-monotonic function of r . More generally, in an environment with other costs of waiting or non-learning investment, the scale of investment is relatively low at both low r and high r , yielding a non-monotonic $I_1(r)$ curve with $I_1(r = 0) > 0$.

To see why the main result does not change when scale choice is permitted, it is helpful to consider the following alternative model of scale choice. Suppose a firm has many projects in which it can invest, some of which have higher probabilities of success than others. The firm must make a binary decision about each individual project but can choose the total number of projects to take up. As the firm raises investment, it is forced to choose projects with lower

probabilities of success, making its profits a concave function of investment, as in the continuous scale-choice model. Since each project decision is made independently, investment decisions are determined exactly as in Lemma 2. Consequently, the total scale of investment by this firm, $I^f(r)$, has the same form as equation (6), the expression for aggregate investment in the original model where several small firms make investment decisions on different projects. Firms are divisible, so total investment is identical if many small firms make decisions about one project each or one big firm makes investment decisions on several projects.

Since $I^f(r)$ has the same form as (6), it follows that it also has the same backward-bending shape. Put differently, the original aggregate model with extensive-margin choices at the microeconomic level effectively contained a scale choice in the aggregate, so it already contained the intensive-margin ("plant fewer trees") effect of increasing r . Modelling this effect at the firm level instead of the aggregate level does not change the result.

3.2 Competition

In the stylized model, investors enjoy pure rents from their investments. While patent and copyright protection limit competition in some cases, in practice most firms face some degree of competition in the long run. Competition reduces the option value of delay, since rents cannot persist indefinitely in equilibrium. Hence, it is important to investigate whether the backward-bending result holds when the returns to investment are determined endogenously in competitive equilibrium.

To model competition, let us return to the setting where firms with different product concepts (e.g. different drugs to treat a specific disease) make binary investment decisions. Firms must make a decision to invest within two periods, indexed by $t = 1, 2$. Each firm enters period 1 with a prior probability of success of λ_0 . There is a distribution of λ_0 s to capture heterogeneous expectations as in the basic model. Each firm receives an independent signal about demand for its product at the end of period 1 which is used to update beliefs.

Investment in period t yields revenues in period $t + 1$. Each firm that invests in period t ends up with either a good product that sells for p_t in period $t + 1$ or a bad product that is worthless (sells for \$0). Prices are determined by cumulative supply. Let I_t denote aggregate investment in period t and I_t^c denote cumulative investment up to and including period t . The inverse-demand function for good products made in period t is given by an arbitrary downward-sloping function $p(I_t^c)$.¹⁵ This implies that the price of the product falls over time ($p_1 > p_2$).

To capture free entry in the long run, assume that profits are bid to zero after the first period in which a particular product is sold. After this point, other firms can replicate the technology, forcing the original firm to sell at cost. A firm that invests in period 1 thus has a chance to earn positive profits in period 2 only; firms that invest in period 2 can earn positive profits in period 3 only. The one-period lag captures adjustment costs which prevent competitors from bidding away infra-marginal quasi-rents (short run surplus) by selling an identical product instantly. Note that this model of competition parallels neoclassical competitive production theory, where producer surplus is positive in the short run and falls to zero in the long run. The pharmaceutical industry is a good example to keep in mind for concreteness: First-movers can earn large profits in the short run (e.g. Aspirin until generics are introduced), while subsequent firms with slightly different products can also earn temporary rents (Tylenol, Advil, Motrin) until their profits are bid away by generics as well.

This setup allows us to write the expected profit from immediate investment (i) and learning (l) for a firm with prior λ_0 as:

$$V(i, \lambda_0) = \frac{\lambda_0 p_1}{1+r} - C$$

$$V(l, \lambda_0) = \frac{1}{1+r} \{ \lambda_0 \beta(z^*) (\frac{p_2}{1+r} - C) + (1 - \lambda_0) \alpha(z^*) (-C) \}$$

I first establish the existence and uniqueness of equilibrium in this model. In equilibrium, all firms with $V(i, \lambda_0) > V(l, \lambda_0)$ at the market price vector (p_1, p_2) invest in period 1, and markets

¹⁵This model is equivalent to one where prices depend on the supply of good products rather than total investment because there is a monotonic link between I_t^c and the supply of good products, and $p(I_t^c)$ is an arbitrary function.

clear in each period. Investment behavior in the economy follows a pattern similar to that in the base case.

Lemma 3 In period 1 equilibrium, there is a unique price vector (p_1, p_2) and threshold λ_0^* at which $V(i, \lambda_0^*) = V(l, \lambda_0^*) > 0$.

Firms with $\lambda_0 < \lambda_0^*$ delay their investment decision.

Firms with $\lambda_0 \geq \lambda_0^*$ invest in period 1.

The key point of Lemma 3 is that the marginal investor in period 1 equilibrium earns *strictly positive* expected profits from immediate investment. Unlike in the neoclassical model of competition, profits are not driven to zero at the margin in the period 1 equilibrium. To understand this result, first consider period 2 decisions. Since there is no further option to delay, a firm invests in period 2 if its expected return to investment at the market-clearing price exceeds the cost of investment. Consequently, there is a threshold value λ_1^* such that only firms with updated probabilities of success $\lambda_1 > \lambda_1^*$ invest in period 2. The marginal firm with belief λ_1^* earns zero profits in equilibrium. But the infra-marginal firms who have higher λ_1 s earn positive profits in expectation. These firms are able to earn short-run quasi-rents despite being in a competitive market because they have a better technology (higher λ_1) that cannot be instantly replicated by other firms. For example, they might have access to more fertile land, a better chemical compound, or better human capital that gives them a short-run advantage. However, after one period passes, other firms are able to observe the technology of successful firms, and free entry leads to zero profits.

Now turn to period 1 behavior. There is some probability that the marginal investor in period 1 will be one of the infra-marginal investors in period 2. Hence the value of postponing must be strictly positive for this indifferent firm. The reason that NPV is not driven to zero in the period 1 equilibrium is again heterogeneity in success probabilities. Other firms are free to

enter the market and try to capture the positive rents, but they have lower probabilities of success than the indifferent firm, and therefore opt to delay instead.

Since the option value of delaying is positive for the marginal period 1 investor in equilibrium, changes in r continue to affect that firm's behavior via both an NPV and learning effect, as in the basic model. The existence of these two opposing forces suggests that period 1 investment demand, $I_1(r)$, may be non-monotonic in competitive equilibrium. This result cannot be established as easily as in the baseline model because there is now a non-interest cost to waiting, so $I_1(r = 0)$ is no longer 0.¹⁶ Even at a zero interest rate, the most confident (highest λ_0) firms will invest immediately to take advantage of the high initial price they can extract. Nonetheless, one can obtain a simple condition under which the investment demand curve in this model is upward-sloping at low r .

Proposition 3 Let λ_0^* and λ_1^* denote the success probabilities of the marginal (indifferent)

investors in periods 1 and 2, respectively. Then $\partial I_1 / \partial r(r = 0) > 0$ if at $r = 0$,

$$\lambda_0^* \beta(\lambda_0^*) > \lambda_1^* \quad (9)$$

This condition requires that the marginal investor in period 1 have a significantly higher success probability than the marginal investor in period 2. Since the marginal investor in period 2 earns zero profits in equilibrium, this condition guarantees that the marginal investor in period 1 can gain substantial rents by delaying and investing in period 2, since he is likely to be an infra-marginal investor in that period.

To understand why (9) is required intuitively, it is helpful to consider two extreme examples. First, suppose signals are perfect, so that $\beta(\lambda_0^*) = 1$. In this case, the distribution of λ_1 is a degenerate two-point distribution, and if supply is sufficiently large, price is driven down to $p_2 = C$ in the second period. Since firms cannot earn any profits if they delay investment, the option to delay is worthless. The model collapses into the conventional single-period model,

¹⁶However, at very high r , it remains the case that investment is suboptimal for all firms, so aggregate investment falls to zero as $r \rightarrow \infty$. Hence, $I(r)$ must have a downward-sloping segment in the competitive model.

where r has only a conventional cost-of-capital (scale) effect and $I(r)$ is strictly downward sloping. Correspondingly, (9) does not hold in this case because $\lambda_1^* = 1$. This example illustrates that the "timing effect" of r can emerge only if delaying is a "real option" that has value in equilibrium. Condition (9) essentially guarantees that the option to delay has value.

Now consider a second example, where signals are imperfect. Suppose the demand curve for the good product is

$$p_t = f(I_t^c) + K$$

where K is a constant and $\partial f / \partial I_t^c < 0$ so that demand is downward-sloping. Suppose the cost of investment is

$$C = C_0 + \frac{1}{2}K$$

In this example, K controls the variance of payoffs: High K yields higher profits in the good state but a bigger loss in the bad state. The following result establishes that (9) holds when payoff uncertainty is sufficiently high, implying that $I(r)$ is upward-sloping at low r :

Corollary to Proposition 3 For K sufficiently large, $\partial I / \partial r(r = 0) > 0$

The mechanics driving this result are straightforward. As K becomes large, the threshold for investment in period 2 approaches $\lambda_1^* = \frac{1}{2}$ because firms earn approximately the same amount in the good state ($\frac{K}{2}$) as they lose in the bad state. In period 1, increased uncertainty makes delay more attractive for each firm, raising the threshold for investment λ_0^* . Therefore, as the amount of uncertainty grows larger, λ_0^* and consequently $\beta(\lambda_0^*)$ approach 1 while λ_1^* approaches $\frac{1}{2}$, so that (9) is eventually satisfied. Intuitively, in a very risky environment, the incentive to delay and acquire information is large, so only the most confident investors take advantage of high equilibrium prices in period 1. However, once there is no further opportunity to learn, many lower-capability firms are willing to take risky but positive NPV risks, creating large infra-marginal rents in the second period for the marginal period 1 investor. These large rents become less valuable when interest rates rise, compelling the marginal firm to start investing

immediately when r rises from $r = 0$, and raising aggregate investment in competitive equilibrium.

The model of competition analyzed here is obviously quite stark, but the results can be extended to richer settings where entry dynamics are endogenous and prices fall gradually as competitors enter the market. The main conclusion to be drawn from this analysis is that $I(r)$ is upward-sloping at small r in competitive equilibrium if firms can earn significant temporary rents. More generally, as long as the non-interest costs of delay -- whether from competitive forces, research costs, or other sources -- are small relative to the rents earned by the marginal period 1 investor, $I(r)$ is non-monotonic. The shape of $I(r)$ thus depends on whether firms actually earn temporary rents in practice and treat delaying as a valuable option. Anecdotal evidence suggests that many successful firms do earn revenues far above costs at least in certain industries with large fixed costs and substantial uncertainty, such as software, pharmaceuticals, apparel, and media. More systematic evidence that firms value the option to delay at the margin is given in section 4.

