Review of Finance (2015) vol: 1–44
doi: doi
Advance Access publication: dd month YYYY

Corporate Investment Over the Business Cycle

Thomas Dangl[1], Youchang Wu[2]
[1] *Vienna University of Technology;* [2] *University of Wisconsin-Madison*

Abstract. The average capital growth rate across firms declines sharply during a recession, and recovers only slowly. We provide a micro-founded explanation for this and several new stylized facts of investment asymmetry. Our investment model features various degrees of reversibility, cyclical macroeconomic shocks, and uncertainty about the state of the economy. Model simulations replicate strikingly different empirical patterns of capital growth rates at the aggregate and firm levels, featuring no slope asymmetry and a positive level asymmetry at the firm level, negative slope and level asymmetries at the aggregate level, and a positive relation between the industry-level slope asymmetry and asset illiquidity.

JEL Classification: G31, G30

Keywords: Investment, Real option, Learning, Business Cycle

Capital investment is one of the most important corporate decisions. It determines a firm's long-term prospects and shareholder value creation. Capital investment is also a fundamental driver of economic growth. Although

---

We thank Thierry Foucault (Editor) and an anonymous referee for thoughtful comments and suggestions. We are also grateful to David Brown, Bjorn Eraker, Michael Gofman, Richard Green, Christopher Hennessy, Oliver Levine, Xiaoji Lin, Antonio Mello, Erwan Quintin, Mark Ready, Suresh Sundaresan, Luke Taylor, Sheridan Titman, Grigory Vilkov, Neng Wang, Toni Whited, Randy Wright, Wei Zhang, and seminar audiences at the 2014 American Economic Association meetings, the 2013 European Winter Finance Summit, the 2012 Financial Intermediation Research Society meetings, the 2012 Swiss Society for Financial Market Research meetings, the 2011 China International Conference in Finance, University of Vienna, Vienna University of Economics and Business, University of Wisconsin-Madison, Cheung Kong Graduate School of Business, University of Hong Kong, University of Frankfurt, and Karlsruhe University. Part of this work is done when Youchang Wu is a visiting financial economist at the U.S. Securities and Exchange Commission. The Commission, as a matter of policy, disclaims responsibility for any private publication or statement by any of its employees. The views expressed herein are those of the author and do not necessarily reflect the views of the Commission or of the author's colleagues on the staff of the Commission.
This work is supported by a grant from the University of Wisconsin-Madison Graduate School [MSN126496].

the literature on this subject is vast, what drives capital investment remains an elusive question. What is particularly puzzling is that capital investment tends to decline sharply when an economy enters a recession, and to recover only slowly as economic conditions improve. During the recent financial crisis, the U.S. gross private domestic investment quantity index dropped sharply from 141 in 2007 to 100 in 2009. In 2013, the index was still below its pre-crisis level, despite several years of record-low interest rates, rising stock markets, and strong corporate earnings after the crisis.[1] The weak recovery of corporate investment activity has been a topic of much discussion in the post-crisis policy debates and media reports.[2]

We propose a micro-founded explanation for the sharp decline and slow recovery of capital investment, accounting also for several new patterns of investment asymmetry that we uncover. Using the quarterly Compustat-CRSP merged database, we first document strikingly different patterns of investment asymmetry at the aggregate and firm levels. Confirming the existence of a slope asymmetry at the aggregate level, we find that the decline in the average capital growth rate across firms during the recession is much faster than the recovery during the expansion. Furthermore, the first-order difference of the average capital growth rate is negatively skewed, especially when the growth rate is measured over three or four quarters. In addition to the evidence of a slope asymmetry, we also find the average capital growth rate itself to be negatively skewed, suggesting the existence of a level asymmetry. However, neither of these asymmetries holds at the individual firm level. Instead, firm-level capital growth rates are strongly positively skewed, and their first-order differences are basically symmetric. These strikingly different patterns of asymmetries impose important restrictions on a structural explanation of the slow recovery of investment.

Our model of optimal investment captures three important features of the real world: (1) installed capital is illiquid, resalable only at a discount (i.e., costly reversibility) or not resalable at all (i.e., complete irreversibility); (2) the profitability of an individual firm is strongly influenced by macroeconomic conditions; and (3) firms face substantial uncertainty about the true state of the economy. We consider a cross-section of risk-neutral firms with an infinite horizon. Each firm faces its own business conditions, summarized by a random demand factor. The expected growth rate of a firm's demand factor depends on the state of the economy, which shifts between a high-growth state (expansion) and a low-growth state (recession) at random

[1] Index values for the years 2010 through 2013 were 113, 119, 130, and 136, respectively. See Bureau of Economic Analysis National Income and Product Accounts Tables 5.2.3.

[2] See, for example, the Wall Street Journal article, "Investment falls off a cliff: U.S. companies cut spending plans amid fiscal and economic uncertainty," December 2, 2012.

times. The true state is not directly observable. Firms update their beliefs continuously by observing their own operating profits and a public signal. Capital stock can be expanded instantaneously at a constant marginal cost, but the resale price of capital is lower than its purchase price.

We calibrate the parameters of our model, and derive numerically the optimal investment policy of a typical firm. Under costly reversibility, the policy is characterized by two reflecting boundaries: an upper bound and a lower bound on the firm's capital stock, normalized by its current demand factor. The lower bound represents the firm's optimal capacity, and the upper bound its maximum tolerated capacity. Both bounds are functions of the firm's belief about the current state of the economy. The firm takes no action as long as its normalized capital stock is between these two boundaries, and invests (disinvests) instantaneously once it hits the lower (upper) bound.

Since demand grows faster during an expansion, the firm's optimal normalized capacity increases in general with the posterior probability that the economy is in an expansion. Furthermore, because the option value of waiting increases with uncertainty about the true state, the relation between optimal capacity and the firm's belief is convex, especially when signals are very informative. When the current posterior probability of expansion is low, positive signals lead to only a relatively modest increase of optimal capacity because of increased uncertainty. When this posterior probability is high, bad signals lead to a sharper decline of optimal capacity. This arises because the decline in the expected growth rate is accompanied by increased uncertainty.

After deriving the optimal investment/disinvestment policy, we simulate a large number of firms following this policy. Firms experience the same sequence of macroeconomic shocks as empirically observed. They receive a common stream of public signals and heterogeneous shocks to their own business conditions. We examine the patterns of capital growth rates at both the firm and aggregate levels. Our simulated data match the empirical data remarkably well. The average capital growth rate declines fast and recovers slowly. Both its level and its slope are negatively skewed. At the same time, capital growth rates of individual firms are positively skewed in level and symmetric in slope.

These contrasting features of asymmetry arise because aggregate investment depends not only on the scale of investment by any given firm but also, and more crucially, on the number of firms investing at a given time. Costly reversibility implies that a firm disinvests rarely, invests intermittently, and remains inactive in most of the time. Therefore firm-level capital growth rates are positively skewed. Yet because an individual firm invests and returns to inactivity at a similar speed, firm-level capital growth rates have a symmetric slope.

At the aggregate level, because idiosyncratic shocks in the cross-section are averaged out, the distribution of the average capital growth rate is determined largely by the relative length of expansions and recessions. Since recessions are generally shorter than expansions, there are relatively few observations drawn from recession periods. These observations form a long left tail of the distribution of the average growth rate, making it negatively skewed. More important, the speed of adjustment is also asymmetric, due to the endogenous distribution of firms relative to their optimal capacities. During a recession, a high proportion of firms are far from their investment boundaries, because of low demand growth. The excess capacities these firms have cumulated over time prevent them from investing even when their beliefs change significantly upon the arrival of a positive signal. During an expansion, however, a large proportion of firms are pushed to the investment boundary by high demand growth. When a negative signal arrives, these firms cease investment simultaneously, causing a sharp decline in the average capital growth rate, even if no firm actively disinvests.

Our analysis shows that both incomplete information and costly reversibility are key to explain the slow recovery. If information is complete, both the decline and the rebound of the average capital growth rate will be fast and abrupt. If investment is perfectly reversible, then the average capital growth rate simply mirrors the demand growth rate, and its slope is symmetric. More generally, our simulation shows that as reversibility increases, the negative slope asymmetry of the average capital growth rate becomes weaker, or even disappears. In support of this prediction, we find that capital growth rates of industries with higher capital asset liquidity, measured by the activeness of capital reallocation relative to capital investment, following Eisfeldt and Rampini (2006), exhibit less slope asymmetry.

We extend our model by allowing for procyclical depreciation and countercyclical financing frictions. We find that procyclical depreciation further slows down the recovery of the average capital growth rate, and magnifies its slope asymmetry. This arises because the higher depreciation rate during expansions reduces the marginal value of capital, thus dampening the incentive to invest as economic conditions improve. On the other hand, we find that countercyclical financing frictions per se do not increase the slope asymmetry at the aggregate level.

Our work is closely related to the study of Guo et al. (2005) on optimal irreversible investment under random regime shifts, but with some significant differences. First, while Guo et al. (2005) focus on the theoretical properties of the optimal investment policy of an individual firm, we study a large cross-section of firms facing both common and heterogeneous shocks, and calibrate our model to replicate the major empirical features at both the aggregate and firm levels. Second, while they assume regime shifts to be perfectly observable, we assume that firms can only infer the true regime from noisy signals. The incomplete information setup provides a natural framework to examine the impacts of endogenously determined time-varying uncertainty. It is also key for our model to match the empirical data. Complete information implies abrupt changes in investment following a regime shift in either direction, which does not conform to empirical observations. Third, while they assume investment to be completely irreversible, we allow for different degrees of reversibility, encompassing complete reversibility and irreversibility as two special cases. This allows us to investigate the economic impacts of different degrees of reversibility, and link investment to cross-sectional measures of reversibility, such as capital asset liquidity.

Our work extends the literature on the real options approach to investment by considering firms' optimal investment behavior over the business cycle and its implications for aggregate investment.[3] Bloom (2009) shows that macro uncertainty shocks can generate sharp recessions and recoveries through their impacts on firms' investment and hiring decisions. Unlike us, he takes the degree of uncertainty as exogenous. Capital investment under incomplete information has been studied by Alti (2003), Decamps et al. (2005), Klein (2009), and Grenadier and Malenko (2010). These authors do not consider the cyclical features of investment, which is the focus of our study.

Our work also contributes to the literature on business cycle asymmetry. There is a long debate about whether economic downturns are more abrupt and violent than upturns. The empirical evidence is somehow mixed. While Neftci (1985) and Sichel (1993) find evidence supporting the asymmetry hypothesis, Falk (1986) and DeLong and Summers (1986) conclude that there is very little evidence of asymmetry. These studies focus on macroeconomic series such as GNP, industry production, or unemployment. Van Nieuwerburgh and Veldkamp (2006) show evidence of negative skewness in the growth rate of several macroeconomic series, including aggregate investment.[4] Our empirical analysis of the firm-level data provides strong evidence of asymmetry in corporate investment. To the best of our knowledge, we are the first to examine the asymmetry of the *change* in the capital growth rate,

[3] See Dixit and Pindyck (1994) for a review of the capital investment literature using the real options approach.

[4] Explanations for business cycle asymmetry have been advanced by Chamley and Gale (1994), Gale (1996), Acemoglu and Scott (2003), Veldkamp (2005), and Van Nieuwerburgh and Veldkamp (2006).

and to contrast and reconcile the different patterns of asymmetries of capital growth rates at the firm and aggregate levels.[5]

The paper is organized as follows. Section 1 presents empirical evidence for the asymmetry of capital growth rates at the firm and aggregate levels. Section 2 presents the model. Section 3 characterizes the optimal investment/disinvestment policy. Section 4 compares the paths of simulated and empirical capital growth rates over the business cycle. Section 5 compares the skewness patterns of simulated and empirical data, and examine the effects of different degrees of reversibility. Section 6 considers two extensions of the model to allow for procyclical depreciation and countercyclical financing frictions. Section 7 concludes.

1. Asymmetries of Capital Growth Rates: Empirical Facts

1.1 CAPITAL GROWTH RATES OVER THE BUSINESS CYCLE

To examine the empirical patterns of firms' investment behavior, we use the quarterly Compustat-CRSP merged database from 1975 through 2011 (quarterly data on capital stock are sparse before 1975), excluding financial firms (SIC codes between 6000 and 6999), utilities (SIC codes between 4900-4999), and government entities (SIC codes of 9000 or above). We measure a firm's net capital stock by its net property, plant, and equipment (PPENT in Compustat), and measure its investment by the continuously compounded growth rate of net capital.[6] All nominal values are converted into year 2005 dollars using the quarterly GDP deflator. A firm is included in the sample if its PPENT reaches the $1 million threshold in the current year or any previous year.

Since we are interested in capital growth arising from physical investment, which is different from growth through mergers and acquisitions, we exclude firms heavily involved in M&A activities. For this purpose, we match our sample to the SDC Mergers and Acquisitions database of Thomson Reuters, which has a comprehensive coverage of M&A deals from 1980 through 2011. We exclude fiscal years in which the total transaction value of a firm's M&A

[5] Albuquerque (2012) documents and provides a theory to reconcile the negative skewness of aggregate stock market returns and the positive skewness of individual stock returns. His explanation hinges on the heterogeneity of firms' announcement events.

[6] A simple percentage growth rate is inherently asymmetric, since it cannot go below -100% due to the non-negativeness of capital stock. We use the continuously compounded growth rate to avoid such mechanical asymmetry. Measuring investment by the gross investment rate, defined as capital expenditures minus sales of capital, deflated by the lagged gross capital, leads to very similar results for the time period 1984 through 2011 (quarterly capital expenditures data are not available for earlier years).

deals exceeds 20% of its total assets at the prior fiscal year end. Furthermore, we require the difference between the increase in net capital and net investment (defined as capital investment minus sales of capital and depreciation) to be less than 20% of the net capital at the prior fiscal year end. In the absence of M&As, this difference should be zero by accounting identity.

