Biased beliefs, costly external finance, and firm behavior: A Unified theory
Biased Beliefs, Costly External Finance, and Firm Behavior: A Unified Theory

Delong Li, Lei Lu, Congming Mu, and Jinqiang Yang

August 22, 2019

Abstract

Overconfidence and overextrapolation are two behavioral biases that are pervasive in human thinking. A long line of research documents that such biases influence business decisions by distorting managers’ expected productivity. We propose a new mechanism in which the biases change firms’ precautionary motives when external financing is costly, finding that the influences of biases on investment, payouts, and refinancing are stronger for financially weaker firms. Moreover, biased and rational firms display differential responses to economic booms and busts holding financial positions constant. Our work illustrates that managerial traits, when interacting with imperfect capital markets, drive firm dynamics in business cycles.

Keywords: costly external finance; investment; liquidity management; overconfidence; overextrapolation; payout.

1 Introduction

In standard finance theory, a firm’s investment relies on the expected productivity of capital. A long line of research finds that overconfidence and overextrapolation are two behavioral biases that are pervasive in human thinking and can affect business decisions by influencing how a firm’s manager processes information to form expectations.\footnote{Using U.S. confidential survey data, Barreros (2018) finds empirical evidence indicating that firm managers have overconfident and overextrapolative biases. Alti and Tetlock (2014) study how these two biases influence firm investment and asset return predictability. For studies of overconfidence, see also Oskamp (1965); Fischhoff, Slovic, and Lichtenstein (1977); Alpert and Raiffa (1982); Lichtenstein, Fischhoff, and Phillips (1982); Bondt and Thaler (1995); Malmendier and Tate (2005); Ben-David, Graham, and Harvey (2013). For overextrapolation, see also Kahneman and Tversky (1972, 1974); Grether (1980); Bloomfield and Hales (2002); Hirshleifer (2001); Barberis and Thaler (2003); Fuster, Laibson, and Mendel (2010); Greenwood and Hanson (2014); Greenwood and Shleifer (2014).} In this paper, we introduce a new mechanism in which such information processing biases interact with precautionary motives when external financing is costly and study the biases’ impacts on investment, payouts, refinancing, and risk management in a unified model.

Suppose that the manager receives two pieces of information at each point in time, based on which they form expectations of future productivity: (i) a partially informative soft signal and (ii) realized past output.\footnote{We follow the benchmark setup in Alti and Tetlock (2014), assuming that for a given firm, there is a representative decision-maker—referred to as the manager—who determines both optimal policies and asset prices. Alti and Tetlock (2014) discuss the possibilities of separating the two roles but find that the representative-agent setup best fits the data.} The manager can be biased by overly relying on one of the two sources of information. Overconfidence bias refers to the situation in which the manager overestimates the precision of the signal. Overextrapolation bias instead refers to the situation when the manager overestimates the persistence of past output.

Besides the expected productivity, firm behavior in our model also relies on financial positions. The manager possesses a precautionary motive to hoard cash to avoid costly liquidation, which limits investment and dividend payouts.\footnote{See, for example, Daniel, Denis, and Naveen (2007) for empirical evidence.} The firm may also have to refinance from external investors when its cash stock is insufficient. Seeing that external finance is costly, the firm optimally chooses the refinancing time and size by balancing the precautionary motive and financing costs.
We find that both overconfidence and overextrapolation biases can affect the strength of precautionary motives by changing the manager’s perceived risk ex ante. An overconfident manager overestimates the precision of the signal, thus underestimating the risk, and possesses weaker precautionary motives consequently. In contrast, an overextrapolative manager overly relies on past output to forecast future productivity. When past output is low, the biased manager tends to overestimate the chances of low productivity, resulting in greater perceived risk and stronger precautionary motives, and vice versa.

The biased and rational managers’ differential precautionary motives lead to distinct firm decisions on investment, payouts, refinancing, and risk management, holding their expected productivity constant, which is dubbed as the “precautionary channel” hereafter. For instance, an overconfident manager tends to overinvest due to his/her weaker precautionary motives. More importantly, the magnitudes of such influences of biases rely on the firms’ financial positions. Precautionary motives in general play a larger role in influencing financially weaker firms’ decisions; thus, these firms are affected more significantly by the biases through the precautionary channel. As financial positions improve, the precautionary motives play a lesser role in the managers’ decision processes and, as a consequence, the biases’ impacts through changing the precautionary motives diminish. Therefore, our results demonstrate that firms lacking financial resources are more vulnerable to managerial biases and, on the other hand, liquidity management can be useful to mitigate the negative consequences caused by biased managers. These conclusions cannot be reached without considering the interaction between information processing biases and costly external finance.

We further show that biased and rational firms with same financial positions differ in their responses to common signal and output shocks. Take investment policy as an example.

---

4 The risk ex ante refers to the inferred possibilities of liquidation or the situation in which the firm is forced to pay external financing costs to replenish its cash stock before the signal realizes.

5 This is in the same spirit as Gervais, Heaton, and Odean (2011), where overconfidence makes risk-averse managers less conservative in taking risky investment projects. In our model, managers are risk-neutral; however, they act as a risk-averse agent because of costly external finance and liquidation. Overconfidence makes managers less conservative in the sense that the precautionary-saving motives become weaker.

6 This is a standard result in the literature of precautionary saving; see, for example, Carroll (1997) and Bolton, Chen, and Wang (2013).
The overconfident firm cuts investment more strongly than the rational firm when receiving a negative signal, given that the biased manager overestimates the informativeness of the signal.\textsuperscript{7} The overextrapolative firm instead responds more aggressively when the observed output deviates from past trends.

A direct implication of this theoretical finding is that firms with different degrees of overconfidence or overextrapolation will display distinct dynamics in business cycles, even when they have same financial positions and are faced with same external shocks. This result offers a new explanation for why individual firms cut investment to different extents in the 2008-2009 global financial crisis, for example. Much of the existing literature suggests that this phenomenon occurred because firms differed in their financial positions or exposures to external shocks.\textsuperscript{8} Instead, we argue that firms can process information differently and thus have distinct beliefs about future productivity, leading to differential investment decisions.

This paper is the first to offer a unified theory of firm investment, payouts, refinancing, cash holding, and risk management under both information processing biases and financial-market imperfections. Our modeling framework is, therefore, of independent interest itself. It can be extended to study many other important questions in corporate finance. For instance, Malmendier and Tate (2005) document that overconfident firms display a greater sensitivity of investment to cash flows; see also Malmendier and Tate (2015). Our model can relate to this discussion because it is equipped with both behavioral biases and costly external finance that generates nonzero investment-cash flow sensitivity.

Another example deals with dividend payouts. The previous literature finds that dividend policy varies over time and across firms and that neither investment nor financing decisions alone can uniquely determine dividends.\textsuperscript{9} Among others, Ben-David, Graham, and

\textsuperscript{7}The opposite is true if the overconfident firm sees a positive signal.

\textsuperscript{8}See, for instance, Duchin, Ozbas, and Sensoy (2010) for a study of U.S. firms, and Li, Magud, and Valencia (2019) for a study of firms in emerging markets. Other references include, but are not limited to, Almeida, Campello, Laranjeira, and Weisbenner (2009), Iyer and Peydro (2011), Carvalho, Ferreira, and Matos (2015), and Kalemli-Ozcan, Kamil, and Villegas-Sanchez (2016).

\textsuperscript{9}For instance, see Fama and French (2001); Ben-David, Graham, and Harvey (2007); Deshmukh, Goel, and Howe (2013); Adam, Fernando, and Golubeva (2015); Bliss, Cheng, and Denis (2015); Lambrecht and Myers (2017).
Harvey (2007), and Deshmukh, Goel, and Howe (2013) link dividend policy to managerial overconfidence. In our unified model, dividend policy is simultaneously determined by investment and financing plans, and all of them are influenced by both overconfidence and overextrapolation. In Section 5, we discuss how our model’s predictions about payouts are in line with what people observed in the post-2008-recession period.

