NBER WORKING PAPER SERIES

WHAT DRIVES CORPORATE SAVINGS

Nan Chen
Xavier Giroud
Ling Qin
Neng Wang

Working Paper 34622
http://www.nber.org/papers/w34622

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2026

We thank Patrick Bolton, Joao Cocco, Christopher Hennessy, Harrison Hong, Jennifer Huang, Jussi Keppo, David Lando, Ben Moll, Anna Pavlova, Shuaijie Qian, Thomas Sargent, Henri Servaes, Laura Starks, Mindy Xiaolan, as well as seminar participants at CKGSB, Copenhagen Business School, London Business School, National University of Singapore, and Stockholm School of Economics for helpful comments and suggestions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

What Drives Corporate Savings
Nan Chen, Xavier Giroud, Ling Qin, and Neng Wang
NBER Working Paper No. 34622
January 2026
JEL No. C61, D92, G31, G32

ABSTRACT

We study the determinants of corporate cash holdings by extending the standard q theory of investment with financing frictions and productivity shocks. While existing models predict a low propensity to hold cash, we show that realistic cash holdings arise only when three ingredients are combined: 1.) costly external financing, 2.) persistent productivity shocks, and 3.) contemporaneous productivity shocks. With costly external financing, persistent productivity shocks generate predictable cash flows and investment opportunities, but yield little need for savings since the internally generated cash flows are aligned with the investment needs. Contemporaneous shocks make internally generated cash flows random, inducing firms to hold cash at levels consistent with those observed empirically.

Nan Chen
The Chinese University of
Hong Kong (CUHK)
Faculty of Engineering
Systems Engineering and
Engineering Management
nchen@se.cuhk.edu.hk

Xavier Giroud
Columbia University
Columbia Business School
and NBER
xavier.giroud@gsb.columbia.edu

Ling Qin
ShanghaiTech University
qinling@shanghaitech.edu.cn

Neng Wang
Cheung Kong Graduate School of Business
nwang@ckgsb.edu.cn

1 Introduction

Cash management is at the very core of corporate finance. Companies hold cash for at least two reasons. First, cash provides funds for firms to finance their investments ("investment motive"). Second, cash provides a buffer that helps insulate firms from negative shocks ("precautionary saving motive"). Underlying both motives is the notion of financing flexibility. By holding cash, firms can respond to changes in their environment without having to issue external financing. Such flexibility is especially valuable when external financing is costly. Because of the importance of cash management, it is often seen as one of the main tasks of the Chief Financial Officer (CFO), as well as the Chief Executive Officer (CEO) when it comes to financial management (e.g., Brealey, Myers, Allen, and Edmans, 2025). Empirically, firms hold substantial amounts of cash. In 2024, the average cash-to-asset ratio across all U.S. publicly-traded companies in Compustat was 22.7%. That is, cash accounts for more than 20% of the average firm's balance sheet.[1]

Despite the importance of cash management, existing models in the corporate finance literature typically predict that companies hold low levels of cash, which is at odds with what is observed in the data. This discrepancy between theory and empirics is often referred to as the cash puzzle.[2] In this paper, we develop a dynamic model of investment, financing, and cash management for a firm that faces costly external finance. Using this model, we characterize the economic forces that are needed to generate realistic levels of cash holdings. In a nutshell, the main insight from our model is that, in a standard q-theory model of investment, three forces are needed: 1.) costly external financing, 2.) persistent productivity shocks, and 3.) contemporaneous productivity shocks.

The backbone of our model is the standard q-theory of investment as in Hayashi (1982) and Abel and Eberly (1994). This is a well-established framework in both the finance and macroeconomics literature, and hence a natural starting point for our analysis. In this framework, to generate persistent corporate earnings, uncertainty is often modeled in the form of a persistent productivity shock—e.g., as a first-order autoregressive (AR(1)) process

[1] This phenomenon is not unique to the U.S. As Pinkowitz, Stulz, and Williamson (2016) show, large cash holdings are also observed in public companies throughout the world.

[2] See Denis and Wang (2024) for a recent survey of the literature on corporate cash holdings and a discussion of the cash puzzle.

in which productivity in one period is given by the previous period's productivity plus the persistent productivity shock. However, in such a model, cash holdings are indeterminate, as firms are neither better off nor worse off by holding any cash. Put differently, as firms do not face any financing frictions, they can always invest in positive-NPV projects.

To account for financing frictions and generate economically meaningful financing policies, we augment the standard q-theory framework with costly external financing. We model the cost of external financing in the same way as in the literature, e.g., Hennessy and Whited (2007), Riddick and Whited (2009), and Bolton, Chen, and Wang (2011), by assuming that firms can issue equity and, when they do, they incur both fixed and marginal costs. In this setting, firms have incentives to hold cash to avoid (or delay) costly equity issuance and to finance their future investment needs, particularly when productivity is high.

While intuitive, we show that this specific cash holding motive is quantitatively negligible. This is because, when firms are hit by positive persistent productivity shocks, they know that their cash flows will be higher as well. As we show, these higher cash flows are sufficient to finance the desired level of investment. Similarly, when firms are hit by negative persistent productivity shocks, their cash flows are lower but sufficient since their investment needs are lower as well. This underscores the important, yet subtle distinction between *cash financing* (when firms use their cash holdings to finance investments) and *cash flow financing* (when firms use their internally generated cash flows to finance investments). Effectively, in a q-theory model with costly external financing and persistent productivity shocks, firms have little need to hold cash—their internally generated cash flows provide a source of financing that is sufficient to finance their desired level of investments. Accordingly, such a model yields a solution that is very close to the first-best Modigliani Miller (MM) solution.

Rather, for firms to hold realistic amounts of cash, we show that a third ingredient is needed: contemporaneous productivity shocks. We model these shocks as in Bolton, Chen, and Wang (2011) by assuming that a firm's productivity is subject to contemporaneous *i.i.d.* shocks. As mentioned above, persistent productivity shocks have little bearing on cash management since they give rise to internally generated cash flows that are sufficient to finance investment. However, with contemporaneous shocks, these internally generated cash flows are uncertain and hence need not be sufficient to finance the desired level of investment. Moreover, the cash balance one period ahead is uncertain as well. This induces firms to increase their precautionary savings to reduce their future reliance on costly external financing. Accordingly, both the investment motive and the precautionary saving motive induce firms to maintain cash buffers. In our quantitative analysis, we show that, under reasonable parameter values, such a model generates cash holdings that are consistent with those that are observed in the data.

In sum, the key takeaway from our analysis is that q-theory models need three economic forces to generate realistic levels of cash holdings: 1.) costly external financing, 2.) persistent productivity shocks, and 3.) contemporaneous productivity shocks. All three of them are needed. First, without costly external financing, there is no need to hold cash, as any financing need can be satisfied by raising new capital at no cost. Second, persistent productivity shocks are needed to generate persistent cash flows and a realistic investment dynamic. Third, contemporaneous productivity shocks are needed to generate sufficient uncertainty in future cash flows, and hence in the firms' ability to use their internally generated cash flows as a source of financing. If only persistent shocks are considered, the high predictability of future cash flows would trivialize the need to maintain cash buffers.

Another way to see why contemporaneous productivity shocks are crucial in inducing firms to hold cash—and why persistent productivity shocks alone are not sufficient—is by considering the timing of these shocks. In the q-theory framework, persistent productivity shocks are typically modeled as AR(1) innovations and hence affect cash flows with a $\mathrm{d}t$ lag. That is, from the perspective of time t, the internally generated cash flows over the interval $(t, t + \mathrm{d}t)$ are deterministic and can be used to finance investment. Accordingly, firms have little need to hold cash. However, with contemporaneous shocks, these internally generated cash flows have a stochastic component that cannot be anticipated at time t. This induces firms to hold cash to avoid costly equity issuance.

In Figure 1, we illustrate this key insight by using binomial trees to compare our model with standard models. Panel A depicts the timeline in standard models without contemporaneous shocks such as Hennessy and Whited (2007) and Riddick and Whited (2009). The firm enters period t with cash holding W_t, observes the persistent productivity shock z_t, and chooses investment I_t. Because z_t is known, the firm can perfectly predict its internally generated cash flows this period and thus its next period's cash holding, W_{t+1}. Accordingly, the firm can choose a level of investment I_t that is fully funded by internally generated cash flows, thereby avoiding costly equity issuance. Since the benefit of avoiding costly equity issuance is large relative to the cost of a one-period under-investment, the firm has little incentive to carry cash into period t.

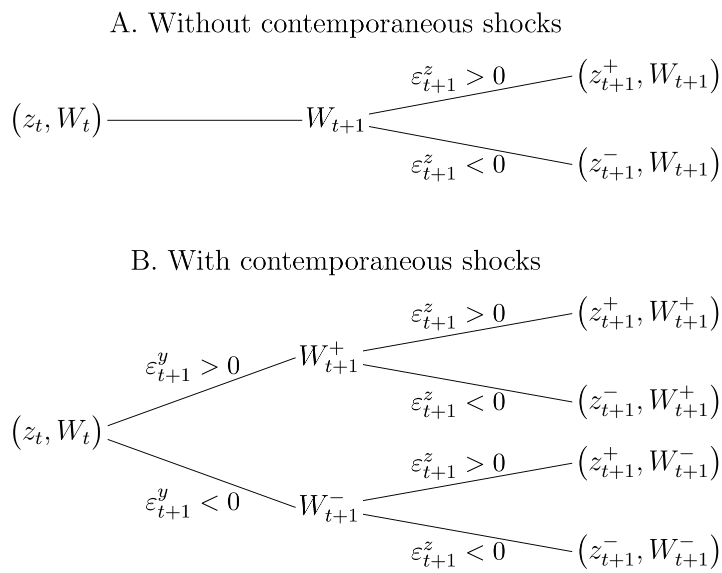

Figure 1: **Using binomial trees to illustrate the effects of contemporaneous shocks on savings.** Panels A and B depict the evolution of key state variables in standard models and our model with contemporaneous shocks, respectively.

Entering period $t + 1$ with cash holding W_{t+1}, the firm observes the realized persistent productivity shock ε^z_{t+1}, causing z_{t+1} to equal either z^+_{t+1} or z^-_{t+1}, and then chooses its investment, I^+_{t+1} or I^-_{t+1}, respectively. Again by choosing its investment I_{t+1} after observing the realized z_{t+1}, the firm can avoid costly equity issuance in period $t+1$, by the same reasoning as in period t. Iterating this reasoning forward, we thus conclude that there is little need, if any, for the firm to save starting from the beginning ($t = 0$).

However, this iterative reasoning no longer applies in the model with contemporaneous shocks illustrated in panel B. This is because the internally generated cash flows in period t are no longer deterministic, as captured by the two potential realizations of the contemporaneous shock ε^y_{t+1}. As a result, these cash flows may no longer suffice to finance a non-state-contingent level of investment I_t. Anticipating this randomness, the firm holds cash buffers at time t to avoid operating losses in the bad state (where $\varepsilon_{t+1}^{y} < 0$), which could trigger a costly equity issuance.

In summary, starting in a given period, the firm anticipates four states in the next period when contemporaneous shocks are present, but only two states otherwise (compare the last columns of panels A and B). This is because only with contemporaneous shocks is the next period's cash holding *random*.

The use of continuous-time modeling also helps understand why contemporaneous shocks are crucial to generate realistic levels of cash holdings. Contemporaneous shocks make cash accumulation a diffusion process, which in turn gives rise to two additional terms in the Bellman equation—the second derivative P_{WW} and the cross-hedging term P_{zW}, where W is cash holdings and z captures the firm's persistent productivity—that would otherwise be missing.[3] The first term (P_{WW}) captures the precautionary saving motive, while the second term (P_{zW}) captures the investment motive since it ties directly to the (persistent) productivity shocks. These terms add considerable curvature to the value function, which ultimately increases the need to hold cash.

As the above considerations illustrate, contemporaneous shocks are crucial in explaining firms' cash holding policies. The importance of contemporaneous shocks is underscored by practitioners. Indeed, anecdotal evidence abounds with examples of managers highlighting the challenges of forecasting cash flows. For example, in a recent interview, Jamie Dimon, CEO and Chairman of JP Morgan Chase, noted that "*[q]uarterly earnings: they're a function of the weather, commodity prices, volumes, competitor pricing. And you don't really control that as CEO. [...] Sometimes you're just like the cork in the ocean*" (CNBC, 2018). Similarly, a recent report by the "big-four" accounting firm Ernst & Young that analyzed 2,400 of the largest global companies from 2017-2023 finds that companies miss their quarterly earnings forecasts by at least 10% more than 70% of the time. The report concludes that "*[u]nfortunately, most companies are not very good at it [cash flow forecasting]. Many have responded instead by increasing balance sheet buffers to soften the impact of forecasting variances*" (Ernst & Young, 2024), explicitly linking the difficulty of forecasting cash flows

[3]For example, in Hennessy, Levy, and Whited (2007), the absence of contemporaneous shocks implies that cash accumulation is locally deterministic rather than a diffusion process.

with companies' policy of maintaining higher cash holdings, as predicted by our model.

The latter is further substantiated by empirical studies. Indeed, the empirical literature documents that firms face substantial cash flow volatility, and that firms with higher cash flow volatility tend to hold more cash on their balance sheet (e.g., Bates, Kahle, and Stulz, 2009; Duchin, 2010; Graham and Leary, 2018). This is consistent with the predictions of our model and confirms the importance of considering contemporaneous shocks when modeling firms' cash holding policies.

Related literature. Our characterization of the economic forces that induce firms to hold cash helps explain why existing models in the literature typically predict that companies hold low levels of cash. As mentioned above, three ingredients are needed—1.) costly external financing, 2.) persistent productivity shocks, and 3.) contemporaneous productivity shocks. As we show, if any of these three is missing, the model cannot deliver realistic predictions for cash holdings. This is precisely the challenge with existing models, as they feature at most two of these forces. In particular, models in the spirit of Hennessy, Levy, and Whited (2007), Hennessy and Whited (2007), and Riddick and Whited (2009) only consider the first two forces. Dynamic investment models such as Bolton, Chen, and Wang (2011) only include the first and third forces. Finally, q-theory models of investment in the tradition of Hayashi (1982) and Abel and Eberly (1994) do not include costly external financing. Loosely speaking, while many insights could be gained from these models, they all miss one ingredient in order to explain the large cash holdings that are observed in the data. In this regard, our model helps connect the dots between different types of models from the perspective of liquidity management.[4]

The relationship between our paper and the work by Hennessy, Levy, and Whited (2007), Hennessy and Whited (2007), and Riddick and Whited (2009) warrants more discussion. Their models—which feature both costly external financing and persistent productivity shocks—are often seen as workhorse models of firm policies, and do offer predictions in terms of cash dynamics. However, it is worth stressing that these models do not aim to

[4]Also related is the model of Décamps, Gryglewicz, Morellec, and Villeneuve (2017) that distinguishes between permanent and transitory cash flow shocks. However, their setup is quite different from ours, as it does not feature investment dynamics.

explain the level of cash holdings that is observed in the data. In fact, the authors explicitly acknowledge this limitation, noting that "*we do not tackle directly the issue of the high level of corporate cash holdings*" (Riddick and Whited, 2009, p. 1729).

As in Hennessy, Levy, and Whited (2007), we specify our model in continuous time and assume that the production function features constant returns to scale.[5] This is a standard assumption in q-theory models and investment-based asset pricing models (e.g., Lucas and Prescott, 1971; Hayashi, 1982; Liu, Whited, and Zhang, 2009). An alternative would be to assume decreasing returns to scale as in the discrete-time settings of Hennessy and Whited (2007) and Riddick and Whited (2009). Having decreasing returns to scale would matter for the cash policies of small firms. Indeed, under decreasing returns to scale, smaller firms have stronger incentives to save since they have higher marginal products of capital from investing. This yields some saving demand among smaller firms, but only to a limited extent since it is mitigated by the higher costs they face in raising their cash holdings.[6] Importantly, this channel does not explain the saving behavior of larger firms, nor does it explain the large cash savings observed in the data, as emphasized in Riddick and Whited (2009, cf. above quote).[7]

More broadly, our paper is related to three strands of the literature. First, our work is related to the literature that uses dynamic models of firms facing costly external financing to characterize their investment, financing, and risk management decisions. In addition to the articles referenced above—which are directly related to our model—this literature features a wide set of other models (e.g., Cooley and Quadrini, 2001; Gomes, 2001; Almeida, Campello, and Weisbach, 2004; Moyen, 2004; DeAngelo, DeAngelo, and Whited, 2011; Décamps, Mar-

[5]Hennessy, Levy, and Whited (2007) incorporate credit rationing and convex equity issuance costs into the continuous-time q-theoretic investment model with exogenous perpetual risky debt developed in Hennessy (2004) to study the effect of debt overhang on corporate investment and firm value.