3.3 Investment Composition Decisions

We have assumed thus far that firms make a one-dimensional decision about the scale of investment. However, in practice, firms make many choices about projects beyond scale. For instance, they may choose technologies for construction, speed of delivery to market, etc. To see how these "investment composition" decisions affect the shape of $I(r)$, consider a model where firms can choose between two construction methods, A and B. Method A requires the use of expensive building materials and is fast (e.g. 1 year to build). Method B involves less real investment but is slower (e.g. 2 years to build). At $r = 0$, time is costless, so the firm will use only method B. When r is very high, time is precious, and the firm will use only method A. For intermediate interest rates, the firm will use a combination of these two methods. Since method A involves more real investment than method B, the composition effect, holding scale fixed,

makes $I(r)$ strictly upward sloping. As the scale effect dominates at high r -- for sufficiently high r , it is best not to invest with any technology -- $I(r)$ is downward-sloping for high r when scale is endogenous. However, composition effects lengthen the upward-sloping segment of $I(r)$ and raise the investment-maximizing r^* generated by the basic model with only learning effects.

In the tree-cutting and planting analogy of section 2, composition choices are the kinds of trees one plants (oak or apple). Increases in the interest rate have three effects in this environment: (1) Plant fewer trees; (2) Cut trees later; (3) Plant trees that mature later. Effects 2 and 3 act to make $I(r)$ upward-sloping at low interest rates. Generalizing the model to allow composition choices thus reinforces the main result, and illustrates that learning is not the only reason that firms may reduce current investment in response to an interest rate cut.

4 Evidence for Timing Effects

The upward-sloping portion of $I(r)$ emerges only if firms actually compare the value of immediate investment with the discounted value of future investment when making decisions in practice (and do not simply follow NPV rules). Hence, a natural place to start in assessing the empirical relevance of this model of investment is to ask whether companies take into account the option to delay investment. There is now considerable evidence supporting this hypothesis. For example, Coy (1999) describes several examples where companies such as Airbus, Hewlett-Packard, Chevron, and Enron made investment decisions using an explicit real-options analysis. Summers (1987) and Poterba and Summers (1995) surveyed Fortune 1,000 CEOs about their business practices and found that hurdle rates for immediate investment are 2-3 times higher than the user cost of capital. McDonald (2000) argues that the hurdle rate methods used by many companies serve as good rules-of-thumb for real-options calculations. These studies suggest that most firms do not simply maximize NPV, and at least implicitly value the option to wait and acquire more information.

Several studies also give direct evidence that firms actively choose the *time* of investment. Hellman and Puri (2000) find that the mean time-to-market for startup firms in Silicon Valley is 2.6 years, but various factors such as sources of finance affect this time significantly. In addition, microeconomic and macroeconomic studies have shown that firms wait longer to invest in more uncertain environments, as the model of this paper and real options theory predicts (e.g. Pindyck and Solimano 1993; Bulan, Mayer, and Somerville 2003).

The very fact that the value of delay is taken into account indicates that the option to delay is not worthless at the margin in competitive equilibrium. In addition, interest rate changes must affect the timing of investment at least for firms that explicitly compare the discounted values of investment options. To complement this reasoning, I now describe a set of recent studies which provide direct evidence on the timing effects of interest rate changes.

In recent work independent of this paper, Jovanovic and Rousseau (2001, 2004) construct a model of initial public offerings (IPOs) which predicts that the time of IPO investment should be positively related to r at low levels of the interest rate and negatively related to r at higher levels of the interest rate.¹⁷ They use time-series variation over the 20th century in the interest rate in the United States, and document a backward-bending relationship between r and time to IPO, precisely as the model predicts.¹⁸

Moel and Tufano (2002) also examine the effect of time-series variation in r on investment decisions. They study the timing of gold mine closures in North America between 1988 and 1997. Since closing a mine involves a large fixed cost, profit-maximizing firms with the option to delay should take the discounted value of expected future cashflows into account when making such decisions. When interest rates are high, the future matters less relative to the

¹⁷The intuition underlying the Jovanovic-Rousseau result is somewhat similar to that here, but the premises and structure of the two models are quite different. Given their focus on the timing of IPOs in particular, their paper does not model learning, heterogeneity across firms, and aggregation as in the model of investment here. The papers also differ in the discussion of extensions and additional comparative statics.

¹⁸Jovanovic and Rousseau do not address the potential endogeneity of the interest rate explicitly. However, they show that r has little effect on incumbent firms' investment in the time series. Since IPOs are a small fraction of total investment, autonomous shocks to IPO investment are unlikely to affect r . Given that total investment is uncorrelated with r in the time series, r can arguably be taken as exogenous to changes in IPO investment.

present, and the immediate costs of closing a mine are more likely to outweigh the potential benefits. Hence, if firms take interest rates into account when timing decisions, mine closures should be less frequent when interest rates are high. Moel and Tufano find that firms are indeed more likely to keep mines open in periods when interest rates are high.

Capozza and Li (2001) test a model of real-estate development where interest-rate changes have non-monotonic effects on construction for reasons related to those described above.¹⁹ They provide cross-sectional evidence for timing effects using data on building permits in the United States from 1981 to 1989. They first estimate the interest-elasticity of building permits for each metropolitan area during this period. The interest-elasticities vary substantially across metropolitan areas. The elasticities are most negative in areas in with minimal price volatility and most positive in the most risky metropolitan areas. Hence, timing effects -- which drive up the interest elasticity -- seem to be most important in risky environments. This finding is consistent with the model's prediction that timing effects should be larger when firms face uncertainty and the option to delay is very valuable (see section 5.1).

In another cross-sectional study, Hellman and Puri (2000) examine variation in the time to market for startup firms in Silicon Valley in the 1990s. They show that firms backed by venture capitalists, who charge a high effective rate of interest, rush their product to market much more quickly than firms financed through other traditional means. Hellman and Puri conduct several tests using data on ex-ante characteristics to rule out the hypothesis that this difference in time-to-invest is a selection effect caused by differences in the types of firms that seek VC funding vs. other sources of funding. They conclude that VC funding -- which comes with a high interest rate -- has a causal effect on time to invest, again supporting the main timing intuition of the present model.

¹⁹Capozza and Li study a continuous-time model with stochastic price that builds on Arnott and Lewis's (1979) model of land development and other real options models. They show that the effect of changes in r on the speed of investment is indeterminate in this setting. The most important difference between these papers and the present study is that they do not show that $I(r)$ has a backward-bending shape. In addition, they do not model learning, aggregate over heterogenous firms, or consider the extensions such as competition and other comparative statics studied here.

5 Additional Empirical Implications

This section presents a set of testable comparative statics that could be used to test the empirical relevance of the model for aggregate investment more systematically in future work. These predictions are derived in the basic model of section 2 for simplicity.

5.1 Potential to Learn

I first show that an increase in r is most likely to increase investment in environments with a high potential to learn. A formal definition of changes in the "potential to learn" is necessary to operationalize the comparative statics analysis. Intuitively, a firm can learn more rapidly if "signal noise" is lower, i.e. if it is easier to distinguish whether z is drawn from f or g . Recall that any manager's second period decision is the outcome of a hypothesis test. We will say that "signal noise" rises if the power of the test, $\beta(x) = \int_x^\infty g(z)dz$, falls while the type 1 error rate, $\alpha(x) = \int_x^\infty f(z)dz$, rises for *all* cutoff values x below the point at which f and g are indistinguishable. Formally, let $s(f, g)$ denote the level of signal noise with densities f and g , and x' the unique point at which $\frac{g(x')}{f(x')} = 1$. Then

$$s(f_1, g_1) > s(f_2, g_2) \text{ if } \beta_2(x) > \beta_1(x) \text{ and } \alpha_1(x) > \alpha_2(x) \forall x < \min(x'_1, x'_2) \quad (10)$$

where x'_j is s.t. $\frac{g(x'_j)}{f(x'_j)} = 1$. Note that this definition is an incomplete ordering since it does not rank all distributions in terms of signal noise. A leading example of an increase in signal noise according to this definition is a rightward shift of $g(z)$ or a leftward shift of $f(z)$ in Figure 1.

Before turning to the relationship between signal noise and $\frac{\partial I}{\partial r}$, it is useful to first establish the connection between signal noise and the level of I itself.

Lemma 4 An increase in signal noise increases current investment

$$s(f_1, g_1) > s(f_2, g_2) \Rightarrow I(f_1, g_1) > I(f_2, g_2)$$

When signal noise rises, a firm's ability to learn about the true value of μ by waiting is reduced. This reduces the value of delaying investment, making aggregate investment rise. Cukierman (1980) gives an analogous result: Increases in the variance of earnings reduce current investment by raising the value of delay.

How does an increase in signal noise affect the shape of $I(r)$? To build intuition, consider the extreme case of totally uninformative signals ($f = g$). In this case, the model collapses into the neoclassical model and the $I(r)$ curve is downward-sloping, i.e. $r^* = 0$. This observation suggests that the potential to learn should be positively associated with r^* ; that is, the upward-sloping segment of the investment-demand curve should be larger in industries or times where there is more to be learned. The following proposition establishes that this is indeed the case provided that the payoff in the bad state is sufficiently low, or, equivalently, the variance of returns is sufficiently high relative to the expected return.

Proposition 4 $\exists \overline{R}_0 > 0$ s.t. if $R_0 < \overline{R}_0$, a reduction in signal noise raises r^* :

$$s(f_2, g_2) < s(f_1, g_1) \Rightarrow r_2^* > r_1^*$$

To see the intuition for this result, observe that changes in signal noise affect only $V(l)$, leaving $V(i)$ unaffected for each firm. An increase in r is more likely to raise aggregate investment if it tends to reduce $V(l)$ more than $V(i)$, making immediate investment preferable, for a given set of parameters. When signal uncertainty is lowered, $V(l)$ changes in two ways. First, firms have a higher probability of investing in the good state in period 2 (β rises). Second, firms have a lower probability of investing in the bad state (α falls). The first effect makes expected period 2 profits more sensitive to the interest rate, since there is a higher probability of earning revenues in the good state. The second effect goes in the opposite direction, since there is a *lower* probability of earning revenues in the bad state. If R_0 is small, the second effect is

small in magnitude relative to the first, so $V(l)$ is more sensitive to r overall. For instance when $R_0 = 0$, an increase in r has no effect at all on revenues in the bad state. Therefore, provided that R_0 is small, an increase in r is more likely to reduce $V(l)$ relative to $V(i)$ for each firm and thereby raise aggregate investment when signal uncertainty is lower.