In addition to the literature mentioned above, our paper also relates to works studying financial constraints and firm behavior. Recent papers include Denis and Sibilkov (2009), Bolton, Chen, and Wang (2011, 2013), Décamps, Mariotti, Rochet, and Villeneuve (2011), Gamba and Triantis (2014), Hugonnier, Malamud, and Morellec (2015), Décamps, Gryglewicz, Morellec, and Villeneuve (2017) and others. None of these studies take into account managerial traits—to that end, our paper relates to the literature on behavioral corporate finance. Relevant work includes, but is not limited to, Hackbarth (2008) on capital structure, Malmendier and Tate (2008) on value-destroying acquisitions, Gervais, Heaton, and Odean (2011) on capital budgeting, Malmendier, Tate, and Yan (2011) on financing policies, Huang, Tan, and Faff (2016) on maturity decisions, Otto (2014) and Humphery-Jenner, Lisic, Nanda, and Silveri (2016) on compensation structure, Ho, Huang, Lin, and Yen (2016) on banks’ risk-taking behavior, Barreros (2018) on hiring decisions, and Malmendier (2018) for a recent survey paper.¹⁰

The remainder of the paper is organized as follows. Section 2 presents the model setup, and Section 3 presents the solution method. The numerical analysis is in Section 4 and the simulated firm dynamics in response to signal and output shocks are given in Section 5. Section 6 considers a model extension with financial hedging, and Section 7 concludes.

2 Model setup

2.1 Production technology

A firm uses capital goods as the only input and $AK$ production technology to produce the output given by

$$d\Pi_t = K_t dA_t,$$

where $dA_t$ is the marginal output of capital over the time interval $(t, t+dt)$, and $K_t$ represents capital stock. Following the neoclassical investment literature (e.g., Hayashi, 1982; Abel and Eberly, 1994), $K_t$ evolves according to

$$dK_t = (I_t - \delta K_t)dt,$$

where $I_t$ denotes the firm’s instantaneous gross investment, and $\delta$ is the depreciation rate. Investment features adjustment costs, $G(I, K)$, given by

$$G(I, K) = g(i)K,$$

where $i = I/K$; $g(i)$ follows a quadratic form, $g(i) = \frac{\theta i^2}{2}$ (DeMarzo, Fishman, He, and Wang, 2012; Pindyck and Wang, 2013); and $\theta > 0$ captures the degree of adjustment costs.

The firm’s marginal output $dA_t$ evolves with the following dynamics:

$$dA_t = u_t dt + \sigma_A d\mathcal{Z}_t^A,$$

where $\mathcal{Z}_t^A$ is a standard Brownian motion; its innovations, $d\mathcal{Z}_t^A$, represent output disturbances; and $\sigma_A > 0$ is the volatility. Moreover, the drift term $u_t$ follows an Ornstein-
Uhlenbeck process given by (e.g., Alti and Tetlock, 2014)

\[ du_t = \kappa(u_t - \bar{u})dt + \sigma_u dZ^u_t, \]  

(5)

where \( \kappa \) captures the extent of mean reversion to the long-run mean \( \bar{u} \); \( \sigma_u > 0 \) is the volatility; and \( Z^u \) is a standard Brownian motion independent of \( Z^A \).

### 2.2 Information processing and biased beliefs

Hereafter, we refer to the drift term \( u_t \) as the underlying productivity, which is unobservable by the firm’s manager. The manager needs to estimate \( u_t \) based on models (4) and (5) they have in mind and the following two pieces of information available at time \( t \) (Alti and Tetlock, 2014). First, the manager observes the realized (marginal) output \( dA_t \). Second, the manager receives a soft signal \( s_t \) correlated with \( u_t \). Specifically,

\[ ds_t = \eta dZ^u_t + \sqrt{1-\eta^2} dZ^s_t, \]  

(6)

where \( dZ^u_t \) is the innovation in \( u_t \) given by equation (5), and \( \eta \in [0,1] \) is the correlation between \( dZ^u_t \) and \( ds_t \) that captures the informativeness of the signal; \( Z^s \) is a standard Brownian motion independent of both \( Z^A \) and \( Z^u \), with \( dZ^s_t \) representing the noise in the signal process. Define the signal precision to be \( \nu = \frac{\eta}{\eta + \sqrt{1-\eta^2}} \), which increases with \( \eta \).

We consider overconfidence and overextrapolation biases when the manager estimates \( u_t \). An overconfident manager believes that the soft signal is more informative than it actually is; namely, \( \eta_{oc} > \eta \). As a result, the manager’s perceived signal precision is higher than the actual precision (\( \nu_{oc} > \nu \)).\(^{12}\) In contrast, an overextrapolative manager believes the firm’s productivity to be more persistent than it actually is; that is, the manager’s mean-reverting

\(^{11}\) Allowing for a nonzero correlation between \( Z^u \) and \( Z^A \) complicates the mathematics without generating new insights.

\(^{12}\) This definition of overconfidence is also referred to as “overprecision” by the existing literature. In other contexts, overconfidence can be defined in alternative ways; see Malmendier (2018).
parameter is smaller than the true value, $\kappa_{oe} < \kappa$.

Let $x_t$ be the mean of the manager’s estimates of $u_t$ conditional on the information up to time $t$—hereafter referred to as the expected productivity—and $\gamma$ be the steady-state variance of the estimation error, $\mathbb{E}[(x_t - u_t)^2]$. From the manager’s perspective, the marginal output evolves according to

$$dA_t = x_t dt + \frac{dA_t - x_t dt}{\sigma_A},$$  \hspace{1cm} (7)$$

where $d\tilde{Z}_t^A \equiv (dA_t - x_t dt)/\sigma_A$ is the perceived output shock at time $t$ and $\tilde{Z}_t^A$ is also a standard Brownian motion under the manager’s beliefs. In response to the signal and output shocks, the manager updates $x_t$ based on the standard filter theory (Liptser and Shiryaev, 2001):

$$dx_t = \kappa_i (\bar{u} - x_t) dt + \eta_i \sigma_u ds_t + \gamma_i \frac{dA_t - x_t dt}{\sigma_A},$$  \hspace{1cm} (8)$$

with

$$\gamma_i = \sigma_A \left(-\kappa_i \sigma_A + \sqrt{\kappa_i^2 \sigma_A^2 + (1 - \eta_i^2) \sigma_u^2}\right),$$

where $\kappa_i \in \{\kappa, \kappa_{oe}\}$, $\eta_i \in \{\eta, \eta_{oc}\}$, $\gamma_i \in \{\gamma, \gamma_{oe}, \gamma_{oc}\}$.

The first term in (8) captures the mean reversion of the expected productivity. If the current $x_t$ is higher than the long-run mean $\bar{u}$, the manager adjusts the expectation downward, $dx_t < 0$, and vice versa. The last two terms describe the manager’s responses to signal and output shocks, respectively. When there is a positive signal shock, $ds_t > 0$, or when the observed output is higher than expected,
\( dA_t > x_t dt \), the manager adjusts the expectation upward, \( dx_t > 0 \); and vice versa.

Managers who receive the same information \( \{dA_t, ds_t\} \) but differ in information processing (\( \eta_i \) and \( \kappa_i \)) update their expected productivity differently. In equation (8), an overconfident manager overreacts to signal shocks \( ds_t \) due to a higher \( \eta_{oe} \). Moreover, given that \( \gamma_i \) decreases in \( \eta_i \), the overconfident manager has a smaller \( \gamma_{oe} \), thus underreacting to output shocks \( dA_t \). In contrast, a manager with overextrapolation bias possesses a smaller \( \kappa_{oe} \) and is therefore more reluctant to mean-revert when \( x_t \) is different from \( \bar{u} \). The overextrapolative manager also reacts more aggressively to output shocks, provided that \( \gamma_{oe} \) is greater than that of the rational manager.\(^\text{15}\) In addition, overextrapolation does not change the manager’s beliefs about signal precision; thus, the manager behaves rationally in response to signal shocks.

2.3 Financing costs and managers’ optimization

Based on the expected productivity \( x_t \), the firm’s free cash flows are given by

\[
dY_t = K_t dA_t - \left( I_t + G(I_t, K_t) \right) dt.
\]

If the free cash flows are positive, the firm can either retain this amount of money or pay it out as dividends. If the free cash flows are negative, for instance, when the firm experiences operating losses \( (dA_t < 0) \) or when investment exceeds operating income, the firm must cover the gap using either internal funding or external financing; otherwise, the firm has to cut investment. Internal funding comes from the firm’s cash inventory \( W_t \), which is accumulated from the firm’s past operations. In terms of external financing, the firm can raise funds from equity markets. However, it has to pay a fixed cost \( \phi_0 K \) and a proportional cost \( \phi_1 \) on the amount financed, which are referred to as the financing costs.\(^\text{16}\)

\(^{15}\)Because \( \gamma_i \) decreases with \( \kappa_i \), a smaller \( \kappa_{oe} \) leads to a bigger \( \gamma_{oe} \).