[6]Due to the fixed costs of equity issuance, smaller firms effectively face higher (shadow) cost of equity financing, consistent with the empirical evidence (e.g., Gustafson and Iliev, 2017). Moreover, as smaller firms have higher marginal returns of investment, their opportunity costs of holding cash (rather than investing) are also higher.

[7]Another distinction between the models of Hennessy and Whited (2007) and Riddick and Whited (2009) and ours is their use of discrete time, while our model is set in continuous time. Although our conclusions are not sensitive to this modeling choice, we note that their use of discrete time increases the need to hold cash. This is because firms can only respond to shocks at discrete frequencies (e.g., once a year) and hence have additional incentives to maintain cash buffers. In practice, however, shocks can occur anytime and firms are unlikely to delay their response until the next discrete time period.

iotti, Rochet, and Villeneuve, 2011; Hugonnier, Malamud, and Morellec, 2015; Nikolov, Schmid, and Steri, 2019; Abel and Panageas, 2023; Dai, Giroud, Jiang, and Wang, 2024; Hartman-Glaser, Mayer, and Milbradt, 2025). We complement this literature by highlighting the economic forces that are needed to generate realistic levels of cash holdings. In doing so, we describe how existing models need to be extended accordingly.

Second, our paper is related to the large empirical literature that studies the determinants of cash holdings (e.g., Harford, Mansi, and Maxwell, 2008; Bates, Kahle, and Stulz, 2009; Duchin, 2010; Pinkowitz, Stulz, and Williamson, 2016; Graham and Leary, 2018). This literature further highlights the high level of cash holdings observed in the data, and the lack of guidance from the theoretical literature to explain this phenomenon (e.g., Denis and Wang, 2024). Our paper aims to fill this gap by outlining the economic forces that models need to include in order to generate levels of cash holdings that align with those observed empirically.

Third, while our model is as parsimonious as possible—building on the standard q-theory framework—the three economic forces present in the model make it a mathematically challenging problem to solve. Unlike Hennessy, Levy, and Whited (2007), who use a Hamilton-Jacobi-Bellman (HJB) equation and associated first-order conditions (e.g., for equity issuance) to tease out empirical predictions, we fully characterize the firm's value function and optimality. Specifically, we use the variational-inequality methodology to analytically and intuitively characterize Tobin's average q and optimal state-contingent policies, via four mutually exclusive regions. The four-region solution further gives rise to an efficient, accurate numerical algorithm for a set of equations including a nonlinear partial differential equation (PDE) for Tobin's q and first-order conditions for investment, payouts, savings, and liquidation. Our model is the first to use a variational inequality and apply the associated numerical solution method to q-theoretic models with persistent shocks.[8]

The remainder of this paper is organized as follows. Section 2 presents the baseline model. Section 3 characterizes the first-best solution. Section 4 describes the model solution. Section

[8]Using the variational-inequality approach, Bolton, Wang, and Yang (2019) analyze real-options exercising and valuation for a firm facing costly external financing, and Dai, Giroud, Jiang, and Wang (2024) study internal capital markets for a financially constrained conglomerate subject to independently and identically distributed shocks.

5 provides a quantitative analysis. Section 6 examines the comparative statics. Section 7 provides an extension of our baseline model with debt financing and shows that our results continue to hold in this more general setting. Finally, Section 8 concludes.

2 Model

In this section, we develop a dynamic model of corporate investment, financing, and valuation with the following key features. First, firms face persistent productivity shocks as in the classic q theory of investment, e.g., Abel and Eberly (1994), and dynamic corporate finance models, e.g., Hennessy (2004), Hennessy, Levy, and Whited (2007), Hennessy and Whited (2005, 2007), and Riddick and Whited (2009). This shock is important to generate persistent corporate earnings. Second, firms face contemporaneous (transitory) shocks as in continuous-time corporate finance models, e.g., Bolton, Chen, and Wang (2011) and Décamps, Mariotti, Rochet, and Villeneuve (2011), and dynamic contracting models, e.g., DeMarzo and Sannikov (2006) and DeMarzo, Fishman, He, and Wang (2012). Third, firms face costly external financing as in the dynamic corporate finance literature, e.g., Hennessy and Whited (2005, 2007), Riddick and Whited (2009), Bolton, Chen, and Wang (2011), Décamps, Mariotti, Rochet, and Villeneuve (2011), and Hugonnier, Malamud, and Morellec (2015).[9] Finally, following Lucas and Prescott (1971) and Hayashi (1982), we make assumptions for both production and financing sides so that our model has the homogeneity property.

As we show later, a key result of our model is that the firm saves only when facing contemporaneous shocks. Even when facing persistent productivity shocks, the firm can effectively finance its investment using internally generated cash flows in almost all periods without tapping costly external financing. As a result, the firm effectively has no incentive to save absent contemporaneous shocks.

Capital accumulation and adjustment costs. The firm uses capital for production. Let K_t and I_t denote the firm's capital stock and gross investment at time t, respectively.

[9] In this section, we focus on the case where external financing is in the form of equity financing. In Section 7, we extend our baseline model to allow the firm to borrow and show that our results continue to hold.

Starting with an initial level of $K_0 > 0$, the firm accumulates capital as follows:

$$\mathrm{d}K_t = (I_t - \delta K_t)\mathrm{d}t, \quad t \geq 0, \tag{1}$$

where $\delta \geq 0$ is a constant depreciation rate of capital. When accumulating capital, the firm incurs a cost C_t, which we specify as a function of I_t and K_t, $C(I_t, K_t)$, per unit of time including a purchase cost of capital (if $I_t > 0$) and various capital adjustment costs. As in Lucas and Prescott (1971), Hayashi (1982), Hennessy, Levy, and Whited (2007), and Liu, Whited, and Zhang (2009), we assume that $C(I_t, K_t)$ is homogeneous with degree one in I_t and K_t, which allows us to write $C(I_t, K_t) = c(i_t)K_t$ where $i_t = I_t/K_t$. We discuss details of the $c(i)$ specification in subsection 5.1 where we conduct our quantitative analysis. The firm can also liquidate itself at any time. Let η_ℓ denote the liquidation time.

Productivity and output. The firm's output over a small time interval $(t, t + \mathrm{d}t)$, $\mathrm{d}Y_t$, is proportional to its time-t capital stock K_t so that we can express its net operating profit over this time interval as:

$$\mathrm{d}Y_t = A_t K_t(\mathrm{d}t + \sigma_y \mathrm{d}\mathcal{B}_t^y) - C(I_t, K_t)\mathrm{d}t, \tag{2}$$

where $\{A_t\}$ is the firm's (stochastic) productivity process and $\{\mathcal{B}_t^y\}$ is a standard Brownian motion. A key parameter of our analysis is $\sigma_y > 0$, which measures the impact of the contemporaneous shock, $d\mathcal{B}_t^y$, on the firm's output, given by the first term in (2). If A_t were constant over time, then (2) would become the model analyzed in Bolton, Chen, and Wang (2011), where firms are subject to contemporaneous productivity shocks, drawn from a normal distribution with mean μdt and variance $\sigma_y^2 \mathrm{d}t$ over a small time increment $\mathrm{d}t$. These shocks are independent over time. Importantly, unlike Bolton, Chen, and Wang (2011), the firm in our model is also subject to a persistent productivity shock A_t, as in the neoclassical q theory of investment, see, e.g., Abel and Eberly (1994), Caballero (1999), and Cooper and Haltiwanger (2006).

To ease comparison of our model with widely used corporate finance models, e.g., the discrete-time formulations of Hennessy and Whited (2005, 2007) and Riddick and Whited (2009), and the continuous-time formulation of Hennessy, Levy, and Whited (2007), we assume that $\{A_t\}$ is a diffusion process and work with a monotone and smooth transformation of A_t: $z_t = f(A_t)$, where $f(\,\cdot\,)$ is twice continuously differentiable and $f'(\,\cdot\,) > 0$. Using Ito's Lemma, we know that $\{z_t\}$ is also a diffusion process, which we write as:

$$\mathrm{d}z_t = \mu_z(z_t)\mathrm{d}t + \sigma_z(z_t)\mathrm{d}\mathcal{B}_t^z, \tag{3}$$

where $\mu_z(\,\cdot\,)$ and $\sigma_z(\,\cdot\,) > 0$ are the drift and volatility functions, respectively, and $\{\mathcal{B}_t^z\}$ is another standard Brownian motion. Let ρ denote the (constant) correlation between the two Brownian shocks: $\{\mathcal{B}_t^y\}$ and $\{\mathcal{B}_t^z\}$. In subsection 5.1, we choose a mean-reverting process for z and specify $z = f(A) = \ln A$ so that the implied discrete-time productivity process in our model is the same as in Hennessy and Whited (2007) and Riddick and Whited (2009).

A key feature of the output process (2) is that the net operating profit, $\mathrm{d}Y_t$, over a small time interval $(t, t + \mathrm{d}t)$ is stochastic only when $\sigma_y > 0$. Absent this contemporaneous transitory shock (i.e., when $\sigma_y = 0$), (2) would have been locally deterministic as in widely used discrete-time models, e.g., Hennessy and Whited (2007) and Riddick and Whited (2009). Later we show quantitatively that whether shocks to $\mathrm{d}Y_t$ are contemporaneous (i.e., when $\sigma_y > 0$) or not (i.e., when $\sigma_y = 0$) makes a fundamental difference for cash holding dynamics.

External equity financing and cash management. The firm can raise external equity at any time t. Much evidence shows that external equity financing is costly due to imperfections such as information asymmetries or managerial incentive problems, as in, e.g., Jensen and Meckling (1976), Leland and Pyle (1977), and Asquith and Mullins (1986). The fixed cost is assumed to be proportional to the firm's capital stock K_t, i.e., $\Phi_t = \phi_0 K_t$ for some constant $\phi_0 \geq 0$. By modeling the fixed financing costs as proportional to firm size, we ensure that the firm does not grow out of the fixed costs.[10]

Let H_t and X_t denote the firm's cumulative external financing and cumulative equity issuance costs up to time t, respectively, with $H_0 = X_0 = 0$. Given the cumulative external financing $\{H_t\}$ and capital stock $\{K_t\}$ processes, the cumulative equity issuance costs X_t can be characterized as follows:

$$\mathrm{d}X_t = \phi_0 K_t \mathbf{1}_{\{\mathrm{d}H_t > 0\}} + \phi_1 \mathrm{d}H_t\,, \tag{4}$$

[10] Indeed, this is a common assumption in the investment literature. See Cooper and Haltiwanger (2006) and Riddick and Whited (2009), among others. If the fixed cost is independent of firm size, it will not matter when firms become sufficiently large in the long run.

where $\mathbf{1}_A$ is an indicator function that equals one if and only if event A occurs. Because equity issuance is costly, the firm will tap equity markets only intermittently.[11]

Because issuing external equity is costly, when earning profits, the firm may want to save by retaining a fraction of its profits over time to lower its external equity issuance costs in the future. This implies the following dynamics for its cash balance W_t starting with $W_0 \geq 0$:

$$\mathrm{d}W_t = (r - \lambda)W_t \mathrm{d}t + \mathrm{d}Y_t + \mathrm{d}H_t - \mathrm{d}U_t, \ t \geq 0\,. \tag{5}$$

The first term $(r-\lambda)W_t\mathrm{d}t$ on the right side of (5) is the return on the firm's cash inventory W_t at time t, where the rate of return is given by the interest income (risk-free rate $r > 0$) net of the cash-carrying cost $\lambda > 0$. The second term $\mathrm{d}Y_t$ represents the firm's net operating profit given by (2). The third and the last terms are the equity issuance amount and dividend payout amount over time interval $(t, t + \mathrm{d}t)$, respectively. Here U_t denotes the cumulative (non-decreasing) dividend payment to the shareholders up to time t (with $U_0 = 0$) and is a stochastic control chosen by the firm.

The firm is also subject to a death shock, which arrives at a constant probability $\zeta\mathrm{d}t$ over $\mathrm{d}t$, so that the exogenous death time, η_d, is exponentially distributed with a mean of $1/\zeta > 0$. Finally, when the firm is hit by a death shock at η_d or liquidates itself at η_ℓ, it is worth $W_t + L_t$ where $t = \eta_\ell \wedge \eta_d$: the sum of its cash holding W_t and the liquidation value of its capital, $L_t = \ell K_t$, where $\ell \in [0, 1)$ is the recovery fraction of its capital upon liquidation.

Firm optimization. The firm chooses investment $\{I_t\}$, payout $\{U_t\}$, the external financing $\{H_t\}$, and liquidation time η_ℓ to solve the following problem:

$$\max \ \mathbb{E}_t \left[\int_t^{\eta_\ell \wedge \eta_d} e^{-r(s-t)} (\mathrm{d}U_s - \mathrm{d}H_s - \mathrm{d}X_s) + e^{-r(\eta_\ell \wedge \eta_d - t)} \left(L_{\eta_\ell \wedge \eta_d} + W_{\eta_\ell \wedge \eta_d} \right) \right], \tag{6}$$

[11] Whether ϕ_0 is strictly positive or zero has very different implications in terms of predictions. If $\phi_0 > 0$, conditioning on issuing equity, the firm issues a lumpy amount in order to capitalize on the fixed cost. Technically, this is an impulse control problem. Otherwise, conditioning on issuing equity, the firm issues an infinitesimal amount as there is no fixed cost, but only a positive proportional issuance cost ϕ_1. Technically, the latter case features a single control. We specify equity issuance costs using a combination of fixed and proportional equity issuance costs. If, instead, we assumed that issuance costs were convex in the net amount of equity issued—as in Hennessy, Levy, and Whited (2007)—the firm would issue external equity more frequently, using it as the marginal source of financing, with each issuance being small to smooth out the convex costs. Empirically, however, equity issuance tends to be infrequent and lumpy (e.g., DeAngelo, DeAngelo, and Stulz, 2010).

subject to its capital accumulation process (1), the persistent productivity process (3), and the cash management equation (5). There are two terms in (6). The first term represents the discounted value of net cash flows to equityholders before the firm liquidates or dies, and the second term is the discounted value at liquidation or death, whichever is earlier.

Firm value depends on productivity A (equivalently z), capital stock K, and cash holding W. Let $P_t = P(z_t, K_t, W_t)$ denote the value function for the optimization problem (6). Next, we establish the following key property of the firm's optimization problem defined above.

Homogeneity property: $P(z_t, \beta K_t, \beta W_t) = \beta P(z_t, K_t, W_t)$ for any constant $\beta > 0$. Let $w_t = W_t/K_t$ denote the firm's cash-capital ratio at t. We can also express the homogeneity property using $P(z_t, K_t, W_t) = p(z_t, w_t)K_t$.[12] Tobin's average q at t, q_t, is the enterprise value, which is the firm's value net of its cash holdings W_t, $P(z_t, K_t, W_t) - W_t = (p(z_t, w_t) - w_t)K_t$, divided by the firm's capital stock K_t:

$$q_t = \frac{P(z_t, K_t, W_t) - W_t}{K_t} = p(z_t, w_t) - w_t := q(z_t, w_t)\,. \tag{7}$$

Applying Ito's Lemma to $w_t = W_t/K_t$ and using (5) for $\mathrm{d}W_t$ and (1) for $\mathrm{d}K_t$, we obtain the following $\{w_t\}$ process:

$$\mathrm{d}w_t = (r-\lambda)w_t\mathrm{d}t + A_t(\mathrm{d}t + \sigma_y \mathrm{d}\mathcal{B}_t^y) - c(i_t)\mathrm{d}t - (i_t - \delta)w_t\mathrm{d}t - \mathrm{d}u_t + \mathrm{d}h_t\,, \tag{8}$$

where $i_t = I_t/K_t$ is the investment-capital ratio at t. The first term on the right side of (8) represents the net interest income, the second term is the firm's realized profit per unit of capital, the third term is the cost of investment, the fourth term is the effect of the net investment on the cash-capital ratio, the fifth term, $du_t = dU_t/K_t$, is the payout-capital ratio, and the last term, $dh_t = dH_t/K_t$, is the net equity issuance scaled by K_t, respectively.