The final sample consists of 484,824 observations of quarterly capital growth rate at the firm level, with an average of 3276 firms in each quarter. To limit the impact of extreme outliers or potential data errors, we winsorize the firm-level growth rates at the 2.5th and the 97.5th percentiles of the cross-sectional distribution of each quarter. We use the cross-sectional average of quarterly capital growth rates (equally weighted) to measure investment at the aggregate level.

We first investigate the behavior of the average capital growth rate at the turning points of business cycles identified by the National Bureau of Economic Research (NBER). According to the NBER, there are six troughs and six peaks during our sample period.[7] A trough marks the end of a contraction and the start of an expansion, while a peak marks the end of an expansion and the start of a contraction. The left diagram in Figure 1 shows the crash of the average capital growth rate from one quarter before the peak through six quarters after it. The right diagram shows the slow recovery from one quarter before the trough through 12 quarters after it. As one can see, the downward shift is very steep, while the upward move is much slower. It takes six quarters for the average capital growth to move from the peak to the bottom, and twice as long for it to get back to the peak. This is consistent with the business cycle asymmetry documented by Neftci (1985), Sichel (1993), and Van Nieuwerburgh and Veldkamp (2006) using macroeconomic data.

1.2 SKEWNESS OF CAPITAL GROWTH RATES

To quantify asymmetry, we use a standard measure in statistics, skewness, which is defined as the third central moment of a random variable, normalized by the third power of standard deviation. A negative skewness value indicates that the left tail of a distribution is longer than the right tail and that the bulk of the observations lie to the right of the mean. We call this a negative asymmetry. A positive skewness value indicates the opposite. A symmetric distribution has zero skewness.

Following Sichel (1993), we distinguish between two types of asymmetry. Level asymmetry (deepness) refers to the characteristic that troughs are

[7] The trough quarters are 1975 Q1, 1980 Q3, 1982 Q4, 1991 Q1, 2001 Q4, and 2009 Q2. The peak quarters are 1973 Q4, 1980 Q1, 1981 Q3, 1990 Q3, 2001 Q1, and 2007 Q4.

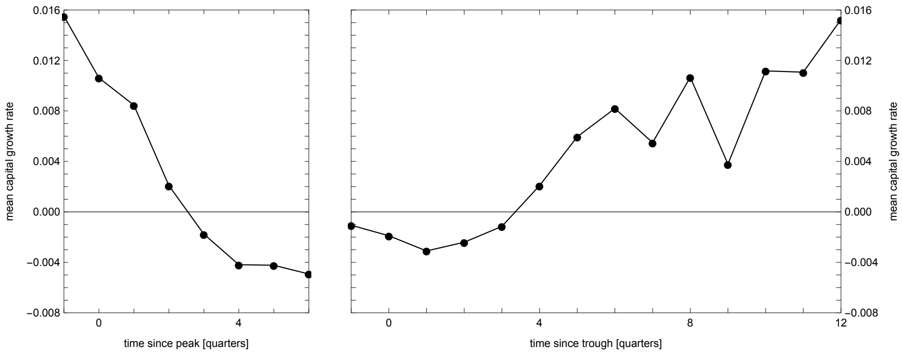
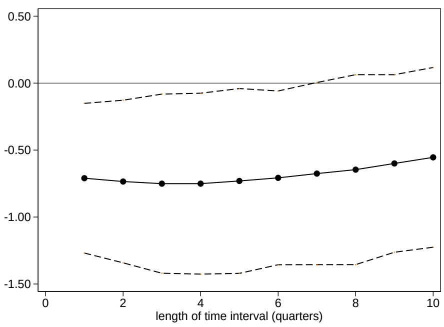


Fig. 1: **Mean net capital growth rate over the business cycle.** The left diagram shows the average capital growth rate from quarter -1 to quarter 6, where quarter 0 is the business cycle trough. The right diagram shows the average net capital growth rate from quarter -1 to quarter 12, where quarter 0 is the business cycle peak. Both trough and peak quarters are dated by the NBER. The capital growth rate is averaged first cross firms and then across cycles during 1975 Q1—2011 Q4.

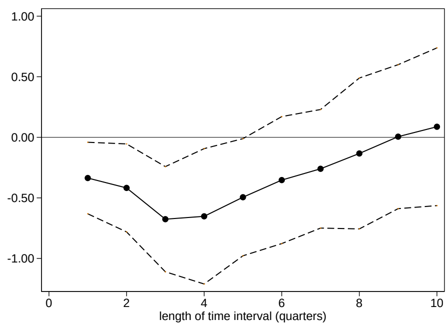
farther below the trend than peaks are above. Slope asymmetry (steepness) refers to the characteristic that downturns are steeper than upturns, as we see in Figure 1. We use the skewness of the capital growth rate to measure level asymmetry, and the skewness of the first-order difference of the capital growth rate to measure slope asymmetry. To capture the asymmetries of capital growth over different time intervals, we calculate growth rates over various time horizons, following Van Nieuwerburgh and Veldkamp (2006). An n-quarter growth rate measured in quarter t, $g_{t,n}$, is defined as the sum of n continuously compounded quarterly growth rates from quarters $t-n+1$ through t. The first-order difference, i.e., the slope, of the n-quarter growth rate is then defined as $g_{t,n} - g_{t-n,n}$.

Panels (a) and (b) of Figure 2 show, respectively, the estimated skewness values for the levels and slopes of average capital growth rates calculated over various time intervals, from one to ten quarters. To test whether the values are statistically different from zero, we also plot the 90% confidence interval of each estimate.[8]

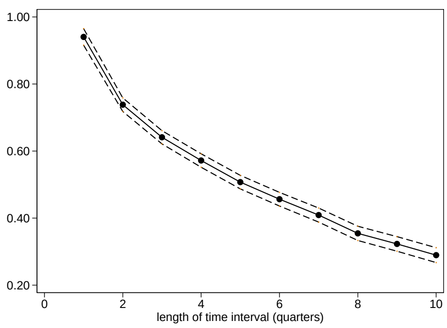
[8] Since the average n-quarter capital growth rate is autocorrelated because of overlapping time intervals, the standard error of the skewness estimate cannot be computed using the standard method. We therefore follow the Monte Carlo procedure of DeLong and Summers (1986). First, we estimate a fifth-order autoregressive model for each time series of the growth rate (higher orders of the autoregressive model do not change the significance of our



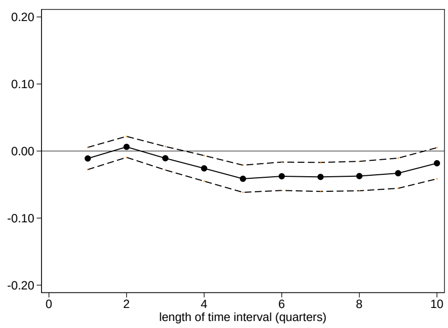
(a) Skewness of average capital growth rates



(b) Skewness of differenced average growth rates



(c) Skewness of individual capital growth rates



(d) Skewness of differenced individual growth rates

Fig. 2: **Skewness of capital growth rates: Empirical estimates.** Panels (a) and (b) show the skewness values for the levels and the slopes of average capital growth rates measured over various time intervals (from one quarter to ten quarters), respectively. Panels (c) and (d) show the average skewness values for the levels and slopes of firm-level capital growth rates measured over various time intervals, respectively. The dashed curves show the upper and lower bounds of the 90% confidence interval of each skewness estimate.

Consistent with the sharp decline and slow recovery observed in Figure 1, the average capital growth rates exhibit both level asymmetry and slope

results). Second, we use the estimated model to generate 1000 artificial series for the sample period under the assumption that shocks to the autoregressive process were independent and normally distributed. Third, we use the standard deviation of the skewness values across the simulated series as the standard error of the skewness estimate under the null hypothesis of no asymmetry.

asymmetry. The estimated values of skewness are below -0.5 for the levels of growth rates measured over all time intervals in Panel (a).[9] For measurement intervals up to six quarters, they are negative at the 90% significance level. This indicates that troughs are deeper below than peaks are above the mean, which is evidence of level asymmetry.

The first-order differences of the average capital growth rates are also negatively skewed, except when the measurement interval is longer than two years. This indicates a steeper slope of the downturn than of the upturn, i.e., a slope asymmetry. The estimated skewness value is non-monotonic in the length of the measurement interval. It starts at -0.34 for the growth rate measured over one quarter, declines to a bottom of -0.68 for the growth rate measured over three or four quarters, and rises again for growth rates measured over longer time intervals. This suggests that the decline over three or four consecutive quarters generates more extreme outliers than the decline over one quarter does. As the measurement interval lengthens further, however, the slope asymmetry is gradually smoothed out due to averaging across quarters.

Are the level and slope asymmetries of the average capital growth rate simply a carryover of asymmetries at the firm level? To answer this question, we examine skewness at the firm level. For each firm with at least 40 observations, we calculate the skewness values of the levels and slopes of its capital growth rates over various time horizons, and then take a simple average across firms. The cross-sectional averages of the skewness estimates, together with their 90% confidence intervals, are plotted in Panels (c) and (d) of Figure 2.[10] The patterns are strikingly different from those at the aggregate level. The levels of various capital growth rates are strongly positively, instead of negatively, skewed. The average skewness value starts at 0.94 for the quarterly growth rate, deceases steadily as the measurement interval lengthens, and remains positive even for the ten-quarter growth rate (Panel (c)). At the same time, the slope asymmetry is very small economically. The point estimates of the skewness for the slopes of capital growth rates never go below -0.05.

These results suggest that investment at the aggregate level behaves very differently from that at the firm level. In the next section, we present a dynamic model that generates a slow recovery, and at the same time, reconciles the different patterns of asymmetries at the aggregate and firm levels.

---

[9] Van Nieuwerburgh and Veldkamp (2006) report a skewness value of -0.72 for the quarterly growth rate of per capita real investment from 1952 through 2002.

[10] The standard error of the firm-level skewness estimates is the standard deviation across firms divided by the square root of the number of firms.

2. A Dynamic Model of Investment

2.1 SETUP

We consider a cross-section of risk-neutral firms with an infinite time horizon. The firms are identical ex ante, but face both common and heterogeneous shocks. A typical firm operates in an environment similar to that of Guo et al. (2005), but with incomplete information. Time is continuous. Investment is incremental. It is either completely irreversible, as in Guo et al. (2005), or costly reversible. Each firm's cash flow is driven by a distinct stochastic factor. Depending on the state of the macroeconomy, the expected growth rate of this factor shifts between a high level and a low level at random times. We describe the economic environment and the optimization problem from the perspective of a typical firm in the cross-section.

Cash Flows. The operating income (before depreciation) of the firm is assumed to be given by a linearly homogeneous function $f : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ satisfying:

$$f(x_t, k_t) = \frac{1}{1-\alpha} x_t^{\alpha} k_t^{1-\alpha}, \tag{1}$$

where $(k_t)_{t\geq 0}$ represents the process of the firm's net capital stock, and $(x_t)_{t\geq 0}$ represents the process of a demand factor.[11] Assuming that the firm's output is nonstorable, Equation (1) can be interpreted as the profit of either a price-taking firm with decreasing returns to scale, or a monopolist facing constant returns to scale and a constant elasticity demand curve (see Abel and Eberly (1996) and Morellec (2001)).

Demand Shocks. Assume that the demand factor for the firm, x_t, evolves according to the stochastic differential equation:

$$d\ \ln(x_t) = \mu_t dt + \sigma_x dW_{xt},\ \ x_0 > 0, \tag{2}$$

where W_{xt} is a standard Wiener process, and $\ln(x)$ is the natural logarithm of x.[12] The volatility σ_x is a known constant. The expected (continuously-compounded) growth rate of the demand factor, μ_t, is determined by macroeconomic conditions, and is identical for all firms in the economy. It is low in a recession ($\mu_t = \mu_l$) and high in an expansion ($\mu_t = \mu_h > \mu_l$). Within a

---

[11] The factor x_t can be generally interpreted as an index of a firm's business conditions, which reflects demand, productivity, and costs of factors other than capital.

[12] We write the stochastic differential equation in terms of $d\ln(x)$ instead of dx so that the expected continuously compounded growth rate is not affected by the volatility. Note also that σ_x is the same in both states. Since the volatility of a process can be estimated almost instantaneously in continuous time, firms would learn the true state of the economy almost instantaneously if σ_x differs across states.

given state of the economy, the demand factor follows a standard geometric Brownian motion.

The macroeconomic condition switches between expansions and recessions at random times. Correspondingly, μ_t switches randomly between μ_h and μ_l. More specifically, we assume that μ_t is driven by a continuous-time Markov jump process with the transition probabilities

$$P(\Delta t) = \mathbb{I} + \begin{pmatrix} -\lambda_{h,l} & +\lambda_{h,l} \\ +\lambda_{l,h} & -\lambda_{l,h} \end{pmatrix} (\Delta t + o(\Delta t)), \tag{3}$$

where $\mathbb{I}$ is the identity matrix and $\lambda_{h,l}$ and $\lambda_{l,h}$ are, respectively, the constant intensities of transition from μ_h to μ_l and vice versa.[13] The magnitudes of $\lambda_{h,l}$ and $\lambda_{l,h}$ determine the persistence of the expansion and the recession, respectively. The lower the transition intensity, the higher the persistence.

Investment, Disinvestment, and Depreciation. As in Abel and Eberly (1996), a firm can add capital incrementally and instantaneously at a constant marginal cost, which is normalized to be one. Due to the specificity of physical capital and frictions in secondary capital asset markets, the resale price of capital, $b \in [0, 1)$, is lower than the purchase price.[14] The wedge between the purchase and resale prices of capital, $1 - b$, can be interpreted as a bid-ask spread, as in the standard asset pricing literature. This wedge captures the illiquidity of capital assets, which is the essence of costly reversibility of investment. $b = 0$ indicates complete irreversibility, which is embedded in our model as a special case.[15]

Installed capital depreciates at a constant rate of ξ.