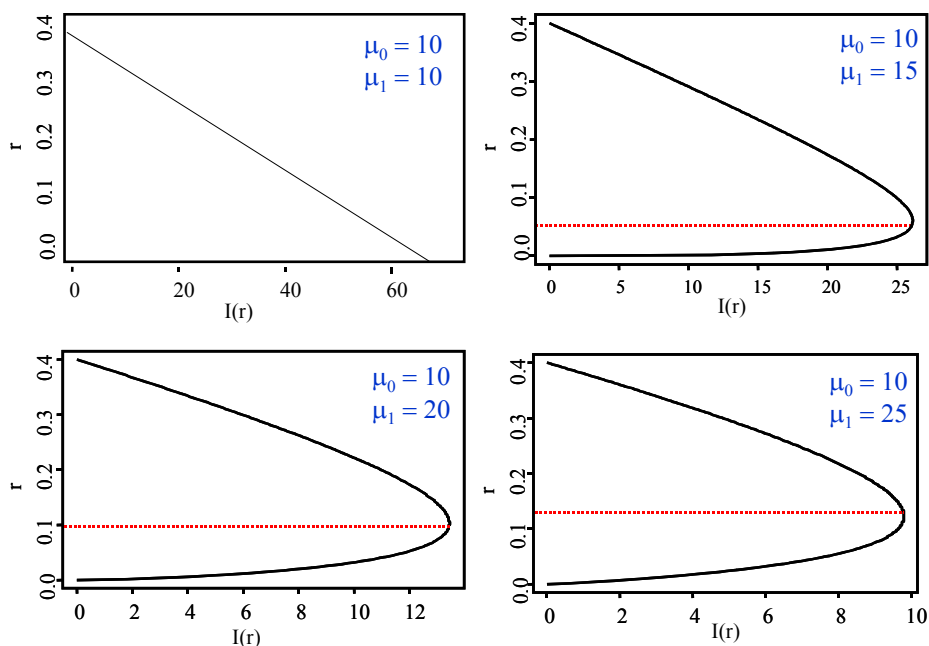


FIGURE 4. SIGNAL NOISE AND THE SHAPE OF $I(r)$
 This figure shows $I(r)$ for four pairs of signal distributions f and g . The distributions are Normal with a mean of μ_0 for f and μ_1 for g and a standard deviation of 16. As $\mu_1 - \mu_0$ rises, signal uncertainty falls.

The low R_0 condition on the result requires that the variance of earnings be high relative to the mean profit rate, which is essentially a requirement that good information about the state of the world is valuable. The variance of profits is typically quite high in practice: Investments usually either have very large payoffs or are complete failures. Hence, the important empirical implication is that interest rate increases are more likely to stimulate investment in industries or times where the potential to learn is greater, as shown in Figure 4.

5.2 Short Run vs. Long Run

I now turn to the effects of changes in r on total investment over a longer horizon, taking into account changes in investment behavior beyond the current period. In this case, it is necessary to consider the more general T period formulation of the model instead of the two period special case discussed above. In this more general model, the firm has the option to delay investment in every period from 1 to $T - 1$.²⁰ As noted earlier, the preceding results apply when $T > 2$ as well. In the T period model, total investment from period 1 to t is

$$I_{1,t} = \sum_{s=1}^t I_s = \int_{\lambda_0^*}^1 C d\eta(\lambda_0) + \sum_{s=2}^t \int_0^{\lambda_0^*} P_1(I_s|\lambda_0) C d\eta(\lambda_0) \quad (11)$$

where $P_1(I_s|\lambda_0)$ is the probability that a firm with prior λ_0 ends up investing in period s . The next proposition analyzes the relationship between $I_{1,t}$ and r .

Proposition 5

(i) $I_{1,t}(r)$ is a backward-bending function of $r \forall t < T$:

$$r_{1,t}^* \equiv \operatorname{argmax}_r I_{1,t}(r) > 0 \text{ and } r > r_{1,t}^* \Rightarrow \frac{\partial I_{1,t}}{\partial r} < 0$$

(ii) The upward sloping portion of the $I_{1,t}(r)$ curve becomes smaller at t rises:

$$r_{1,t}^* > r_{1,t+1}^*$$

The first part of the proposition is driven by the same two effects that make the response of investment demand in period 1 to a change in r non-monotonic. For any $t < T$, a decrease in r creates two opposing forces: first, the cost of capital falls, compelling more entrepreneurs to invest at any given time; second, the cost of waiting falls, encouraging investors to wait until they are more certain that they will make money. The waiting effect reduces total investment between periods 1 and t by causing investors to postpone investment beyond t . As explained above, the waiting effect dominates when the interest rate is very low and the cost of capital effect dominates when the interest rate is higher, resulting in a backward bending curve. One

²⁰See Lemma 1 in the appendix for details of the T period model.

way of seeing why the backward-bending result emerges here as well is to observe that if $r = 0$, all the firms will postpone their decision until T and $I_{1,t}(r = 0) = 0$; but if $r > 0$, some firms will find it optimal to invest.

The second part of the proposition arises from the fact that the learning effect has a smaller impact on the interest elasticity of investment demand over longer horizons. The growth in profits from delay diminishes over time because the marginal return to information falls as more knowledge is accumulated. When r is lowered, less confident investors may delay investment for a few periods to acquire information. But some of these firms will eventually decide to invest, since there is still a non-zero cost to delay and the potential benefits for further information have fallen. A reduction in r thus leads to temporary delays via the learning effect, making investment fall more in the short run relative to the long run. The conventional cost of capital effect therefore starts to dominate at lower levels of r in the long run, and the investment-maximizing $r_{1,t}^*$ falls.²¹

The important implication of this result is that the long run elasticity of investment demand is more negative than the short run elasticity of investment demand when firms learn over time. If the near-zero existing estimates of the short-run interest elasticity of investment demand are due to learning effects, interest rate reductions from policies that stimulate savings could nonetheless increase investment over a longer horizon.²²

5.3 Average Profit Rates

In the neoclassical model, a higher interest rate increases the average rate of return of investments that are undertaken by driving out low-NPV ventures. This result also breaks down when firms learn over time, providing another ancillary test of the learning model.

²¹Unlike learning effects, which die away in the long run as firms acquire perfect information, composition effects may never subside. Hence, when composition effects are permitted, the long-run investment demand curve may continue to exhibit a substantial upward-sloping segment.

²²Interestingly, Caballero (1994) and Caballero, Engel, and Haltiwanger (1995) find that a higher cost of capital reduces investment much more in the long run than the short run in the U.S.

To analyze the average rate of return, we must identify the level of ex-post profitable ($\mu=1$) and ex-post unprofitable ($\mu=0$) investment by specifying how frequently a project that a manager expects to succeed with probability λ_0 actually does succeed. A natural benchmark is rational expectations: $P_{\lambda_0}[\mu = 1] = \lambda_0$. In this case, the average (net) profit rate among investments that are undertaken in period 1 is given by:

$$\rho(r) = \frac{\int_{\lambda_0^*}^1 \lambda_0 (R_1 - C) + (1 - \lambda_0) (R_0 - C) d\eta(\lambda_0)}{\int_{\lambda_0^*}^1 C d\eta(\lambda_0)} \quad (12)$$

Proposition 6 The average profit rate ρ is a backward-bending function of the interest rate:

$$r < r^* \Rightarrow \frac{\partial \rho}{\partial r} < 0$$

As established in Proposition 1, when $r < r^*$, an increase in the interest rate draws the marginal investor with prior $\lambda_0^*(r)$ into the period 1 pool of investors. This firm has the lowest probability of success among the set of firms who are investing. Consequently, it pulls down the average rate of return in the overall pool. Conversely, when $r > r^*$, an increase in r eliminates the marginal investor with prior $\lambda_0^*(r)$, who has the lowest probability of success in the pool of investors, increasing the average rate of return.

The average observed profit rate on current investment is thus a backward-bending function of r . Building on earlier results, an increase in the interest rate is more likely to lower the average observed rate of return when the potential to learn is greater and in the short run relative to the long run.

5.4 Temporary Interest Rate Changes

The results above relate to permanent changes in the interest rate. Variation in the "permanent" interest rate can only be obtained cross-sectionally, e.g. by comparing across

countries, or using low-frequency variation in the time series, as in Jovanovic and Rousseau (2004). In this section, I discuss a few comparative statics for temporary changes in the interest rate that may be easier to test. Proofs (available by request) are omitted since these results are simple extensions of the preceding propositions.

First consider the effect of a temporary *unanticipated* temporary increase in the interest rate. Let r_{1t} denote the per-period interest rate between periods 1 and $t < T$. An unanticipated increase in r_{1t} (holding fixed r in all other periods) is more likely to reduce current investment than a permanent increase in r because one can take advantage of lower future costs of capital by delaying investment. If the potential to learn is sufficiently high, $I(r_{1t})$ is backward-bending, with a smaller upward-sloping segment than $I(r)$. When the potential to learn is low, $I(r_{1t})$ is strictly downward-sloping. The longer the duration of an interest rate change, the less the incentive to postpone investment following a temporary increase, and the larger the range of parameters over which $I(r_{1t})$ is upward-sloping.

Now consider the effect of a temporary *anticipated* change in the interest rate that begins in period $s > 1$ and lasts until period $t > s$. An anticipated increase in r_{st} is more likely to raise current investment than a permanent change in r because one can take advantage of lower current costs of capital by investing immediately. In fact, if the change is anticipated sufficiently far in advance, current investment may be a strictly upward-sloping function of r_{st} . In contrast, future investment falls when r_{st} rises because of the intertemporal substitution.

Together, these results indicate that the shape of the yield curve on bonds should have a significant effect on current and future investment patterns. The yield curve embodies investors' expectations of current and future interest rates. When the yield curve becomes steeper, current investment should rise relative to subsequent investment. Hence, variations in the shape of the yield curve -- especially when coupled with cross-industry variation in learning potential -- can provide powerful tests of model beyond the basic backward-bending $I(r)$ result.

A final set of predictions relates to the effect of tax changes. In the neoclassical model, taxes matter only through the user cost of capital. In the present model, both the user cost and

the discount rate matter, and tax policies may affect these two quantities differently. For example, accelerated depreciation provisions change the user cost but need not change the discount rate (there is no additional incentive to delay from accelerated depreciation itself). Hence, they should unambiguously raise investment.²³ In contrast, changes in the tax treatment of capital income can affect equilibrium interest rates, thereby changing the discount rate and user cost simultaneously. Hence, these changes could generate non-monotonic investment responses. A more thorough examination of the relationship between tax policies and investment behavior is left to future work.