3 First Best

In this section, we present the solution for the first-best setting where external financing is costless: $\mathrm{d}X_t = 0$ for all $t \geq 0$. Under First Best, the Modigliani-Miller (MM) Theorem states that financing policies are indeterminate. Therefore, the firm's problem is to choose

[12]By setting $\beta = 1/K_t$, we obtain $P(z_t, 1, w_t) = P(z_t, K_t, W_t)/K_t := p(z_t, w_t)$ for some function $p(\,\cdot\,,\cdot\,)$.

investment $\{I_t; t < \eta_\ell \wedge \eta_d\}$ and liquidation time η_ℓ to solve the problem defined in (6) where $\mathrm{d}U_s - \mathrm{d}H_s - \mathrm{d}X_s = \mathrm{d}Y_s$.[13]

Using the model's homogeneity property, we can show that the firm's enterprise value, $V^{fb}(z, K)$, is proportional to its contemporaneous capital stock K: $V^{fb}(z, K) = q^{fb}(z)K$, where $q^{fb}(z)$ is Tobin's average q. Intuitively, compared to the recovery value ℓ per unit of capital, if a firm's productivity $A(z)$ is so low that its Tobin's q, $q^{fb}(z)$, is lower than ℓ, the firm should optimally liquidate itself. Because the firm's enterprise value shall increase with its productivity $A(z)$, there exists an endogenous cutoff threshold of z, denoted by $\underline{z}$, above which the firm operates as a going-concern and below which the firm is liquidated.

This suggests that in the region where $z > \underline{z}$ the following ordinary differential equation (ODE) shall hold for Tobin's q, $q^{fb}(z)$:

$$rq^{fb} = \max_i \quad \mu_z(z)q_z^{fb} + \frac{1}{2}(\sigma_z(z))^2 q_{zz}^{fb} + (i-\delta)q^{fb} + (A(z) - c(i)) + \zeta\left(\ell - q^{fb}\right) . \tag{9}$$

The first term (q_z^{fb}) on the right side of (9) represents the marginal effect of the productivity shock on Tobin's q. The second term (q_{zz}^{fb}) reflects the impact of the volatility of the persistent productivity shock on Tobin's q. The third term (q^{fb}) accounts for the marginal effect of the net investment $(i - \delta)$ on Tobin's q. The fourth term is the expected free cash flow, and the last term captures the exogenous death effect on enterprise value.

By choosing i to maximize the right side of (9), the firm optimally equates the total expected benefit per unit of time of investing with the annuitized Tobin's q, rq^{fb}, where the risk-free rate r is the proper discount rate. This maximization yields the following expression for the optimal investment-capital ratio, $i^{fb}(z) = I^{fb}(z, K)/K$:

$$i^{fb}(z) = \arg\max_i \quad q^{fb}(z)i - c(i). \tag{10}$$

The expression following the $\arg\max$ operator in (10) is the NPV of investing i units, which equals $q^{fb}(z)i$ minus the cost of investing i units, $c(i)$. This is because the value of one unit of capital equals Tobin's average q: $q^{fb}(z)$. In subsection 5.2, we further discuss the tradeoff characterized in (10) with a functional form specification of $c(\,\cdot\,)$.

[13] Of course, we can also show that the first-best model is a special case of the firm's problem defined in the preceding section. When issuing external equity is costless, the firm carries no cash, $W_t = 0$ for all t, as doing so is costly. Then, (5) is simplified to $\mathrm{d}U_t - \mathrm{d}H_t = \mathrm{d}Y_t$. Neither external equity financing $\mathrm{d}H_t$ nor equity payouts $\mathrm{d}U_t$ can be uniquely determined.

In the region where $z \le \underline{z}$, the firm is liquidated and

$$q^{fb}(z) = \ell\,. \tag{11}$$

As firm liquidation is endogenous, the following smooth-pasting condition holds when $z = \underline{z}$:

$$q_z^{fb}(\underline{z}) = 0\,. \tag{12}$$

4 Solution

In this section, we solve the firm's problem (6) when $\sigma_y > 0$ and external equity has both fixed and proportional issuance costs. We show that the firm's optimal strategy and value function are characterized by four (endogenously) determined regions: savings region, payout region, external financing region, and liquidation region.

Savings. In this region, firm value $P(z,K,W)$ solves the following partial differential equation (PDE):

$$\begin{aligned} rP \;=\; \max_I \;\; & \mu_z(z)P_z + \frac{1}{2}(\sigma_z(z))^2 P_{zz} + (I-\delta K)P_K \\ & + \big((r-\lambda)W + A(z)K - C(I,K)\big)P_W + \frac{1}{2}\sigma_y^2(A(z))^2K^2P_{WW} \\ & + \rho\sigma_z(z)\sigma_y A(z)KP_{zW} - \zeta(P-(\ell K+W))\,. \end{aligned} \tag{13}$$

Seven terms on the right side of (13) describe how various factors affect firm value. The first term (P_z) and the second term (P_{zz}) represent the drift and volatility effects, respectively, of the persistent shock z. The third term (P_K) represents the marginal effect of the net investment ($I - \delta K$). The fourth term (P_W) captures the effect of the firm's expected savings. The fifth term (P_{WW}) is the effect of the volatility of the contemporaneous shocks to operating profits, the sixth term (P_{zW}) captures the cross effect of the cash saving and productivity, and the last term describes the expected value change due to exogenous death. In sum, the firm optimally chooses I to equate the sum of these seven terms to the annuitized firm value, rP.

The optimal investment policy, $I^*(z,K,W)$, is given by:

$$I^*(z,K,W) = \arg\max_I \;\; \frac{P_K(z,K,W)}{P_W(z,K,W)}I - C(I,K). \tag{14}$$

Substituting (7) into (14) and using $i^*(z,w) = I^*(z,K,W)/K$, we can simplify (14) to:

$$i^*(z,w) = \arg\max_i \ \left(\frac{q(z,w)+w}{q_w(z,w)+1} - w\right) i - c(i). \tag{15}$$

The firm will evaluate whether capital adjustments are beneficial. If an adjustment is considered beneficial, a first-order condition (FOC) that determines the optimal investment $i^*(z,w)$ will arise. Note that (15) is simplified to (10) when equity financing is costless, as $q_w(z,w) = 0$ (implied by $P_W = 1$).

Substituting (7) into the PDE (13) and using the homogeneity property,[14] we obtain the following PDE for Tobin's average q in the savings region:

$$\mathcal{L}q(z,w) = 0\,, \tag{16}$$

where $\mathcal{L}q(z,w)$ is the operator (infinitesimal generator) that describes the savings region:

$$\begin{aligned}\mathcal{L}q(z,w) = \mu_z(z)q_z(z,w) &+ \frac{1}{2}(\sigma_z(z))^2 q_{zz}(z,w) + (i^*(z,w) - (r+\delta+\zeta))\, q(z,w)\\ &+ ((r-\lambda-i^*(z,w)+\delta)w + A(z) - c(i^*(z,w)))\, q_w(z,w)\\ &+ \frac{1}{2}\sigma_y^2(A(z))^2 q_{ww}(z,w) + \rho\sigma_z(z)\sigma_y A(z) q_{zw}(z,w)\\ &+ A(z) - c(i^*(z,w)) - \lambda w + \zeta\ell \end{aligned} \tag{17}$$

and $i^*(z,w)$ is the optimal investment-capital ratio given in (15). When the firm does not save, $\mathcal{L}q(z,w) < 0$ must hold.[15] Since the firm either saves or not, for all (z,w), the following inequality must hold:

$$\mathcal{L}q(z,w) \ \leq 0\,. \tag{18}$$

As we show later, a key determinant of corporate savings is the $q_{ww}(z,w)$ term, which only exists in the presence of contemporaneous shocks: $\sigma_y > 0$. With $\sigma_y = 0$, the firm effectively has no precautionary savings demand. This is because the internally generated cash flows can almost always finance the firm's investment needs.

Payouts. When a firm makes an equity payout, its enterprise value cannot change. This follows from the accounting identity: $P(z_t,K_t,W_t) = P(z_t,K_t,\overline{W}_t) + W_t - \overline{W}_t$, where $\overline{W}_t =$

[14]This property implies $P_K(z,K,W) = q(z,w) - wq_w(z,w)$, $P_W(z,K,W) = q_w(z,w)+1$, $P_{WW}(z,K,W) = q_{ww}(z,w)/K$, $P_z(z,K,W) = q_z(z,w)K$, $P_{zz} = q_{zz}(z,w)K$, and $P_{zW}(z,K,W) = q_{zw}(z,w)$.

[15]This inequality states that, for the given (z,w), retaining profits inside the firm lowers its value.

$\overline{W}(z_t, K_t)$ denotes the firm's cash holding immediately after it makes a payment of $W_t - \overline{W}_t$ to shareholders.[16] Using our model's homogeneity property, we can equivalently express this value-continuity condition as $q(z, w) = q(z, \overline{w}(z))$, where $\overline{w}(z) = \overline{W}(z, K)/K$.

In summary, in the payout region, Tobin's q does not vary with w conditional on z, and thus $q_w(z, w) = 0$ where $w \geq \overline{w}(z)$. When a firm chooses to save inside the firm rather than pay out to shareholders, its (net) marginal value of savings must be positive: $q_w(z, w) > 0$. Since the firm either makes a payout to shareholders or not, the following inequality must hold for all admissible (z, w):

$$-q_w(z, w) \ \leq 0\,. \tag{19}$$

External financing. When raising external equity at t, the firm optimally chooses its equity issuance size, $M_t = \mathrm{d}H_t$, to maximize its shareholders' value. Because shareholders' value must be continuous at the moment of issuance, we have:

$$P(z_t, K_t, W_t) = \max_{M_t > 0} P(z_t, K_t, W_t + M_t) - M_t - \big(\phi_0 K_t + \phi_1 M_t\big). \tag{20}$$

Letting $m_t = M_t/K_t$ and using (7) to simplify (20), we obtain:

$$q(z, w) = \mathcal{M}q(z, w) \tag{21}$$

where

$$\mathcal{M}q(z, w) := \max_{m > 0} \ q(z, w + m) - (\phi_0 + \phi_1 m) \tag{22}$$

is the operator describing the firm's equity issuance policy. When a firm chooses not to issue equity, the inequality $\mathcal{M}q(z, w) < q(z, w)$ must hold, as the left side of this inequality is the firm's value conditional on issuing equity. Combining the issuance and no issuance scenarios, the following inequality must hold for all admissible (z, w):

$$\mathcal{M}q(z, w) - q(z, w) \ \leq 0\,. \tag{23}$$

Liquidation. When the firm liquidates its capital, Tobin's average q equals the firm's liquidation recovery ℓ so that the following holds in the liquidation region:

$$q(z, w) = \ell\,. \tag{24}$$

[16] As our model is time homogenous, the firm's payout policy is the same conditional on (z, W, K).

When the firm operates as a going concern, $q(z,w) > \ell$ must hold. Combining the liquidation and going-concern cases, the following inequality must hold for all admissible (z, w):

$$\ell - q(z,w) \ \leq 0\,. \tag{25}$$

Boundary conditions at $w = 0$. With $\sigma_y > 0$, $\{w_t\}$, given in (8), is locally random and evolves as a controlled diffusion process. Therefore, when $w = 0$, the only way for the firm to avoid liquidation is to issue equity. The firm issues equity at a cost if $\mathcal{M}q(z,0) > \ell$. Mathematically, combining the inequality (23) for equity issuance with (25) for liquidation, we obtain the following boundary condition for $q(z,w)$ at $w = 0$:

$$\max\left\{\mathcal{M}q(z,0) - q(z,0),\ \ell - q(z,0)\right\} = 0\,. \tag{26}$$

Having characterized all the state-contingent policies, we can now summarize the model's four-region solution.

Summary. Consolidating the four inequalities (18), (19), (23), and (25), we can characterize Tobin's average q, $q(z,w)$, using the following variational inequality:

$$\max\left\{\underbrace{\mathcal{L}q(z,w)}_{\text{Savings}},\ \underbrace{-q_w(z,w)}_{\text{Payout}},\ \underbrace{\mathcal{M}q(z,w) - q(z,w)}_{\text{Equity issuance}},\ \underbrace{\ell - q(z,w)}_{\text{Liquidation}}\right\} = 0\,, \tag{27}$$

subject to the boundary condition (26) at $w = 0$. As the firm is active along at least one of the four decision margins, *saving, payout, external financing, and liquidation,* at least one of the four terms inside the max operator in (27) equals zero. When $\mathcal{L}q(z,w) = 0$, the firm saves and its investment policy satisfies (15). In Appendix A, we state a verification theorem and provide a proof. For completeness, we also characterize the solution for 1.) the special case with only proportional equity issuance costs ($\phi_0 = 0$ and $\phi_1 > 0$) in Appendix B and 2.) the special case without contemporaneous shocks ($\sigma_y = 0$) in Appendix C.

5 Quantitative Analysis

In this section, we examine our model's quantitative predictions and further explore our model's mechanism. A key, robust takeaway is that only in models with contemporaneous shocks ($\sigma_y > 0$) do firms save.

5.1 Functional Forms, Parameter Choices, and Calibration

As in Hennessy and Whited (2007) and Riddick and Whited (2009), we assume that the logarithmic productivity shock, $z_t := f(A_t) = \ln A_t$, follows a first-order autoregressive (AR(1)) process. The continuous-time counterpart of an AR(1) process is an Ornstein-Uhlenbeck (OU) process (see, e.g., Karatzas and Shreve, 2012) with the following drift and volatility functions:

$$\mu(z) = \kappa(\mu - z) \quad \text{and} \quad \sigma_z(z) = \sigma_z\,. \tag{28}$$

In (28), $\kappa > 0$ is the constant rate of mean reversion, μ is the constant long-run average of z_t, and $\sigma_z > 0$ is the constant volatility.

To facilitate comparisons with the literature, we use the same parameter values for the logarithmic productivity process, z, as in Hennessy and Whited (2007) and Riddick and Whited (2009). Converting the parameter values for an AR(1) process into their counterparts for the corresponding continuous-time OU process, we set 1.) the annual mean-reversion rate of log productivity (z) to $\kappa = 0.380$, so that the AR(1) coefficient per annum for the persistent productivity shock z is $e^{-\kappa} = 0.684$ per annum; 2.) the volatility parameter of z to $\sigma_z = 0.141$, so that the annualized (conditional) volatility of the productivity shock z is $\sigma_z\sqrt{\frac{1-e^{-2\kappa}}{2\kappa}} = 0.118$; and 3.) the (unconditional) long-run average of z to $\mu = -1.217$, so that the long-run average of productivity $A = e^z$, μ_A, equals 0.3 using $\mu = \ln \mu_A - \frac{\sigma_z^2}{4\kappa} = -1.217$.

We set the annual risk-free rate to $r = 0.04$ as in Riddick and Whited (2009) and the annual depreciation rate of capital stock to $\delta = 0.1$ as in Stokey and Rebelo (1995), which are commonly used values in macro and finance. We set the (annual) cash-carry cost to $\lambda = 0.01$, the marginal cost of issuing equity to $\phi_1 = 0.09$, and the liquidation recovery of capital stock to $\ell = 0.9$. These numbers are within the range commonly used in the literature, e.g., Hennessy and Whited (2007), Riddick and Whited (2009), and Bolton, Chen, and Wang (2011). The firm's exogenous (annual) death rate is set to $\zeta = 0.04$, which aligns with estimates from the firm dynamics literature, e.g., Hackbarth and Johnson (2015) and Ai, Kiku, Li, and Tong (2021). We set the correlation between the permanent and transitory shocks to zero, $\rho = 0$.

Next, we specify $c(i)$. Following Riddick and Whited (2009), we choose

$$c(i_t) = \theta_0 \mathbf{1}_{\{i_t \neq 0\}} + i_t + \frac{\theta_2}{2} i_t^2 \,, \tag{29}$$

where $\theta_0 \geq 0$, $\theta_2 > 0$, and $\mathbf{1}_{\{i_t \neq 0\}}$ is an indicator function that equals one if $i_t \neq 0$ and zero otherwise. Economically speaking, θ_0 is the fixed adjustment cost (per unit of capital) that the firm pays if and only if it invests ($i_t > 0$) or divests ($i_t < 0$), and θ_2 measures the degree of (quadratic) adjustment costs.