Information. The firm can observe its own realized demand factor x_t through its realized operating profit f_t, but not its expected growth rate μ_t. In other words, the true state of the economy that determines μ_t is a hidden process. This implies that W_{xt} is not observable as well. When the firm observes a certain increase or decrease in operating profit, it does not know whether it comes from the drift μ_t or from the noise terms W_{xt}. Yet the distinction between these possible sources is of core relevance for the firm's investment decision. Since disinvestment is costly, investment depends not only on the current demand factor, but also on expectations about its future growth rate.

Other information available to the firm regarding the macroeconomic state is summarized by a publicly observable signal, s_t, which evolves according

[13] The assumption of only two states is for simplicity and clarity of economic intuition. The model can be extended to allow for a finite number of states.

[14] See, for example, Eisfeldt (2004), for a model with frictions in the secondary markets of capital assets.

[15] The case of perfect reversibility, $b = 1$, is discussed in Section 3.2.

to the stochastic differential equation:

$$d\ \ln(s_t) = \mu_t dt + \sigma_s dW_{st},\ \ s_0 > 0; \tag{4}$$

where W_{st} is a standard Wiener process, and σ_s is a publicly known constant representing the volatility of the signal process. The parameter σ_s measures the noisiness of the public signal. It characterizes (inversely) the accuracy of the publicly available data about the economy. We assume that dW_{st} and dW_{xt} are jointly normally distributed, with a known instantaneous correlation coefficient $\rho \in [-1, 1]$. Note that the drift terms in equations (2) and (4) are identical. This assumption is made for simplicity. Our main results remain unchanged if we allow for imperfect correlation between the expected growth rates of x_t and s_t.

Allowing the firm to observe a public signal about the macroeconomic state not only makes our model more realistic, but also allows us to investigate the effect of information quality while keeping the volatility of the demand factor constant. Without the public signal, information quality is directly tied to the volatility of the demand factor σ_x. Since σ_x affects investment not only through the information channel, as we show in Section 3.2, we cannot separate the information quality effect from the demand volatility effect in the absence of s_t.

2.2 LEARNING ABOUT THE STATE OF THE ECONOMY

By observing the demand factor x_t and the signal s_t over the time interval $[0, t)$, the firm can continuously update its belief about the state of the economy. To formalize the rational learning rule, we denote by $\mathcal{F}_t$ the canonical nondecreasing filtration jointly created by x_t and s_t, and by π_t the probability that $\mu_t = \mu_h$ conditional on $\mathcal{F}_t$ and a prior π_0. The conditional mean of the continuously compounded growth rate at time t is then $\pi_t \mu_h + (1 - \pi_t)\mu_l$. Unexpected changes in x_t and s_t give rise to an update of the belief.

Learning under this circumstance is a standard nonlinear filtering problem and can be characterized by the proposition below.[16]

Proposition 1. *The optimal updating of the belief satisfies the stochastic differential equation:*

$$d\pi_t = [-\pi_t \lambda_{h,l} + (1 - \pi_t)\lambda_{l,h}]dt + (\mu_h - \mu_l)\pi_t(1 - \pi_t)\mathbf{1}'(\Phi')^{-1}\mathbf{dW_t^{\mathcal{F}}}, \tag{5}$$

[16] See David (1997) for an early application of this filter in finance.

where $\mathbf{W}_\mathbf{t}^\mathcal{F}$ *is a two-dimensional independent Wiener process with respect to* $\mathcal{F}_t$ *defined as*

$$\mathbf{dW}_\mathbf{t}^\mathcal{F} \equiv \begin{pmatrix} dW_{xt}^\mathcal{F} \\ dW_{st}^\mathcal{F} \end{pmatrix} \equiv \Phi^{-1} \begin{pmatrix} d\ \ln(x_t) - E(\mu_t|\mathcal{F}_t)dt \\ d\ \ln(s_t) - E(\mu_t|\mathcal{F}_t)dt \end{pmatrix}, \tag{6}$$

Φ *is a* 2×2 *matrix satisfying*

$$\Phi\Phi' = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_s \\ \rho\sigma_x\sigma_s & \sigma_s^2 \end{pmatrix},$$

and $\mathbf{1}$ *is a two-dimensional column vector with both elements equal to 1.*

Proof 1. See Theorem 9.1 in Liptser and Shiryaev (2001) for the basic filtering equation, and Equation (A1) in Veronesi (2000) for an extension to the vector case. Equation (5) is obtained by applying Equation (A1) in Veronesi (2000). $dW_{xt}^\mathcal{F}$ and $dW_{st}^\mathcal{F}$ are uncorrelated because

$$dW_{xt}^\mathcal{F} dW_{st}^\mathcal{F} = (1,0)\,\Phi^{-1}(\Phi\Phi' dt)(\Phi^{-1})'(0,1)' = 0.$$

This, plus the assumption of joint normality, implies that $dW_{xt}^\mathcal{F}$ and $dW_{st}^\mathcal{F}$ are independent. ■

The diffusion process π_t is bounded between 0 ($\mu_t = \mu_l$, almost sure) and 1 ($\mu = \mu_h$, almost sure). The drift term in Equation (5) indicates that in the absence of information shocks, there is a tendency for the belief to revert toward the unconditional mean:

$$\bar{\pi} = \frac{\lambda_{l,h}}{\lambda_{h,l} + \lambda_{l,h}}, \tag{7}$$

which satisfies $[-\pi_t\lambda_{h,l} + (1-\pi_t)\lambda_{l,h}] = 0$. Therefore, the impact of any particular information shock decays gradually over time.

The diffusion term in Equation (5) characterizes the response of the belief to unexpected changes in the realized demand factor x_t and the signal s_t. $E(\mu_t|\mathcal{F}_t)dt$ represents the best forecast of $d\ln(x_t)$ and $d\ln(s_t)$ conditional on the information set $\mathcal{F}_t$, and $d\mathbf{W}_t^\mathcal{F}$ represents the standardized forecast errors. It is straightforward to see from the equation that the belief is more sensitive to the forecast errors the greater the difference between the two possible growth rates, $(\mu_h - \mu_l)$, and the higher the uncertainty about the state of the economy, captured by the conditional variance of the belief, $\pi_t(1-\pi_t)$. These results are quite intuitive. When the growth rates in the two states do not differ much, or when the firm is very sure about the true state (i.e., π_t is close to one or zero), unexpected changes in the signals do not have much impact on the belief.

An alternative formulation of the optimal updating rule, Equation (A1) in Appendix A.1, provides some further insights into the learning process. It suggests that the optimal learning from these two jointly normal signals can proceed in two steps. One first forms a minimum variance "portfolio" of both signals, and then updates the belief using this compound signal. The information quality of this compounded signal is measured by the inverse of its variance, $\frac{1}{\sigma^2} = \mathbf{1}'(\Phi\Phi')^{-1}\mathbf{1}$. Equation (A1) shows that when the compound signal is more precise (i.e., when σ^2 is small), the firm's belief responds to the compound signal more strongly.

The fact that $\mathbf{W}_t^{\mathcal{F}}$ is a Wiener process with respect to the filtration $\mathcal{F}_t$ means that all information about the state of the economy available at time t is incorporated in the current belief π_t. In other words, any information relevant for the future is used immediately to update the current belief. Therefore, the belief π follows a $\mathcal{F}_t$-Markov process. Using $dW_{xt}^{\mathcal{F}}$ and $dW_{st}^{\mathcal{F}}$ defined by Equation (6), we can rewrite the joint dynamics of x_t and s_t in terms of unexpected changes with respect to $\mathcal{F}_t$:

$$\begin{pmatrix} d\,\ln(x_t) \\ d\,\ln(s_t) \end{pmatrix} = [\pi_t\mu_h + (1-\pi_t)\mu_l] \begin{pmatrix} 1 \\ 1 \end{pmatrix} dt + \Phi \begin{pmatrix} dW_{xt}^{\mathcal{F}} \\ dW_{st}^{\mathcal{F}} \end{pmatrix}, \tag{8}$$

where π_t is updated as stated in Equation (5).

2.3 FIRM VALUE DYNAMICS AND BOUNDARY CONDITIONS

Since there is no fixed adjustment cost, and the marginal cost of capital is constant, the firm's optimal investment policy can be characterized by two reflecting boundaries, which split the state space into an investment region, an inaction region, and a disinvestment region. The firm remains inactive in the interior area of the inaction region, and increases or decreases capital instantaneously by an infinitesimal amount dk whenever it hits the boundaries.[17] We first derive the firm value dynamics in the inaction region, and then specify the boundary conditions that characterize the optimal investment/disinvestment policy.

Since the Bayesian belief π_t follows an $\mathcal{F}_t$-Markov process, the firm value is fully determined by the current capital stock k_t, the current demand factor x_t, and the current belief π_t. The value function V, which represents the present value of operating profit flow under the optimal investment policy,

[17] For the special case of complete irreversibility ($b = 0$), the upper boundary is infinity, and the disinvestment region is empty.

can be written as $V(k_t, x_t, \pi_t)$. Let r denote the instantaneous riskless rate of interest. Proposition 2 describes the dynamics of the firm value V:

Proposition 2. *The firm value can be written as* $V(k, x, \pi) = xv(h, \pi)$, *where* $h \equiv \frac{k}{x}$ *and* $v \equiv \frac{V}{x}$ *represent the capital stock and firm value, respectively, normalized by the demand factor. Furthermore, in the inaction region, the normalized firm value,* v, *has to satisfy the partial differential equation (Hamilton-Jacobi-Bellman equation):*

$$
\begin{aligned}
[r - (\pi\mu_h + (1-\pi)\mu_l + \frac{1}{2}\sigma_x^2)]v = {} & \frac{h^{1-\alpha}}{1-\alpha} - h[\pi\mu_h + (1-\pi)\mu_l + \frac{1}{2}\sigma_x^2 + \xi]\frac{\partial v}{\partial h} \\
& + \frac{1}{2}\sigma_x^2 h^2 \frac{\partial^2 v}{\partial h^2} + [-\pi\lambda_{hl} + (1-\pi)\lambda_{lh}]\frac{\partial v}{\partial \pi} \\
& + [\pi(1-\pi)(\mu_h - \mu_l)]\frac{\partial v}{\partial \pi} \\
& + \frac{[\pi(1-\pi)(\mu_h-\mu_l)]^2}{2\sigma^2}\frac{\partial^2 v}{\partial \pi^2} \\
& - h\pi(1-\pi)(\mu_h - \mu_l)\frac{\partial^2 v}{\partial h \partial \pi},
\end{aligned}
\tag{9}
$$

with $\sigma^2 \equiv \frac{1}{\mathbf{1}'(\Phi\Phi')^{-1}\mathbf{1}}$, *as defined in Equation (A2).*

Proof 2. A proof is provided in Appendix A.2. ■

From Proposition 2, it follows that the normalized firm value, v, depends only on the ratio of installed capital to the demand factor and the belief about the state of the economy.[18]

The partial differential equation (9) has to be solved under proper boundary conditions. As the marginal value of capital decreases with installed capital and increases with the demand factor, the inaction region where the firm neither invests nor disinvests is associated with intermediate values of $h = k/x$. When installed capital stock is low relative to demand, the marginal value of installed capital increases. Therefore, as h reaches a certain lower bound, the firm will invest. This critical threshold h_i^* forms the investment boundary. It specifies the optimal capital stock relative to the demand factor, and therefore can be interpreted as the optimal normalized

[18] This homogeneity property simplifies the solution of the valuation equation because it reduces the dimensionality of the problem. It arises from the homogeneity property of the operating profit ($f(x_t, k_t) = xf(1, k_t/x_t)$), and the absence of fixed costs. Note that the homogeneity property also allows us to solve the problem in terms of Tobin's average Q, which can be written as $Q(x_t/k_t, \pi_t) = V(k_t, x_t, \pi_t)/k_t$, and has to satisfy a partial differential equation similar to (9) in x_t/k_t and π_t.

capacity. At this boundary, every positive demand shock (which represents a negative shock to h) is offset by an appropriate increase in capital. Since the firm never enters the interior area of the investment region, this boundary is called a reflecting boundary. It is a function of the belief π, because the marginal value of invested capital depends on the growth prospects of the firm.

When installed capital is high relative to current demand and, thus, the marginal value of capital is low, the firm has an incentive to sell capital at a discounted price b per unit. Therefore, there exists an upper threshold h_d^* at which any negative shock in demand is accommodated by proper disinvestment. This threshold forms the disinvestment boundary. Like the investment boundary, h_d^* is also a function of the belief π. The normalized capital never exceeds h_d^*.

Since the marginal cost of capital is normalized to be one, a rational valuation of the firm implies that at the investment and disinvestment boundaries, the firm value has to satisfy the value-matching conditions

$$V(x, k, \pi) = V(x, k + dk, \pi) - dk,$$

and

$$V(x, k, \pi) = V(x, k - dk, \pi) + b\,dk,$$

respectively. These conditions can be written in derivative form as

$$\lim_{k \to x\, h_i^*} \frac{\partial V}{\partial k} = 1, \quad \lim_{k \to x\, h_d^*} \frac{\partial V}{\partial k} = b.$$

These value-matching conditions arise because firm value today fully reflects future investment/disinvestment activity. They require that the marginal value of capital, i.e., Tobin's marginal Q, be equal to the constant marginal cost of adding capital at the investment boundary, and equal to the constant resale price of capital at the disinvestment boundary. Using the homogeneity feature of the value function, we can rewrite these boundary conditions as

$$\lim_{h \to h_i^*} \frac{\partial v(h, \pi)}{\partial h} = 1, \quad \lim_{h \to h_d^*} \frac{\partial v(h, \pi)}{\partial h} = b. \tag{10}$$

To ensure the optimality of the endogenously determined boundaries, we also require smoothness of the marginal value of capital at the boundaries, which implies the following super-contact (or smooth-pasting) conditions at both boundaries (see Dumas (2001)):

$$\frac{\partial^2 V}{\partial k \partial x} = 0, \quad \frac{\partial^2 V}{\partial k^2} = 0, \quad \frac{\partial^2 V}{\partial k \partial \pi} = 0.$$

Table 1: **Parameter values.** This table summarizes the parameter values for the base case scenario.