6 Conclusion

One of the central questions for tax and monetary policy makers is, "How will a policy change that increases interest rates affect real investment?" This paper has explored this question in an environment where firms making irreversible investments learn over time. In this setting, real investment is a backward-bending function of the interest rate. At low interest rates, an increase in r raises the cost of learning and *increases* investment by enlarging the set of projects for which the interest rate exceeds the rate of return to delay.

Empirical studies of firm behavior have found that interest rate changes cause some firms to change the timing of investments as the model predicts. However, additional tests of the model's auxiliary predictions are needed to determine whether it can help explain why empirical estimates of the short-run interest elasticity of aggregate investment demand are so low. Testing whether temporary and permanent interest rates changes affect the timing of investment in sectors with a high potential to learn may be a promising direction for future work.

²³This point is empirically relevant because the clearest evidence that the cost of capital affects investment behavior comes from tax reforms (e.g. Cummins, Hassett, and Hubbard (1994)). Most of the reforms are changes in depreciation allowances and investment tax credits, which arguably do not change the cost of delay.

References

- Alchian, Armen, "Costs and Outputs," in *The Allocation of Economic Resources*, edited by Moses Abramovitz (1959).
- Arnott, R. and F. Lewis, 1979, "The Transition of Land to Urban Use," *Journal of Political Economy* 87, 161-169.
- Bertola, G. and R.J. Caballero, 1994. "Irreversibility and Aggregate Investment." *Review of Economic Studies* 61(2): 223-246.
- Blanchard, O., 1986. "Investment, output, and the cost of capital: a comment", *Brookings Papers on Economic Activity* 1986(1): 153-158.
- Bulan, L., C. Mayer, and T. Somerville., 2004. "Irreversible Investment, Real Options, and Competition: Evidence From Real Estate Development," Columbia Univ. mimeo.
- Caballero, R.J., 1994. "Small sample bias and adjustment costs", *Review of Economics and Statistics* 76(1): 52-58.
- Caballero, R.J., E. Engel and J. Haltiwanger, 1995. "Plant-level adjustment and aggregate investment dynamics", *Brookings Papers on Economic Activity* 1995(2): 1-54.
- Caballero, R. J., 1999. "Aggregate investment" in *Handbook of Macroeconomics*, Vol. 1B, eds. Taylor, J. P. and Woodford, M., 813–862, Elsevier.
- Capozza, D. and Y. Li, 2001 "Residential Investment and Interest Rates: An Empirical Test of Development as a Real Options," *Real Estate Economics* 29:3.
- Chetty, R., 2000. "Interest Rates, Learning, and Business Investment Under Uncertainty" UC-Berkeley mimeo.
- Chirinko, R., 1993a. "Business Fixed Investment Spending: A Critical Survey of Modelling Strategies, Empirical Results, and Policy Implications." *Journal of Economic Literature* 31: 1875-1911.
- Chirinko, R., 1993b. *Econometric Models and Empirical Findings for Business Investment*. New York: Basil Blackwell.
- Coy, P. "Exploiting Uncertainty," *Business Week*, June 7, 1999
- Cukierman, A., 1980. "The Effects of Uncertainty on Investment Under Risk Neutrality with Endogenous Information." *Journal of Political Economy* 88(3): 462-275.

Cummins, J., K. Hassett, and G. Hubbard, 1994, "A Reconsideration of Investment Behavior Using Tax Reforms as Natural Experiments," *Brookings Papers on Economic Activity*, Vol. 2, 1-74.

Demers, M., 1991. "Investment under uncertainty, irreversibility and the arrival of information over time." *Review of Economic Studies* 58, 333–350.

Dixit, A. and Pindyck, R. 1994. *Investment Under Uncertainty*. Princeton: Princeton Univ. Press.

Doms, M., and T. Dunne, 1993. "An investigation into capital and labor adjustment at the plant level." Center for Economic Studies, Census Bureau mimeograph.

Eisner, R. and Nadiri, R. I. 1968. "Investment behavior and neoclassical theory." *Review of Economics and Statistics* 50:369-82.

Feldstein, M.S. and Flemming, J.S. 1971. "Tax policy, corporate saving, and investment behaviour in Britain." *Review of Economic Studies* 38:415-434

Goolsbee, A. and D. Gross, 1997. "Estimating Adjustment Costs with Data on Heterogeneous Capital Goods." NBER Working Paper 6342.

Haavelmo, T., 1960 *A Study in the Theory of Investment*, Chicago: Chicago University Press.

Hassett, K. and G. Hubbard, 2002, "Tax Policy and Business Investment," in *Handbook of Public Economics* Vol. 3, edited by A. Auerbach and M. Feldstein. Elsevier.

Hellman, T. and M. Puri, "The interaction between product market and financing strategy: The role of venture capital," *Review of Financial Studies*, 2000, 13(4), 959-984.

Jorgenson, D. 1963. "Capital Theory and Investment Behavior." *American Economic Review* 53(2), 247-259.

Jovanovic, B. and P. Rousseau, 2001. "Why Wait? A Century of Life before IPO." *American Economic Review Papers and Proceedings* 91, no. 2, May 2001, pp 336-341.

Jovanovic, B. and P. Rousseau, 2004. "Interest Rates and Initial Public Offerings." NBER Working Paper 10298.

Pindyck R.S. and A. Solimano. *Economic Instability and Aggregate Investment*. NBER Macroeconomics Annual. 8, 1993, 259-303.

Poterba, James M. and Lawrence H. Summers, "A CEO Survey of U.S. Companies' Time Horizons and Hurdle Rates," *Sloan Management Review*, Fall 1995, Vol. 37 No. 1, pp. 43–53.

Marglin, S., 1967. *Public investment criteria; benefit-cost analysis for planned economic growth*. Cambridge: MIT Press.

Marglin, S., 1970. "Investment and Interest: A Reformulation and Extension of Keynesian Theory" *The Economic Journal* 80(320), 910-931.

McDonald, R., 2000. "Real Options and Rules of Thumb in Capital Budgeting" in M. J. Brennan and L. Trigeorgis (eds.), *Project Flexibility, Agency, and Competition*. London: Oxford University Press.

McDonald, R. and D. Siegel, 1986. "The Value of Waiting to Invest." *Quarterly Journal of Economics* 101(4), 707-728.

Moel, A. and P. Tufano, 2002. "When are real options exercised? An empirical study of mine closings," *Review of Financial Studies* 15, 35-64.

Pindyck, R., 1988. "Irreversible Investment, Capacity Choice, and the Value of the Firm," *American Economic Review* 78 (5) 969-985.

Pindyck, R. 1991. "Irreversibility, Uncertainty, and Investment." *Journal of Economic Literature* XXIX:1110-1148

Summers, L. H., 1987. "Investment Incentives and the Discounting of Depreciation Allowances," in M. Feldstein (ed.), *The Effects of Taxation on Capital Accumulation*. Chicago: University of Chicago Press.

Stiglitz, J. and A. Weiss, 1981, "Credit Rationing in Markets with Imperfect Information", *American Economic Review*.

Appendix A: Proofs for Basic Model

All proofs below apply for the general model with arbitrary decision horizon T . The corresponding results discussed in the text are for the $T = 2$ case, unless otherwise noted. Proofs for section 3 (Extensions) are given in a separate appendix. Note that λ_{t-1} denotes the firm's prior in period t .

Lemma 1 In period T , the firm invests iff $z_{T-1} > z_{T-1}^*$ where z_{T-1}^* satisfies

$$\frac{g(z_T^*)}{f(z_T^*)} = \frac{1 - \lambda_{T-1} - \pi_T(0)}{\lambda_{T-1} \pi_T(1)}$$

In any period $t < T$, the firm invests iff $V_t(i) > V_t(l)$.

$$V_t(i) = \lambda_{t-1} \pi_t(1) + (1 - \lambda_{t-1}) \pi_t(0)$$

$$V_t(l) = \sum_{s=t+1}^T \lambda_{t-1} P_t(I_s | \mu = 1) \pi_s(1) + (1 - \lambda_{t-1}) P_t(I_s | \mu = 0) \pi_s(0)$$

$$\text{where } \frac{\lambda_t}{1 - \lambda_t} = \frac{\lambda_0}{1 - \lambda_0} \frac{g(z_1, \dots, z_t)}{f(z_1, \dots, z_t)} = \frac{\lambda_0}{1 - \lambda_0} \frac{g(z_1) \cdots g(z_t)}{f(z_1) \cdots f(z_t)}$$

$\forall t > 0$: $z_t^*(z_1, \dots, z_{t-1})$ is uniquely defined by $V_t(i, z_t^*, \lambda_{t-1}) = V_t(l, z_t^*, \lambda_{t-1})$,

$$P_t(I_{t+1} | \mu = 1) = \int_{z_t^*}^{\infty} g(z_t) dz_t, P_t(I_{t+1} | \mu = 0) = \int_{z_t^*}^{\infty} f(z_t) dz_t$$

$\forall s \in \{t + 2, \dots, T\}$:

$$P_t(I_s | \mu = 1) = \int_{-\infty}^{z_t^*} \int_{-\infty}^{z_{t+1}^*} \cdots \int_{-\infty}^{z_{s-2}^*} \int_{z_{s-1}^*}^{\infty} g(z_{s-1}) dz g(z_{s-2}) dz \cdots g(z_{t+1}) dz g(z_t) dz$$

$$P_t(I_s | \mu = 0) = \int_{-\infty}^{z_t^*} \int_{-\infty}^{z_{t+1}^*} \cdots \int_{-\infty}^{z_{s-2}^*} \int_{z_{s-1}^*}^{\infty} f(z_{s-1}) dz f(z_{s-2}) dz \cdots f(z_{t+1}) dz f(z_t) dz$$

Proof

First, $\frac{\lambda_t}{1 - \lambda_t} = \frac{\lambda_0}{1 - \lambda_0} \frac{g(z_1, \dots, z_t)}{f(z_1, \dots, z_t)} = \frac{\lambda_0}{1 - \lambda_0} \frac{g(z_1) \cdots g(z_t)}{f(z_1) \cdots f(z_t)}$ follows from Bayes rule and $z_t \perp z_s$ for $t \neq s$.