We consider two versions of the model, with and without contemporaneous shocks (that is, $\sigma_y > 0$ and $\sigma_y = 0$, respectively). For the $\sigma_y > 0$ version of the model, we calibrate the fixed adjustment cost parameter (θ_0), the quadratic adjustment cost parameter (θ_2), the fixed equity financing cost (ϕ_0), and σ_y, by matching our model-implied moments of the cash-capital ratio, Tobin's average q, the investment-capital ratio, and the equity issuance-capital ratio to their respective empirical moments.[17]

The calibrated fixed adjustment cost parameter is $\theta_0 = 0.05$, the quadratic adjustment cost parameter is $\theta_2 = 7.78$, and the fixed financing cost parameter is $\phi_0 = 0.01$. These are close to the values used in the literature, e.g., Zhang (2005), Eberly, Rebelo, and Vincent (2008), and Bolton, Chen, and Wang (2011). The calibrated volatility of the contemporaneous shock is $\sigma_y = 0.45$. For the restricted version of the model where $\sigma_y = 0$, we re-calibrate the other three parameters (θ_0, θ_2, and ϕ_0) and obtain similar values.

In Table 1, we compare the empirical moments with the model-generated moments (panel A) and report the calibrated parameters (panel B) for the two versions of the model with

[17]To compute empirical moments, we use Compustat data from 2010 to 2021. We first exclude all firm-year observations with missing data, as well as those with negative values for total assets, capital expenditures, or cash and short-term investments. Next, we retain only the longest consecutive time series for each firm and remove firms with only a single observation. We then trim the top and bottom 2% of all variables to mitigate the influence of outliers. The investment-capital ratio is measured using capital expenditures (Compustat item *capx*) scaled by lagged total assets (*at*). The cash-capital ratio is calculated using cash and short term investment (*che*) scaled by total assets. The market value of assets, which is used to compute Tobin's average q, is computed as the book value of assets (*at*) plus the market value of common equity ($prcc\_f \times csho$) minus the book value of common equity (*ceq*) and deferred taxes (*txdb*). Tobin's average q is then measured as the market value of assets divided by a weighted sum of 90% book value of assets and 10% market value of assets as in Duchin (2010). Finally, we adopt the empirical mean of the equity issuance-capital ratio from Hennessy and Whited (2007). To match the model-implied moments with empirical moments, we follow common practice in the literature (e.g., Hennessy and Whited, 2007). Specifically, we simulate 10 panels of data. Each panel consists of 1,000 independent and identically distributed firms. For each firm, we simulate 200 years of monthly data, discard the first 188 years, and use the remaining 12 years of data for our moment matching calculations.

$\sigma_y > 0$ and $\sigma_y = 0$. As can be seen, the model does a better job matching empirical moments when allowing for contemporaneous shocks. In particular, it does a good job matching the average cash-capital ratio of 20.98% that is observed in the data (the model-based moment is 20.53%). This is in sharp contrast to the trivial cash-capital ratio (0.00%) that is obtained when the model does not include contemporaneous shocks. This illustrates the main finding of this paper: without contemporaneous shocks, the model cannot explain the high levels of cash holdings that we observe in the data.

Table 1: Moments Comparison and Calibrated Parameters

Panel A. Moments			
	Data	$\sigma_y > 0$	$\sigma_y = 0$
Cash-capital ratio	0.2098	0.2053	0.0000
Equity issuance-capital ratio	0.0892	0.0913	0.0000
Tobin's average q	1.9085	1.8725	1.8874
Investment-capital ratio	0.0413	0.0412	0.0413
Panel B. Calibrated Parameters			
Parameter	Symbol	$\sigma_y > 0$	$\sigma_y = 0$
Contemporaneous shock volatility	σ_y	0.452	0
Fixed adjustment cost	θ_0	0.052	0.057
Quadratic adjustment cost	θ_2	7.775	7.391
Fixed refinancing cost	ϕ_0	0.011	0.011

Notes. This table compares empirical moments with model-simulated moments (panel A) and reports the corresponding calibrated parameter values (panel B) for the models with contemporaneous shocks ($\sigma_y > 0$) and without ($\sigma_y = 0$).

5.2 Quantifying the Effects of Contemporaneous Shocks

In this subsection, we show that the presence of contemporaneous shocks is key to drive corporate savings. To show this result, we first analyze the solutions of our two calibrated models (with $\sigma_y > 0$ and $\sigma_y = 0$) and compare them.[18] Then, we show that the solution for the $\sigma_y = 0$ case is indistinguishable from the First-Best solution, further indicating the importance of incorporating $\sigma_y > 0$ into a dynamic corporate finance model for financially

[18]We describe the numerical solution method in Appendix D.

constrained firms.[19]

Solution under costly external equity. Figure 2 characterizes the solution regions and the optimal $i^*(z, w)$ in the (A, w) space for the cases with $\sigma_y > 0$ (panel A) and $\sigma_y = 0$ (panel B). The color bar on the right-hand side indicates the level of $i^*(z, w)$.

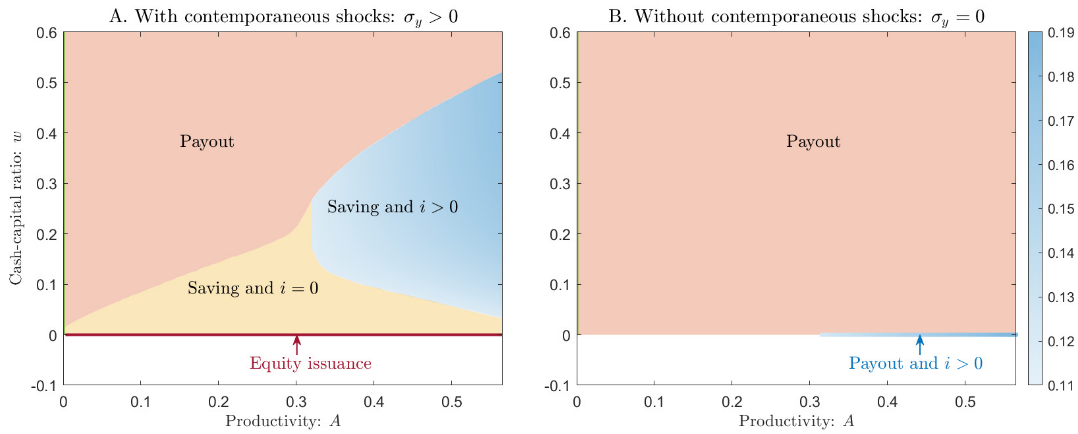

Figure 2: **Comparing the solution for the $\sigma_y > 0$ case with that for the $\sigma_y = 0$ case under costly external equity.** Panel A plots the solution regions for the $\sigma_y > 0$ model. The color bar on the right-hand side indicates the level of i^*. Panel B plots the solution regions for the $\sigma_y = 0$ model, which shows that the firm only invests when $w = 0$ and $A > 0.31$. Parameter values for panels A and B of this figure are given in the $\sigma_y > 0$ and $\sigma_y = 0$ columns, respectively, of panel B of Table 1.

When facing contemporaneous shocks ($\sigma_y > 0$, panel A), the firm pays out to shareholders only when its savings are sufficiently high for a given level of productivity A. Moreover, for a given level of cash-capital ratio w, higher productivity reduces teh firm's willingness to pay dividends, since cash is needed to finance investment. This explains the upward-sloping payout boundary in the (A, w) space. Below this payout boundary, there are two saving regions divided by a downward-sloping boundary. To the right of this boundary the firm saves and invests (since its productivity is sufficiently high), as the firm has both precautionary and investment motives. The color map in this region darkens as we move toward either the north or east, as $i^*(z, w)$ increases with both z and w. In contrast, to the left of this boundary the firm saves but does not invest as its productivity is not high enough and the

[19]Throughout the analysis, unless indicated otherwise, we use the parameter values described in subsection 5.1, namely, $\kappa = 0.38$, $\sigma_z = 0.14$, $\mu_A = 0.30$, $r = 0.04$, $\delta = 0.1$, $\lambda = 0.01$, $\phi_1 = 0.09$, $\ell = 0.9$, $\zeta = 0.04$, and $\rho = 0$ for both the $\sigma_y > 0$ and $\sigma_y = 0$ cases, along with the case-specific parameter values listed in Table 1.

firm saves primarily for its precautionary needs.

Note that the firm only issues equity when it runs out of cash. That is, the scaled endogenous external financing boundary, denoted by $\underline{w}(z)$, lies on the horizontal axis: $\underline{w}(z) = 0$. This result also holds in Décamps, Mariotti, Rochet, and Villeneuve (2011) and Bolton, Chen, and Wang (2011), where the firm faces only *i.i.d.* contemporaneous shocks. The intuition is as follows. First, had the firm issued equity before exhausting its cash, it would lose the interest income that it could earn from the funds used to pay for the equity issuance costs. Second, the funds raised via equity issuance that are not used immediately earn a lower rate of return due to a positive cash-carrying cost, which is suboptimal.

Finally, we turn to the liquidation region, located to the left of the optimal liquidation boundary $\underline{A}(w)$, which is barely visible. This is because $\underline{A}(w)$ is steep (slightly downward sloping) and lies very close to the vertical axis. Numerically, the firm liquidates itself when $w = 0$ and its productivity A falls below $\underline{A}(0) = 0.0024$, which is almost indistinguishable from the liquidation boundary $\underline{A} = 0.0022$ when equity issuance is costless, i.e., under the First Best.[20]

In sharp contrast to the $\sigma_y > 0$ case (shown in panel A), panel B shows that absent contemporaneous shocks ($\sigma_y = 0$), the firm has *no* savings demand. Indeed, the two savings regions observed in panel A disappear.[21] On the horizontal axis there exists a segment where the firm invests and pays out its remaining profits to shareholders. The higher the firm's productivity A the more it invests, yet the firm does not need savings to finance the desired level of investment. The intuition is as follows. Because the cash flows are (locally) deterministic, the firm faces virtually no precautionary savings motive and can fund investment directly from internally generated cash flows when productivity is high, while it chooses not to invest when productivity is low. This flexibility to adjust investment in response to productivity shocks provides a natural hedge, leaving little incentive for the firm to save.

Importantly, this reasoning no longer holds when the firm is subject to contemporaneous

[20] Mathematically, the firm's liquidation decision in our model depends on its cash holdings. Indeed, the liquidation boundary is downward sloping. This is different from Bolton, Chen, and Wang (2011), where liquidation occurs only when the firm runs out of cash. However, this effect is quantitatively negligible.

[21] The liquidation region is very similar to that in the $\sigma_y > 0$ case.

shocks. The intuition is as follows. When $\sigma_y > 0$, the firm's (scaled) cash holding w is volatile, and hence the q_{ww} term appears in the HJB equation (as discussed in Section 4 when introducing the savings region). As a result, the firm has precautionary savings motives to 1.) avoid costly external equity issuance and/or costly liquidation, and 2.) fund its desired level of investment. In contrast, when $\sigma_y = 0$, w_t is locally deterministic and the q_{ww} term in the HJB equation (16)-(17) disappears. As a result, the firm's precautionary saving motives—to delay costly equity issuance, avoid costly liquidation, or mitigate under-investment—are negligible. This is because liquidation and equity issuance become remote possibilities, and the firm can almost always rely on internally generated cash flows to finance the desired level of investment. Note that this intuitive reasoning is fully aligned with the mathematical structure of the problem solutions, as the PDEs with and without contemporaneous shocks are fundamentally different.

Next, we turn to the firm's equity issuance. When $\sigma_y = 0$, the firm has no incentive to save, and hence equity issuance plays no economically meaningful role. When $\sigma_y > 0$, as we discussed earlier, the firm only issues equity when its w_t reaches zero. How much does a firm issue? Figure 3 shows that the size of the equity issuance m_t increases with productivity A_t. Intuitively, because of the fixed issuance costs, the higher its productivity the more funds the firm raises for its investment needs.

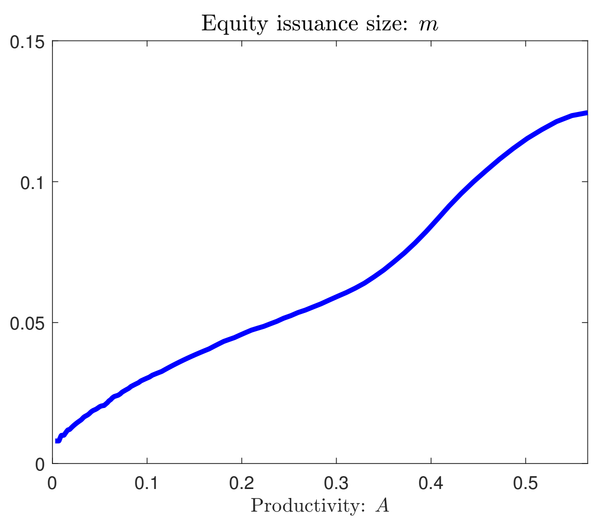

Figure 3: **Equity issuance size m as a function of productivity A for the $\sigma_y > 0$ case.** The firm only issues equity when it runs out of cash. Parameter values are given in the $\sigma_y > 0$ column of panel B of Table 1.

Finally, we show that absent contemporaneous shocks, costly external equity has quantitatively negligible effects on corporate investment and valuation. Indeed the solution for the $\sigma_y = 0$ case is quantitatively indistinguishable from the first-best solution. In what follows, we show that this is the case by summarizing the first-best solution and comparing the two.

First-best. We solve for the first-best investment-capital ratio, $i^{fb}(z)$, in two steps. First, conditional on paying a fixed cost, the investment-capital ratio, $i(z)$, satisfies the following first-order condition:

$$c'(i(z)) = q^{fb}(z)\,. \tag{30}$$

Second, we check whether the net present value of investing $i(z)$ for each unit of capital is positive, that is, whether $q^{fb}(z)i(z) - c(i(z)) > 0$ holds. If so, $i(z)$ is the first best $i^{fb}(z)$. Otherwise, $i^{fb}(z) = 0$. Note that in the first-best scenario, the marginal q equals the average q as in Hayashi (1982). Panel A of Figure 4 shows that Tobin's average q under First Best, q^{fb}, is increasing and concave in productivity A (dashed blue line). Moreover, panel B shows that, under First Best, the firm invests when its productivity $A > 0.31$ and does not invest when $A \in [0.0022, 0.31]$. When $A < 0.0022$ (corresponding to $z < -6.12$), the firm liquidates its capital and $q^{fb}(z) = \ell = 0.9$. Note that there is a discontinuity in the investment function $i(z)$ at $i = 0$ due to the fixed investment cost.

Recall that when there are no contemporaneous shocks ($\sigma_y = 0$), even in the presence of costly external equity, the firm retains no cash ($w = 0$) as it finances its investments through internally generated cash flows and pays out the remaining profits to shareholders. In Figure 4, we plot the firm's average q and its investment-capital ratio as functions of productivity (solid red lines) and compare them with the first-best solution (dashed blue lines). As can be seen, the solutions for the two models coincide. That is, not only does the firm have no motive to hold cash, but its solution is also quantitatively indistinguishable from the first-best solution. In other words, in order for costly external equity to have a quantitatively meaningful effect on corporate policies and valuation, it is necessary for firms to face contemporaneous shocks.

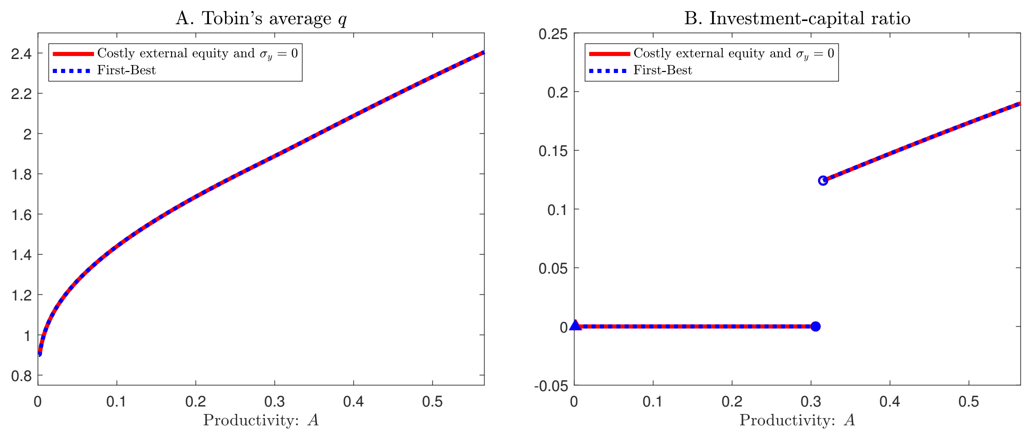

Figure 4: **Tobin's average q and the investment-capital ratio: comparing the $\sigma_y = 0$ case and the first-best.** This figure shows that, when $\sigma_y = 0$, the solution under costly external equity is indistinguishable from the first-best solution, implying that a financially constrained firm has no savings demand even in the presence of costly external equity. The firm invests when $A > 0.31$, is inactive when $A \in [0.0022, 0.31]$, and is liquidated when $A < 0.0022$. Parameter values are given in the $\sigma_y = 0$ column of panel B of Table 1.