Notation	Economic meaning	Value
α	operating profit parameter	0.74
μ_h	expected continuously compounded growth rate in expansion	0.0669
μ_l	expected continuously compounded growth rate in recession	-0.1122
$\lambda_{h,l}$	transition intensity from expansion to recession	0.2557
$\lambda_{l,h}$	transition intensity from recession to expansion	0.7163
ξ	depreciation rate	0.1162
σ_x	instantaneous volatility of demand factor x_t	0.3115
σ_s	instantaneous volatility of signal s_t	0.25
b	resale price of one unit of capital	0.80
r	risk-free rate	0.05
ρ	instantaneous correlation between dW_{st} and dW_{xt}	0.05

These super-contact conditions translate into conditions for $v(h,\pi)$ as follows:

$$\begin{aligned} \lim_{h\to h_i^*} \frac{\partial^2 v(h,\pi)}{\partial h^2} &= \lim_{h\to h_i^*} \frac{\partial^2 v(h,\pi)}{\partial h \partial \pi} = 0, \\ \lim_{h\to h_d^*} \frac{\partial^2 v(h,\pi)}{\partial h^2} &= \lim_{h\to h_d^*} \frac{\partial^2 v(h,\pi)}{\partial h \partial \pi} = 0. \end{aligned} \tag{11}$$

3. Optimal Investment/Disinvestment Policy

Our model does not have an analytical solution, so we solve it numerically. Table 1 summarizes the parameter values for our base case scenario. Parameters with four digits after the decimal point are estimated using the annual Compustat-CRSP merged database over 1950-2011. Appendix A.3 details our calibration procedure. To solve the Hamilton-Jacobi-Bellman equation (9) along with the boundary conditions (10) and (11), we apply the approach derived in Nelson and Ramaswamy (1990) to map the dynamics of the belief π onto a recombining tree. We then use a two-dimensional tree, as outlined in Boyle et al. (1989), to jointly determine the firm value and the optimal investment/disinvestment boundaries, $h_i^*(\pi)$ and $h_d^*(\pi)$. A detailed description of the numerical procedure is in Appendix A.4.

3.1 INVESTMENT/DISINVESTMENT BOUNDARIES

Figure 3 shows the investment (Panels (b), (d), (f)) and disinvestment (Panels (a), (c), (e)) boundaries as a function of the current belief π. The area between the boundaries is the inaction region. The solid curves represent the boundaries for our base case. The dashed curves represent cases in which one of the parameter values deviates from the base case.

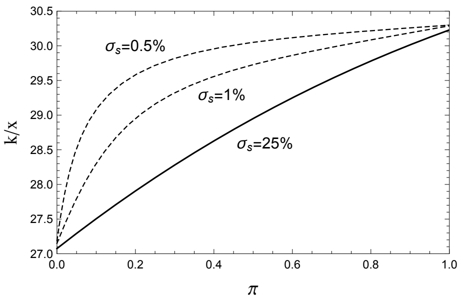
Consider first the base case investment boundary in Panel (b). This boundary represents the firm's optimal normalized capital stock. Not surprisingly, this boundary increases with the belief π. When the firm believes that the economy is in an expansion ($\pi \to 1$), it invests earlier, i.e., at a higher boundary, than when it believes the economy is in a recession ($\pi \to 0$). Therefore, investment occurs when k/x hits the boundary either from above (as capital depreciates or demand increases), or from the left (as π increases).

Notably, the investment boundary is convex in the belief π, indicating that the firm's investment decision is relatively insensitive to changes in the belief when π is low. This results from interaction of the expected growth rate and uncertainty about the growth rate. When π is close to zero, i.e., when the firm is almost sure of being in the low-growth state, a positive signal increases the expected growth rate of future demand. At the same time, it also increases uncertainty about the current state, captured by a higher value of conditional variance, $\pi(1-\pi)$, thus increasing the option value of waiting. As a result, the firm is reluctant to invest. If, however, a negative signal is received when π is high, both the resulting lower expected growth rate and greater uncertainty diminish the firm's incentive to invest. Therefore the investment boundary drops sharply.[19]

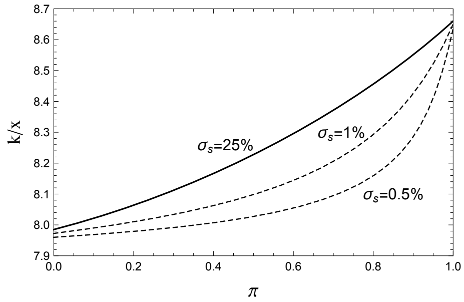
The base case disinvestment boundary, shown in Panel (a), also increases with π. This boundary characterizes the maximum tolerated normalized capacity k/x. When k/x hits this boundary, either from below (x decreases) or from the right(π decreases), disinvestment takes place, which pushes k/x back toward the inaction region. When π is high, expected demand growth is high, so a firm allows the normalized capacity k/x to reach a higher threshold.

Unlike the investment boundary, the disinvestment boundary is a concave instead of a convex function of the belief π. This concavity arises again from the effects of signals on both expected demand growth and uncertainty. When a firm receives a bad signal in a good time, the resulting decrease in expected demand growth induces an incentive to disinvest, yet the increased

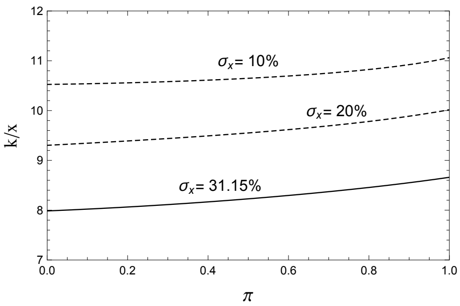
[19] Veronesi (1999) derives a similar asymmetry in the responses of stock prices to signals. He does not consider firm investment. The asymmetry arises in his model due to risk aversion rather than the option value of waiting.



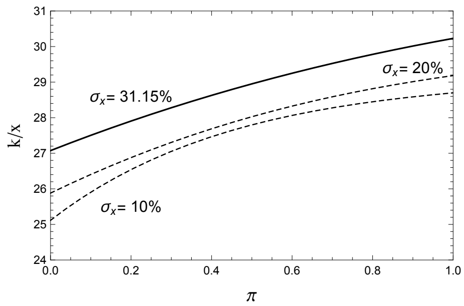
(a) Disinvestment boundary and σ_s



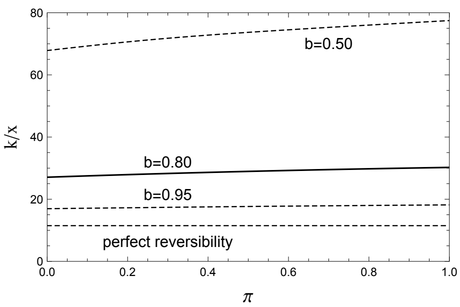
(b) Investment boundary and σ_s



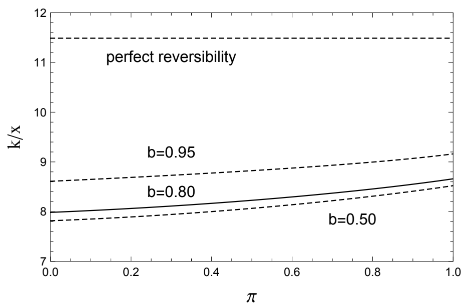
(c) Disinvestment boundary and σ_x



(d) Investment boundary and σ_x



(e) Disinvestment boundary and b



(f) Investment boundary and b

Fig. 3: **Optimal investment and disinvestment boundaries.** This figure depicts the optimal investment boundary (Panel (b), (d), and (f)) and disinvestment boundary (Panel (a), (c), and (e)) as a function of the belief π. The solid curves represent the base case parameterization specified in Table 1. The dashed curves represent cases in which one of the parameters (marked in graphs) deviates from its base case value while all others remain unchanged.

uncertainty induces an incentive to wait. As a result, the maximum tolerated capacity decreases only slightly. By contrast, when a firm receives a good signal in a bad time, both the higher expected growth rate and increased uncertainty weaken the firm's incentive to disinvest, therefore the maximum tolerated capacity increases significantly.

3.2 COMPARATIVE STATICS

Information Quality. By varying the volatility σ_s of the public signal s_t while keeping the other parameters at the base case values, we can examine the effects of information quality on investment/disinvestment policy. A lower σ_s implies faster learning, and therefore lower uncertainty about the macroeconomic state.

Comparing the boundaries under different values of σ_s in Panels (a) and (b) of Figure 3, we find that more precise signals make the investment boundary more convex and the disinvestment boundary more concave. Other things equal, the more precise the information, the stronger the response of the belief to signals, as we have noted in Section 2.2. This leads to a more volatile belief, and thus a higher option value of waiting. Higher information quality therefore amplifies the convexity of the boundary. When σ_s is sufficiently small, the firm always invests conservatively as if it were in a recession, unless it is almost sure of being in an expansion. As a result, the investment boundary is almost flat for π below one and increases sharply as π approaches one. In the limit, as σ_s goes to zero, our model converges to the complete information model of Guo et al. (2005) (with the extension to allow for disinvestment). The investment boundary is then characterized by two distinct threshold values of k/x, a higher one for the good state and a lower one for the bad state. A similar intuition explains why the disinvestment boundary becomes more concave as σ_s decreases.

Instantaneous Demand Volatility. Panels (c) and (d) show the effects of instantaneous demand volatility (σ_x) on the disinvestment and investment boundaries, respectively. An increase in σ_x has two effects, both confirmed by the figures. First, it reduces the information quality of an individual firm's demand as a signal. This is similar to the effects of an increase in σ_x. It reduces the nonlinearity of both boundaries. Second, a lower volatility of demand implies that capital stock needed today is less likely to turn into excess capacity tomorrow, and vice versa. Therefore, the firm is less hesitant to both invest and disinvest. As a result, the investment boundary shifts upward, the disinvestment boundary shifts downward, and the inaction region becomes narrower.

Reversibility. The degree of investment reversibility is determined by the gap between the purchase price and the resale price of capital. Since we fix the purchase price of capital at 1, the resale price, b, characterizes the liquidity of capital assets and the reversibility of investment, with $b = 1$ defining perfect reversibility and $b = 0$ defining perfect irreversibility.

The case of perfect reversibility serves as a good benchmark. If investment is perfectly reversible, firms adjust capital stock instantaneously to any level appropriate for the realized demand factor, because both upward and downward adjustments are costless. The belief about future growth rates becomes irrelevant. The optimal capacity is determined by the classical optimality condition of Jorgenson (1963), i.e., the equality of the marginal revenue product of capital ($\frac{\partial f}{\partial k}$) and the user cost of capital, which in our case is simply $r + \xi$. Using Equation (1), this optimality condition leads to:

$$\frac{k}{x} = (r + \xi)^{-\frac{1}{\alpha}}. \tag{12}$$

Since firms maintain capacity constantly at this optimal ratio, the capital growth rate is identical to the growth rate of the demand factor.

Panels (e) and (f) of Figure 3 show the effects of reversibility on the disinvestment and investment boundaries, respectively. Not surprisingly, with a higher degree of reversibility (a high b), firms invest and disinvest more actively. Therefore, the disinvestment boundary shifts downward and the investment boundary shifts upward, leading to a narrower inaction region. Furthermore, both boundaries become flatter, suggesting less sensitivity of the investment/disinvestment decision to the state of the economy. For the perfect reversibility case ($b = 1$), investment and disinvestment boundaries collapse into one, and the combined boundary is perfectly flat (i.e., independent of the belief), at a value implied by Equation (12).[20]

4. Capital Growth Rates: Simulated versus Empirical

4.1 SIMULATION APPROACH

The investment/disinvestment boundaries, together with a firm's position within these boundaries, determine the reaction of its normalized capacity to innovations in demand and beliefs. At the aggregate level, the speed of capital adjustment depends crucially on the distribution of firms' normalized capacities relative to the boundaries, which is endogenously determined by the history of demand shocks and firms' investment activities.

[20] Based on the values of r, ξ and α in Table 1, the optimal normalized capacity (k/x) in this case is 11.30.

To examine the dynamics of the capital growth rate in our model, we use Monte Carlo simulations. We simulate a panel of 1000 firms over a time horizon of 57 years.[21] We drop the first 20 years so that the results are not affected by initial conditions (firms are distributed evenly on the investment boundary initially), and use the remaining 37 years for our analysis. We choose such a time horizon to match the length of our empirical sample, which covers the time period from 1975 through 2011. We assume the latent macroeconomic regime follows the sequence of business cycles observed over the 57 years ending in 2011, according to the NBER. Effectively, we examine how firms behave in our model if the macroeconomic regime follows the particular sample path that we observe ex post for the 1955–2011 period.

Each firm observes its own sequence of randomly generated demand factor x_t, and a common public signal process s_t, forms its beliefs about the macroeconomic state, and invests/disinvests according to the optimal policy derived in Section 3.. The demand factors are correlated across firms due to their correlations with the common signal. Their innovations are jointly normally distributed. The parameter values are taken from Table 1 unless otherwise noted. We then examine the path of the average capital growth rate around the turning points of the business cycle, and calculate various skewness measures of the capital growth rate at the firm and aggregate levels. We repeat our simulations for 100 rounds, and compare the average results across these rounds to the empirical data.