We begin by analyzing behavior in period T . In period T , the payoff to investing $V_T(i)$ is computed using the updated belief $\lambda_T(z)$ about the probability with which $\mu = 1$ occurs. The firm invests iff $V_T(i) > 0$, where

$$V_T(i) = (1 - \lambda_1(z)) \pi_2(0) + \lambda_1(z) \pi_2(1)$$

So the firm invests iff $\frac{\lambda_T(z)}{1 - \lambda_T(z)} > \frac{-\pi_T(0)}{\pi_T(1)}$

By Bayesian updating, if a signal z is observed at the end of period $T - 1$,

$$\frac{\lambda_T(z)}{1-\lambda_T(z)} = \frac{\lambda_{T-1}}{1-\lambda_{T-1}} \times \frac{g(z)}{f(z)} \Rightarrow \lambda_T(z) = \frac{\lambda_{T-1}g(z)}{\lambda_{T-1}g(z)+(1-\lambda_{T-1})f(z)}$$

Since $\frac{g(z)}{f(z)}$ is monotonically increasing, the period T decision rule is:

$$\text{Invest if } z_{T-1} > z_{T-1}^*, \text{ where } z_{T-1}^* \text{ is defined by } \frac{g(z_{T-1}^*)}{f(z_{T-1}^*)} = \frac{1-\lambda_{T_{D-1}}-\pi_T(0)}{\lambda_{T_{D-1}}\pi_T(1)}$$

The remainder of the proof is done by backward induction starting with period $T - 1$, where the investor is faced with a 2 period decision problem. We first show that the V_{T-1} functions have the form claimed above. In period $T - 1$, there are 2 possible actions: i and l . $V_{T-1}(i)$ is computed by taking an expectation over the π_{T-1} function:

$$V_{T-1}(i) = \lambda_0\pi_{T-1}(1) + (1 - \lambda_0)\pi_{T-1}(0)$$

To compute $V_{T-1}(l)$, integrate the expected payoff in period T over the prior density of z . The payoff in period T depends upon the action taken in period T , which follows the decision rule derived above:

$$V_{T-1}(l) = \int_{-\infty}^{\infty} \text{Max}(V_T(d_1), V_T(i)) dm(z) = \int_{-\infty}^{z^*} V_T(d_1) + \int_{z^*}^{\infty} V_T(i) dm(z)$$

where $dm(z) = \lambda_0g(z) + (1 - \lambda_0)f(z)$ is the unconditional density on z .

$$\Rightarrow V_{T-1}(l) = \lambda_0\pi_T(1)\int_{z_{T_{D-1}}^*}^{\infty} g(z)dz + (1 - \lambda_0)\pi_T(0)\int_{z_{T_{D-1}}^*}^{\infty} f(z)dz$$

We also need to show that the firm follows a threshold rule for investment in period $T - 1$:

\exists unique $z_{T-2}^*(z_1, \dots, z_{T-3})$ defined by $V_{T-1}(i, z_{T-2}^*, \lambda_{T-3}) = V_{T-1}(l, z_{T-2}^*, \lambda_{T-3})$ s.t. that investing is optimal iff $z_{T-2} > z_{T-2}^*$.

It is sufficient to show that \exists unique λ_{T-2}^* s.t. $V_{T-1}(i, \lambda_{T-2}^*) = V_{T-1}(l, \lambda_{T-2}^*)$ and that $\lambda_{T-2} > \lambda_{T-2}^*$ makes investing optimal because by monotonicity of the likelihood ratio, this implies that there is a unique z_{T-2}^* that satisfies the given expression, conditional on λ_{T-3} .

To see that there is a unique λ_{T-2} : rewrite

$$V_{T-1}(i) = \lambda_{T-2}b + (1 - \lambda_{T-2})a$$

$$V_{T-1}(l) = \lambda_{T-2}b' + (1 - \lambda_{T-2})a'$$

where $b = \pi_{T-1}(1)$ $b' = P_{T-1}(I_T|\mu = 1)\pi_T(1)$

$$a = \pi_{T-1}(0) \quad a' = P_{T-1}(I_T|\mu = 0)\pi_T(0)$$

Now, $V_{T-1}(l, \lambda_{T-1} = 0) = 0 > V_{T-1}(i, \lambda_{T-1} = 0)$

By assumption, $\exists \lambda_{T-1}$ s.t. $V_{T-1}(l, \lambda_{T-1}) \leq V_{T-1}(i, \lambda_{T-1})$

By the Intermediate Value Theorem (IVT), $\exists \lambda_{T-2}^*$ s.t. $V_{T-1}(i, \lambda_{T-2}^*) = V_{T-1}(l, \lambda_{T-2}^*)$

$$\frac{\partial V_{T-1}(l)}{\partial \lambda_{T-2}} > 0 \Rightarrow V_{T-1}(l, \lambda_{T-2}^*) > 0$$

$$\frac{\partial \{V_{T-1}(i) - V_{T-1}(l)\}}{\partial \lambda_{T-2}} \Big|_{\lambda_{T-2}^*} = b - a + a' - b > 0$$

because $V_s(i, \lambda_{T-2}^*) = V_s(l, \lambda_{T-2}^*) \Leftrightarrow$

$$\lambda_{T-2}^* b + (1 - \lambda_{T-2}^*) a = \lambda_{T-2}^* b' + (1 - \lambda_{T-2}^*) a' > 0$$

Therefore, at any λ_{T-2}^* , we must have $\frac{\partial \{V_{T-1}(i) - V_{T-1}(l)\}}{\partial \lambda_{T-2}} \Big|_{\lambda_{T-2}^*} > 0 \Rightarrow \lambda_{T-2}^*$ is unique.

Hence $V_{T-1}(i, \lambda_{T-2}) > V_{T-1}(l, \lambda_{T-2})$ iff $\lambda_{T-2} > \lambda_{T-2}^*$.

Having characterized behavior in period $T - 1$, we proceed to the general step:

Given that $V_{t+1}(-)$ has the form claimed, we will prove that $V_t(-)$ also has the same form.

First, clearly $V_t(i) = \lambda_{t-1}\pi_t(1) + (1 - \lambda_{t-1})\pi_t(0)$.

The expected payoff to learning, $V_t(l)$, is computed by recognizing that one will maximize profits in the next period. By the inductive assumption, the firm will invest in period $t + 1$ iff $z_t > z_t^*$, where z_t^* has already been computed.

$$\begin{aligned} V_t(l) &= \int_{z_t^*}^{\infty} [\lambda_t \pi_{t+1}(1) + (1 - \lambda_t) \pi_{t+1}(0)] [\lambda_{t-1} g(z_t) + (1 - \lambda_{t-1}) f(z_t)] dz_t + \\ &\quad \int_{-\infty}^{z_t^*} V_{t+1}(l, \lambda_t) [\lambda_{t-1} g(z_t) + (1 - \lambda_{t-1}) f(z_t)] dz_t \\ &= \lambda_{t-1} \int_{z_t^*}^{\infty} g(z_t) dz_t \pi_{t+1}(1) + (1 - \lambda_{t-1}) \int_{z_t^*}^{\infty} f(z_t) dz_t \pi_{t+1}(0) + \\ &\quad \lambda_{t-1} \sum_{s=t+2}^T \pi_s(1) \int_{-\infty}^{z_t^*} P_{t+1}(I_s|\mu = 1) g(z_t) dz + \\ &\quad (1 - \lambda_{t-1}) \sum_{s=t+2}^T \pi_s(0) \int_{-\infty}^{z_t^*} P_{t+1}(I_s|\mu = 0) f(z_t) dz \\ &= \lambda_{t-1} P_t(I_{t+1}|\mu = 1) \pi_{t+1}(1) + (1 - \lambda_{t-1}) P_t(I_{t+1}|\mu = 0) \pi_{t+1}(0) + \\ &\quad \sum_{s=t+2}^T \lambda_{t-1} P_t(I_s|\mu = 1) \pi_s(1) + (1 - \lambda_{t-1}) P_t(I_s|\mu = 0) \pi_s(0) \\ \therefore V_t(l) &= \sum_{s=t+1}^T \lambda_{t-1} P_t(I_s|\mu = 1) \pi_s(1) + (1 - \lambda_{t-1}) P_t(I_s|\mu = 0) \pi_s(0) \end{aligned}$$

Finally, to complete the induction, we need to show that \exists unique z_{t-1}^* s.t. $V_{t-1}(i, z_{t-1}^*, \lambda_{t-2}) = V_s(l, z_{t-1}^*, \lambda_{t-2})$ and that if $z_{t-1} > z_{t-1}^*$ the investor will find it optimal to invest in period t . Again, it is sufficient to show that \exists unique λ_{t-1}^* s.t. $V_{T-1}(i, \lambda_{t-1}^*) = V_{T-1}(l, \lambda_{t-1}^*)$ and that $\lambda_{t-1} > \lambda_{t-1}^* \Leftrightarrow \text{Invest}$.

As the proof is virtually identical to that above, we are brief:

$$V_t(l, \lambda_{t-1} = 0) = 0 > V_t(i, \lambda_{t-1} = 0)$$

$$\frac{\partial V_t(l)}{\partial \lambda_{t-1}} > 0 \Rightarrow V_t(l, \lambda_{t-1}^*) > 0$$

Continuity + IVT $\Rightarrow \lambda_{t-1}^*$ exists

$$\frac{\partial \{V_t(i) - V_t(l)\}}{\partial \lambda_{t-1}} \Big|_{\lambda_{t-1}^*} = b - a + a' - b > 0 \text{ follows from using the condition that defines } \lambda_{t-1}^*$$

$$\begin{aligned} \text{where } b &= \pi_t(1) & b' &= \sum_{s=t+1}^T P_t(I_s | \mu = 1) \pi_s(1) \\ a &= \pi_t(0) & a' &= \sum_{s=t+1}^T P_t(I_s | \mu = 0) \pi_s(0) \end{aligned}$$

The combination of arguments above implies that λ_{t-1}^* and consequently z_{t-1}^* are unique. QED.

Lemma 2 There is a unique λ_0^* at which the value of investing equals that of postponing.

In period 1, firms with $\lambda_0 < \lambda_0^*$ postpone their investment decision.

Firms with $\lambda_0 \geq \lambda_0^*$ invest in period 1.

Proof

The proof follows directly from the second step of Lemma 1. It was shown that in any period t ,

$$\exists \text{ unique } \lambda_{t-1}^* \text{ s.t. } V_t(i, \lambda_{t-1}^*) = V_t(l, \lambda_{t-1}^*) \text{ and that } \lambda_{t-1} > \lambda_{t-1}^* \Leftrightarrow \text{Invest.}$$

Applying this result to $t = 1$ gives the result.