\(^{16}\)As in Bolton, Wang, and Yang (2014) and Décamps, Gryglewicz, Morellec, and Villeneuve (2017), the fixed cost is scaled by firm size, preventing the firm from growing out from the fixed cost.
We let $dH_t$, $dU_t$ and $dD_t$ represent the net proceeds from external financing, the total financing costs, and the dividend payout in $(t, t+dt)$, respectively. Further suppose that the rate of return the firm earns on its cash inventory equals the risk-free rate $r$ minus the cost of carrying cash $\lambda$. Then, the firm’s cash inventory evolves according to

$$dW_t = (r - \lambda)W_t dt + dY_t + dH_t - dD_t,$$

where the first term indicates that cash inventory increases with the interest earned, net of the carrying cost; the second term is the free cash flows; the third term is the net proceeds from external financing; and the last one is the dividend payout.\(^\text{17}\)

At any time, the manager can choose to stop operating the firm and liquidate all the capital stock. The firm’s capital is sold at a discount price $l < 1$, where 1 is the purchasing price of capital and the gap $(1 - l)$ represents the deadweight loss in liquidation (Bernanke, Gertler, and Gilchrist, 1999). The liquidation value and cash inventory will be paid out as dividends. The manager chooses the optimal policies of investment $I$, payout $D$, external financing $H$, and liquidation time $\tau$ to maximize the firm value:

$$P(x, W, K) = \max_{I, D, H, \tau} \mathbb{E}_t^M \left[ \int_0^\tau e^{-rt} (dD_t - dH_t - dU_t) + e^{-r\tau} (lK_\tau + W_\tau) \right],$$

where $\mathbb{E}_t^M[\cdot]$ denotes the manager’s conditional expectations under their beliefs, which can be either rational or biased. Here, we follow the benchmark setup in Alti and Tetlock (2014), assuming that the firm’s manager is the representative decision-maker who determines both optimal policies and asset prices.

\(^\text{17}\)The net proceeds from external financing equal the amount of money financed minus the financing costs. Therefore, the net proceeds term $dH_t$ in equation (10) has already incorporated the contribution of the financing costs $dU_t$.\(^\text{9}\)
3 Model solution

A firm’s optimal behavior is determined by the manager’s expected productivity \( x_t \), the cash stock \( W_t \), and existing capital \( K_t \). The firm hoards cash precautionarily due to costly external financing and liquidation. Furthermore, the cash carrying cost \( \lambda \) encourages the firm to pay out dividends once its cash stock reaches a certain level. Therefore, like Bolton, Chen, and Wang (2011), the firm’s cash policy has two endogenous barriers: \( \bar{W}(x, K) \) as the upper barrier, and \( W(x, K) \) as the lower barrier.

The firm pays out dividends when cash reserves exceed \( \bar{W}(x, K) \); thus, the upper barrier is also referred to as the “payout barrier.” In contrast, the firm either liquidates its capital stock or raises external financing when the cash stock is below \( W(x, K) \), so the lower barrier is dubbed as the “liquidation/refinancing barrier.” Whether to liquidate or to refinance at the lower barrier is determined by the manager’s expected productivity: the manager will liquidate when faced with low productivity, and refinance when expecting high productivity. In addition, when the firm’s cash is between the two barriers, it operates based on internal cash inventory, referred to as the “internal funding region.” We analyze the firm’s optimal policies in the three cash regions separately.

3.1 Internal funding region

We first consider the case in which the manager operates the firm using internal funds, that is, without dividend payout or external financing. The firm value \( P(x, W, K) \) satisfies the following Hamilton-Jacobi-Bellman (HJB) equation:\(^{18}\)

\[
\begin{align*}
    rP(x, W, K) &= \max_f \left\{ (I - \delta K)P_K + \left[ xK - I - G(I, K) + (r - \lambda)W \right] P_W \\
    &\quad + \frac{\sigma_A^2 K^2}{2} P_{WW} + \kappa(\bar{w} - x)P_x + \frac{1}{2} \left( \eta^2 \sigma_u^2 + \frac{\gamma^2}{\sigma_A^2} \right) P_{xx} + \gamma KP_{Wx} \right\}. 
\end{align*}
\]

\(^{18}\)Rational, overconfident, and overextrapolative managers are faced with the same optimization problems except that they feature differential \( \kappa_i, \eta_i \) and \( \gamma_i \). Hereafter, we omit the subscript \( i \) when possible for succinctness.
The left side of the equation represents the required rate of return for holding the firm’s shares. The right side is the expected change in the firm’s value in the time interval, \((t, t + dt)\), when it operates using internal funds. The first term represents the marginal effect of net investment. The second and third terms capture the marginal effects of saving and the volatility of cash holdings. These three terms are standard in the continuous-time model with financial frictions, e.g., Bolton, Wang, and Yang (2014). The fourth term denotes how the manager’s expected productivity \(x_t\) affects the firm value, and the fifth term represents the effects of the volatility of \(x_t\). These two terms also appear in Alti and Tetlock (2014), capturing the manager’s needs to update their beliefs about productivity according to available information. The last one is new in our model, which characterizes the interaction between cash and expected productivity.

Taking the first-order condition with respect to investment \(I\) yields

\[
1 + \theta i = \frac{P_K(x, W, K)}{P_W(x, W, K)},
\]

which states that the marginal cost of adjusting capital stock, \(1 + \theta i\), equals the ratio of marginal \(q\), \(P_K(x, W, K)\), to the marginal cost of financing, \(P_W(x, W, K)\).

We define the cash-to-capital ratio as \(w = W/K\). Due to the homogeneity of the AK production technology and the quadratic adjustment cost \(G(I, K) = Kg(i)\), the firm value can be written as \(P(x, W, K) = KP(x, w)\). Plugging it into equation (13) gives the optimal investment plan

\[
i(x, w) = \frac{1}{\theta} \left( \frac{p(x, w)}{p_w(x, w)} - w - 1 \right).
\]

Given the expected productivity \(x\), investment \(i(x, w)\) increases with internal cash \(w\), which is a standard result in the literature of financial frictions.\(^{19}\) Moreover, investment can be

\(^{19}\)In fact, investment is below the first-best level due to the financial constraint imposed by the internal funding region. More cash would ease the financial constraint and allow the firm to invest more. Therefore, investment increases with cash in the internal funding region.
either positive or negative (disinvestment), and there exists a cash-to-capital ratio \( w_{i_0}^* \) such that \( i(x, w_{i_0}^*) = 0 \). The firm invests if \( w > w_{i_0}^* \), while it disinvests if \( w < w_{i_0}^* \). By construction, \( w_{i_0}^* \) depends on the expected productivity \( x \), and hence can be written as \( w_{i_0}^* = w_{i_0}^*(x) \). We refer to it as the “zero-investment boundary” hereafter.

### 3.2 Liquidation/refinancing region

The fixed cost of equity issuance makes it optimal for the firm to delay issuing equity until the cash buffer is depleted (Bolton, Chen, and Wang, 2011; Décamps, Gryglewicz, Morellec, and Villeneuve, 2017). The same argument applies to the liquidation decision, in which a fraction \((1-l)\) of the firm’s capital value is destroyed. That is, if refinancing or liquidating activities take place, it must be when the firm’s cash drops to zero; i.e., \( w(x) = W(K; x)/K = 0 \) for all \( x \).

Whether to liquidate the capital stock or to refinance when running out of cash relies on the manager’s expected productivity \( x \). When cash is zero, the value function satisfies

\[
p(x, 0) = \max \left\{ \sup_{m \geq 0} \left[ p(x, m) - \phi_0 - (1 + \phi_1)m \right], l \right\},
\]

where \( m = m(x) \) is the optimal amount of money to raise in refinancing. The \( \sup[\cdot] \) term is the continuation value (per unit of capital goods) under optimal refinancing; \( l \) is the liquidation value. The firm prefers whichever is higher.

It is notable that the continuation value increases with \( x \). Therefore, given the fixed cost of financing \( \phi_0 \), the marginal cost \( \phi_1 \), and the liquidation value \( l \), there exists a boundary \( x_{m_0}^* \) such that

\[
\sup_{m \geq 0} \left\{ p(x, m) - \phi_0 - (1 + \phi_1)m \right\} \leq l, \quad \text{for all } x \leq x_{m_0}^*.
\]

Put differently, if \( x \leq x_{m_0}^* \), the firm prefers to liquidate the capital stock rather than refinancing when cash is depleted because the manager’s expected productivity is too low to
cover the financing costs. Later, we refer to \( x_{r_0} \) as the firm’s “no-refinancing boundary.”