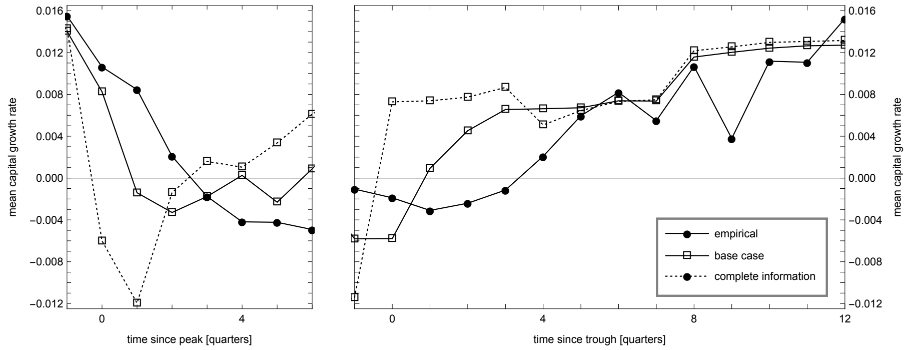
4.2 AVERAGE CAPITAL GROWTH RATE OVER THE BUSINESS CYCLE

Figure 4 shows the simulated equal-weighted average capital growth rate during the recession and the expansion, along with the empirical estimate reproduced from Panel (a) of Figure 1.[22] Panel (a) shows the results for the base case, as well as the case of (almost) complete information, in which all parameter values are at the base level except that the volatility of the public signal, σ_s, equals 0.005 instead of 0.25.

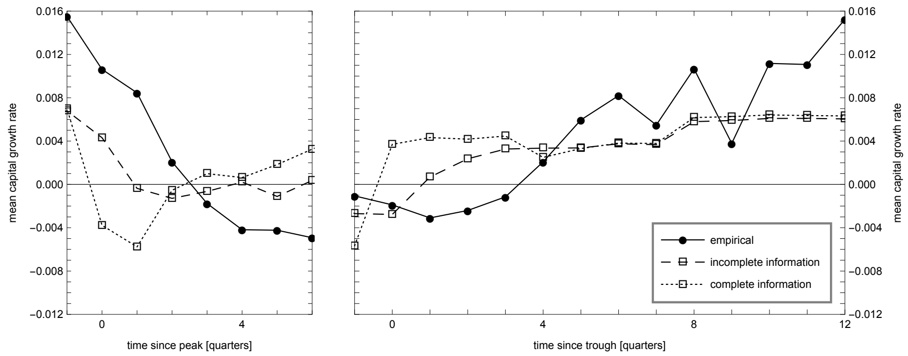
Our simulated capital growth rate under the baseline parameterization features a sharp decline at the beginning of the recession, and a slower recovery during the expansion. Although the recovery is still faster than empirically observed, overall the simulated curve matches the empirical curve reasonably well. The simulation under complete information shows a more dramatic decline and an immediate rebound, both much too sharp to be reconciled with the empirical curve. When information is complete, optimal

[21] The results are similar if we increase the number of firms in the simulation.

[22] As in Figure 1, the capital growth rate is averaged first across firms and then across business cycles.



(a) Average capital growth rate: Base case vs. complete information



(b) Average capital growth rate: Lower gap between μ_h and μ_l

Fig. 4: **Mean capital growth rate over the business cycle.** Panel (a) shows the simulated average quarterly capital growth rate during the recession (left diagram) and the expansion (right diagram) for both the base case and the complete information case, along with the empirical growth rate. In Panel (b), we set μ_h and μ_l at half of their base values and keep other parameter values unchanged (σ_s equals 0.25 in the incomplete information case, and 0.005 in the complete information case.) Quarter 0 is the business cycle peak in the left diagrams, and is the trough in the right diagrams.

capacity jumps immediately from one end of the investment boundary to the other as the macroeconomic regime switches. The sudden change of the optimal capacity leads to a dramatic change in the average capital growth rate at both regime switching points. When information is incomplete, the capital growth rate is smoother because of the time needed for learning.

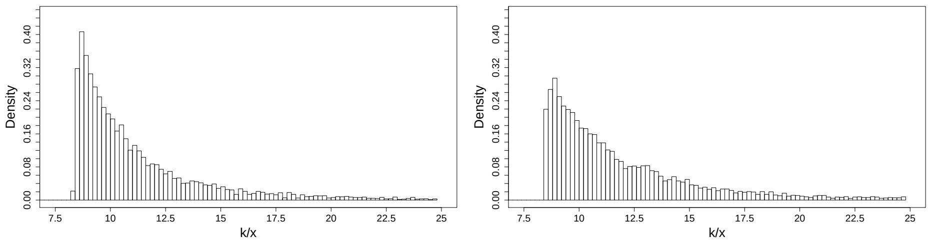
The difference in the capital growth rate between recession and expansion is obviously driven by the gap between the expected demand growth rates of these two states, $\mu_h - \mu_l$. A smaller gap makes the transition less dramatic, and potentially allows the complete information model to fit the empirical data better. To explore this possibility, we set μ_h and μ_l at half of their empirically estimated base values, and keep other parameter values unchanged. The results, plotted in Panel (b) of Figure 4, show that this alternative parameterization leads to an average capital growth rate that is too smooth over the business cycle compared to the empirically observed one. In particular, it is too low at the peak, with or without complete information. Furthermore, even with the smaller gap between $\mu_h - \mu_l$, the transition between business cycle regimes is still too abrupt under complete information.

4.3 DISTRIBUTION OF NORMALIZED CAPITAL OVER THE BUSINESS CYCLE

The reason for the asymmetry of decline and recovery in our baseline model is the endogenous distribution of firms relative to their optimal capacities. Since demand grows fast during the expansion, a large proportion of firms are pushed to the investment boundary at the end of an expansion. These are the marginal firms that react to changes in beliefs. When a negative signal comes, their optimal capacities decrease, and they stop investment immediately, generating a sharp decline in the average capital growth rate.

By contrast, at the end of a recession, not only is the optimal capacity relatively insensitive to a modest improvement in the belief, the proportion of firms that are close to the investment boundary is low because of the low demand growth. For firms that are far from the investment boundary, investment is not optimal even if their beliefs change significantly upon the arrival of a positive signal — the capacity they have accumulated during the recession is too high even compared to the high end of the investment boundary. Consequently, the overall reaction to the positive signal is small, and the recovery is slow.

To illustrate this point more clearly, we present histograms of simulated normalized capital (k/x) at two turning points of a business cycle in Figure 5. Panels (a) and (b) show the distributions at the end of an expansion and a recession, respectively. One can see clearly that in Panel (a) firms are more concentrated along the investment boundary, while in Panel (b) they are much more dispersed, indicating a larger proportion of firms with excess capacity. The proportion of firms at the disinvestment boundary is very small in both panels, indicating that firms in general are very reluctant to disinvest due to the discounted capital resale price.



(a) At the end of an expansion

(b) At the end of a recession

Fig. 5: **Histograms of normalized capital: Expansion vs. recession.** This figure shows the histograms of simulated normalized capital (k/x) at two turning points of a business cycle: at the end of an expansion (Panel (a)) and at the end of a recession (Panel (b)). The parameter values used for the simulation are in Table 1.

5. Skewness of Capital Growth Rates

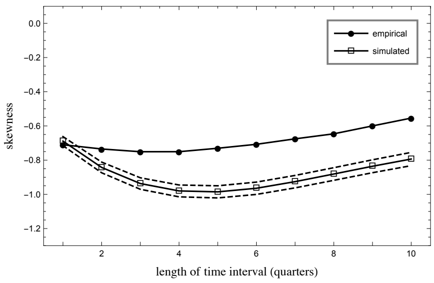
To further gauge the empirical plausibility of our model, we examine the skewness of capital growth rates, both at the individual firm level and in the aggregate.

5.1 BASE CASE

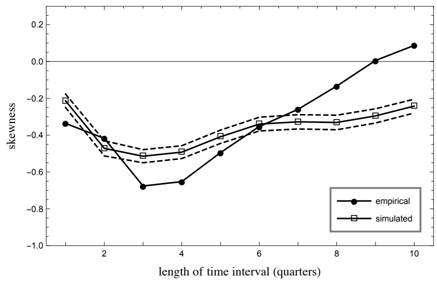
Figure 6 compares the simulated results in the base case with their empirical counterparts (reproduced from Figure 2). Panels (a) and (b) show, respectively, the skewness values for the levels and slopes of average capital growth rates measured over time intervals varying from one to ten quarters. Panels (c) and (d) show the average skewness values for the levels and slopes, respectively, of capital growth rates of individual firms. To gauge the magnitude of simulation errors, we plot both the point estimates of the skewness values and their 90% confidence intervals for the simulated data.[23]

At the aggregate level, our model generates the negative skewness in both the level (Panel (a)) and the slope (Panel (b)) of the average capital growth rate observed in the data. The estimated skewness values for both the level and the slope are non-monotonic in the length of time interval over which the growth rate is measured, first decreasing and then slowly increasing, as in

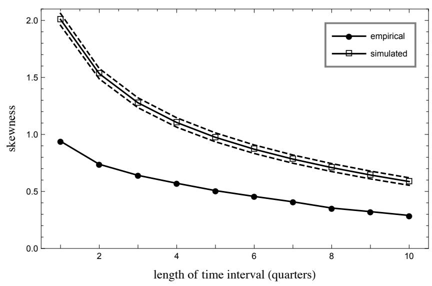
[23] The point estimate is the simple average across 100 rounds of simulations, and the standard error is the standard deviation divided by $\sqrt{100}$.



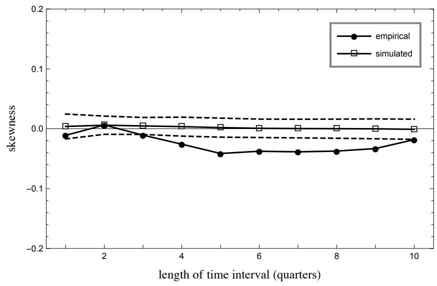
(a) Skewness of average capital growth rates



(b) Skewness of differenced average growth rates



(c) Skewness of individual capital growth rates



(d) Skewness of differenced individual growth rates

Fig. 6: **Skewness of capital growth rates: Empirical vs. simulated.** Panels (a) and (b) show, respectively, the skewness values for the levels and slopes of average capital growth rates measured over time intervals varying from one to ten quarters. Panels (c) and (d) show, respectively, the average skewness values for levels and slopes of capital growth rates of individual firms. The dashed lines indicate the 90% confidence intervals of the skewness estimates for the simulated data. The parameter values used for the simulation are in Table 1.

the empirical data. At the firm level, however, the simulated capital growth rate is strongly positively skewed in the level (Panel (c)), and has virtually no asymmetry in the slope (Panel (d)), again mimicking what we empirically observe. The positive level asymmetry is highest when the growth rate measured over one quarter, and deceases steadily as the measurement interval lengthens, but remains positive even when the measurement interval is ten quarters. These results demonstrate that our model is able to replicate the main empirical features of capital growth rates at both the firm and aggregate levels.

The positive level skewness at the firm level is a natural outcome of costly reversibility, which makes firms reluctant to disinvest.[24] Capital is built up for specific uses, and is therefore often highly illiquid in nature. Frictions in the secondary markets further increase the difficulty of reallocating capital from one firm to another. As a result, on the upside, firms may expand capital stock rapidly; yet on the downside, the decrease of capital stock normally occurs only at the rate of depreciation due to firms' unwillingness to sell capital at a discount. This leads to the positive skewness of the firm-level capital growth rate. However, the slope is largely symmetric, because an individual firm hits the investment boundary and returns to the inaction region at the same speed.

The negative level asymmetry of average capital growth rates arises because expansions usually last longer than recessions. This asymmetry of duration implies that most of the short-run growth rate observations are drawn from expansion periods, which lie to the right of the mean. The relatively small number of observations drawn from recession periods then form a long tail at the left side of the distribution, resulting in negative skewness.[25]

The negative slope asymmetry of the average capital growth rate is a key indicator of the sharp-decline-slow-recovery feature of corporate investment. As we discuss in Section 4.3, this results from the endogenous distribution of firms relative to their optimal capacities over the business cycle. More firms are close to the investment boundary during an expansion than during a recession. Therefore, more firms react to a negative signal arriving during an expansion than to a positive signal arriving during a recession. As a result, downturns are sharper upturns at the aggregate level, even though the slope at the firm level is symmetric.

5.2 REVERSIBILITY AND INDUSTRY INVESTMENT SLOPE ASYMMETRY

A key friction in our model is costly reversibility of capital investment, i.e., the illiquidity of capital assets. If capital assets can be liquidated at low costs (i.e., b is close to 1), then firms will not build up much excess capacity in the recession. This would allow them to respond quickly as economic signals

[24] Quantitatively, the positive skewness of the simulated firm-level capital growth rate appears too high. This is due to our assumption that firms can add capital instantaneously with zero adjustment costs. Introducing some adjustment costs for capital expansion will make the positive skewness less pronounced.

[25] A more technical explanation for the opposite signs of the skewness values of the average and the firm-level capital growth rates is provided in Appendix A.5.

become more positive. Therefore, we expect the decline and recovery of the average capital growth rate to be more symmetric if capital is more liquid.