Lemma 4 An increase in signal noise increases current investment

$$s(f_1, g_1) > s(f_2, g_2) \Rightarrow I(f_1, g_1) > I(f_2, g_2)$$

Proof

Take any $(f_1, g_1), (f_2, g_2)$ s.t. $s(f_1, g_1) > s(f_2, g_2)$.

For the firm with $\lambda_0' = \lambda_0^*(f_1, g_1)$,

$$V_1(l) = \sum_{s=2}^T \lambda_0 P_1(I_s | \mu = 1) \pi_s(1) + (1 - \lambda_0) P_1(I_s | \mu = 0) \pi_s(0)$$

$$V(i, \lambda'_0) = \lambda'_0 \left\{ \frac{R_1}{1+r} - C \right\} + (1 - \lambda'_0) \left\{ \frac{R_0}{1+r} - C \right\}$$

The shift from (f_1, g_1) to (f_2, g_2) only affects $V(l, \lambda'_0)$. It follows from the definition of a reduction in signal noise that for any s , $P_1(I_s | \mu = 1)$ is higher under (f_2, g_2) than under (f_1, g_1) , while $P_1(I_s | \mu = 0)$ is lower. Hence,

$$V(l, \lambda'_0; f_2, g_2) > V(l, \lambda'_0; f_1, g_1) = V(i, \lambda'_0)$$

By Lemma 2, it follows that $\lambda_0^*(f_2, g_2) > \lambda'_0 = \lambda_0^*(f_1, g_1)$. As $\frac{\partial I}{\partial \lambda_0^*} < 0$, the result is established.

Proposition 1 Investment demand is a backward-bending function of the interest rate.

$$(i) I(r = 0) = 0 \text{ and } \lim_{r \rightarrow 0} \frac{\partial I}{\partial r}(r) = +\infty$$

$$(ii) r^* \equiv \operatorname{argmax}_r I(r) > 0 \text{ and } r \leq r^* \Rightarrow \frac{\partial I}{\partial r} \geq 0$$

Proof

(i) Recall

$$V(i) = \lambda_0 \pi_1(1) + (1 - \lambda_0) \pi_1(0)$$

$$V(l) = \sum_{s=2}^T \lambda_0 P_1(I_s | \mu = 1) \pi_s(1) + (1 - \lambda_0) P_1(I_s | \mu = 0) \pi_s(0)$$

$$\text{where } \pi_t(\mu) = \frac{R_\mu}{(1+r)^t} - \frac{C}{(1+r)^{t-1}}$$

Suppose $r = 0$. Then

$$V(i) = \lambda_0 \{R_1 - C\} + (1 - \lambda_0)(R_0 - C)$$

$$V_1(l, \lambda_0) = \lambda_0 \{R_1 - C\} \sum_{t=2}^T P_1(I_t | \mu = 1) + (1 - \lambda_0) \{R_0 - C\} \sum_{t=2}^T P_1(I_t | \mu = 0)$$

Now, for $\lambda_0 = 1$, the definitions in Lemma 2 imply that $z_t^* = \infty$ for $t < T$ and $z_T^* = -\infty$

$$\Rightarrow \sum_{t=2}^T P_1(I_t | \mu = 1) = 1$$

$$\lambda_0 = 1 \Rightarrow V(i, \lambda_0) = (\pi_1 - \delta) = V(l, \lambda_0) \Rightarrow \lambda_0^* = 1$$

$$\therefore I(r = 0) = \int_1^1 C(\lambda_0^i) d\eta(\lambda_0^i) = 0$$

To establish that $\lim_{r \rightarrow 0} \frac{\partial I}{\partial r}(r) = +\infty$, note first that

$$\frac{\partial I}{\partial r} = \frac{\partial I}{\partial \lambda_0^*} \times \frac{\partial \lambda_0^*}{\partial r}.$$

By the IFT, $\frac{\partial \lambda_0^*}{\partial r} = -\frac{\partial V(i, \lambda_0)/\partial r - \partial V(l, \lambda_0)/\partial r}{\partial V(i, \lambda_0)/\partial \lambda_0 - \partial V(l, \lambda_0)/\partial \lambda_0} \Big|_{\lambda_0^*}$

$$\Rightarrow \frac{\partial \lambda_0^*}{\partial r} = \frac{N(r)}{D(r)}$$

After some algebra, it can be shown that $N(r=0) = C - R_1 < 0$ and $D(r) > 0 \forall r > 0$.

Since $D(r=0) = 0$ and $\frac{\partial I}{\partial \lambda_0^*} = -\lambda_0^* d\eta(\lambda_0^*) < 0$ it follows that $\lim_{r \rightarrow 0} \frac{\partial I}{\partial r}(r) = +\infty$

(ii) Note that $C < R_1 \Rightarrow \exists r' \text{ s.t. } \frac{R_1}{1+r'} = C$

Then $r \geq r' \Rightarrow V(i, \lambda_0) < 0 \forall \lambda_0 \leq 1 \Rightarrow \lambda_0^*(r) = 1 \Rightarrow I(r) = 0$.

$I(r)$ is continuous because it is a composition of continuous functions. Since $I(0) = I(r') = 0$ and $[0, r']$ is compact, $I(r)$ has an interior maximum $r^* \in (0, r')$.

To prove uniqueness of r^* , recall $\frac{\partial I}{\partial r} = \frac{\partial I}{\partial \lambda_0^*} \times \frac{\partial \lambda_0^*}{\partial r} = \frac{\partial I}{\partial \lambda_0^*} \frac{\partial V(l, \lambda_0)/\partial r - \partial V(i, \lambda_0)/\partial r}{\partial V(i, \lambda_0)/\partial \lambda_0 - \partial V(l, \lambda_0)/\partial \lambda_0}$

I continuous and $\frac{\partial V(i, \lambda_0) - \partial V(l, \lambda_0)}{\partial \lambda_0}(\lambda_0^*) > 0 \forall r > 0 \Rightarrow N(r) \equiv \frac{\partial V(l) - \partial V(i)}{\partial r} = 0$ at any critical r .

A sufficient condition for r^* to be unique is that $N(r)$ has a unique root. We will establish this by showing that $N(r) = 0 \Rightarrow \frac{\partial N}{\partial r}(r) > 0$.

Since $\frac{\partial V(l) - \partial V(i)}{\partial r}(r^*) = 0$ and $V(l, \lambda_0^*(r^*)) = V(i, \lambda_0^*(r^*))$,

$$\frac{\partial N}{\partial r}(r^*, \lambda_0^*(r^*)) = \partial \left[\frac{\partial V(l) - \partial V(i)}{\partial r} \right] / \partial r = \frac{1}{1+r^*} \partial \left[\frac{(1+r)(\partial V(l) - \partial V(i))}{\partial r} \right] / \partial r = \frac{1}{1+r^*} \partial [M(r)] / \partial r$$

$$\text{where } M(r) = \frac{\partial((1+r)V(l) - \partial((1+r)V(i))}{\partial r}$$

Let $\beta \equiv \sum_{t=2}^T \frac{P_1(I_t | \mu=1)}{(1+r)^{t-1}}$ and $\alpha \equiv \sum_{t=2}^T \frac{P_1(I_t | \mu=0)}{(1+r)^{t-1}}$

By definition, $\frac{\partial \lambda_0^*}{\partial r}(r^*) = 0$, so we can take λ_0^* as fixed when differentiating w.r.t. r :

$$\begin{aligned} M(r) &= \partial(\lambda_0 \beta(z^*) (\frac{R_1}{1+r} - C) + (1 - \lambda_0) \alpha(z^*) \{ \frac{R_0}{1+r} - C \}) / \partial r \\ &\quad - (\lambda_0 (R_1 - C(1+r)) + (1 - \lambda_0) (R_0 - C(1+r))) / \partial r \\ &= -\lambda_0 \beta(z^*) \frac{R_1}{(1+r)^2} - (1 - \lambda_0) \alpha(z^*) \frac{R_0}{(1+r)^2} + C \end{aligned}$$

Now differentiate $M(r)$ to obtain the second derivative:

$$\partial [M(r)] / \partial r = \lambda_0 \beta(z^*) \frac{2R_1}{(1+r)^3} + (1 - \lambda_0) \alpha(z^*) \frac{2R_0}{(1+r)^3} - \frac{\partial \beta}{\partial r} \frac{R_1}{(1+r)^2} - \frac{\partial \alpha}{\partial r} \frac{R_0}{(1+r)^2}$$

To sign this expression, note that $\frac{\partial \alpha}{\partial r}(r^*) < 0$ and $\frac{\partial \beta}{\partial r}(r^*) < 0$.

This is easiest to see when $T = 2$, where

$$\begin{aligned} \frac{\partial \beta}{\partial r} &= -\frac{\partial z^*}{\partial r} g(z^*) \text{ and } \frac{\partial \alpha}{\partial r} = -\frac{\partial z^*}{\partial r} f(z^*), \\ \frac{\partial z^*}{\partial r} > 0 \text{ follows from } \frac{g(z^*)}{f(z^*)} &= \frac{1-\lambda_0}{\lambda_0} \frac{C-R_0/(1+r)}{R_1/(1+r)-C} \text{ and } \partial[\frac{g(z)}{f(z)}]/\partial z > 0 \\ &\Rightarrow \frac{\partial \alpha}{\partial r} < 0 \text{ and } \frac{\partial \beta}{\partial r} < 0. \end{aligned}$$

When $T > 2$, this last step can be established using $\frac{\partial P_1(I_t|\lambda_0)}{\partial r}(r^*) < 0 \forall t > 1$, which follows from $\frac{\partial^2\{V(l,\lambda_0)-V(i,\lambda_0)\}}{\partial \lambda_0 \partial r} < 0$ (see proof of Proposition 4 below) and the definition of $P_1(I_t|\lambda_0)$.

Since $\frac{\partial \alpha}{\partial r} < 0$ and $\frac{\partial \beta}{\partial r} < 0$, it follows that

$$\partial M(r)/\partial r(r^*) > 0 \Rightarrow \partial N(r)/\partial r(r^*) > 0.$$

QED

Proposition 4 $\exists \bar{R}_0 > 0$ s.t. if $R_0 < \bar{R}_0$, a reduction in signal noise raises r^* :

$$s(f_2, g_2) < s(f_1, g_1) \Rightarrow r_2^* > r_1^*$$

$$\text{Let } \beta \equiv \sum_{t=2}^T \frac{P_1(I_t|\mu=1)}{(1+r)^{t-1}} \text{ and } \alpha \equiv \sum_{t=2}^T \frac{P_1(I_t|\mu=0)}{(1+r)^{t-1}}$$

Take (f_1, g_1) and (f_2, g_2) s.t. $s(f_2, g_2) < s(f_1, g_1)$. Let r_1^* and r_2^* denote the investment-maximizing values of r in each case, and let $\lambda'_0 = \lambda_0^*(r = r_1^*, f_1, g_1)$ and $\lambda''_0 = \lambda_0^*(r = r_1^*, f_2, g_2)$.