### 3.3 Payout region

Although hoarding cash helps the firm reduce the likelihood of liquidation or of being forced to raise costly external finance, the marginal value of cash diminishes as the firm accumulates cash. Moreover, retaining cash involves a carrying cost \( \lambda \). Therefore, there exists a target cash-holding barrier \( \bar{w}(x) = \bar{W}(K; x)/K \), such that the firm opts to pay out dividends if \( w \geq \bar{w}(x) \). In this region, we have

\[
p(x, w) = p(x, \bar{w}(x)) + (w - \bar{w}(x)), \quad \text{for all } w \geq \bar{w}(x),
\]

where the first term on the right represents the firm value of the payout boundary; the second term refers to the dividends paid out. Note that subtracting \( p(x, \bar{w}(x)) \) from both sides of equation (17), dividing it by \( w - \bar{w}(x) \), and taking the limit as \( w \) tends to \( \bar{w}(x) \) yields \( p_w(x, \bar{w}(x)) = 1 \). Therefore, the value of an additional dollar added to the cash inventory on the payout boundary is one, so the firm is indifferent between retaining this one dollar and paying it out. In addition, following Dumas (1992), the payout boundary \( \bar{w}(x) \) satisfies a super-contact condition \( p_{ww}(x, \bar{w}(x)) = 0 \). Applying this condition together with equation (17), we can solve out \( \bar{w}(x) \).

### 4 Numerical analysis

In this section, we show that cash inventory and expected productivity simultaneously influence a firm’s optimal behavior. Therefore, separately studying cash hoarding or information processing biases, such as what the previous research has done, cannot produce a full picture of firm dynamics. Due to its complexity, our model does not have a closed-form solution. Instead, we discuss a firm’s optimal policies for investment, payout, refinancing, and liquidation numerically. Parameter values are summarized in Table 1: \( l \) is obtained from
Carlstrom and Fuerst (1997); \( \sigma_u \) is from Cujean and Hasler (2017); and the rest are from Bolton, Chen, and Wang (2011).\(^{20}\)

### 4.1 Rational benchmark

We choose the true signal precision \( \nu = 0.5 \) (Alti and Tetlock, 2014) and the true mean-reverting rate \( \kappa = 0.2 \) (Dumas, Kurshev, and Uppal, 2009). Figure 1 presents the stationary distribution of the expected productivity \( x \) (panel A) and the optimal policy rules (panel B) of a rational manager who updates his/her beliefs over \( x \) using the true values of \( \nu \) and \( \kappa \). Panel A shows that most of the time, \( x \) falls into the region \([0.13, 0.23]\), with its mean equal to \( \bar{u} = 0.18 \).\(^{21}\)

Panel B shows the rational firm’s optimal policies in the \( x-w \), or \(-m(x)\), diagram. The solid line is the payout barrier \( \bar{w}(x) \) and the dotted, vertical line represents its lower bound—hereafter referred to as \( x^*_{w_0} \)—below which, the firm’s payout barrier equals zero; that is, \( \bar{w}(x) = 0 \) for any \( x \leq x^*_{w_0} \). When the expected productivity is sufficiently low, the firm pays out all the cash immediately. We interpret this result together with the optimal refinancing size \( m(x) \) in the dash-dotted line, where the vertical part depicts the no-refinancing boundary \( x^*_{m_0} \) (defined in section 3.2). Because \( x^*_{w_0} < x^*_{m_0} \), according to (16), at \( x^*_{w_0} \), the firm not only pays out all the cash but also liquidates its capital stock. Therefore, \( x^*_{w_0} \) acts as the effective “shut-down boundary”, below which the manager finds it optimal to stop operating the firm. Under the parameters in Table 1, this boundary for the rational firm is \( x^*_{w_0} = 0.036 \).

---

\(^{20}\)Parameter values are rounded to two decimal places.

\(^{21}\)The probability of \( x \) falling into this region is 99.87\%. 
Notes: Panel A reports the stationary distribution of expected productivity $x$ for the rational manager. Panel B describes the rational manager’s optimal policies.

Moreover, the dashed line denotes the zero-investment boundary $w^*_i(x)$ (defined in section 3.1). Together with the dotted line and the solid line, this boundary determines the firm’s optimal policies by dividing the $x$-$w$ diagram into four regions: in region A, the firm shuts down; in region B, the firm pays out dividends until its cash stock returns to $\bar{w}(x)$; in region C, the firm invests a positive amount, while in region D, the firm disinvests.\footnote{In region D, when $x$ is between $x^*_w$ and $x^*_m$, the firm still operates as long as its cash inventory is not depleted; if the firm has instead exhausted its cash, it liquidates all the capital immediately and stops operating. In contrast, when $x$ is greater than $x^*_m$, the firm chooses to refinance by raising external equity of an amount $m(x)$, when its cash stock drops to zero.}

The payout barrier $\bar{w}(x)$ increases with $x$ (solid line) because, when expecting higher productivity, the manager tends to retain more cash inside the firm to prepare for investing later while simultaneously avoiding costly external finance. Additionally, the zero-investment boundary $w^*_i(x)$ decreases in $x$ (dashed line). This finding indicates that, when expecting higher productivity, the manager is more likely to engage in positive investment and less likely to disinvest because a larger $x$ leads to a larger marginal $q$ in equation (13).\footnote{Particularly, when $x$ is sufficiently high, the firm never engages in disinvestment because selling assets is more costly than issuing equity. The firm instead chooses to tap the equity market if it needs to replenish its cash balance.}

In addition, the optimal refinancing size $m(x)$ increases with $x$ (dash-dotted line). The
higher the expected productivity is, the more funds the firm will raise in external financing when its cash inventory is depleted because, as productivity is expected to grow, the manager plans to invest more and thus needs more cash. The manager also wants to avoid paying the fixed cost of financing again in the future and therefore raises more funds each time he/she issues equity; see also Décamps, Gryglewicz, Morellec, and Villeneuve (2017).

4.2 Overconfident belief

An overconfident manager believes that the soft signal is more precise than it actually is, $\nu_{oc} > \nu$. We use $\nu_{oc} = 0.85$ (Alti and Tetlock, 2014) as an example to illustrate the impacts of overconfidence on firm behavior. All the other parameters are set according to Table 1.

First, an overconfident manager is less likely to shut down the firm—the shut-down boundary is smaller, $x^*_{w_0} = 0.034$, than that of the rational firm. This finding indicates that the overconfident manager is willing to run the firm at a lower expected productivity level when the rational firm stops operating, because the manager’s overconfidence in signal precision makes the manager believe that he/she can forecast future productivity more precisely and operate the firm more efficiently. All else being equal, this enhances the manager’s perceived continuation value in (15) and delays shutting down.

Panels A, B, and C in Figure 2 contrast the overconfident firm’s payout barrier $\bar{w}(x)$, the optimal refinancing size $m(x)$, and the zero-investment boundary $w^*_{i_0}(x)$, respectively, with those of the rational firm. Panel D instead depicts the differences in the two firms’ investment levels, $(i_{oc} - i_r)$, as a function of expected productivity and cash, $(x, w)$.

In panel A, the overconfident firm displays a higher payout boundary compared to the rational firm, meaning the overconfident firm delays dividend payouts and hoards more cash. A major purpose for the firm to hold cash internally is to prepare for future investment, seeing that external financing is costly. This motive is stronger for the overconfident firm whose manager is overoptimistic about the informativeness of the signal and, therefore, more decisive in planning investment. The manager retains more cash in advance as a preparation.
Figure 2: Optimal policies (overconfident)

Notes: Panel A compares the payout boundary of the overconfident firm to that of the rational firm. Panel B compares the optimal refinancing sizes. Panel C compares the zero-investment boundaries. Panel D depicts the differences between the overconfident firm's investment and the rational firm's investment, $i_{oc} - i_r$.

In panel B, the overconfident firm refinances a larger amount of money when its cash drops to zero. Again, the biased manager is more determined in the firm’s future operating plans due to the perceived higher signal precision. The manager is also willing to run the firm for a longer period of time, given the lower shut-down boundary. Therefore, the overconfident manager expects to need more funding for the firm’s operation later, ceteris paribus. Thus, the manager would rather raise more money at each time when tapping the equity market to avoid paying the fixed financing cost again in the future.

Panel C compares the zero-investment boundary of the overconfident firm to that of the rational firm, demonstrating that the overconfident firm features a lower boundary. Given that firms invest in the top-right regions, the lower boundary indicates that the overconfident firm is more likely to invest and less likely to disinvest. In fact, a crucial reason leading to
disinvestment is the need for replenishing the firm’s cash stock to avoid shutting down. The overconfidence bias induces the manager to believe that the firm is less likely to shut down because the shut-down boundary \( x_{w_0}^{oc} \) is lower and there is, therefore, a lower need to disinvest. The zero-investment boundary shifts down as a result.