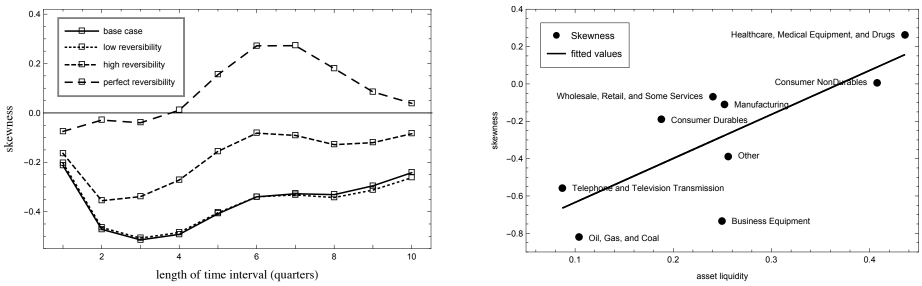
Panel (a) of Figure 7 confirms this intuition. In the figure, we plot the skewness values of the differenced average capital growth rates measured over various time horizons for different values of reversibility: $b = 0.5$, 0.8 (base case), 0.9, and 1.0 (perfect reversibility). The difference between the cases of $b = 0.8$ and $b = 0.5$ is negligible. In fact, even for $b = 0$, in which investment is completely irreversible, the results are still very similar to those in the base case. This suggests that a discount of 20 percents in the resale price of capital is sufficient to capture the main effect of irreversibility in our model. However, as b increases from 0.8 to 0.9, the skewness values become less negative for all time horizons. For $b = 1.0$, the skewness even becomes positive for most time horizons.[26] These results suggest that higher reversibility reduces the slope asymmetry of investment at the aggregate level. They also demonstrate that the slope asymmetry does not arise mechanically from the assumed Markov regime-switching process. A certain degree of irreversibility is necessary to generate the fast decline and slow recovery in our model.

Panel (b) of Figure 7 provides some empirical support for the connection between the aggregate level slope asymmetry and the degree of irreversibility, using data from nine of the ten Fama-French industries[27] We measure investment reversibility at the industry level by the activeness of secondary market capital reallocation. The assumption is that if secondary market reallocation is active, then capital is more liquid, and investment reversibility is high. Following Eisfeldt and Rampini (2006), we measure industry-wide capital reallocation by the sum of acquisitions and sales of property, plant, and equipment (Compustat items AQC and SPPE, respectively), and scale it by the industry capital expenditures (CAPX in Compustat). We calculate this reallocation ratio for each industry year by year from 1971 to 2011 (AQC and SPPE data are not available from Compustat for earlier years), and use the time series average as a measure of industry capital asset liquidity.

To measure the industry-level investment slope asymmetry, we calculate industry-level average capital growth rates over different time intervals, from one to five quarters, using quarterly data from 1975 through 2011. We then take a first-order difference of each series and estimate the skewness of the differenced series. We use the average of skewness values estimated from

[26] Appendix A.5 shows analytically why the slope of the average capital growth rate is symmetric under perfect reversibility using a discrete version of our model.

[27] We exclude the utility industry, and exclude financial firms from the "Other" industry. We use the ten-industry instead of a finer classification to make sure that each industry has a sufficiently large number of firms to reveal the aggregation effects.



(a) Skewness and reversibility: Simulated

(b) Asset liquidity and skewness across industries

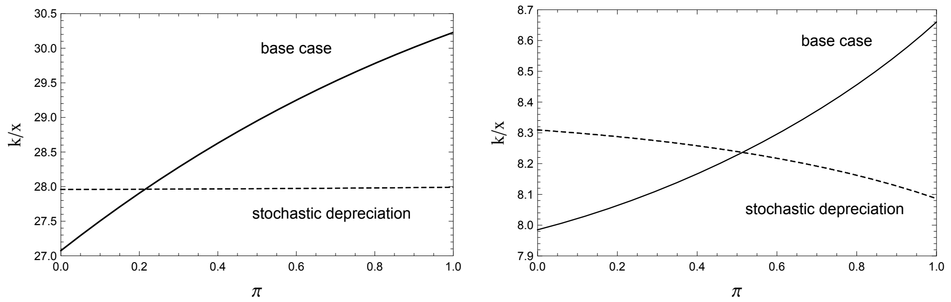
Fig. 7: **Reversibility and industry-level slope asymmetry.** Panel (a) shows skewness values of the simulated differenced average capital growth rate, measured over various time horizons, for different degrees of reversibility. The resale price of capital (b) varies from 0.5 (low reversibility), 0.8 (base case), 0.9 (high reversibility), to 1.0 (perfect reversibility). Other parameter values are given in Table 1. Panel (b) shows the empirical relation between the skewness of the differenced industry-level capital growth rate and asset liquidity for nine of the ten Fama-French industries.

the five differenced series as a summary measure of the industry-level slope asymmetry.

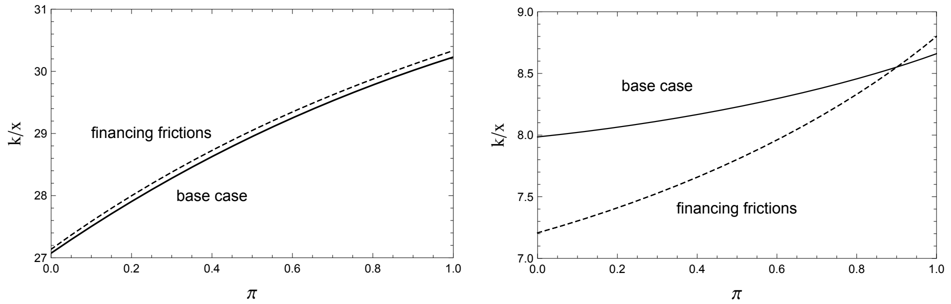
Panel (b) of Figure 7 plots the average skewness value against asset liquidity industry by industry. Consistent with the prediction of our model, the graph shows clearly that the slope asymmetry is most prominent in industries with low asset liquidity (Oil, Gas and Coal; Telephone and Television Communication). For industries with high asset liquidity (Consumer Nondurables and Healthcare), the skewness of the slope is either zero or positive. A univariate regression of the slope skewness on asset liquidity produces a positive coefficient of 2.36 with a t-statistic of 3.18.

6. Extensions

We extend our model to allow for two possibilities, procyclical depreciation and countercyclical financing frictions.



(a) Disinvestment boundary and depreciation (b) Investment boundary and depreciation

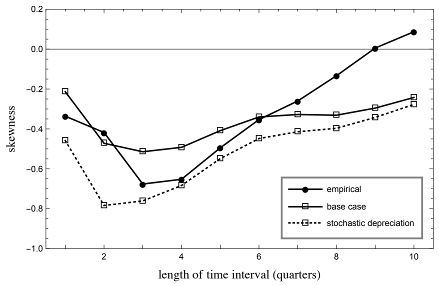


(c) Disinvestment boundary and financing frictions (d) Investment boundary and financing frictions

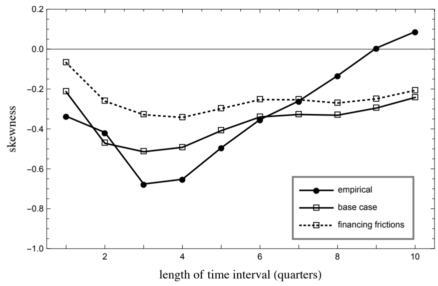
Fig. 8: **Investment/disinvestment boundaries: Stochastic depreciation and financing frictions.** Panels (a) and (b) show, respectively, the disinvestment and investment boundaries for the procyclical depreciation case ($\xi_0 = 10\%$, and $a_\xi = 3\%$). Panels (c) and (d) show, respectively, the disinvestment and investment boundaries for the countercyclical financing frictions case ($a_\xi = 0.025$). For comparison, the base case results (solid curves) are also shown in each panel.

6.1 PROCYCLICAL DEPRECIATION RATE

Capital depreciation rates, especially in economic instead of accounting terms, may be higher in expansion than in recession, because technologies develop faster when the economy is booming. To allow for the possibility of a procyclical depreciation rate, we specify the following functional form for



(a) Stochastic depreciation and slope asymmetry



(b) Financing frictions and slope asymmetry

Fig. 9: **Slope asymmetry in the aggregate: Stochastic depreciation and financing frictions.** This figure shows the skewness of the first-order differences of average capital growth rates measured over different time horizons (from one quarter to ten quarters). Panel (a) compares the results from the stochastic depreciation case ($\xi_0 = 10\%$, and $a_\xi = 3\%$) with those from the benchmark case, as well as the empirical estimates. Panel (b) compares results from the countercyclical financing frictions case ($a_\xi = 0.05$) with those from the benchmark case, as well as the empirical estimates.

capital depreciation:

$$\xi = \xi_0 + a_\xi \pi, \tag{13}$$

where ξ_0 is the depreciation rate in a recession, and $a_\xi > 0$ is a constant governing the sensitivity of depreciation to the belief π, which is stochastic and path-dependent. For the purpose of illustration, we set $\xi_0 = 10\%$, and $a_\xi = 3\%$, and keep other parameters at their base values. Since no additional state variable is introduced, this extended model can be solved with minor modifications of our numerical procedure.

Panels (a) and (b) of Figure 8 show that when the depreciation rate is procyclical, the disinvestment boundary becomes flat and the investment boundary becomes a downward sloping function of the belief π. This is because higher depreciation reduces the marginal value of capital during an expansion. Therefore, the optimal capacity, normalized by the realized demand factor, is lower in expansion than in recession. Similarly, the firm is now more willing to reduce capital during an expansion than in the base case.

Panel (a) of Figure 9 shows the effects of the procyclical depreciation rate on the slope asymmetry of the average capital growth rate. It shows that the slope asymmetry is stronger in the extended model than in either the base case or the empirical data. This is because the higher depreciation associated with a higher π reduces firms' incentives to invest at the beginning of an expansion, which further slows down the recovery, and amplifies the slope asymmetry.[28]

6.2 FINANCING FRICTIONS

Another important factor potentially affecting the dynamics of investment is the costs of raising external finance. Financing frictions are especially severe during a recession, when both financial intermediaries and investors are reluctant to lend or invest. To investigate the impact of financing frictions as parsimoniously as possible, we introduce an equity issuance cost ϕ, which is a function of a firm's path-dependent belief π. We normalize this cost to zero during an expansion, and assume that it increases linearly with the posterior probability of being in a recession $(1-\pi)$. Specifically,

$$\phi(\pi) = a_\phi(1-\pi), \tag{14}$$

where a_ϕ is a constant specifying the cost of raising \$1 equity during a recession. Since firms in our model do not hold any cash, the effective cost of investing is $1+\phi(\pi)$, i.e., the normalized price of capital plus the financing costs. Following Bolton et al. (2011), the value-matching condition at the investment boundary then becomes

$$\lim_{k \to x\, h_i^*} \frac{\partial V}{\partial k} = 1 + a_\phi(1-\pi),$$

or equivalently,

$$\lim_{h \to h_i^*} \frac{\partial v(h,\pi)}{\partial h} = 1 + a_\phi(1-\pi).$$

This extended model can also be solved relatively easily, because h and π are still the only two state variables. For the purpose of illustration, we set $a_\phi = 0.025$, and keep the other parameters at their base values. Figure 8 shows that with financing frictions, the investment boundary (Panel (d)) increases more sharply with the belief π than in the base case, while the disinvestment boundary (Panel (c)) shifts upward slightly. Financing

[28] Because the investment boundary is downward sloping in π in the extended model with a procyclical depreciation rate, we find that under complete information, the average capital growth rate, counterfactually, shoots up sharply at the beginning of a recession, and jumps down abruptly at the beginning of an expansion.

frictions further reduce firms' incentives to invest during a recession, generating a steeper investment boundary. At the same time, financing frictions also make firms slightly more reluctant to cut capital. By tolerating a larger amount of excess capital, firms can reduce the probability of paying financing costs in the future.

While these results on the effects of financing frictions make perfect economic sense, the extended model does not seem to generate a better empirical fit. As Panel (b) of Figure 9 shows, in general the model with countercyclical financing frictions generates less asymmetry in the slopes of the average capital growth rate, especially when the growth rate is measured over relatively short time horizons.

Why do countercyclical financing frictions not help to explain the slope asymmetry in our model? One reason is that they increase and decrease at the same speed. If we assume that financing frictions increase sharply as the economy enters a recession, and decrease slowly during a recovery, the slope asymmetry of the aggregate investment would be amplified. While this assumption may be realistic, it is not very appealing as it is linked almost mechanically to the empirical facts we set out to explain. We therefore leave it for future research.

7. Conclusion

The sharp decline of corporate investment after a negative macroeconomic shock and its slow recovery thereafter are at the center of many recent policy discussions. We provide a micro-founded explanation for this asymmetry, accounting for new empirical evidence on the asymmetries of capital growth rates at the aggregate and firm levels. Our dynamic investment model features various degrees of reversibility, cyclical macroeconomic shocks, and uncertainty about the true state of the economy.

In our baseline model, we show that a firm's optimal investment threshold, defined in terms of a lower bound on the firm's capital normalized by the demand factor, is a convex function of its posterior belief that the economy is in an expansion. Monte Carlo simulations of a large panel of firms facing heterogeneous shocks generate patterns that match the empirical data remarkably well at both the firm and aggregate levels. In particular, the simulated average capital growth rate features both a negative slope asymmetry and a negative level asymmetry, while the capital growth rate at the firm level features no slope asymmetry and a positive level asymmetry. Furthermore, the aggregate level slope asymmetry is more severe when the degree of investment reversibility is lower.

Our model provides a new perspective on understanding firms' investment behavior, and has clear policy implications. A key friction leading to slow recovery is incomplete information about the macroeconomic state. Policies that help to reduce such uncertainty can boost investment by reducing the option value of waiting. One example is a central bank's announced commitment to a certain interest rate policy. Other examples include legislation reducing the uncertainty of future government budget and tax policies. Another key friction in our model is costly reversibility, i.e., the illiquidity of physical capital. Policies that facilitate the reallocation of capital in the secondary markets, either through merger and acquisition or through asset sales, can reduce the excess capacity of firms facing unfavorable shocks, which in turn would boost investment as the macroeconomic climate improves.

One limitation of our model is that it abstracts from the feedback effects of firm investment on product and capital prices. Analysis based on reduced-form specifications of capital price shows that the recovery of investment activity in the upturn is further delayed if the price of capital is procyclical in our model. Therefore, accounting for the general equilibrium effect of investment on capital price can potentially improve the empirical fit, but is unlikely to change our basic results. We view this line of extension as a fruitful avenue for future research.