By Lemma 4, $\lambda''_0 > \lambda'_0$.

It can be shown that $\frac{\partial^2\{V(l,\lambda_0)-V(i,\lambda_0)\}}{\partial \lambda_0 \partial r} =$

$$\partial\left\{\left(\frac{R_1}{(1+r)^2} - \frac{C}{1+r}\right)\beta(z^*) - \left(\frac{R_0}{(1+r)^2} - \frac{C}{1+r}\right)\alpha(z^*) - \left(\left(\frac{R_1}{1+r} - C\right) - \left(\frac{R_0}{1+r} - C\right)\right)\right\}/\partial r < 0$$

Since $\frac{\partial V(l,\lambda'_0;f_1,g_1)}{\partial r} = \frac{\partial V(i,\lambda'_0;f_1,g_1)}{\partial r}$ at $r = r_1^*$, it follows that

$$\frac{\partial V(l,\lambda''_0;f_1,g_1)}{\partial r} < \frac{\partial V(i,\lambda''_0;f_1,g_1)}{\partial r} \text{ at } r = r_1^*$$

Using logic similar to that in Lemma 4, $s(f_2, g_2) < s(f_1, g_1) \Rightarrow \beta_2 > \beta_1$ and $\alpha_2 < \alpha_1$.

$$\text{Now, } \frac{\partial V(l)}{\partial r} = -\lambda_0 \beta \left\{ \frac{2R_1}{(1+r)^3} - \frac{C}{(1+r)^2} \right\} - (1-\lambda_0) \alpha \left\{ \frac{2R_0}{(1+r)^3} - \frac{C}{(1+r)^2} \right\}$$

If $R_0 < C/2$, both terms become more negative when β rises and α falls; hence

$$\frac{\partial V(l, \lambda_0''; f_2, g_2)}{\partial r} < \frac{\partial V(l, \lambda_0''; f_1, g_1)}{\partial r} < \frac{\partial V(i, \lambda_0''; f_1, g_1)}{\partial r} = \frac{\partial V(i, \lambda_0''; f_2, g_2)}{\partial r}$$

$$\Rightarrow \frac{\partial I}{\partial r}(r = r_2^*; f_1, g_1) = \frac{\partial I}{\partial \lambda_0^*} \times \frac{\partial V(l, \lambda_0)/\partial r - \partial V(i, \lambda_0)/\partial r}{\partial V(i, \lambda_0)/\partial \lambda_0 - \partial V(l, \lambda_0)/\partial \lambda_0} \Big|_{\lambda_0''} > 0$$

$\therefore R_0 < C/2 \Rightarrow r_1^* > r_2^*$ by Proposition 2.

Proposition 5 Let $I_{1,t}(r)$ denote total investment from periods 1 to t

(i) $I_{1,t}(r)$, is a backward bending function of r :

$$r_{1,t}^* \equiv \operatorname{argmax}_r I_{1,t}(r) > 0 \text{ and } r > r_{1,t}^* \Rightarrow \frac{\partial I_{1,t}}{\partial r} < 0$$

(ii) The upward sloping portion of the $I_{1,t}(r)$ curve becomes smaller at t rises:

$$r_{1,t}^* > r_{1,t+1}^*$$

Proof

(i) For any $t < T$, $r = 0 \Rightarrow \lambda_0^* = 1$ and $z_1^* = z_2^* = \dots = z_t^* = \infty$ by the equations in Lemma 1.

So $I_{1,t}(r = 0) = 0$. As above, $I_{1,t}(r) = 0$ for $r > r' > 0$ for r' s.t. $\frac{R_1}{1+r'} = C$

By the Intermediate Value Theorem, $\exists r_{1,t}^*$ that is a critical point and global maximum for $I_{1,t}$.

Uniqueness of $r_{1,t}^*$ follows from an argument equivalent to that given in Proposition 1 to show

that $r < r_{1,t}^* \Leftrightarrow \frac{\partial I_{1,t}}{\partial r} > 0$.

(ii) The claim is established by showing that for any r , $\frac{\partial I_{1,t}}{\partial r}(r) < 0 \Rightarrow \frac{\partial I_{1,t+1}}{\partial r}(r) < 0$. Hence, if

$I_{1,t}(r)$ is downward-sloping, $I_{1,t+1}(r)$ must also be downward-sloping at that r . Given that $r_{1,t}^*$ is

unique, it will follow immediately that $r_{1,t}^* > r_{1,t+1}^*$. We now calculate $\frac{\partial I_{1,t}}{\partial r}$:

$$I_{1,t}(r) = \sum_{s=1}^t I_s(r) = \int_{\lambda_0^*}^1 C d\eta(\lambda_0) + \sum_{s=2}^t \int_0^{\lambda_0^*} P_1(I_s | \lambda_0) C d\eta(\lambda_0)$$

$$\frac{\partial I_{1,t}}{\partial r} = -\frac{\partial \lambda_0^*}{\partial r} C d\eta(\lambda_0^*) \left\{ 1 - \sum_{s=2}^t \frac{1}{(1+r)^{s-1}} P_1(I_s | \lambda_0^*) \right\} +$$

$$C \int_0^{\lambda_0^*} \left[\sum_{s=2}^t \frac{\partial P_1(I_s | \lambda_0)}{\partial r} \right] d\eta(\lambda_0)$$

Now, suppose $r > r^*$, the investment-maximizing interest rate for period 1. Then $\frac{\partial \lambda_0^*}{\partial r} > 0$

implies the first term is negative in the expression above. To evaluate the sign of the second

term, note that $\frac{\partial P_1(I_t|\lambda_0)}{\partial r} < 0 \forall \lambda_0 < \lambda_0^*$ when $r > r^*$. This follows from $\frac{\partial^2\{V(i,\lambda_0)-V(i,\lambda_0)\}}{\partial \lambda_0 \partial r} < 0$

and the definition of $P_1(I_t|\lambda_0)$.

Since $P_1(I_t|\lambda_0) > 0 \forall t > 1$ and $\frac{\partial P_1(I_t|\lambda_0)}{\partial r} < 0 \forall t > 1$, it follows by inspection that

$$\frac{\partial I_{1,t}}{\partial r}(r) < 0 \Rightarrow \frac{\partial I_{1,t+1}}{\partial r}(r) < 0. \text{ QED.}$$

Proposition 6 The average rate of return ρ is a backward-bending function of the interest rate:

$$r < r^* \Rightarrow \frac{\partial \rho}{\partial r} < 0$$

Proof

$$\rho(r) = \int_{\lambda_0^*}^1 \lambda_0 (R_1 - C) + (1 - \lambda_0)(R_0 - C) d\eta(\lambda_0) / \int_{\lambda_0^*}^1 C d\eta(\lambda_0)$$

$$\partial \rho / \partial r = \left(- \frac{\partial \lambda_0^*}{\partial r} \right) \frac{C d\eta(\lambda_0^*)}{[\int_{\lambda_0^*}^1 C d\eta(\lambda_0)]^2} A \text{ where}$$

$$A = \int_{\lambda_0^*}^1 [\lambda_0^* (R_1 - C) + (1 - \lambda_0^*)(R_0 - C)] - [\lambda_0 (R_1 - C) + (1 - \lambda_0)(R_0 - C)] d\eta(\lambda_0)$$

Since $\int_{\lambda_0^*}^1 (\lambda_0^* - \lambda_0) d\eta(\lambda_0) < 0$, it follows that $A < 0$

$$\text{Hence } r < r^* \Rightarrow \frac{\partial \lambda_0^*}{\partial r} < 0 \Rightarrow \frac{\partial \rho}{\partial r} < 0$$

Appendix B: Proofs for Extensions (Section 3)

Proposition 2 (Scale Choice)

Let $\bar{p} = \lambda_0 p_1 + (1 - \lambda_0) p_0$ denote the expected price ex-ante. Let us first characterize the firm's optimal investment policy for a given interest rate r . In the second period, conditional on observing z , the firm updates its beliefs to $\lambda_1(z)$. It then chooses I_2 to

$$\max\{\lambda_1(z)p_1 + (1 - \lambda_1(z))p_0\} \left\{ \sum_{t=1}^{T_P+1} \frac{f(I_1 + I_2) - f(I_1)}{(1+r)^{t+1}} \right\} - \frac{I_2}{1+r}$$

At an interior optimum, I_2 must satisfy the following first order condition:

$$f'(I_1 + I_2) \{\lambda_1(z)p_1 + (1 - \lambda_1(z))p_0\} \sum_{t=1}^{T_P+1} \frac{1}{(1+r)^t} = 1$$

If there is no positive $I_2 > 0$ that satisfies this condition, the concavity of f guarantees that the optimal $I_2 = 0$. Let $I^*(z) = I_1 + I_2^*(z)$ denote the optimal total level of investment conditional on observing a signal z . Then we can write $I_2^*(z, I_1) = \max(I^*(z) - I_1, 0)$.

It is easy to show that the optimal investment level $I_2^*(z)$ is monotonically increasing in z . Hence, there is an optimal cutoff value z^* such that $z < z^*$ implies $I_1 > I^*(z)$. Let $\sigma(I_1, z^*)$ denote the ex-ante probability that $z > z^*$. Note that the rational expectations condition implies

$$\int \lambda_1(z)p_1 + (1 - \lambda_1(z))p_0 dm(z) = \bar{p}$$

where $dm(z)$ denotes the marginal distribution of z as in the text. With this notation, the value function in period 2 is given by

$$V_2(I_1) = \bar{p} \int_{z=-\infty}^{\infty} \left\{ \sum_{t=1}^{T_P+1} \frac{f(I^*(z)) - f(I_1)}{(1+r)^{t+1}} \right\} dm(z) - \int_{z>z^*}^{\infty} \frac{I^*(z)}{1+r} dm(z) - \frac{\sigma I_1}{1+r}$$

Now consider the firm's value function in period 1:

$$\begin{aligned}
V_1 &= \bar{p} \left(\sum_{t=0}^{T_P-1} \frac{f(I_1)}{(1+r)^{t+1}} \right) - I_1 + V_2(I_1) \\
&= \bar{p} \left(\sum_{t=0}^{T_P-1} \frac{f(I_1)}{(1+r)^{t+1}} - \frac{f(I_1)}{(1+r)^{t+2}} \right) - \left(1 - \frac{\sigma}{1+r}\right) I_1 \\
&\quad + \bar{p} \int_{z=-\infty}^{\infty} \left[\sum_{t=1}^{T_P+1} \frac{f(I^*(z))}{(1+r)^{t+1}} \right] dm(z) - \int_{z>z^*}^{\infty} \frac{I^*(z)}{1+r} dm(z)
\end{aligned}$$

Claim (i): $I_1(r=0) = 0$

Using the definition of V_1 , it follows that at $r=0$,

$$V_1(r=0) = -(1-\sigma)I_1 + K$$

where K is a term that does not depend on I_1 . Since $\sigma < 1$ if $\lambda_0 < 1$, it follows that $I_1^*(r=0) = 0$ as claimed.