Panel D depicts the differences between the overconfident firm’s investment and that of the rational firm, \( i_{oc} - i_r \). Notably, the investment gap is positive; that is, given the same expected productivity and cash, the overconfident firm overinvests compared to the rational firm.\(^{24}\) The reason can be attributed to the differences in the two firms’ precautionary motives, wherein firms limit their investment to save cash to avoid shutting down. Given the same \((x, w)\), the overconfident firm is further away from the shut-down boundary compared to the rational firm, provided that \( x_{w_0}^{oc} < x_{w_0}^{r} \). Thus, the overconfident firm possesses a weaker precautionary motive and invests more aggressively.\(^{25}\)

The investment gap \((i_{oc} - i_r)\) largely decreases with \( w \) while holding \( x \) constant because both the overconfident firm and the rational firm gradually stop worrying about shutting down as they accumulate cash. Thus, the differences in their precautionary motives become less significant and their investment converges.\(^{26}\) On the other hand, holding \( w \) constant, the investment gap grows larger as \( x \) increases because the two firms’ investment grows with \( x \) at differential paces, given the same amount of cash stock. The overconfident firm possesses a weaker precautionary motive and its investment grows more quickly with \( x \) compared to

---

24 The shaded region represents the area in which both firms invest. It is precisely the overlapping area of the two firms’ region C-s, where for each firm, region C is defined in Figure 1. For this overlapping area, the upper bound of cash \( w \) is the rational firm’s payout barrier, given that this barrier is lower than the payout barrier of the overconfident firm.

25 This result is in the same spirit as Gervais, Heaton, and Odean (2011), wherein overconfidence makes a risk-averse manager less conservative in implementing risky investment projects; see also Hirshleifer, Low, and Teoh (2012). Our model instead studies continuous investment decisions, and managers are risk-neutral. The fixed costs of external finance and the deadweight loss in liquidation make the manager act as a risk-averse agent. Overconfidence encourages the manager to be less conservative in the sense that the manager’s precautionary motives become weaker.

26 The investment gap increases slightly with \( w \) when \( w \) is close to its upper bound in panel D. In fact, the upper bound is the payout barrier of the rational firm (see footnote 24). At these \( w \) values, the rational firm’s investment is close to optimal, thus increasing more slowly with \( w \). In contrast, provided that \( \bar{w}_{oc}(x) > \bar{w}_r(x) \), the overconfident firm’s investment is further below optimality, thus growing more quickly with \( w \). Consequently, the investment gap becomes larger as \( w \) increases. This effect is quantitatively small.
the rational firm; consequently, the investment gap widens as \( x \) rises.\(^{27}\)

### 4.3 Overextrapolative belief

An overextrapolative manager believes that productivity is less mean-reverting, in other words, more persistent than it actually is, \( \kappa_{oe} < \kappa \). In this section, we use \( \kappa_{oe} = 0.154 \) as an example to illustrate the impacts of overextrapolation on firm behavior, where the difference \( (\kappa - \kappa_{oe}) = 0.046 \) captures the degree of overextrapolation and is taken from Alti and Tetlock (2014). All the other parameters are set according to Table 1.

First, the overextrapolative manager is more likely to liquidate the firm when faced with low productivity compared to the rational manager—the shut-down boundary is higher, \( x^{* \text{oe}}_{w_0} = 0.052 \). This effect is a result of the biased manager overestimating the persistence of low productivity and, hence, being too pessimistic to keep operating the firm.

In Figure 3, we compare the two firms’ optimal policy rules on payouts, refinancing, and investment. In panel A, the payout barrier of the overextrapolative firm is below that of the rational firm when the expected productivity \( x \) is low and above that of the rational firm when \( x \) is high. In the low-\( x \) region, the firm expects future productivity to continue to be low and, as a result, is over-conservative about future investment plans. The firm chooses to hold less cash internally, as it does not plan to invest much in the future. The payout barrier is therefore lower. On the other hand, when \( x \) is high, the overextrapolation bias makes the manager overoptimistic about future productivity, expecting the high productivity to last longer. Consequently, the manager is more aggressive in holding cash internally to prepare for investment, which delays payouts, leading to a higher payout barrier.

\(^{27}\)In equation (13), the denominator (the marginal value of cash) is smaller for the overconfident firm due to its weaker precautionary motive; thus, its investment is more sensitive to marginal \( q \) and, in turn, associated with \( x \) to a greater extent. Furthermore, the investment gap \( (i_{\text{oe}} - i_r) \) can decrease with \( x \) when \( w \) is fixed at a large enough value close to the upper boundary of the shaded area in panel D. In fact, the upper boundary is the payout barrier of the rational firm (footnote 24) at which the rational firm is financially unconstrained; the denominator in (13) for the rational firm equals 1, smaller than that of the overconfident firm. Consequently, the rational firm’s investment is more sensitive to \( x \) compared to the overconfident firm, and therefore, the investment gap decreases as \( x \) grows. We do not focus on this case because we are interested in both firms being financially constrained, that is, when \( w \) is relatively low.
Figure 3: Optimal policies (overextrapolative)

Panel A compares the payout boundary of the overextrapolative firm to that of the rational firm. Panel B compares the optimal refinancing sizes. Panel C compares the zero-investment boundaries. Panel D depicts the differences between the overextrapolative firm’s investment and the rational firm’s investment, \( i_{oe} - i_{r} \).

Panel B shows that the overextrapolative manager tends to raise less money when refinancing in the low-\( x \) region and more money in the high-\( x \) region. The reason is similar to that in panel A. When \( x \) is low, the overextrapolation bias induces the manager to believe that low productivity will persist. Consequently, the manager plans not to invest much in the future and raises less money to save the proportional cost. On the other hand, when \( x \) is high, the manager believes that high productivity will last longer and tends to be more aggressive in planning investment. As a result, the manager expects a larger funding need and raises more money when tapping the equity market to avoid paying the fixed financing cost again in the future.

In panel C, the zero-investment boundary of the overextrapolative firm is higher than that of the rational firm in the low-\( x \) region and lower than that in the high-\( x \) region. A
higher boundary means that the overextrapolative firm is more likely to disinvest because the biased manager believes that the low productivity will persist and there is a higher likelihood that the firm will have to liquidate or to pay external financing costs. The manager would rather sell capital goods earlier to replenish the cash stock to avoid liquidation or paying financing costs. The low $x$ also makes capital less productive and thus “cheaper” to give up in disinvestment. In contrast, when $x$ is high, the manager believes that the productivity will remain high and there will be fewer needs to disinvest. It is also more costly to disinvest because of the high marginal product of capital that the firm must give up; if necessary, the firm would rather raise external equity. Consequently, the manager delays disinvestment, in line with a lower zero-investment boundary.

In panel D, we present the differences in investment between the overextrapolative firm and the rational firm, $i_{oe} - i_{r}$, as a function of $(x, w)$, where there exist both a positive part and a negative part. The positive part represents the region in which the overextrapolative firm invests more than the rational firm does, which occurs when the current expected productivity $x$ is high and the overextrapolation bias induces the manager to expect higher productivity in the future. In contrast, the negative part is the area where $x$ is low, indicating that the overextrapolative firm underinvests compared to the rational firm because the biased manager underestimates the future productivity.\footnote{For conciseness, we plot the negative part as a flat region here; the actual shape is symmetric to the positive part and shown in the appendix.}

The positive part of $i_{oe} - i_{r}$ relies on $(x, w)$ in the same way as panel D in Figure 2. In fact, in this relatively high-$x$ region, the overextrapolative manager’s behavior is similar to that of the overconfident manager because both biases lead to overoptimism about the firms’ future. Compared to the rational firm, the biased firm cares less about risk and possesses a weaker precautionary motive. The biased firm’s investment therefore grows more quickly with $x$ (see footnote 27), enlarging the investment gap as $x$ rises while holding $w$ equal. On the other hand, for any given $x$, the importance of the precautionary motive diminishes with
$w$, meaning the biased and the rational firms’ investment converges as $w$ rises.\textsuperscript{29}

5 Simulated firm dynamics

Our model can explain why firms behave differently in business cycles. A more general question is why they differ in investment, saving, and payouts when receiving common external shocks. Much of the existing literature focuses on firms’ heterogeneous financial positions—some firms are more financially constrained and thus change their behavior more strongly when an external shock influences the availability of funds.\textsuperscript{30} We show in this section a new mechanism: firms with various information processing biases display distinct dynamics when controlling for financial positions because they view the same shock as having different information content and accordingly form different productivity expectations.