Appendices

A.1 ALTERNATIVE FORMULATION OF THE OPTIMAL UPDATING RULE

Equation (5) can be rewritten as:

$$\begin{aligned} d\pi_t = {} & [-\pi_t\lambda_{h,l} + (1-\pi_t)\lambda_{l,h}]dt \\ & + \frac{(\mu_h - \mu_l)\pi_t(1-\pi_t)}{\sigma^2}\mathbf{w}\begin{pmatrix} d\ln(x_t) - E_t(\mu_t|\mathcal{F}_t)dt \\ d\ln(s_t) - E_t(\mu_t|\mathcal{F}_t)dt \end{pmatrix} \end{aligned} \tag{A1}$$

where

$$\sigma^2 \equiv \frac{1}{\mathbf{1}'(\Phi\Phi')^{-1}\mathbf{1}}, \tag{A2}$$

$$\mathbf{w} \equiv \frac{\mathbf{1}'(\Phi\Phi')^{-1}}{\mathbf{1}'(\Phi\Phi')^{-1}\mathbf{1}} = \left(\frac{\sigma_s^2 - \rho\sigma_x\sigma_s}{\sigma_x^2 + \sigma_s^2 - 2\rho\sigma_x\sigma_s}, \quad \frac{\sigma_x^2 - \rho\sigma_x\sigma_s}{\sigma_x^2 + \sigma_s^2 - 2\rho\sigma_x\sigma_s}\right). \tag{A3}$$

Note that $\Phi\Phi'$ is simply the instantaneous variance-covariance matrix of $d\ln(x_t)$ and $d\ln(s_t)$. Readers familiar with the classic mean-variance portfolio analysis will immediately recognize that σ^2 is the minimum instantaneous variance that can be obtained using all possible linear combinations of $d\ln(x_t)$ and $d\ln(s_t)$, while $\mathbf{w}$ specifies the weights of each individual signal in the minimum variance combination. This formulation thus reveals an important feature of the Bayesian learning process. When there are multiple jointly normally distributed signals, the agent can form a minimum variance "portfolio" of all the available signals, and base learning on this compound signal.

The standard deviation of this optimally constructed compound signal, σ, measures the noisiness of the overall information of all the signals. The learning Equation (5) thus implies a stronger response to forecasting errors is when signals are more precise. Furthermore, by the nature of the minimum variance portfolio, the optimal weighting vector $\mathbf{w}$ assigns more weight to the signal with lower variance, indicating that agents pay more attention to the signal that has less noise. In particular, when $\sigma_s = \rho\sigma_x$, the optimal weight of $d\ln(x_t)$ in the compound signal is zero. Learning is based entirely on the signal $d\ln(s_t)$.

A.2 PROOF OF PROPOSITION 2

Consider firm value V as a claim on the firm's operating profit as a function of its current demand factor x, the installed capital k, and the belief about the current state of the economy π, i.e., $V = V(x, k, \pi)$. For the risk-neutral decision maker, the value function must satisfy the following Hamilton-Jacobi-Bellman equation:

$$rV(x,k,\pi)dt = f(x,k)\ dt + E(dV(x,k,\pi)|\mathcal{F}_t) \tag{A4}$$

The expectation of dV is to be determined by Itô's lemma using the $\mathcal{F}_t$-dynamics of the state variables x, k, and π. Note that since $dW_{xt}^{\mathcal{F}}$ and $dW_{st}^{\mathcal{F}}$ are uncorrelated, we have

$$\begin{aligned}
dx &= x(\pi\mu_h + (1-\pi)\mu_l + \frac{1}{2}\sigma_x^2)dt + x\sigma_x dW_{xt}^{\mathcal{F}},\\
(d\pi)^2 &= d\pi(d\pi)'\\
&= [\pi(1-\pi)(\mu_h-\mu_l)]^2 \left[\mathbf{1}'(\Phi')^{-1}\mathbf{dW}_t^{\mathcal{F}}\right]\left[\mathbf{1}'(\Phi')^{-1}\mathbf{dW}_t^{\mathcal{F}}\right]'\\
&= [\pi(1-\pi)(\mu_h-\mu_l)]^2\frac{1}{\sigma^2}dt,\\
dx\,d\pi &= dx\,(d\pi)'\\
&= x\pi(1-\pi)(\mu_h-\mu_l)\left[(1,0)\,\Phi\mathbf{dW}_t^{\mathcal{F}}\right]\left[\mathbf{1}'(\Phi')^{-1}\mathbf{dW}_t^{\mathcal{F}}\right]'\\
&= x\pi(1-\pi)(\mu_h-\mu_l)dt.
\end{aligned}$$

Therefore, in the inaction region, where $dk = -\xi k dt$, we have:

$$\begin{aligned}E[dV(x,k,\pi)|\mathcal{F}_t] = \Bigg[& -\frac{\partial V}{\partial k}\xi k + \frac{\partial V}{\partial x}x[\pi\mu_h + (1-\pi)\mu_l + \frac{1}{2}\sigma_x^2] \\ & + \frac{1}{2}\frac{\partial^2 V}{\partial x^2}x^2\sigma_x^2 + \frac{\partial V}{\partial \pi}[-\pi\lambda_{h,l} + (1-\pi)\lambda_{l,h}] \\ & + \frac{1}{2}\frac{\partial^2 V}{\partial \pi^2}[\pi(1-\pi)(\mu_h-\mu_l)]^2\frac{1}{\sigma^2} \\ & + \frac{\partial^2 V}{\partial x \partial \pi}x\pi(1-\pi)(\mu_h-\mu_l)\Bigg] dt.\end{aligned}$$

Substituting this expression into the Hamilton-Jacobi-Bellman Equation (A4) and dropping dt from both sides of the equation yields

$$\begin{aligned}rV = & \frac{1}{1-\alpha}x^{\alpha}k^{(1-\alpha)} - \frac{\partial V}{\partial k}\xi k + x[\pi\mu_h + (1-\pi)\mu_l + \frac{1}{2}\sigma_x^2]\frac{\partial V}{\partial x} \\ & + \frac{1}{2}x^2\sigma_x^2\frac{\partial^2 V}{\partial x^2} + [-\pi\lambda_{h,l} + (1-\pi)\lambda_{l,h}]\frac{\partial V}{\partial \pi} \\ & + \frac{[\pi(1-\pi)(\mu_h-\mu_l)]^2}{2\sigma^2}\frac{\partial^2 V}{\partial \pi^2} + x\pi(1-\pi)(\mu_h-\mu_l)\frac{\partial^2 V}{\partial x\partial \pi}\end{aligned} \tag{A5}$$

The last part of the proof to show that writing $V(x,k,\pi)$ as $V = xv(h,\pi)$ with $h = \frac{k}{x}$ gives Equation (9). This is done by substituting the partial derivatives below in Equation (A5):

$$\begin{aligned}\frac{\partial V}{\partial k} &= \frac{\partial v(h,\pi)}{\partial h}, \\ \frac{\partial V}{\partial x} &= v(h,\pi) - h\frac{\partial v(h,\pi)}{\partial h}, \\ \frac{\partial^2 V}{\partial x^2} &= \frac{1}{x}h^2\frac{\partial^2 v(h,\pi)}{\partial h^2}, \\ \frac{\partial V}{\partial \pi} &= x\frac{\partial v(h,\pi)}{\partial \pi}, \\ \frac{\partial^2 V}{\partial \pi^2} &= x\frac{\partial^2 v(h,\pi)}{\partial \pi^2}, \\ \frac{\partial^2 V}{\partial x\partial \pi} &= \frac{\partial v}{\partial \pi} - h\frac{\partial^2 v(h,\pi)}{\partial x\partial \pi}.\end{aligned}$$

A.3 CALIBRATION

We use the annual Compustat-CRSP merged database over 1950-2011 to estimate most of our paramters. Note that from Equation (1) we can back out a firm's demand factor x_t using its operating profit $f(x_t, k_t)$ and capital stock k_t:

$$x_t = \left[\frac{(1-\alpha)f(x_t, k_t)}{k_t^{1-\alpha}}\right]^{1/\alpha}. \tag{A6}$$

This formula allows us to compute x_t for each individual firm in each year. We measure $f(x_t, k_t)$ by operating income before depreciation (OIBDP), and k_t by operating assets (defined as total assets (AT) minus cash and short-term investment (CHE)), all converted into year 2005 dollar values using the annual GDP deflator.[29]

Using the estimated x_t series, we calculate the continuously compounded annual growth rate of x_t at the firm level, and winsorize it at the top and bottom 2.5% of each sample year. We then average it across firms, weighted by the lagged operating asset value. This gives us a time series of 61 annual observations. We apply the Expectation-Maximization (EM) algorithm developed by Dempster et al. (1977) to estimate the parameters of the hidden Markov chain: $\lambda_{h,l}$, $\lambda_{l,h}$, μ_h, μ_l.

Our estimates of the transition intensities imply an expected length of 3.9 years ($= 1/0.2557$) for an expansion and 1.4 years ($= 1/0.7163$) for a recession. The estimated π_t series matches the NBER-dated business cycles very well, as shown in Figure 10, where we plot the time series of the estimated average demand factor growth rate and the posterior beliefs π_t. The bars indicate the historical recession periods according to the NBER.

The depreciation rate ξ is estimated using the median value of the ratio of depreciation of tangible fixed assets (DFXA) to the lagged PPENT. For the instantaneous volatility of the demand factor, σ_x, we regress the annual demand growth rate on an NBER recession dummy firm by firm (requiring firms to have at least 25 annual observations), and use a simple average of the root mean squared errors of these regressions as its estimate.

The other parameter values are set as follows. We set the instantaneous volatility, σ_s, of the public signal to be 0.25. This allows us to capture the idea that there is a significant amount of uncertainty about the true economic state. If this value is too small, then learning will be too fast to

[29] We set $\alpha = 0.74$, along the line of Guo et al. (2005). The operating profit function (1) approximates the following specification: (1) Constant returns to scale Cobb-Douglas production function with labor and capital: $q = \lambda L^{\phi} K^{1-\phi}$; (2) Isoelastic demand function given by the inverse demand curve: $p = x^{1-\theta} q^{\theta-1}$, where $0 < \theta < 1$. It follows from this specification that the share of profits going to capital depends on θ and ϕ through the relation: $1 - \alpha = (1-\phi)\theta/(1-\theta\phi)$. Labor's share of national income in the U.S. since World War II is relatively stable at $\phi = 0.64$. Assuming $\theta = 0.5$, we get $\alpha \approx 0.74$. Guo et al. (2005) obtain $\alpha \approx 0.53$ due to a typo in their expression for $1 - \alpha$.

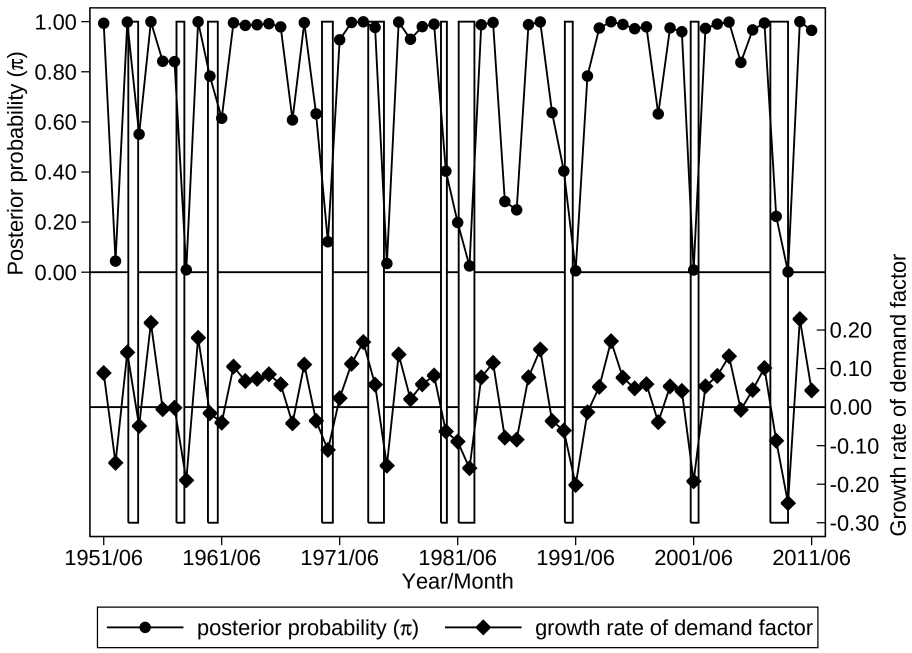


Fig. 10: **Growth rate of demand factor and posterior probability of expansion.** The upper part shows the time series of the posterior belief π estimated using the EM algorithm. The bottom part (axis on the right) shows the average continuously-compounded annual growth rate of the demand factor x, estimated using the Compustat-CRSP merged database. The bars indicate the recession periods dated by the NBER.

have a significant impact. The risk-free rate, r, is 0.05. The resale price of capital, b, is 0.80. The instantaneous correlation, ρ, between the public signal and the firm's demand factor is 0.05. Experiments show that the main features of our results are fairly robust to alternative values of these parameters.

A.4 NUMERICAL OPTIMIZATION OF BOUNDARIES

We solve the Hamilton-Jacobi-Bellman equation (9) with the boundary conditions by solving numerically the underlying stochastic dynamic programming problem. We discretize (A4) as

$$V(x, f, \pi)\Delta t = f(x, k)\Delta t + e^{-r\Delta t}E[V(x + \Delta x, k + \Delta k, \pi + \Delta\pi)],$$

and use its homogeneity property in k to write the program in terms of Tobin's average Q:

$$V(x,t,\pi) = kV(x/k,1,\pi) = kQ(g,\pi),$$

with $g = x/k$. Since in the inaction region, where the firm neither invests nor disinvests, capital, k, is constant, g_t is just a scaled version of x_t there.