6 Comparative Statics

In this section, we report comparative-static results for our model-predicted average cash-to-capital ratio w under various parameter choices. For each panel in Figures 5 and 6, the values of the parameters other than the one being analyzed in the respective panel are the same as in our baseline calculations in Section 5.

In panel A of Figure 5, we vary the volatility of the contemporaneous shock, σ_y. An increase in σ_y raises the contemporaneous uncertainty in the operating cash flows $A_t K_t(\mathrm{d}t + \sigma_y \mathrm{d}\mathcal{B}_t^y)$ and hence the likelihood that internally generated cash flows will be insufficient to finance the desired level of investment.[22] This increases the value of corporate liquidity as a buffer against costly external financing. This intuition is reflected in panel A, which shows that the average w rises monotonically with σ_y. Also, the figure confirms our main result that firms do not hold cash when $\sigma_y = 0$.

Recall that, importantly, since non-trivial cash holdings only arise when contemporaneous shocks are present, all other comparative-static results refer to the case with $\sigma_y > 0$. In panel

[22]Over a small time interval $\mathrm{d}t$, the cash flow volatility effect dominates the drift effect. In diffusion models, the (diffusion) random component of the internally generated cash flows, which is of $\sqrt{\mathrm{d}t}$ magnitude, dominates the (drift) expected component, which is of $\mathrm{d}t$ magnitude as $\sqrt{\mathrm{d}t} >> \mathrm{d}t$ for a small $\mathrm{d}t$.

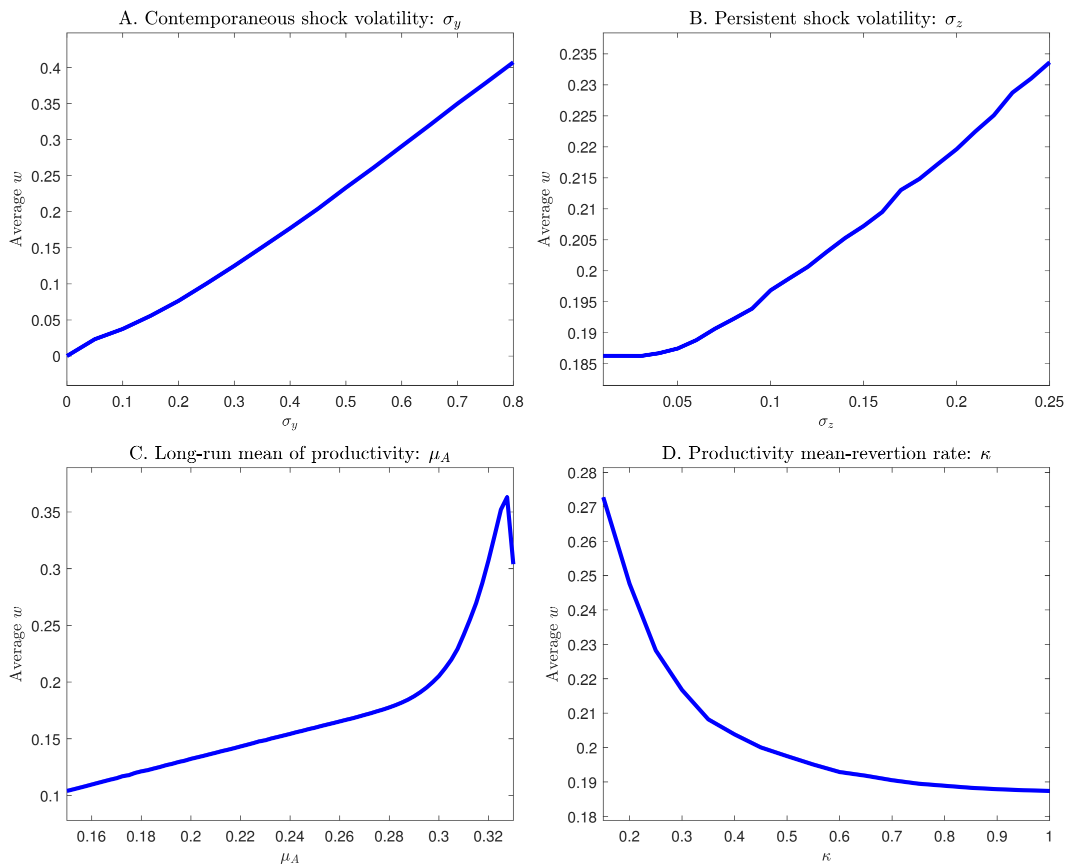

Figure 5: **Comparative statics for the average model-predicted cash-capital ratio w with respect to the volatility of contemporaneous shocks σ_y, volatility of productivity shocks σ_z, long-run productivity μ_A, and mean reversion of z, κ.**

B of Figure 5, we vary the volatility of the persistent shock z, σ_z. Cash holdings increase monotonically with σ_z. This is because higher volatility of the persistent productivity shock z makes future desired investment levels more uncertain, increasing the firm's savings demand today. Persistent and contemporaneous shocks therefore reinforce each other, inducing the firm to hold more cash to mitigate future underinvestment.

In panel C of Figure 5, we vary the long-run mean of the productivity shock $A_t = e^{z_t}$: μ_A.[23] Increasing μ_A has two opposite effects on cash holdings. On the one hand, more productive firms require additional funding to support higher investment, which induces them to hold more cash. On the other hand, more productive firms generate larger internal

[23]Recall that the long-run average of productivity $A = e^z$, μ_A, is related to the long-run average of z, μ, via $\mu = \ln \mu_A - \frac{\sigma_z^2}{4\kappa}$. See, e.g., Karatzas and Shreve (2012), and our discussion in subsection 5.1.

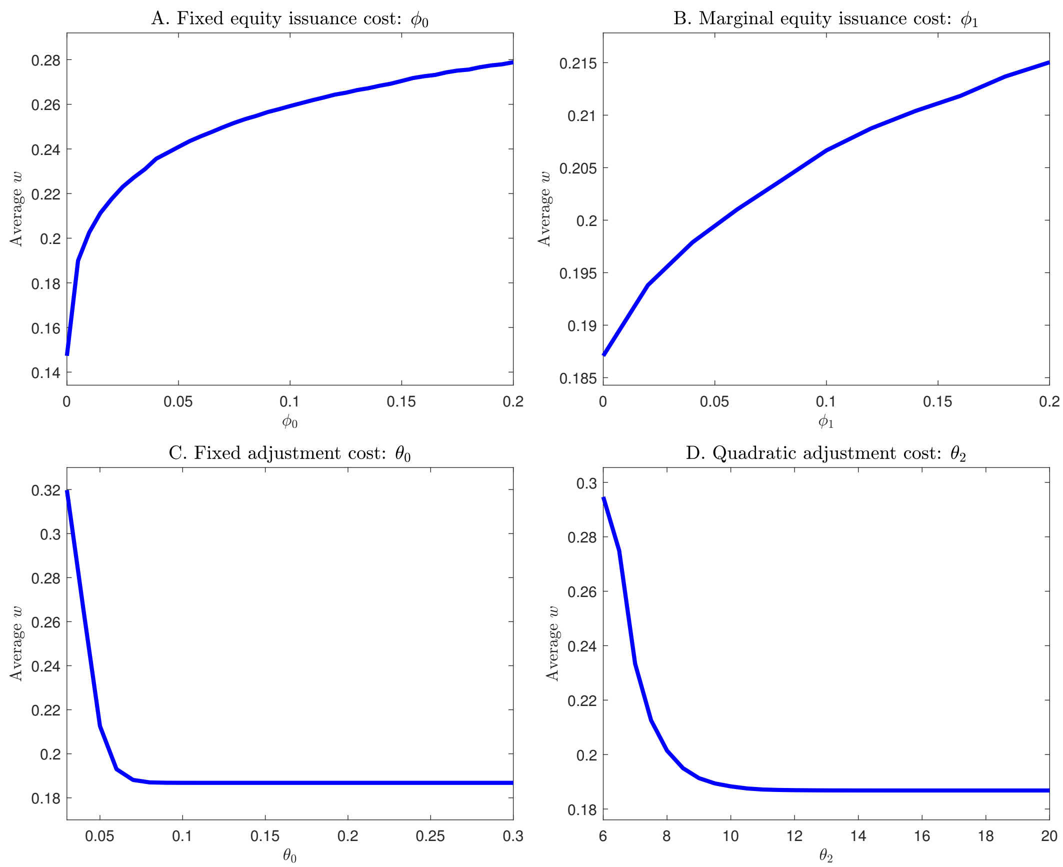

Figure 6: **Comparative statics for the average model-predicted cash-capital ratio w with respect to the equity financing costs (ϕ_0, ϕ_1) and capital adjustment costs (θ_0, θ_2).**

cash flows, which reduces their need for cash. As can be seen from the figure, the former effect dominates except when μ_A becomes very large, in which case the latter effect prevails.

In panel D of Figure 5, we vary the mean-reversion rate of the productivity shock z, κ. Increasing κ reduces the persistence of productivity shocks, making them more transitory and thereby lowering the uncertainty in the firm's cash flows.[24] Accordingly, firms have less need to hold cash, as reflected in this panel where w decreases monotonically with κ.

Next, we turn to our comparative-static analyses with respect to equity financing costs (ϕ_0 and ϕ_1) and capital adjustment cost parameters (θ_0, θ_2). In Figure 6, we vary the cost of external financing—specifically the fixed equity issuance cost, ϕ_0, in panel A and the

[24] Recall that the corresponding annualized AR(1) coefficient in a discrete-time setting is $e^{-\kappa}$. Therefore, increasing κ lowers the AR(1) coefficient. Also the annualized (conditional) volatility of the productivity shock z is $\sigma_z\sqrt{\frac{1-e^{-2\kappa}}{2\kappa}}$, which implies that this also decreases with κ.

marginal equity issuance cost, ϕ_1, in panel B. Not surprisingly, since higher external financing costs increase the appeal of internal financing, we find that w increases monotonically with both ϕ_0 and ϕ_1. Finally, we vary the capital adjustment costs—specifically the fixed adjustment cost, θ_0, in panel C and the quadratic adjustment cost, θ_2, in panel D. Effectively, higher investment costs make firms less productive, which reduces their investment needs and hence their need to hold cash. Consistent with this intuition, we find that w decreases monotonically with both θ_0 and θ_2.

7 Extension with Debt Financing

In this section, we extend the baseline model of Section 2 to allow the firm not only to save but also to borrow, and we show that our main results continue to hold in this more general setting. Specifically, the firm borrows by issuing debt with instantaneous maturity as in Abel (2018).[25] If the firm issues an amount D_t at time t, it must repay creditors at time $t + \mathrm{d}t$, where $\mathrm{d}t$ is infinitesimal. Then at $t + \mathrm{d}t$, the firm chooses a new amount $D_{t+\mathrm{d}t}$ to issue and repays the new creditors at time $t+2\mathrm{d}t$, and so on. This rollover process continues indefinitely. Short-term debt in our framework can equivalently be interpreted as the balance on a credit-line facility provided by a long-term creditor, which the firm can continuously draw down and repay. Although the institutional details of instantaneous debt rollover and credit-line financing differ, the two are economically and mathematically equivalent in our model. Below we first describe our extended model with debt financing (subsection 7.1) and then solve the model and analyze the main results (subsection 7.2).

7.1 An Extended Model

In addition to retaining all the assumptions from the baseline model of Section 2, we introduce the following additional assumptions. As in Hennessy and Whited (2005), at any time t, the firm can borrow an amount of D_t up to bK_t where $b > 0$ measures debt capacity. Moreover, the firm's debt D_t is fully collateralized by its capital stock ($b < \ell$) so that debt is risk-free

[25]Debt with instantaneous maturity is widely used in asset pricing, e.g., Black and Scholes (1973) and Merton (1973). See Duffie (2010) for a textbook treatment.

in our model.[26] The firm pays an exogenous spread $\alpha > 0$ over the risk-free rate on its debt balance, D_t, which we interpret as an intermediation cost as in Bolton, Chen, and Wang (2011). When $W_t < 0$, debt is the marginal source of financing and $D_t = -W_t > 0$. When $W_t > 0$, cash is the marginal source of financing as in our baseline model. We describe both scenarios by formulating a single $\{W_t\}$ process as follows (with $D_t = -W_t$ when $W_t < 0$):

$$\mathrm{d}W_t = (r - \lambda \mathbf{1}_{\{W_t>0\}} + \alpha \mathbf{1}_{\{W_t<0\}})W_t \mathrm{d}t + \mathrm{d}Y_t + \mathrm{d}H_t - \mathrm{d}U_t,\ t \geq 0, \tag{31}$$

where $\mathrm{d}Y_t$ is the net operating profit over the time interval $(t, t + \mathrm{d}t)$ given in (2), and $\mathrm{d}H_t$ and $\mathrm{d}U_t$ are the instantaneous equity issuance and payout amounts, respectively. Appendix E provides the solution for the extended model.

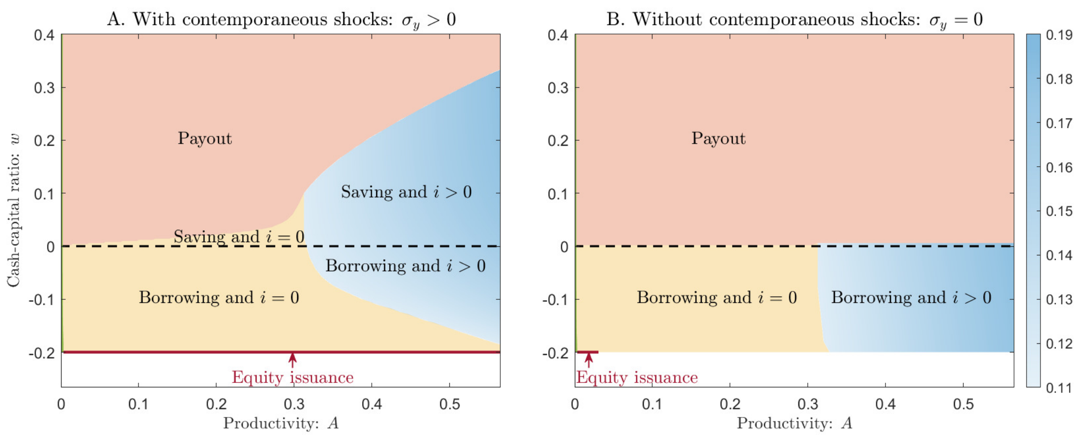

Figure 7: **Effects of contemporaneous shocks in an extended model with debt financing on the solution regions.** Panels A and B plot the solution regions for the $\sigma_y > 0$ and $\sigma_y = 0$ cases, respectively. The color bar on the right-hand side indicates the level of i^*. Debt capacity is $b = 0.2$ and the interest spread over the risk-free rate r is $\alpha = 1.5\%$. Other than σ_y, all parameter values for both cases are given in the $\sigma_y > 0$ column of panel B of Table 1.

7.2 Solution and Main Results

Figure 7 characterizes the solution regions and the optimal investment-capital ratio i^* in the (A, w) space for the cases with $\sigma_y > 0$ (panel A) and $\sigma_y = 0$ (panel B). The color bar on the right-hand side of the figure indicates the level of i^*.

[26] Instantaneous debt in our continuous-time setting corresponds to one-period debt in discrete-time models, e.g., Hennessy and Whited (2005). In this regard, it is worth noting that the focus of Hennessy and Whited (2005) is on debt financing (the title of their paper is "Debt Dynamics"), not cash holdings.