To this end, we rewrite (8) as

$$dx_t = \kappa_i(\pi - x_t)dt + \eta_i\sigma_u d\tilde{Z}_t^s + \frac{\gamma_i}{\sigma_A} d\tilde{Z}_t^A,$$

(18)

where

$$d\tilde{Z}_t^s = ds_t \quad \text{and} \quad d\tilde{Z}_t^A = \frac{dA_t - x_t dt}{\sigma_A}$$

are the managers’ observed signal and output shocks. We simulate the impulse responses to a negative shock in $d\tilde{Z}_t^s$ or $d\tilde{Z}_t^A$.\textsuperscript{31} Suppose that, at $t = 0$, firms are at the target cash levels (i.e., payout barriers), and their managers’ expected productivity is at the long-run mean $\bar{u}$.

We introduce a one-unit negative shock at $t = 2dt$ by setting $d\tilde{Z}_t^s = -1$ or $d\tilde{Z}_t^A = -1$. Firms respond by updating their expected productivity $x_t$ as in equation (18) and changing their

\textsuperscript{29}For the same reasons as in panel D of Figure 2, here, the investment gap slightly increases with $w$, when $x$ is held constant and $w$ is close to its upper boundary, which is the payout barrier of the rational firm. Additionally, the investment gap can decrease with $x$ when $w$ is fixed at a sufficiently large value close to its upper boundary.


\textsuperscript{31}We include the similar, but opposite, impulse responses to positive shocks in the appendix.
behavior accordingly. Further suppose that there are no additional shocks after that. At each $t = ndt$ ($n = 0, 1, 2, ...$), we plot the impulse responses of the expected productivity, cash stock, payout, and investment in their percentage changes from the baseline case without such shocks.

5.1 Overconfident vs. rational firms

We show the impulse responses of the overconfident firm and the rational firm in Figure 4, where panel A is for the negative output shock and panel B is for the negative signal shock. The dynamics of the rational firm are represented by the “stars”, and that of the overconfident firm is represented by the “circles.” All the parameters are set according to Table 1 and Section 4.2.

In panel A, we can see that compared to the rational firm, the overconfident firm under-reacts in adjusting its expected productivity downward when faced with a negative output shock. This finding is in line with equation (18) that the response $dx$ to an output shock is proportional to $\gamma$, and the overconfident firm features a smaller $\gamma_{oc}$. Both firms’ cash drops immediately because the negative output shock reduces free cash flows, as suggested by equation (9). Consequently, both firms cut dividends to save cash. In the following periods, the rational firm’s cash and dividends recover more quickly than the overconfident firm because the overconfident firm has a higher cash target level (panel A of Figure 2), thus requiring more time to accumulate cash. Moreover, both firms pay out dividends before the expected productivity fully recovers because the lower expected productivity brings down cash target levels, triggering dividend payouts.

Both firms invest less because (i) the expected productivity $x$ is lower and (ii) they become financially constrained due to the drops in $w$. The overconfident firm has a smaller jump-down in investment compared to the rational firm because the biased firm’s $x$ drops less. The higher expected productivity of the overconfident firm also leads its investment to recover more quickly in the subsequent periods.
Panel A. Negative output shock.

Panel B. Negative signal shock.

Note: Panel A shows the impulse responses to a negative output shock of the overconfident firm ("circles") and those of the rational firm ("stars"). The shock $d\tilde{Z}_A = -1$ is introduced at $t = 2\Delta t$, supposing that there are no additional shocks. At each $\Delta t$, the impulse responses are calculated as the percentage changes (in decimal) of the expected productivity, cash stock, payout, and investment from the baseline case without introducing the shock. Panel B instead shows the impulse responses to a negative signal shock $d\tilde{Z}_s = -1$. The amount of “cash” is counted before investment and dividend expenditures in the corresponding period.

Panel B describes the impulse responses to a negative signal shock. The overconfident firm adjusts its expected productivity downward more aggressively compared to the rational firm because the biased manager overreacts to the bad signal ($\eta_{oc} > \eta$ in equation (18)). Even though both firms cut investment, the overconfident firm cuts more, which in turn leads to larger free cash flows and internal piles of cash. Meanwhile, the lower expected productivity reduces the value of holding cash to prepare for future investment and thus encourages dividend payouts. The overconfident firm pays out more due to its larger free cash flows (caused by investing less) and lower expected productivity. As the expected productivity recovers, all the variables return to normal.
Figure 5: Impulse responses (overextrapolative vs. rational)

Panel A. Negative output shock.

Panel B. Negative signal shock.

Note: Panel A shows the impulse responses to a negative output shock of the overextrapolative firm ("circles") and those of the rational firm ("stars"). The shock $d\tilde{Z}^A = -1$ is introduced at $t = 2dt$, supposing that there are no additional shocks. At each $dt$, the impulse responses are calculated as the percentage changes (in decimal) of the expected productivity, cash stock, payout, and investment from the baseline case without introducing the shock. Panel B instead shows the impulse responses to a negative signal shock $d\tilde{Z}^s = -1$. The amount of "cash" is counted before investment and dividend expenditures in the corresponding period.

5.2 Overextrapolative vs. rational

The impulse responses of the overextrapolative firm are in Figure 5, in “circles”, and those of the rational firm are in “stars”; panel A shows the negative output shock and panel B shows the signal shock. All the parameters are set according to Table 1 and Section 4.3.

In panel A, compared to the rational firm, the overextrapolative firm overreacts to the negative output shock in adjusting the expected productivity downward. In subsequent periods, the overextrapolative firm’s expected productivity also recovers more slowly. The larger jump-down corresponds to the larger $\gamma_{oe}$ in equation (18), and the slower recovery is in line with the smaller $\kappa_{oe}$. Both firms’ cash inventories and dividends drop when hit by the output shock due to the reductions in free cash flows. They also cut investment because of the low expected productivity and lower internal cash as a source of funding. During the
recovery process, the overextrapolative firm keeps underinvesting compared to the rational firm, given the lower expected productivity by the biased manager. The overextrapolative firm, therefore, accumulates more cash and pays out more dividends.

In Panel B, the overextrapolative firm displays the same jump-down in $x_t$ as the rational firm when faced with the negative signal shock. This is because overextrapolation does not change how managers view the soft signal. However, the overextrapolative manager’s expected productivity recovers more slowly because of his/her stronger beliefs in persistence. Both firms cut investment, consequently receiving more free cash flows, accumulating cash, and paying out dividends. The overextrapolative firm invests less, accumulating more cash, and paying more dividends compared to the rational firm because of the biased manager’s lower expected productivity.

5.3 Business cycles

Next, we discuss how overconfidence and overextrapolation can lead to heterogeneous firm dynamics in business cycles. This finding adds to previous studies that focus on differential financial positions, arguing that various information processing biases act as complementary drivers. In our model, a recession can be described by an unobserved shock to $u_t$, $dZ^u_t < 0$ (equation (5)), leading to a drop in the expected productivity, lower output, and underinvestment for a period of time. This type of recession is arguably close to the secular stagnation that the literature has discussed after the 2008 financial crisis. Given that firms cannot see $u_t$, their decisions must rely on the observed changes in output and signals. We then predict that firms behave differently based on the analysis in Sections 5.1 and 5.2.

Firms differ in their responses to $dZ^u_t$ for three reasons. First, they receive distinct signals, as shown in equation (6): for a given $dZ^u_t < 0$, the observed signal $ds_t$ relies on the parameter $\eta$. An overconfident firm has a larger $\eta_{oc}$, thus receiving a stronger signal for the recession. Moreover, after the recession hits the economy, the output $dA_t$ drops, as suggested by equation (4). All the firms can observe this low output, but the overconfident
manager will respond less due to his/her smaller $\gamma_{oc}$, according to equation (18), while the overextrapolative manager will respond more because of his/her larger $\gamma_{oe}$. In addition, overextrapolative firms underestimate the output’s mean reversion, thus recovering more slowly.

A direct implication of the last point is that overextrapolative firms can display low investment and high payout during the recovery process. This prediction is in line with what Stein (2012) documents about the post-2008 crisis period. Firms have taken advantage of the low borrowing costs (caused by the Federal Reserve’s easing monetary policy) to raise more funding. However, instead of investing, they choose to pay out dividends (through share buybacks). Stein explains this phenomenon by arguing that firms are unlikely to alter their investment plans just because borrowing rates have been temporally lowered by monetary policy. Our model makes it explicit that firms’ investment plans are determined by both financing costs and expected productivity. Easing monetary policy may reduce financing costs; however, it plays a limited role in boosting investment if firms believe that the low productivity will persist; for instance, if they have overextrapolation bias.32

6 Risk management

A firm can use financial hedging to reduce its exposure to productivity risk. In this section, we study how hedging interacts with information processing biases. We specifically ask two questions: (1) Do biases change the optimal hedging policy? (2) Does hedging affect how biases influence firm behavior?