In the investment region, we set Tobin's marginal Q equal to 1 (the purchase price of capital), and in the disinvestment region we set it equal to b (the resale price of capital). We derive

$$\frac{\partial}{\partial k}[V(x,h,\pi)] = \frac{\partial}{\partial k}\left[k\,Q\left(\frac{x}{k},\pi\right)\right] = Q(g,\pi) - g\frac{\partial Q(g,\pi)}{\partial g}.$$

Hence, for given investment and disinvestment boundaries $g_i^*(\pi)$ and $g_d^*(\pi)$, respectively, the value function inside the action regions is set according to

$$\begin{aligned}
Q(g,\pi) &= \frac{g}{g_i^*(\pi)}Q(g_i^*(\pi),\pi) - \frac{g-g_i^*(\pi)}{g_i^*(\pi)}\ , \qquad &&\text{for} \quad g > g_i^*(\pi),\\
Q(g,\pi) &= \frac{g}{g_d^*(\pi)}Q(g_d^*(\pi),\pi) - \frac{g-g_d^*(\pi)}{g_d^*(\pi)}b, \qquad &&\text{for} \quad g < g_d^*(\pi).
\end{aligned}$$

Inside the inaction region, we model the joint dynamics of g and π as a two dimensional binomial tree. The difficulty in doing so is that π follows a mean reversion process with non-constant volatility. Therefore, we employ the approach of Nelson and Ramaswamy (1990). Consider the process

$$Z(\pi) = \int_{\frac{1}{2}}^{\pi} \frac{\sigma}{(\mu_h-\mu_l)p(1-p)}dp = \frac{\sigma}{\mu_h-\mu_l}\ln\left(\frac{\pi}{1-\pi}\right),$$

where σ is defined as the positive square root of σ^2 in Equation (A2).

Then the process $z_t = Z(\pi_t)$ has a constant volatility of 1 and follows the dynamics

$$\begin{aligned}
dz &= \underbrace{\left[\mu_\pi\frac{dZ(\pi)}{d\pi} + \frac{1}{2}\sigma_\pi^2\frac{d^2Z(\pi)}{d\pi^2}\right]}_{\mu_z} dt + \underbrace{\sigma\mathbf{1}'(\Phi')^{-1}\mathbf{dW_t^{\mathcal{F}}}}_{dW_{zt}^{\mathcal{F}}},\\
z_0 &= Z(\pi_0),
\end{aligned}$$

where σ_π is the volatility and μ_π is the drift of the belief (see Proposition 1) given by

$$\begin{aligned}
\sigma_\pi &= \frac{(\mu_h-\mu_l)\pi(1-\pi)}{\sigma},\\
\mu_\pi &= -\pi\lambda_{hl} + (1-\pi)\lambda_{lh},
\end{aligned}$$

and $W_{zt}^{\mathcal{F}}$ is a standard Brownian motion. Consequently, z_t and g_t can be modeled on a two-dimensional binomial tree following Boyle et al. (1989), taking into account the correlation of the two processes that is implicitly given. Then, the belief process π results from applying the inverse, $\pi_t = Z^{-1}(z_t)$, which is given by

$$Z^{-1}(z) = \frac{1}{1 + \exp\left\{-\left(\frac{\mu_h - \mu_l}{\sigma}\right) z\right\}}.$$

The optimization of the free boundaries g_i^* and g_d^* is achieved via value function iteration with the goal to have $Q(g, \pi)$ smooth at the boundaries, which implies that Tobin's marginal Q converges to 1 when moving from the inaction region to the investment boundary, and that it converges to b when moving from the inaction region to the disinvestment boundary.

The boundaries discussed in the text, h_i^* and h_d^*, are calculated simply by

$$h_i^* = \frac{1}{g_i^*}, \quad h_d^* = \frac{1}{g_d^*}.$$

A.5 RELATION BETWEEN AVERAGE SKEWNESS AND SKEWNESS OF THE AVERAGE

Let c_i denote the de-meaned capital growth rate of firm i over one period, and assume that the unconditional moments of firms are identical. The skewness of the average over the capital growth rate of $N > 2$ firms can be written as

$$\text{Skew}\left(\frac{1}{N}\sum_{i=1}^{N} c_i\right) = E\left[\frac{(\frac{1}{N}\sum_{i=1}^{N} c_i)^3}{[\sigma(\frac{1}{N}\sum_{i=1}^{N} c_i)]^3}\right]$$

$$= \frac{\frac{1}{N^3}\left[NE[(c_i)^3] + 3N(N-1)E_{i\neq j}[c_i^2 c_j] + 6\frac{N(N-1)(N-2)}{3!}E_{i\neq j\neq k}[c_i c_j c_k]\right]}{\frac{1}{N^3}\left[NE[c_i^2] + N(N-1)E_{i\neq j}[c_i c_j]\right]^{3/2}},$$

where $\sigma(.)$ denotes the standard deviation. This equation shows that the skewness of the average is determined by (1) the skewness of an individual firm $E[(c_i)^3]$; and (2) co-skewness terms $E_{i\neq j}[c_i^2 c_j]$ and $E_{i\neq j\neq k}[c_i c_j c_k]$. As N increases, the first component becomes less and less important, and the second component becomes dominant.

If capital growth rates of individual firms are skewed, but are independent, then the co-skewness terms are zero (i.e., $E_{i\neq j}[c_i^2 c_j] = E_{i\neq j\neq k}[c_i c_j c_k] = 0$), so is $E_{i\neq j}[c_i c_j]$ as well. Asymptotically, the skewness of the average capital growth rate is then

$$\lim_{N\to\infty} \mathrm{Skew}(\frac{1}{N}\sum_{i=1}^{N} c_i) \;=\; \lim_{N\to\infty} \frac{NE[(c_i)^3]}{\left[NE[c_i^2]\right]^{3/2}} = 0. \tag{A7}$$

When capital growth rates are correlated, the asymptotic behavior of the average is determined by the $N(N-1)(N-2)$ co-skewness terms in the numerator, $E_{i\neq j\neq k}[c_i c_j c_k]$, and the $N(N-1)$ covariance terms in the denominator, $E_{i\neq j}[c_i c_j]$:

$$\lim_{N\to\infty} \mathrm{Skew}(\frac{1}{N}\sum_{i=1}^{N} c_i) = \frac{E_{i\neq j\neq k}[c_i c_j c_k]}{\left(E_{i\neq j}[c_i c_j]\right)^{3/2}}. \tag{A8}$$

Thus, when we take an average across a set of positively skewed independent and identically distributed random variables, the skewness of the average vanishes as the number of random variables increases; i.e., the general central limit theorem is at work, and the average becomes more and more normal. To make the average of positively skewed variables negatively skewed, a model needs an appropriate structure that generates negative co-skewness. Our model provides such a structure. Since the business cycle is a common factor and since it is negatively skewed (expansions on average last longer than recessions, implying that recessions are farther below the long-run mean than expansions are above it), the absolute value of $c_i c_j c_k$ is on average bigger when the three terms involved are all negative than when they are all positive, thus the sum of co-skewness terms is negative.

While the argument above explains the opposite level asymmetries at the aggregate and firm levels, it is important to note that imposing a negatively skewed common factor does not necessarily lead to a negative slope asymmetry at the aggregate level. When investment is perfectly reversible, it can be shown that the slope of the capital growth rate is asymptotically symmetric at both the firm and aggregate levels. To get the intuition for this result, consider a discrete version of our model, in which regime switches occur only at the quarter end.

Consider first a single firm. The probability of switching from a high-growth to a low-growth state is $P_{hl} = 1 - e^{-0.25\lambda_{hl}}$, and the probability of the opposite switch is $P_{lh} = 1 - e^{-0.25\lambda_{lh}}$. Since the demand growth rate is normal within each regime, and the local volatility is the same in both regimes, the asymptotic distribution of changes in the quarterly demand growth rate is then a mixture of three normal distributions with the same standard deviation ($\frac{\sqrt{2}}{2}\sigma$) but different means: 0 (no switch), $0.25(\mu_l - \mu_h)$ (a switch from low growth to high growth), and $0.25(\mu_h - \mu_l)$ (an opposite switch). Given the unconditional probability of each regime ($\frac{P_{lh}}{P_{lh}+P_{hl}}$ for high growth, and $\frac{P_{hl}}{P_{lh}+P_{hl}}$ for low growth), the probability of the distribution with mean 0 is $1-\frac{2P_{lh}P_{hl}}{P_{lh}+P_{hl}}$, while the probability of each of the other two is $\frac{P_{lh}P_{hl}}{P_{lh}+P_{hl}}$. This symmetric mixture of three normals is obviously symmetric. Since the capital growth rate is identical to the demand growth rate under perfect reversibility, this result explains why the capital growth rate at the firm level has zero skewness in the slope.

The average capital growth rate is also normal within each regime, and has the same local volatility under both regimes. This is true because within each regime it is a linear combination of N identically distributed, jointly normal random variables. The same argument above implies that the distribution of the change in the average quarterly capital growth rate is also a symmetric mixture of three normals, and is therefore symmetric.

The logic for the perfect reversibility case does not hold when investment is irreversible or reversible only at a cost. In the presence of such frictions, the capital growth rate, either at the firm level or in the aggregate, is no longer a mixture of normals due to the existence of an inaction region, even though the demand growth rate is. A negative slope asymmetry at the aggregate level then emerges as a consequence of firms' optimal investment decisions in recognition of such frictions.

References

Abel, A. B. and Eberly, J. C. (1996) Optimal investment with costly reversibility, *Review of Economic Studies* 63, 581–593.

Acemoglu, D. and Scott, A. (2003) Asymmetric business cycles: Theory and time-series evidence, *Journal of Monetary Economics* 40, 501–533.

Albuquerque, R. (2012) Skewness in stock returns: Reconciling the evidence on firm versus aggregate stock returns, *Review of Financial Studies* 25, 1630–1673.

Alti, A. (2003) How sensitive is investment to cash flow when financing is frictionless? *Journal of Finance* 58, 707–722.

Bloom, N. (2009) The impact of uncertainty shocks, *Econometrica* 77, 623–685.

Bolton, P., Chen, H. and Wang, N. (2011) A unified theory of Tobin's q, corporate investment, financing, and risk management, *Journal of Finance* 66, 1545–1578.

Boyle, P. P., Evnine, J. and Gibbs, S. (1989) Numerical evaluation of multivariate contingent claims, *Review of Financial Studies* 2, 241–250.

Chamley, C. and Gale, D. (1994) Information revelation and strategic delay in a model of investment, *Econometrica* 62, 1065–1085.

David, A. (1997) Fluctuating confidence in stock markets: Implications for returns and volatility, *Journal of Financial and Quantitative Analysis* 32, 427–462.

Decamps, J.-P., Mariotti, T. and Villeneuve, S. (2005) Investment timing under incomplete information, *Mathematics of Operations Research* 20, 472–500.

DeLong, B. and Summers, L. (1986) Are business cycles symmetric? in Robert Gordon, ed.: *American Business Cycle: Continuity and Change*, University of Chicago Press, Chicago.

Dempster, A., Laird, N. and Rubin, D. (1977) Maximum likelihood from incomplete data via the em algorithm, *Journal of the Royal Statistical Society* Series B 39, 1–38.

Dixit, A. and Pindyck, R. (1994) *Investment Under Uncertainty*, Princeton University Press, Princeton, New Jersey.

Dumas, B. (2001) Super contact and related optimality conditions, *Journal of Economic Dynamics and Control* 15, 675–685.

Eisfeldt, A. L., (2004), Endogenous liquidity in asset markets, *Journal of Finance* 59, 1–30.

——— , and Rampini, A. A. (2006) Capital reallocation and liquidity, *Journal of Monetary Economics* 33, 369–399.

Falk, B. (1986) Further evidence on the asymmetric behavior of economic time series over the business cycle, *Journal of Political Economy* 94, 1096–1109.

Gale, D. (1996) Delay and cycles, *Review of Economic Studies* 63, 169–198.

Grenadier, S. R., and A. Malenko, 2010, A Bayesian approach to real options: The case of distinguishing between temporary and permanent shocks, *Journal of Finance* 65, 1949–1986.

Guo, X., Miao, J. and Morellec, E. (2005) Irreversible investment with regime shifts, *Journal of Economic Theory* 132, 37–59.

Jorgenson, D. W. (1963) Capital theory and investment behavior, *American Economic Review Papers and Proceedings* 53, 247–259.

Klein, M. (2009) Comment on investment timing under incomplete information, *Mathematics of Operations Research* 35, 249–254.

Liptser, R. S. and Shiryaev, A. N. (2001) *Statistics of Random Processes I, General Theory*, Springer, Berlin.

Morellec, E. (2001) Asset liquidity, capital structure and secured debt, *Journal of Financial Economics* 61, 173–206.

Neftci, S. N. (1985) Are economic time series asymmetric over the business cycle? *Journal of Political Economy* 92, 307–328.

Nelson, D. B. and Ramaswamy, K. (1990) Simple binomial processes as diffusion approximations in financial models, *Review of Financial Studies* 3, 393–430.

Sichel, D. E. (1993) Business cycle asymmetry: A deeper look, *Economic Inquiry* 31, 224–236.

Van Nieuwerburgh, S. and Veldkamp, L. (2006) Learning asymmetries in real business cycles, *Journal of Monetary Economics* 53, 753–772.

Veldkamp, L. (2005) Slow boom, sudden crash, *Journal of Economic Theory* 124, 230–257.

Veronesi, P. (1999) Stock market overreaction to bad news in good times: A rational expectations equilibrium model, *Review of Financial Studies* 12, 975–1007.

——— (2000) How does information quality affect stock returns? *Journal of Finance* 55, 807–837.