As can be seen, a borrowing region exists regardless of whether the firm faces contemporaneous shocks or not. This is because, after running out of its savings, the firm has no option but to rely on external financing to survive. Since debt financing only triggers an interest spread, while equity financing involves fixed and proportional costs, debt is less costly than equity. This is why the firm chooses debt over equity. Only when the firm hits its debt limit (that is, when $w_t = -b = -0.2$) does it issue equity, which is captured by the horizontal line representing the equity-issuance region in both panels. The intuition is similar to that in our baseline model of Section 2. The firm preserves the option value of issuing costly external equity to avoid inefficient liquidation until it cannot delay any more.[27]

Importantly, while a borrowing region appears in both the $\sigma_y > 0$ and $\sigma_y = 0$ cases, we continue to find that the savings region appears only when contemporaneous shocks are present ($\sigma_y > 0$). This confirms the main result from our baseline analysis: contemporaneous shocks play a crucial role in generating sizeable cash holdings.

Next, we turn to the issuance amount conditional on issuing equity. Figure 8 plots the firm's equity issuance size m as a function of productivity A for the $\sigma_y > 0$ (solid blue line) and $\sigma_y = 0$ (dashed red line) cases. First, we note that the range of A values for which the firm issues equity is substantially larger when contemporaneous shocks are present. Indeed, when $\sigma_y > 0$ (specifically, $\sigma_y = 0.452$), the firm issues equity when $A \in (0.003, 0.566)$. In contrast, when $\sigma_y = 0$, the firm only issues equity when $A \in (0.003, 0.031)$. The intuition is the same as in our baseline model without debt financing. Since contemporaneous shocks introduce short-term uncertainty in the firm's internally generated cash flows (and hence in the firm's ability to rely on internal financing), firms are more likely to resort to equity financing.

Second, when $\sigma_y > 0$, the issuance amount is non-monotonic in A. At low levels of productivity A, equity issuance m decreases as productivity rises. The reason is that lower productivity increases the likelihood of hitting the borrowing limit, which induces the firm to issue more equity to move away from the debt limit and avoid equity issuance costs. This argument only applies at low levels of productivity where the firm is concerned about

[27] This is premised under the condition that the firm prefers issuing equity over liquidation, which holds in our quantitative analysis when the firm's productivity A is larger than 0.003. When $A < 0.003$, which is almost invisible in Figure 7, the firm prefers liquidation to costly equity financing.

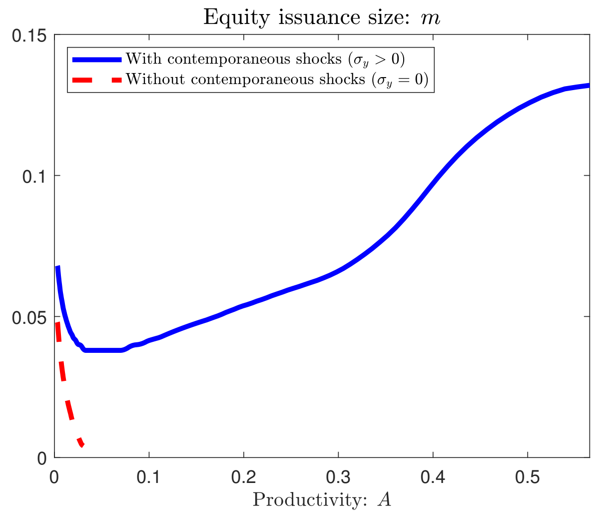

Figure 8: **Effects of contemporaneous shocks in an extended model with debt financing on the equity issuance size m as a function of productivity A.** This figure plots the firm's equity issuance size m for the $\sigma_y > 0$ and $\sigma_y = 0$ cases, respectively. Equity issuance size m is non-monotonic in productivity A when $\sigma_y = 0.452$, but is decreasing in productivity A when $\sigma_y = 0$. In both cases, the firm only issues equity when it runs out of its debt capacity, i.e., when $w = -b$. Debt capacity is $b = 0.2$ and the interest spread over the risk-free rate r is $\alpha = 1.5\%$. Other than σ_y, all parameter values for both cases are given in the $\sigma_y > 0$ column of panel B of Table 1.

survival. At higher levels of productivity, equity issuance m increases as productivity rises, consistent with what we showed in Figure 3 for our baseline model without debt financing. This is because when productivity is high, the firm's precautionary savings are primarily for investment purposes rather than survival. Facing contemporaneous shocks, the benefit of holding a cash buffer to finance investment is more valuable for a productive firm. In contrast, this precautionary motive of holding cash for investment disappears when $\sigma_y = 0$. Again, this is because the locally deterministic cash flows are sufficient to finance the desired level of investment. This is why with $\sigma_y = 0$, the firm only issues equity when productivity is low for survival reasons, and moreover, the issuance size m decreases with productivity.

8 Conclusion

This paper develops a dynamic q-theory model of investment, financing, and cash management to shed light on the cash puzzle in corporate finance—the fact that firms hold substantial amounts of cash, while standard models typically predict low cash holdings. We show that, in order to generate realistic cash holdings, q-theory models need to include three economic forces: 1.) costly external financing, 2.) persistent productivity shocks, and 3.) contemporaneous productivity shocks.

The rationale is as follows. Persistent productivity shocks generate locally deterministic cash flows and stochastic investment opportunities over time, but yield little need for savings since the internally generated cash flows are aligned with the investment needs and hence provide a source of funds that is sufficient to finance the firm's desired level of investment. The combination with contemporaneous shocks introduces short-term uncertainty in these internally generated cash flows—that is, they may no longer suffice as a source of funds—and the firm's future cash balance. This uncertainty, combined with costly external financing, induces firms to hold cash at levels that are consistent with those observed empirically.

By clarifying the economic forces that drive firms' cash holding policies, our framework explains why existing models—featuring at most two of the three forces—typically predict that companies hold low levels of cash. In this regard, our framework connects the dots between different families of models and clarifies how these models can be extended to match real-world cash policies.

Taken together, our findings underscore that cash management is a central element of corporate decision-making that is driven by both the investment and precautionary saving motives. At the core of both motives is the notion of financing flexibility, whose importance was highlighted in the AFA presidential address "Corporate Finance and Reality" of John Graham (2022), who noted that CFOs typically rank financing flexibility as the number one factor underlying companies' financing policies. More broadly, by identifying the economic forces required to generate realistic cash holding behavior, our paper provides a unifying framework on liquidity management that future work can use to study and model firm dynamics in a way that is consistent with the cash holding policies observed in the data.

References

Abel, Andrew B. 2018. Optimal debt and profitability in the trade-off theory. *Journal of Finance* 73 (1):95–143.

Abel, Andrew B. and Janice C. Eberly. 1994. A unified model of investment under uncertainty. *American Economic Review* 84 (5):1369–1384.

Abel, Andrew B. and Stavros Panageas. 2023. Precautionary saving in a financially constrained firm. *Review of Financial Studies* 36 (7):2878–2921.

Ai, Hengjie, Dana Kiku, Rui Li, and Jincheng Tong. 2021. A unified model of firm dynamics with limited commitment and assortative matching. *Journal of Finance* 76 (1):317–356.

Almeida, Heitor, Murillo Campello, and Michael S. Weisbach. 2004. The cash flow sensitivity of cash. *Journal of Finance* 59 (4):1777–1804.

Asquith, Paul and David W. Mullins. 1986. Equity issues and offering dilution. *Journal of Financial Economics* 15 (1–2):61–89.

Bates, Thomas W., Kathleen M. Kahle, and René M. Stulz. 2009. Why do U.S. firms hold so much more cash than they used to? *Journal of Finance* 64 (5):1985–2021.

Black, Fischer and Myron Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81 (3):637–654.

Bolton, Patrick, Hui Chen, and Neng Wang. 2011. A unified theory of Tobin's q, corporate investment, financing, and risk management. *Journal of Finance* 66 (5):1545–1578.

Bolton, Patrick, Neng Wang, and Jinqiang Yang. 2019. Optimal contracting, corporate finance, and valuation with inalienable human capital. *Journal of Finance* 74 (3):1363–1429.

Brealey, Richard A., Stewart C. Myers, Franklin Allen, and Alex Edmans. 2025. *Principles of Corporate Finance*. McGraw Hill.

Caballero, Ricardo J. 1999. Aggregate investment. In *Handbook of Macroeconomics*, edited by John B. Taylor and Michael Woodford, 813–862. Elsevier.

CNBC. 2018. Warren Buffett and Jamie Dimon join forces to convince CEOs to end quarterly profit forecasts. CNBC (June 6, 2018).

Cooley, Thomas F. and Vincenzo Quadrini. 2001. Financial markets and firm dynamics. *American Economic Review* 91 (5):1286–1310.

Cooper, Russell W. and John C. Haltiwanger. 2006. On the nature of capital adjustment costs. *Review of Economic Studies* 73 (3):611–633.

Dai, Min, Xavier Giroud, Wei Jiang, and Neng Wang. 2024. A q theory of internal capital markets. *Journal of Finance* 79 (2):1147–1197.

DeAngelo, Harry, Linda DeAngelo, and René M. Stulz. 2010. Seasoned equity offerings, market timing, and the corporate lifecycle. *Journal of Financial Economics* 95 (3):275–295.

DeAngelo, Harry, Linda DeAngelo, and Toni M. Whited. 2011. Capital structure dynamics and transitory debt. *Journal of Financial Economics* 99 (2):235–261.

Décamps, Jean-Paul, Thomas Mariotti, Jean-Charles Rochet, and Stéphane Villeneuve. 2011. Free cash flow, issuance costs, and stock prices. *Journal of Finance* 66 (5):1501–1544.

Décamps, Jean-Paul, Sebastian Gryglewicz, Erwan Morellec, and Stéphane Villeneuve. 2017. Corporate policies with permanent and transitory shocks. *Review of Financial Studies* 30 (1):162–210.

DeMarzo, Peter M. and Yuliy Sannikov. 2006. Optimal security design and dynamic capital structure in a continuous-time agency model. *Journal of Finance* 61 (6):2681–2724.

DeMarzo, Peter M., Michael J. Fishman, Zhiguo He, and Neng Wang. 2012. Dynamic agency and the q theory of investment. *Journal of Finance* 67 (6):2295–2340.

Denis, David J. and Luxi Wang. 2024. Corporate cash holdings. In *Handbook of Corporate Finance*, edited by David J. Denis, 223–248. Edward Elgar Publishing.

Duchin, Ran. 2010. Cash holdings and corporate diversification. *Journal of Finance* 65 (3):955–992.

Duffie, Darrell. 2010. *Dynamic Asset Pricing Theory.* Princeton University Press.

Eberly, Janice, Sergio Rebelo, and Nicolas Vincent. 2008. Investment and value: A neoclassical benchmark. National Bureau of Economic Research, Working Paper 13866.

Ernst & Young. 2024. Cash forecasting: Difficult, disappointing and more urgent than ever. Ernst & Young (August 21, 2024).

Gomes, Joao F. 2001. Financing investment. *American Economic Review* 91 (5):1263–1285.

Graham, John R. 2022. Presidential address: Corporate finance and reality. *Journal of Finance* 77 (4):1975–2049.

Graham, John R. and Mark T. Leary. 2018. The evolution of corporate cash. *Review of Financial Studies* 31 (11):4288–4344.

Gustafson, Matthew T. and Peter Iliev. 2017. The effects of removing barriers to equity issuance. *Journal of Financial Economics* 124 (3):580–598.

Hackbarth, Dirk and Timothy Johnson. 2015. Real options and risk dynamics. *Review of Economic Studies* 82 (4):1449–1482.

Harford, Jarrad, Sattar A. Mansi, and William F. Maxwell. 2008. Corporate governance and firm cash holdings in the US. *Journal of Financial Economics* 87 (3):535–555.

Hartman-Glaser, Barney, Simon Mayer, and Konstantin Milbradt. 2025. A theory of cash flow-based financing with distress resolution. *Review of Economic Studies* rdaf009.

Hayashi, Fumio. 1982. Tobin's marginal q and average q: A neoclassical interpretation. *Econometrica* 50 (1):213–224.

Hennessy, Christopher A. 2004. Tobin's Q, debt overhang, and investment. *Journal of Finance* 59 (4):1717–1742.

Hennessy, Christopher A. and Toni M. Whited. 2005. Debt dynamics. *Journal of Finance* 60 (3):1129–1165.

———. 2007. How costly is external financing? Evidence from a structural estimation. *Journal of Finance* 62 (4):1705–1745.

Hennessy, Christopher A., Amnon Levy, and Toni M. Whited. 2007. Testing Q theory with financing frictions. *Journal of Financial Economics* 83 (3):691–717.

Hugonnier, Julien, Semyon Malamud, and Erwan Morellec. 2015. Capital supply uncertainty, cash holdings, and investment. *Review of Financial Studies* 28 (2):391–445.

Jensen, Michael C. and William H. Meckling. 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3 (4):305–360.

Karatzas, Ioannis and Steven Shreve. 2012. *Brownian Motion and Stochastic Calculus*. Springer Science & Business Media.

Leland, Hayne E. and David H. Pyle. 1977. Informational asymmetries, financial structure, and financial intermediation. *Journal of Finance* 32 (2):371–387.

Liu, Laura Xiaolei, Toni M. Whited, and Lu Zhang. 2009. Investment-based expected stock returns. *Journal of Political Economy* 117 (6):1105–1139.

Lucas, Robert E. and Edward C. Prescott. 1971. Investment under uncertainty. *Econometrica* 39 (5):659–681.

Merton, Robert C. 1973. An intertemporal capital asset pricing model. *Econometrica* 41 (5):867–887.

Moyen, Nathalie. 2004. Investment–cash flow sensitivities: Constrained versus unconstrained firms. *Journal of Finance* 59 (5):2061–2092.

Nikolov, Boris, Lukas Schmid, and Roberto Steri. 2019. Dynamic corporate liquidity. *Journal of Financial Economics* 132 (1):76–102.

Pinkowitz, Lee, René M. Stulz, and Rohan Williamson. 2016. Do U.S. firms hold more cash than foreign firms do? *Review of Financial Studies* 29 (2):309–348.

Riddick, Leigh A. and Toni M. Whited. 2009. The corporate propensity to save. *Journal of Finance* 64 (4):1729–1766.

Stokey, Nancy L. and Sergio Rebelo. 1995. Growth effects of flat-rate taxes. *Journal of Political Economy* 103 (3):519–550.

Zhang, Lu. 2005. The value premium. *Journal of Finance* 60 (1):67–103.

Appendices

A Verification Theorem

Let SR, PR, ER, and LR denote the following four regions:

$$SR = \{(z,w) : q_w(z,w) > 0,\ q(z,w) > \mathcal{M}q(z,w),\ q(z,w) > \ell\},$$

$$PR = \{(z,w) : q_w(z,w) = 0\},$$

$$ER = \{(z,w) : q(z,w) = \mathcal{M}q(z,w)\},$$

$$LR = \{(z,w) : q(z,w) = \ell\}\,.$$

As indicated in Section 4, these regions map to the savings region (SR), the payout region (PR), the external financing region (ER), and the liquidation region (LR). The variational inequality (27) implies that these four regions span the entire state space $(-\infty,+\infty)\times[0,+\infty)$ for (z,w): $SR\cup PR\cup ER\cup LR = (-\infty,+\infty)\times[0,+\infty)$.

Let $\overline{SR}$ denote the closure of the savings region SR which includes this region's boundary. Then, $\partial P = PR\cap\overline{SR}$ is the payout boundary, $\partial E = ER\cap\overline{SR}$ is the equity issuance boundary, and $\partial L = LR\cap\overline{SR}$ is the liquidation boundary.

Assume that there exists a point $(\tilde{z},\tilde{w})\in(-\infty,+\infty)\times[0,+\infty)$, along with two functions, $\underline{z}(w)$: $[0,+\infty)\mapsto(-\infty,+\infty)$ and $\overline{w}(z)$: $[\tilde{z},\infty)\mapsto[0,+\infty)$, such that

(i) $\underline{z}(w)$ is monotone decreasing for $w\in[0,+\infty)$, with $\underline{z}(w)=\tilde{z}$ for all $w\geq\tilde{w}$;

(ii) $\overline{w}(z)$ is monotone increasing for $z\in[\tilde{z},w)$ and $\overline{w}(\tilde{z})=\tilde{w}$.