We allow firms to use futures contracts on the stock market index to hedge productivity risk. Let $F_t$ denote the futures price at time $t$, given by

$$dF_t = \sigma_m F_t dB_t,$$

32See Cujean and Hasler (2017) for a similar argument.
where $\sigma_m > 0$ is the volatility of index returns and $B_t$ is a standard Brownian motion under the risk-neutral measure. Additionally, suppose that $dF_t$ is correlated with the output $dA_t$, with the correlation coefficient equal to $\rho$, and uncorrelated with the signal $ds_t$.

We denote $\psi_t$ as the fraction of cash that a firm puts in the futures position, i.e., the hedge ratio. We denote $\xi_t$ between 0 and 1 as the fraction of cash held in the required margin account. Suppose that the firm’s cash holdings incur an additional flow cost of $\epsilon$ per dollar because of the margin requirements. Following Bolton, Chen, and Wang (2011), we assume that the firm’s futures position (in absolute value) cannot exceed a constant multiple $\pi$ of the amount of cash in the margin account. Consequently, the upper limit on the firm’s futures position is given by

$$|\psi_t W_t| \leq \pi \xi_t W_t. \quad (19)$$

Without loss of generality, we consider the case $\rho > 0$; thus, the optimal hedge ratio $\psi_t$ is negative. We show in the appendix that the interior solution to the optimal hedge ratio satisfies\(^\text{33}\)

$$\psi_t^*(x, w) = \frac{1}{w} \left( -\frac{\rho \sigma_A}{\sigma_m} - \frac{\epsilon}{\pi \rho_{ww}} \frac{1}{\sigma_m^2} - \frac{\rho \gamma}{\sigma_m \sigma_A \rho_{ww}} \right). \quad (20)$$

The first two terms in the parentheses are also in Bolton, Chen, and Wang (2011);\(^\text{34}\) the last term is new, capturing the effects of belief-updating and its interaction with cash.

Given that $p_{xw} > 0$, $p_{ww} < 0$ and that the whole term in (20) is negative, the overconfident firm holds more futures in hedging compared to the rational firm because of its smaller $\gamma_{oc}$. In contrast, the overextrapolative firm possesses a larger $\gamma_{oe}$ and thus takes a lower futures position in hedging. Here, the key parameter $\gamma$ relates to the correlation of the firm’s free

\(^{33}\)Similar to Bolton, Chen, and Wang (2011), $\psi_t^*$ may take the corner solution equal to $(-\pi)$ in the low-cash region, and zero in the high-cash region.

\(^{34}\)The first term captures the firm hedging against output shocks; the second term represents the firm hedging against the fluctuations in cash.
cash flows $dY_t$ and the innovation in expected productivity $dx_t$. For illustration, equation (9) is rewritten as follows:

$$
\begin{align*}
    dY_t & = K_t dA_t - (I_t + G(I_t, K_t)) dt \\
        & = (x_t K_t - I_t - G(I_t, K_t)) dt + \sigma_A K_t \left( \frac{dA_t - x_t dt}{\sigma_A} \right). \\
    & \quad \text{output shock}
\end{align*}
$$

(21)

Comparing this equation with equation (8), we can see that both $dY_t$ and $dx_t$ are driven by output shocks, and their correlation increases with $\gamma$. Therefore, the overconfident firm has a smaller correlation, and the overextrapolative firm has a larger correlation, which influences their volatilities of free cash flows and, in turn, the optimal hedge ratios.

Take the overconfident firm as an example. Suppose that there is a positive output shock $dA$ that increases the free cash flows $dY$. The firm adjusts its expected productivity upward ($dx > 0$) and invests more. However, because of overconfidence, $dx$ is smaller than that of the rational firm ($\gamma_{oc} < \gamma$). As a result, the overconfident firm underinvests ($I$ being smaller), which further leads to larger free cash flows $dY$.\footnote{This result occurs because investment expenditures are subtracted from $dY$.} In contrast, when there is a negative shock in $dA$, depressing both $dY$ and $dx$, the overconfident firm does not sufficiently cut the expected productivity or investment. The firm then overinvests, which further decreases $dY$. In summary, overconfidence leads to more volatile free cash flows and, consequently, a higher optimal hedge ratio. In the same logic, overextrapolation reduces the volatility of free cash flows and thus decreases the optimal hedge ratio.

Next, we show that allowing for hedging does not qualitatively change the impacts of overconfidence and overextrapolation on firm behavior. First, the shut-down boundaries of the rational, overconfident, and overextrapolative firms are 0.035, 0.032, and 0.051, respectively, under optimal hedging.\footnote{Parameter values are set as before. We also choose additional parameters $\rho = 0.8$, $\sigma_m = 0.2$, $\pi = 5$, and $\epsilon = 0.005$ from Bolton, Chen, and Wang (2011).} Compared to the values reported in section 4, hedging can
Figure 6: Biases and optimal policies with hedging

Notes: Panels A (D), B (E) and C (F) illustrate the payout boundaries, refinancing sizes, and zero-investment boundaries for the rational and overconfident (overextrapolative) firms, respectively.

Figure 7: Investment gaps with hedging

Notes: Panels A (B) depicts the investment gaps between the overconfident (overextrapolative) firm and the rational firm.
reduce the shut-down boundaries for all three types of firms.\textsuperscript{37} However, hedging does not change the relative order of the shut-down boundaries among the three types of firms—the overconfident firm still has a lower shut-down boundary than the rational firm, and the overextrapolative firm has a higher boundary.

We then compare the optimal policies in payout, refinancing, and investment, under optimal hedging, for the overconfident versus rational firms, and for the overextrapolative versus rational firms in Figures 6 and 7. The comparisons are generally similar to those in Section 4, suggesting that hedging does not fundamentally change the influences of managers’ information processing biases on firm behavior.

7 Conclusion

We introduce a unified model of firm investment, payout, refinancing, and risk management in which there are both information processing biases (overconfidence or overextrapolation) and external financing costs. We find that the interaction between the biases and financing costs makes financial positions matter for the biases’ impacts on firm behavior. All else being equal, the biased beliefs influence the managers’ perceived risk and, thus, the precautionary motives. A stronger financial position dwarfs the importance of precautionary motives, therefore reducing the biases’ impacts on firm decisions. An implication is that shareholders can use liquidity management to limit the potential consequences of managers’ biased behavior. This result cannot be reached if the model features solely the information processing biases or only the costly external finance.

Our model also predicts that firms with overconfidence or overextrapolation differ in their responses to economic booms and busts from rational firms. In fact, many shocks driving business cycles can arguably be attributed to affecting the underlying productivity, which

\textsuperscript{37}The lower shut-down boundaries indicate that firms have more operating flexibility in low-productivity regions—and, thus, higher survival probabilities—when they are able to hedge. This result supports the complementarity between risk management and operating flexibility as documented by Mello, Parsons, and Triantis (1995).
is likely to be unobserved in the real world. Firms must therefore form expectations using available information. We abstractly classify the information into (i) soft signals and (ii) past output. Overconfidence and overextrapolation induce firms to overly rely on one of these information sources while undervaluing the other, therefore biasing firms’ behavior. Our results provide a new explanation for why firms cut investment to different extents in recessions and recover in differential paces. We stress the importance of managers’ distinct information processing, while much of the existing literature focuses on the heterogeneity in firms’ financial positions. These two explanations are complementary.

Our unified modeling framework itself is of independent interest. The literature has noted that firm decisions on investment, financing, payout, and liquidity management are largely made at the same time. Having a unified model to study the simultaneous decisions can generate results that are not available if those decisions are considered separately. For instance, we predict that an overextrapolative firm tends to underinvest in the post-recession period, leading to excessive dividend payouts. This payout dynamics are in line with what happened after the 2008 crisis. The model would not be able to generate this result without considering investment, cash holdings, and payouts simultaneously.

Our model can be further extended to allow for endogenous growth, i.e., $u_t$ relying on capital formation. Then, the underinvestment caused by overextrapolation would decrease the underlying, long-run productivity, which in turn leads to more underinvestment. This vicious cycle is able to predict the so-called secular stagnation, in which the productivity would remain low for a lengthened period after a recession. We leave it for future research.
References


Appendix

A. Negative investment gap in panel D of Figure 3.

Figure A1 plots the investment gap between the overextrapolative firm and the rational firm, $i_{oe}(x, w) - i(x, w)$, when this gap is negative. In this figure, $x$ is expected productivity; $w$ is cash stock. Holding $x$ constant, the absolute value of the negative gap first decreases with $w$ and then increases with $w$. Holding $w$ constant, the absolute value of the gap shrinks with $x$.