Claim (ii): $\lim_{r \rightarrow 0} \partial I_1 / \partial r = +\infty$

Since the value function is already optimized over $I^*(z)$ and $I_2^*(z)$ at all z , envelope conditions imply that at an interior optimum, I_1 must satisfy:

$$\frac{\partial V_1}{\partial r} = \bar{p} f'(I_1) \left(\sum_{t=0}^{T_P-1} \frac{1}{(1+r)^{t+1}} - \frac{1}{(1+r)^{t+2}} \right) - \left(1 - \frac{\sigma}{1+r}\right) = 0$$

which implies

$$\bar{p} f'(I_1) = \frac{1+r-\sigma}{1-(1+r)^{-T_P}}$$

Implicit differentiation of this condition yields

$$\frac{\partial I_1}{\partial r} = \frac{1}{\bar{p} f''(I_1)} \frac{1}{1-(1+r)^{-T_P}} \left[1 - \frac{\partial \sigma}{\partial r} - \frac{T_P(1+r-\sigma)}{(1+r)^{T_P} - (1+r)} \right]$$

Note that $\frac{\partial \sigma}{\partial r}$ remains finite as $r \rightarrow 0$ because $\frac{\partial z^*}{\partial r}$ is bounded. Hence

$$\lim_{r \rightarrow 0} \left[1 - \frac{\partial \sigma}{\partial r} - \frac{T_P(1+r-\sigma)}{(1+r)^{T_P} - (1+r)} \right] = -\infty$$

Finally, since $f'' < 0$ it follows that

$$\lim_{r \rightarrow 0} \partial I_1 / \partial r = +\infty$$

These two claims establish that $I_1(r)$ has an upward sloping segment at low r and a downward-sloping segment at high r . The uniqueness of the investment-maximizing interest rate r^* can be established along the lines of part (ii) of Proposition 1, completing the proof.

Lemma 3 (Competition)

In period 2, a firm invests if the expected return to investment at the market-clearing price exceeds the cost of investment:

$$\frac{\lambda_1 p_2}{1+r} > C$$

Assume that $\frac{p_2(I_2=0)}{1+r} > C$ and $\frac{p_2(I_2=1)}{1+r} < C$ to make the problem non-trivial. Since $\frac{\partial p_2}{\partial I_2^c} < 0$, it follows that there is a unique cutoff value λ_1^* and a corresponding price $p_2(I_2^c(\lambda_1^*))$ such that

$$\frac{\lambda_1^* p_2(I_2^c(\lambda_1^*))}{1+r} = C$$

In equilibrium, all firms with $\lambda_1 > \lambda_1^*$ invest, all other firms stay out, and the market clears at the resulting price p_2 by construction. This implies that for any firm with given prior λ_0 , there is a cutoff value $z^*(\lambda_0)$ such that the firm invests in period 1 iff $z > z^*$.

Now consider period 1 investment behavior. Let $\beta(z^*(\lambda_0)) = \int_{z^*}^{\infty} g(z) dz$ and $\alpha(z^*(\lambda_0)) = \int_{z^*}^{\infty} f(z) dz$ denote the unconditional probability of investment in the good and bad states,

respectively, for a firm with prior λ_0 . Firms invest in period 1 if

$$\begin{aligned} V(i, \lambda_0) &= \frac{\lambda_0 p_1}{1+r} - C > \\ V(l, \lambda_0) &= \lambda_0 \beta(z^*(\lambda_0)) \left[\frac{p_2}{(1+r)^2} - \frac{C}{1+r} \right] + (1-\lambda_0) \alpha(z^*(\lambda_0)) \left[-\frac{C}{1+r} \right] \end{aligned}$$

Note that $V(l, \lambda_0)$ is strictly positive in equilibrium if $\lambda_0 > 0$ because there is a non-zero probability that the firm will have a posterior success probability greater than λ_1^* . It follows from Lemma 2 that for a given price vector (p_1, p_2) , there is a unique λ_0^* such that firms with $\lambda_0 > \lambda_0^*$ invest in period 1 and the remainder delay. Since $\partial p_t / \partial I_t^c < 0$, $\partial \lambda_0^* / \partial p_1 < 0$, and $\partial \lambda_0^* / \partial p_2 > 0$ it follows that there is a unique price vector (p_1, p_2) at which all firms are optimizing and markets clear in both periods. Thus, equilibrium investment is characterized by a price vector (p_1, p_2) and a cutoff value λ_0^* such that

$$V(i, \lambda_0^*, p_1) = V(l, \lambda_0^*, p_1, p_2) > 0$$

Proposition 3 (Competition)

Observe that

$$\frac{\partial I_1}{\partial r} = \frac{dI_1}{d\lambda_0^*} \frac{d\lambda_0^*}{dr}$$

The implicit function theorem implies

$$\frac{d\lambda_0^*}{dr} = -\frac{dV(i, \lambda_0)/dr - dV(l, \lambda_0)/dr}{dV(i, \lambda_0)/d\lambda_0 - dV(l, \lambda_0)/d\lambda_0} = -\frac{N}{D}$$

At $r = 0$, the denominator of this expression is:

$$D = p_1 - \beta(\lambda_0^*)p_2 + (\beta(\lambda_0^*) - \alpha(\lambda_0^*))C$$

Since $\beta > \alpha$ and $p_1 > p_2$, it follows that $D > 0$.

The numerator is

$$N = -\frac{\lambda_0 p_1}{(1+r)^2} + \frac{\lambda_0 dp_1/dr}{1+r} - \left\{ \lambda_0 \beta \left(-2 \frac{p_2}{(1+r)^3} + \frac{dp_2/dr}{(1+r)^2} + \frac{C}{(1+r)^2} \right) + (1-\lambda_0) \alpha \left(\frac{C}{(1+r)^2} \right) \right\}$$

Using the fact that $V(i, \lambda_0^*) = V(l, \lambda_0^*)$, N simplifies at $r = 0$ to

$$N(r=0)|_{\text{p-fixed}} = \lambda_0^*(r=0) \beta(\lambda_0^*) p_2 - C$$

if dp_t/dr is zero. I now show that $N(r=0)|_{\text{p-fixed}} \implies \frac{\partial I}{\partial r}(r=0) > 0$ using a proof by contradiction. Suppose that $\frac{\partial I_1}{\partial r}(r=0) < 0$. In this case,

$$dp_1/dr > dp_2/dr > 0$$

because $\partial I_1^c / \partial I_1 = 1$ and $\partial I_2^c / \partial I_1 < 1$ and $\partial p_t / \partial I_t^c < 0$. It follows that $N(r=0) > 0$ when p is variable. Since $N > 0$ and $D > 0$, $\frac{d\lambda_0^*}{dr} < 0$, which implies $\frac{\partial I_1}{\partial r}(r=0) > 0$. But this contradicts the supposition. Therefore $\frac{\partial I_1}{\partial r}(r=0) > 0$ if $N(r=0)|_{\text{p-fixed}} > 0$.

Since $\lambda_1^* p_2 = C(1+r)$ in period 2 equilibrium, $N(r=0)|_{\text{p-fixed}} > 0$ if $\frac{\lambda_0^* \beta(\lambda_0^*)}{\lambda_1^*} > 1$. Hence

$$\frac{\lambda_0^* \beta(\lambda_0^*)}{\lambda_1^*} > 1 \implies \frac{\partial I_1}{\partial r}(r=0) > 0$$

Corollary:

At $r = 0$, the period 2 threshold for investment is characterized by the following condition:

$$\begin{aligned} C_0 + \frac{1}{2}K &= \lambda_1^*(f(I_2^c) + K) \\ \lambda_1^* &= \frac{C_0 + \frac{1}{2}K}{f(I_2^c) + K} \end{aligned}$$

Since I_2^c is bounded, it follows that

$$\lim_{K \rightarrow \infty} \lambda_1^* = \frac{1}{2}$$

I now establish that raising K lowers period 1 investment using a proof by contradiction. Consider the behavior of the marginal period 1 investor, with $\lambda_0 = \lambda_0^*$. At the original equilibrium, the marginal investor has

$$\begin{aligned} V(i, \lambda_0) &= \lambda_0(f(I_1^c) + \frac{1}{2}K - C_0) + (1 - \lambda_0)[-(C_0 + \frac{1}{2}K)] = \\ V(l, \lambda_0) &= \lambda_0\beta(z^*(\lambda_0))[(f(I_2^c) + K) - (C_0 + \frac{1}{2}K)] + (1 - \lambda_0)\alpha(z^*(\lambda_0))[-(C_0 + \frac{1}{2}K)] \end{aligned}$$

Suppose that increasing K raises I_1 . This would require that investment becomes strictly preferable for the marginal investor, i.e. $dV(i, \lambda_0)/dK > dV(l, \lambda_0)/dK$. Recall that $\partial I_1^c/\partial I_1 = 1$ and $\partial I_2^c/\partial I_1 < 1$. Hence, $\partial f/\partial I_1^c < 0$ implies $\partial V(i, \lambda_0)/\partial I_1 < \partial V(l, \lambda_0)/\partial I_1$. In addition, observe that $\partial V(i, \lambda_0)/\partial K < \partial V(l, \lambda_0)/\partial K$ because the $V(l, \lambda_0)$ expressions puts more weight on the first term, where K is positive, since $\beta > \alpha$ when signals are informative. Consequently, both the direct effect of increasing K and the indirect effect of equilibrium price changes lower $V(i, \lambda_0)$ relative to $V(l, \lambda_0)$ for the marginal investor. This makes the marginal firm drop out of the period 1 investment pool, lowering I_1 . But this contradicts the supposition. Therefore increasing K must strictly *lower* I_1 .

Hence, as $K \rightarrow \infty$, $\lambda_0^* \rightarrow 1$, while implies $\beta(\lambda_0^*) \rightarrow 1$. Meanwhile $\lambda_1^* \rightarrow \frac{1}{2}$. It follows that (9) must hold as $K \rightarrow \infty$, as claimed.