Using these functions, we can express the regions as follows:

$$SR = \{(z,w) : z > \underline{z}(w),\ 0 < w < \overline{w}(z)\}, \tag{A.1}$$

$$PR = \{(z,w) : z \geq \tilde{z},\ w \geq \overline{w}(z)\}, \tag{A.2}$$

$$ER = \{(z,w) : z \geq \underline{z}(0),\ w = 0\}, \tag{A.3}$$

$$LR = \{(z,w) : z \leq \underline{z}(w)\}. \tag{A.4}$$

Next, we formally characterize the optimal strategy and establish the connection between the value function and the Hamilton-Jacobi-Bellman (HJB) equation.

Theorem A.1 (Verification Theorem)**.** *Let $q(z,w)$ be a solution to the variational inequality (27) for $w > 0$ with the boundary condition given in (26) on $w = 0$, satisfying certain regularity conditions.*[28] *Then, the firm's initial value is given by $q(z_0, w_0)K_0 + W_0$, where $z_0 = f(A_0)$ and $w_0 = W_0/K_0$. Moreover, the optimal strategy $(I_t^*, U_t^*, H_t^*, \eta_\ell^*)$ is characterized as follows:*

– *Case 1. Starting with $(z_0, w_0) \in \overline{SR}$, the following results hold for $t \geq 0$:*

(a) Investment strategy $I_t^* = i^*(z_t, w_t)K_t$ whenever the state (z_t, w_t) is in the savings region SR, where $w_t = W_t/K_t$ and $i^*(z,w)$ satisfies:

$$i^*(z,w) = \arg\max_i \left(\frac{q(z,w)+w}{q_w(z,w)+1} - w \right) i - c(i); \tag{A.5}$$

(b) Payout strategy U_t^*:[29]

$$U_t^* = \int_0^t \mathbf{1}_{\{(z_s,w_s)\in\partial P\}} dU_s^*; \tag{A.6}$$

(c) External financing strategy H_t^*:

$$H_t^* = \sum_{n=1}^{\infty} m^{n,*} K_{\eta_e^{n,*}} \mathbf{1}_{\{\eta_e^{n,*}\leq t\}}, \tag{A.7}$$

where $\eta_e^{n,*}$ denotes the firm's n-th equity issuance time, and $m^{n,*}$ represents the corresponding capital-scaled issuance amount. These are characterized by

$$\eta_e^{n,*} = \inf \left\{ t \in (\eta_e^{n-1,*}, \eta_\ell^*) : (z_t, w_t) \in ER \right\}, \ n \geq 1, \tag{A.8}$$

$$m^{n,*} = \arg\max_{m>0} \ q(z_{\eta_e^{n,*}}, w_{\eta_e^{n,*}-} + m) - (\phi_0 + \phi_1 m), \tag{A.9}$$

with $\eta_e^{0,*} = 0$;

[28] Specifically, we assume the following regularity conditions: (i) $q(z,w)$ belongs to an appropriate Sobolev space (e.g., $W_p^{1,2}$ for any $p \geq 1$), ensuring that the generalized Ito's formula applies to $q(z,w)$; (ii) the four policy regions can be characterized by the expressions specified in (A.1)–(A.4), and the savings region SR is not empty; (iii) the associated free boundaries (the payout boundary ∂P, the equity issuance boundary ∂E, and the liquidation boundary ∂L) are sufficiently smooth to guarantee that the payout strategy U_t, as given in (A.6), is well-defined; (iv) for all admissible strategies and $t > 0$, $\mathbb{E}\left[\int_0^t e^{-2rs} \left((\sigma_z(z_s)q_z)^2 + (\sigma_y A(z_s)(q_w+1))^2 \right) K_s^2 \mathrm{d}s \right] < +\infty$.

[29] Mathematically, the optimal cumulative cash payout is represented in terms of a local time process. It does not increase until the cash-capital ratio w_t reaches $\overline{w}(z_t)$ from below, and the firm always maintains $W_t \leq \overline{w}(z_t)K_t$ through the cash distribution process.

(d) Liquidation strategy η_ℓ^*:

$$\eta_\ell^* = \inf\{t \geq 0 : (z_t, w_t) \in \overline{SR},\ q(z_t, w_t) = \ell\}\,. \tag{A.10}$$

Finally, the dynamics of K_t and W_t are given in (1) *and* (5)*, respectively, under the strategy* $(I^*, U^*, H^*, \eta_\ell^*)$ *characterized above.*

– Case 2. If starting with $(z_0, w_0) \in PR\backslash\overline{SR}$, the firm makes an initial dividend payout to reach the payout boundary ∂P and subsequently follows the optimal strategy described in Case 1 as the firm is in $\partial P \subset \overline{SR}$.

– Case 3. If starting with $(z_0, w_0) \in LR\backslash\overline{SR}$, the firm immediately liquidates.

Proof. The proof is similar to that in Davis and Norman (1990) and Dai, Qin, and Wang (2023); here we outline the key steps.

Consider Case 1 first. Suppose that $(z_0, w_0) \in \overline{SR}$. Consider any feasible strategy $\{(I_t, U_t, H_t, t \geq 0), \eta_\ell\}$, in which the investment strategy $\{I_t : t \geq 0\}$ is adaptive to the filtration $\{\mathcal{F}_t, t \geq 0\}$ generated by $\{(\mathcal{B}_t^z, \mathcal{B}_t^y), t \geq 0\}$, η_ℓ is a stopping time with respect to the same filtration, U_t is an increasing process, and H_t is another increasing process characterized by the equity issuance time $\{\eta_e^n\}$ and the corresponding capital-scaled issuance amount $\{m^n\}$ for $n \geq 1$, i.e.,

$$H_t = \sum_{n=1}^{\infty} m^n K_{\eta_e^n} \mathbf{1}_{\{\eta_e^n \leq t\}}.$$

Let

$$P(z, K, W) = q(z, w)K + W.$$

Define

$$N_t = \int_0^{t\wedge\eta_d} e^{-rs}(\mathrm{d}U_s - \mathrm{d}H_s - \mathrm{d}X_s) + e^{-rt}P(z_t, K_t, W_t)\mathbf{1}_{\{t<\eta_d\}} + e^{-r\eta_d}\left(L_{\eta_d} + W_{\eta_d}\right)\mathbf{1}_{\{t\geq\eta_d\}}, \tag{A.11}$$

for $t \geq 0$, where

$$\mathrm{d}X_t = \phi_0 K_t \mathbf{1}_{\{\mathrm{d}H_t>0\}} + \phi_1 \mathrm{d}H_t.$$

Let U_t^c be the continuous part of U_t and $\Delta U_t = U_t - U_{t-}$ be the discontinuous part (jump) at time t. Note that it is never optimal for the firm to pay dividends and issue equity simultaneously. Applying the generalized Ito's formula to $e^{-rt}P(z_t, K_t, W_t)$, we obtain:

$$
\begin{aligned}
N_t =& N_0 + \int_0^{t\wedge\eta_d} e^{-rs}(\mathrm{d}U_s - \mathrm{d}H_s - \mathrm{d}X_s) + \left(e^{-r(t\wedge\eta_d)}P(z_{t\wedge\eta_d}, K_{t\wedge\eta_d}, W_{t\wedge\eta_d}) - P(z_0, K_0, W_0)\right) \\
&+ e^{-r\eta_d}\left(L_{\eta_d} + W_{\eta_d} - P(z_{\eta_d}, K_{\eta_d}, W_{\eta_d})\right)\mathbf{1}_{\{t\geq\eta_d\}} \\
=& N_0 + \int_0^{t\wedge\eta_d} e^{-rs}\sigma_z(z_s)K_s q_z \mathrm{d}\mathcal{B}_s^z + \int_0^{t\wedge\eta_d} e^{-rs}\sigma_y A(z_s)K_s(q_w+1)\mathrm{d}\mathcal{B}_s^y + \int_0^{t\wedge\eta_d} e^{-rs}K_s\widetilde{\mathcal{L}}q\mathrm{d}s \\
&+ \left(\int_0^{t\wedge\eta_d} e^{-rs}(-q_w)\mathrm{d}U_s^c + \sum_{0\leq s\leq t\wedge\eta_d:\ \Delta U_s>0} e^{-rs}\left(q(z_s, w_{s-} - \Delta U_s/K_s) - q(z_s, w_{s-})\right)K_s\right) \\
&+ \sum_{n\geq 1:\ \eta_e^n\leq t\wedge\eta_d} e^{-r\eta_e^n}\left(q(z_{\eta_e^n}, w_{\eta_e^n-} + m^n) - q(z_{\eta_e^n}, w_{\eta_e^n-}) - \phi_0 - \phi_1 m^n\right)K_{\eta_e^n} \\
&+ \left(e^{-r\eta_d}\left(\ell - q(z_{\eta_d}, w_{\eta_d})\right)K_{\eta_d}\mathbf{1}_{\{t\geq\eta_d\}} - \zeta\int_0^{t\wedge\eta_d} e^{-rs}\left(\ell - q(z_s, w_s)\right)K_s\mathrm{d}s\right),
\end{aligned}
\tag{A.12}
$$

almost surely, where

$$
\begin{aligned}
\widetilde{\mathcal{L}}q =& \mu_z(z)q_z(z,w) + \frac{1}{2}(\sigma_z(z))^2 q_{zz}(z,w) + \left(i(z,w) - (r+\delta+\zeta)\right)q(z,w) \\
&+ \left((r-\lambda-i(z,w)+\delta)w + A(z) - c(i(z,w))\right)q_w(z,w) \\
&+ \frac{1}{2}\sigma_y^2(A(z))^2 q_{ww}(z,w) + \rho\sigma_z(z)\sigma_y A(z)q_{zw}(z,w) \\
&+ A(z) - c(i(z,w)) - \lambda w + \zeta\ell.
\end{aligned}
$$

We then take expectations on both sides of (A.12). Since $\int_0^{t\wedge\eta_d} e^{-rs}\sigma_z(z_s)K_s q_z \mathrm{d}\mathcal{B}_s^z$ and $\int_0^{t\wedge\eta_d} e^{-rs}\sigma_y(q_w+1)\mathrm{d}\mathcal{B}_s^y$ are martingales, we have

$$
\mathbb{E}\left[\int_0^{t\wedge\eta_d} e^{-rs}\sigma_z(z_s)K_s q_z \mathrm{d}\mathcal{B}_s^z\right] = 0
$$

and

$$
\mathbb{E}\left[\int_0^{t\wedge\eta_d} e^{-rs}\sigma_y(q_w+1)\mathrm{d}\mathcal{B}_s^y\right] = 0.
$$

Using the mean-value theorem, we know that there exists $h_s \in [0, \Delta U_s/K_s]$ for each s such that

$$
q(z_s, w_{s-} - \Delta U_s/K_s) - q(z_s, w_{s-}) = -q_w(z_s, w_{s-} - h_s)\frac{\Delta U_s}{K_s}.
$$

Using the law of iterated conditional expectation, we obtain:

$$
\mathbb{E}\left[e^{-r\eta_d}\left(\ell - q(z_{\eta_d}, w_{\eta_d})\right)K_{\eta_d}\mathbf{1}_{\{t\geq\eta_d\}}\right] = \mathbb{E}\left[\mathbb{E}\left[e^{-r\eta_d}\left(\ell - q(z_{\eta_d}, w_{\eta_d})\right)K_{\eta_d}\mathbf{1}_{\{t\geq\eta_d\}}|(\mathcal{B}_t^z, \mathcal{B}_t^y), t\geq 0\right]\right].
\tag{A.13}
$$

Note that η_d follows an exponential distribution with mean $1/\zeta$ independent of the process $\{(\mathcal{B}_t^z, \mathcal{B}_t^y), t \geq 0\}$. The inner conditional expectation on the right side of (A.13) equals

$$\mathbb{E}\left[e^{-r\eta_d}\left(\ell - q(z_{\eta_d}, w_{\eta_d})\right) K_{\eta_d}\mathbf{1}_{\{t\geq\eta_d\}}|(\mathcal{B}_t^z, \mathcal{B}_t^y), t \geq 0\right] = \int_0^t e^{-rs}(\ell - q(z_s, w_s))K_s \cdot \zeta e^{-\zeta s}\mathrm{d}s, \tag{A.14}$$

where $\zeta e^{-\zeta s}$ is the probability density function of η_d. Moreover,

$$\begin{aligned}
&\zeta\mathbb{E}\left[\int_0^{t\wedge\eta_d} e^{-rs}(\ell - q(z_s, w_s))K_s\mathrm{d}s\right] = \zeta\mathbb{E}\left[\int_0^t e^{-rs}(\ell - q(z_s, w_s))K_s\mathbf{1}_{\{\eta_d>s\}}\mathrm{d}s\right]\\
=&\zeta\mathbb{E}\left[\int_0^t e^{-rs}(\ell - q(z_s, w_s))K_s\mathbb{E}\left[\mathbf{1}_{\{\eta_d>s\}}|(\mathcal{B}_t^z, \mathcal{B}_t^y), t \geq 0\right]\mathrm{d}s\right]\\
=&\zeta\mathbb{E}\left[\int_0^t e^{-rs}(\ell - q(z_s, w_s))K_s \cdot e^{-\zeta s}\mathrm{d}s\right]. \qquad \text{(A.15)}
\end{aligned}$$

Combining (A.13)-(A.15), we obtain:

$$\mathbb{E}\left[e^{-r\eta_d}\left(\ell - q(z_{\eta_d}, w_{\eta_d})\right) K_{\eta_d}\mathbf{1}_{\{t\geq\eta_d\}}\right] = \zeta\mathbb{E}\left[\int_0^{t\wedge\eta_d} e^{-rs}(\ell - q(z_s, w_s))K_s\mathrm{d}s\right]. \tag{A.16}$$

Using these results, we obtain

$$\begin{aligned}
\mathbb{E}\left[N_t\right] = &\mathbb{E}[N_0] + \mathbb{E}\left[\int_0^{t\wedge\eta_d} e^{-rs}K_s\widetilde{L}q\mathrm{d}s\right]\\
&+\mathbb{E}\left[\int_0^{t\wedge\eta_d} e^{-rs}(-q_w)\mathrm{d}U_s^c + \sum_{0\leq s\leq t\wedge\eta_d\colon \Delta U_s>0} e^{-rs}\left(-q_w(z_s, w_{s-} - h_s)\right)\Delta U_s\right]\\
&+\sum_{n\geq1\colon \eta_e^n\leq t\wedge\eta_d}\mathbb{E}\left[e^{-r\eta_e^n}\left(q(z_{\eta_e^n}, w_{\eta_e^n-} + m^n) - q(z_{\eta_e^n}, w_{\eta_e^n-}) - \phi_0 - \phi_1 m^n\right)K_{\eta_e^n}\right].
\end{aligned} \tag{A.17}$$

Note that $q(z, w)$ is the solution to (27) for $w > 0$ subject to the boundary condition (26) on $w = 0$. Using $\tilde{L}q \leq 0$, $-q_w \leq 0$, and

$$q(z, w + m) - q(z, w) - \phi_0 - \phi_1 m \leq 0,$$

we can show that the integrands on the right side of (A.17) are nonpositive. Therefore, for any η_l, we have $\mathbb{E}[N_0] \geq \mathbb{E}[N_{\eta_\ell\wedge\eta_d}]$ and

$$\begin{aligned}
&P(z_0, K_0, W_0) = \mathbb{E}[N_0] \geq \mathbb{E}[N_{\eta_\ell\wedge\eta_d}]\\
=&\mathbb{E}\left[\int_0^{\eta_\ell\wedge\eta_d} e^{-rs}(\mathrm{d}U_s - \mathrm{d}H_s - \mathrm{d}X_s) + e^{-r\eta_\ell}P(z_{\eta_\ell}, K_{\eta_\ell}, W_{\eta_\ell})\mathbf{1}_{\{\eta_\ell<\eta_d\}} + e^{-r\eta_d}\left(L_{\eta_d} + W_{\eta_d}\right)\mathbf{1}_{\{\eta_\ell\geq\eta_d\}}\right]\\
\geq&\mathbb{E}\left[\int_0^{\eta_\ell\wedge\eta_d} e^{-rs}(\mathrm{d}U_s - \mathrm{d}H_s - \mathrm{d}X_s) + e^{-r(\eta_\ell\wedge\eta_d)}\left(L_{\eta_\ell\wedge\eta_d} + W_{\eta_\ell\wedge\eta_d}\right)\right], \qquad \text{(A.18)}
\end{aligned}$$

for any admissible strategy. Also note that (A.17) implies that N_t is a martingale under the strategy $(I^*, U^*, H^*, \eta^*_\ell)$. Therefore, the equality in (A.18) is attained under the strategy $(I^*, U^*, H^*, \eta^*_\ell)$. We thereby establish the stated result.