Figure A1: Negative investment gaps (overextrapolative vs. rational)

---

B. Impulse analysis of positive shocks

Figure A2 shows the impulse responses of the overconfident firm (in “circles”) to a one-unit positive output shock (panel A) and a signal shock (panel B), comparing them to the impulse responses of the rational firm (in “stars”). Panel A shows that the overconfident firm underreacts to a positive output shock in upward adjusting $x_t$ and thus underinvests compared to the rational firm. These results are in line with equation (8) ($\gamma_{oe} < \gamma$). Both firms’ cash inventories increase upon the positive output shock, exceeding the target cash levels. Consequently, both firms pay out dividends, but the overconfident firm pays out more because of its lower investment expenditure. The situation is different for a signal shock (panel B). The overconfident firm upward adjusts its expected productivity to a higher position compared to the rational firm, consistent with equation
(8) ($\eta_{oe} > \eta$). The positive signal induces both firms to increase investment and cut dividends, while the overconfident firm acts more aggressively. Moreover, the higher expected productivity decreases the precautionary motives, so both firms tend to hold less internal cash after the shock hits. All the variables return back to normal as the expected productivity recovers.

Figure A2: Impulse responses (overconfident vs. rational)

Panel A. Positive output shock.

Panel B. Positive signal shock.

Note: Panel A shows the impulse responses to a positive output shock of the overconfident firm (“circles”) and those of the rational firm (“stars”). The shock $dZ = 1$ is introduced at $t = 2\bar{dt}$, supposing that there are no additional shocks. At each $dt$, the impulse responses are calculated as the percentage changes (in decimal) of the expected productivity, cash stock, payout, and investment from the baseline case without introducing the shock. Panel B instead shows the impulse responses to a positive signal shock $dZ^s = 1$.

Figure A3 compares the impulse responses of the overextrapolative firm (in “circles”) with those of the rational firm (in “stars”). In panel A, when there is a (unit and positive) output shock, the overextrapolative firm overreacts in adjusting $x_t$ upward and overinvesting compared to the rational firm. Additionally, in the subsequent periods, the overextrapolative firm’s $x_t$ and investment levels decline back to normal at a slower speed. The larger jump-up corresponds to the larger $\gamma_{oe}$ in equation (8), and the slower decline is in line with the smaller $\kappa_{oe}$. The cash levels of both firms increase.

41
upon the positive output shock, exceeding the target cash levels. Both firms pay out dividends; however, the dividend levels of the two firms become slightly lower than normal in the subsequent periods because of overinvestment. Upon a signal shock (panel B), the overextrapolative firm has the same initial jump-up in $x_t$ as the rational firm does because overextrapolation does not change how managers view the soft signal. However, the overextrapolative manager’s expected productivity returns to normal more slowly because of his/her stronger belief in persistence. The firm with this biased belief therefore invests more and pays out less. The higher expected productivity also reduces the overextrapolative firm’s precautionary motives and, thus, its cash holdings.

Figure A3: Impulse responses (overextrapolative vs. rational)

Panel A. Positive Output Shock.

Panel B. Positive Signal Shock.

Note: Panel A shows the impulse responses to a positive output shock of the overextrapolative firm (“circles”) and those of the rational firm (“stars”). The shock $dZ^A = 1$ is introduced at $t = 2dt$, supposing that there are no additional shocks. At each $dt$, the impulse responses are calculated as the percentage changes (in decimal) of the expected productivity, cash stock, payout, and investment from the baseline case without introducing the shock. Panel B instead shows the impulse responses to a positive signal shock $dZ^s = 1$.

C. Optimal hedge ratio with the interior solution

According to Bolton, Chen, and Wang (2011), it is optimal for a firm to hold the minimum amount of cash in the margin account because it is costless to reallocate the cash when needed.
Thus, according to the upper limit on the firm’s futures position, we have

$$|\psi_t W_t| = \pi \xi_t W_t \text{ for } \epsilon > 0. \tag{A.1}$$

Therefore, the firm’s cash reserves evolve according to

$$dW_t = x_t K_t dt + (r - \lambda) W_t dt - I_t dt - G(I_t, K_t) dt + dH_t$$

$$- dU_t + \sigma_A K_t \frac{dA_t - x_t dt}{\sigma_A} - \epsilon \xi_t W_t dt + \psi_t \sigma_m W_t dB_t. \tag{A.2}$$

Using the same steps as in Section 3, we can derive the scaled value $p(x, w)$ of a firm that engages in hedging, satisfying the following HJB equation in the internal funding region:

$$rp(x, w) = \max_{i, \psi, \xi} \left( i - \frac{\theta}{2} i^2 + (r - \lambda) w - \epsilon \xi w \right) p_w$$

$$+ \kappa(\bar{u} - x) p_x + \frac{1}{2} \left[ \sigma_A^2 + \psi^2 \sigma_m^2 w^2 + 2 \rho \sigma_m \sigma_A \psi w \right] p_{ww}$$

$$+ \frac{1}{2} \left( \eta^2 \mu + \gamma^2 \right) p_{xx} + \left( \gamma + \frac{\rho \gamma \sigma_m \psi}{\sigma_A} w \right) p_{wx}, \tag{A.3}$$

subject to

$$\xi = \min \left\{ \frac{|\psi|}{\pi}, 1 \right\}. \tag{A.4}$$

Without loss of generality, we focus on the case with positive correlation ($\rho > 0$); thus, the firm short sells index futures (i.e., $\psi < 0$) for optimal hedging. We consider the case with an interior solution for $\psi$ such that $\xi = -\psi/\pi < 1$. Taking the first-order condition with respect to $\psi$ for equation (A.3), we can obtain the optimal hedge ratio given by equation (20), i.e.,

$$\psi^*(x, w) = \frac{1}{w} \left( -\frac{\rho \sigma_A}{\sigma_m} - \frac{\epsilon}{p_w} - \frac{1}{\sigma_m^2} - \frac{\rho \gamma \sigma_A}{\sigma_m^2} \right).$$
Money growth and inflation: International historical evidence on high inflation episodes for developed countries

ISBN 978-952-323-257-0, online

IQ, Expectations, and Choice

ISBN 978-952-323-258-7, online

Domestic banks as lightning rods? Home bias and information during Eurozone crisis

ISBN 978-952-323-260-0, online

Model-based regulation and firms' access to finance

ISBN 978-952-323-261-7, online

Assessing U.S. aggregate fluctuations across time and frequencies

ISBN 978-952-323-263-1, online

On the long-run calibration of the Credit-to-GDP gap as a banking crisis predictor

ISBN 978-952-323-264-8, online

The effect of TLTRO-II on bank lending

ISBN 978-952-323-269-3, online

Stagnant wages, sectoral misallocation and slowing productivity growth

ISBN 978-952-323-273-0, online

Can large trade shocks cause crises? The case of the Finnish-Soviet trade collapse

ISBN 978-952-323-274-7, online

Assessing reliability of aggregated inflation views in the European Commission consumer survey

ISBN 978-952-323-277-8, online

U.S. Macroeconomic Policy Evaluation in an Open Economy Context using Wavelet Decomposed Optimal Control Methods

ISBN 978-952-323-279-2, online

The Phillips Curve at 60: time for time and frequency

ISBN 978-952-323-280-8, online

Friends for the benefits: The effect of political ties on sovereign borrowing conditions

ISBN 978-952-323-282-2, online

Predicting systemic financial crises with recurrent neural networks

ISBN 978-952-323-287-7, online
<table>
<thead>
<tr>
<th>Date</th>
<th>Author(s)</th>
<th>Title</th>
<th>ISBN</th>
<th>Online</th>
</tr>
</thead>
<tbody>
<tr>
<td>15/2019</td>
<td>Tuomas Takalo</td>
<td><strong>Switching costs in the Finnish retail deposit market</strong></td>
<td>978-952-323-289-1, online</td>
<td></td>
</tr>
<tr>
<td>16/2019</td>
<td>Sami Alpanda – Eleonora Granziera – and Sarah Zubairy</td>
<td><strong>State dependence of monetary policy across business, credit and interest rate cycles</strong></td>
<td>978-952-323-290-7, online</td>
<td></td>
</tr>
<tr>
<td>17/2019</td>
<td>Gene Ambrocio</td>
<td><strong>Measuring household uncertainty in EU countries</strong></td>
<td>978-952-323-291-4, online</td>
<td></td>
</tr>
<tr>
<td>18/2019</td>
<td>Delong Li – Lei Lu – Congming Mu – Jinqiang Yang</td>
<td><strong>Biased beliefs, costly external finance, and firm behavior: A Unified theory</strong></td>
<td>978-952-323-293-8, online</td>
<td></td>
</tr>
</tbody>
</table>