Turning to Cases 2 and 3, in which $(z_0, w_0) \notin \overline{SR}$, we can write:

$$\begin{aligned} P(z_0, K_0, W_0) &= (L_0 + W_0)\mathbf{1}_{\{(z_0,w_0)\in LR\}} \\ &\quad + \left(P(z_0, K_0, \overline{w}(z_0)K_0) + (w_0 - \overline{w}(z_0))\, K_0\right) \mathbf{1}_{\{(z_0,w_0)\in PR\setminus LR\}} \\ &\quad + \left(P(z_0, K_0, m^1 K_0) - \phi_0 K_0 - (1+\phi_1) m^1 K_0\right) \mathbf{1}_{\{(z_0,w_0)\in ER\setminus LR\}} \\ &= \mathbb{E}\left[\int_0^{\eta^*_\ell \wedge \eta_d} e^{-rs}(\mathrm{d}U_s - \mathrm{d}H_s - \mathrm{d}X_s) + e^{-r(\eta^*_\ell\wedge\eta_d)}\left(L_{\eta^*_\ell\wedge\eta_d} + W_{\eta^*_\ell\wedge\eta_d}\right)\right], \end{aligned} \tag{A.19}$$

by noting that $\{(z_0, \overline{w}(z_0)),\ (z_0, m^1)\} \subset \overline{SR}$.[30] Finally, we complete our proof for Case 3, using a supermartingale argument similar to the one used in our proof for Case 1.

□

B The Case with no Fixed Financing Cost: $\phi_0 = 0$

In this appendix, we characterize the solution for the special case with only proportional equity issuance costs when $\sigma_y > 0$. Without the fixed external financing cost, Tobin's average q, $q(z, w)$, satisfies the following variational inequality:

$$\max\Big\{\underbrace{\mathcal{L}q(z,w)}_{\text{Savings}},\ \underbrace{-q_w(z,w)}_{\text{Payout}},\ \underbrace{q_w(z,w)-\phi_1}_{\text{Equity issuance}},\ \underbrace{\ell - q(z,w)}_{\text{Liquidation}}\Big\} = 0\,, \tag{B.1}$$

subject to the following boundary condition on $w = 0$:

$$\max\{q_w(z,0) - \phi_1,\ \ell - q(z,0)\} = 0\,,$$

where $\mathcal{L}q$ is given by (17) and the optimal investment-capital ratio i^* is given by (15). Absent fixed equity issuance costs, the firm's equity issuance also becomes a singular control problem, which explains the third term in (B.1), which differs from the third term in (27).

[30] $(z_0, m^1) \in \overline{SR}$ follows from the fact that it is suboptimal for the firm to pay dividends and issue equity simultaneously.

C The Case without Contemporaneous Shocks: $\sigma_y = 0$

In this appendix, we characterize the solution for the special case without contemporaneous shocks: $\sigma_y = 0$. In this case, the variational inequality governing Tobin's average q, $q(z,w)$, takes the same form as the one with contemporaneous shocks ($\sigma_y > 0$) for $w > 0$, namely (27), but the operator associated with the savings region is simpler and given by

$$\begin{aligned}\mathcal{L}q(z,w) = \mu_z(z)q_z(z,w) + \frac{1}{2}(\sigma_z(z))^2 q_{zz}(z,w) + (i^*(z,w) - \delta - r - \zeta)q(z,w) \\ + \left((r - \lambda - i^*(z,w) + \delta)w + A(z) - c(i^*(z,w))\right) q_w(z,w) \\ + A(z) - c(i^*(z,w)) - \lambda w + \zeta\ell] \,. \end{aligned} \tag{C.1}$$

It is worth noting that, absent contemporaneous shocks ($\sigma_y = 0$), the second-order derivatives q_{ww} and q_{zw} disappear in (C.1). The optimal investment strategy is given by (15).

The boundary condition for $q(z,w)$ at $w = 0$ is given by

$$\max\{\mathcal{L}q,\ \mathcal{M}q - q,\ \ell - q\} = 0, \tag{C.2}$$

where

$$\begin{aligned}\mathcal{L}q(z,0) \ = \mu_z(z)q_z(z,0) + \frac{1}{2}(\sigma_z(z))^2 q_{zz}(z,0) + (i^{0,*}(z,0) - \delta - r - \zeta)q(z,0) \\ + \left(A(z) - c(i^{0,*}(z,0))\right) q_w(z,0) + \left(A(z) - c(i^{0,*}(z,0)) + \zeta\ell\right) \,. \end{aligned} \tag{C.3}$$

Compared with the condition (26) for the $w = 0$ boundary in the $\sigma_y > 0$ case, there is a new term, $\mathcal{L}q$, in the condition (C.2) for the $w = 0$ boundary in the $\sigma_y = 0$ case: This is because with $\sigma_y = 0$, when the firm's profits are non-negative whenever $A(z) - c(i) \geq 0$, there is no need for the firm to issue equity as internally generated cash flows are sufficient to finance corporate investment. Indeed, when $w = 0$, the optimal investment strategy $i^*(z,0)$ in (C.3) is given by (15), subject to $A(z) - c(i) \geq 0$.

D Numerical Method

In this appendix, we describe the numerical method used for the quantitative analysis of Section 5. Recall that we have specified the logarithmic productivity shock, z_t, as an OU process with the drift and volatility functions given in (28).

D.1 The Penalty Method

We employ the penalty method developed in Dai and Zhong (2010) to solve for the average q, $q(z, w)$, and determine the firm's optimal strategy in the four regions: savings region, payout region, external financing region, and liquidation region. We numerically solve the variational inequality associated with the average q using a penalty approximation as follows:

$$\mathcal{L}q + \beta\left(-q_w\right)^+ + \beta\left(\mathcal{M}q - q\right)^+ + \beta\left(\ell - q\right)^+ = 0, \tag{D.1}$$

where $\beta > 0$ is a large penalty parameter, and the operators $\mathcal{L}$ and $\mathcal{M}$ are given by (17) and (22), respectively.

D.1.1 Computation Domain and Boundary Conditions

For the numerical solution, we fix a finite region in the (z, w) space:

$$D = [z^l, z^u] \times [0, w^u],$$

with $z^u > z^l > 0$ and $w^u > 0$. We select z^u to be sufficiently large and z^l sufficiently close to zero. In addition, we choose w^u to be sufficiently large so that the payout boundary lies below the upper boundary $[z^l, z^u] \times \{w^u\}$. We impose the following boundary conditions:

$$\begin{cases} \max\{-q_w,\ \ell - q\} = 0, & \text{at } w = w^u, \\ \max\{\mathcal{M}q - q, \ell - q\} = 0, & \text{at } w = 0, \end{cases} \tag{D.2}$$

where the first condition corresponds to the optimality of dividend payouts when the firm has ample cash holdings; the second condition reflects the optimality of equity issuance or liquidation.[31]

D.1.2 Numerical Scheme

The numerical scheme is similar to the one used in Dai and Zhong (2010), except that we need to update the optimal investment-capital ratio i^* in each iteration. Small mesh sizes $\Delta z > 0$ and $\Delta w > 0$ are used in the $z-$ and $w-$directions, respectively. The grids are defined as $z^l = z^0 < z^1 < \cdots < z^{n_z} = z^u$ and $0 = w^0 < w^1 < \cdots < w^{n_w} = w^u$, with $z^j = z^l + j\Delta z$ and $w^k = k\Delta w$ for $j = 0, 1, \cdots, n_z$ and $k = 0, 1, \cdots, n_w$. The algorithm proceeds as follows:

[31] We also follow the standard practice in the literature by setting $q_{zz} = q_{zw} = 0$ at $z = z^l$ and $z = z^u$.

1. Initialize the average q, denoted by q^0.

2. Given the average q in the n-th iteration, q^n, update the investment-capital ratio, $i^{*,n}(z^j, w^k)$, to

$$i^{*,n}(z^j, w^k) = \arg\max_i \left(\frac{q^n(z^j, w^k) + w^k}{q^n_{w,c}(z^j, w^k) + 1} - w^k \right) i - c(i), \tag{D.3}$$

for $j = 0, 1, \cdots, n_z$ and $k = 1, 2, \cdots, n_w - 1$. Here $q^n_{w,c}(z^j, w^k)$ denotes the central difference approximation of q^n in the $w-$direction evaluated at (z^j, w^k).

3. Given $i^{*,n}$ and q^n, update q^{n+1} based on (D.1) with boundary conditions (D.2). Specifically, in the interior region of the computational domain, $q^{n+1}(z^j, w^k)$ solves

$$\begin{aligned} &\mathcal{L}^n q^{n+1}(z^j, w^k) + \beta\left(-q^{n+1}_w(z^j, w^k)\right)\mathbf{1}_{\{-q^n_w(z^j,w^k)>0\}} + \beta\left(\ell - q^{n+1}(z^j, w^k)\right)\mathbf{1}_{\{\ell - q^n>0\}} \\ &\qquad + \beta\left(\mathcal{M}^n q^n(z^j, w^k) - q^{n+1}(z^j, w^k)\right)\mathbf{1}_{\{\mathcal{M}^n q^n(z^j,w^k) - q^n(z^j,w^k)>0\}} = 0, \end{aligned} \tag{D.4}$$

where

$$\begin{aligned} \mathcal{L}^n q^{n+1}(z, w) = {} & \mu_z(z) q^{n+1}_z(z, w) + \frac{1}{2}(\sigma_z(z))^2 q^{n+1}_{zz}(z, w) + \left(i^{*,n}(z, w) - (r + \delta + \zeta)\right) q^{n+1}(z, w) \\ & + \left((r - \lambda - i^{*,n}(z, w) + \delta)w + A(z) - c(i^{*,n}(z, w))\right) q^{n+1}_w(z, w) \\ & + \frac{1}{2}\sigma_y^2 (A(z))^2 q_{ww}(z, w) + \rho\sigma_z(z)\sigma_y A(z) q^{n+1}_{zw}(z, w) \\ & + A(z) - c(i^{*,n}(z, w)) - \lambda w + \zeta\ell\,, \end{aligned}$$

and

$$\mathcal{M}^n q^n(z^j, w^k) = \max_{m = w^{k'} - w^k,\ k<k'\le n_w} q^n(z^j, w^k + m) - (\phi_0 + \phi_1 m)\ .$$

We discretize the linear terms (e.g., q^{n+1}_z and q^{n+1}_{zz}) in (D.4) using the upwind finite difference scheme. The boundary conditions are:

   (a) If $k = k_w$, i.e., $w = w^u$,

   $$-q^{n+1}_w(z^j, w^k)\mathbf{1}_{\{-q^n_w(z^j,w^k)\ge \ell - q^n(z^j,w^k)\}} + (\ell - q^{n+1}(z^j, w^k))\mathbf{1}_{\{-q^n_w(z^j,w^k)<\ell - q^n(z^j,w^k)\}};$$

   (b) If $k = 0$, i.e., $w = 0$,

   $$q^{n+1}(z^j, w^0) = \max\{\mathcal{M}^n q^n(z^j, w^0),\ \ell\};$$

   (c) If $0 < k < n_w$ and $j = 0$ or n_z, substitute $q_{zz} = q_{zw} = 0$ into (D.4).

We discretize these boundary conditions using the upwind scheme when applicable.

4. Repeat Steps 2 and 3 until

$$\frac{||q^{n+1} - q^n||}{||q^n||} < \text{Tolerance}, \tag{D.5}$$

where Tolerance is a small positive constant.

E The Case with Debt Financing

Using the extended model's homogeneity property and following essentially the same analysis as that for our baseline model of Section 2, we can show that the firm's average q satisfies the following variational inequality:

$$\max\Bigg\{ \underbrace{\mathcal{L}q(z,w)}_{\text{Savings/borrowing}}, \underbrace{-q_w(z,w)}_{\text{Payout}}, \underbrace{\mathcal{M}q(z,w) - q(z,w)}_{\text{Equity issuance}}, \underbrace{\ell - q(z,w)}_{\text{Liquidation}} \Bigg\} = 0\,, \tag{E.1}$$

subject to the boundary condition at $w = -b$:

$$\max\left\{\mathcal{M}q(z,-b) - q(z,-b),\ \ell - q(z,-b)\right\} = 0\,. \tag{E.2}$$

The first term in (E.1), $\mathcal{L}q(z,w)$, describes the dynamics in the interior region, including both the saving ($W \geq 0$) and borrowing ($W < 0$) subregions, and is given by

$$\begin{aligned}
\mathcal{L}q(z,w) = {} & \mu_z(z)q_z(z,w) + \frac{1}{2}(\sigma_z(z))^2 q_{zz}(z,w) + (i^*(z,w) - (r+\delta+\zeta))\,q(z,w) \\
& + \left((r - \lambda\mathbf{I}_{\{w>0\}} + \alpha\mathbf{I}_{\{w<0\}} - i^*(z,w) + \delta)w + A(z) - c(i^*(z,w))\right) q_w(z,w) \\
& + \frac{1}{2}\sigma_y^2(A(z))^2 q_{ww}(z,w) + \rho\sigma_z(z)\sigma_y A(z) q_{zw}(z,w) \\
& + A(z) - c(i^*(z,w)) - \lambda w\mathbf{I}_{\{w>0\}} + \alpha w\mathbf{I}_{\{w<0\}} + \zeta\ell\,,
\end{aligned} \tag{E.3}$$

where i^* is given in (15) and $\mathcal{M}q$ associated with equity financing is defined by (22). As the firm is active along at least one of the four decision margins (savings/borrowing, payout, external financing, and liquidation), at least one of the four terms inside the max operator in (E.1) equals zero. When $\mathcal{L}q(z,w) = 0$, the firm either saves (if $w > 0$) or borrows (if $w < 0$), and its optimal investment policy is characterized by (15). Compared to $\mathcal{L}q(z,w)$ in the variational inequality (27) for the baseline model (without debt), (E.3) contains two new terms, $\alpha w q_w(z,w)\mathbf{I}_{\{w<0\}}$ and $\alpha w\mathbf{I}_{\{w<0\}}$, capturing the effect of the borrowing cost α.

We have fully described the variational inequality for the extended model with debt financing, which allows us to globally solve Tobin's average q and characterize firm optimality.

Next we consider the special case without contemporaneous shocks ($\sigma_y = 0$). In this case, the variational inequality is the same as that given in (E.1), but with a simpler expression for $\mathcal{L}q(z,w)$ as the q_{ww} and q_{zw} terms disappear. But the boundary condition that characterizes the firm's behavior when it exhausts its debt capacity, i.e., $w_t = -b$, is more involved. Specifically, when $w_t = -b$, Tobin's average q has to satisfy the following condition:

$$\max\left\{\mathcal{L}q(z,-b),\ \mathcal{M}q(z,-b) - q(z,-b),\ \ell - q(z,-b)\right\} = 0\,. \tag{E.4}$$

Compared with the boundary condition (E.2) at the debt-limit boundary ($w_t = -b$), there is an additional term, $\mathcal{L}q(z,-b)$. This new term appears in (E.4) only because the firm faces no contemporaneous shocks ($\sigma_y = 0$). That is, when the firm reaches its debt capacity, if equity issuance and liquidation are both suboptimal, the firm must return to the interior region by choosing its investment so that the drift of w_t is weakly positive, moving w_t away from the boundary $w = -b$. This implies the following constraint when $w_t = -b$: $A(z) - c(i) \geq (r + \alpha + \delta - i)b$ (see (E.3) with $w = -b$).

Mathematically, if $\mathcal{L}q(z,-b) = 0$ holds when $w_t = -b$, the optimal investment-capital ratio must satisfy (15) subject to $A(z) - c(i) \geq (r + \alpha + \delta - i)b$.

References

Dai, Min, Cong Qin, and Neng Wang. 2023. Dynamic trading with realization utility. *Journal of Finance*, forthcoming.

Dai, Min and Yifei Zhong. 2010. Penalty methods for continuous-time portfolio selection with proportional transaction costs. *Journal of Computational Finance* 13(3):1–31.

Davis, Mark H. A. and Andrew R. Norman. 1990. Portfolio selection with transaction costs. *Mathematics of Operations Research* 15(4):676–713.