Moral Hazard and Investment-Cash-Flow Sensitivity

Hengjie Ai, Kai Li, and Rui Li*

March 5, 2017

Abstract

We develop a dynamic model of investment with moral hazard to provide a micro-foundation for financing constraints. In the model, standard investment-cash-flow sensitivity regressions will find a small coefficient on Tobin's Q and a large and significant coefficient on cash flow. Our calibration replicates the empirical fact that larger and more mature firms are *less* financially constrained but have *higher* investment-cash-flow sensitivity. Our theory therefore resolves the long-standing puzzle of the existence of the investment-cash-flow sensitivity and the seemingly weak relationship between investment-cash-flow sensitivity and the severity of financing constraints documented by Kaplan and Zingales (1997) and many others.

Keywords: Financing constraints, dynamic moral hazard, Q theory, investment-cash-flow sensitivity

*Hengjie Ai (hengjie@umn.edu) is at the Carlson School of Management of University of Minnesota; Kai Li (kaili@ust.hk) is associated with Hong Kong University of Science and Technology; and Rui Li (abclirui@gmail.com) is associated with University of Massachusetts Boston. We thank Murray Frank, participants of the Finance Workshop at the University of Minnesota and brownbag workshop at the Federal Reserve Bank of Minneapolis for their helpful comments. The usual disclaimer applies.

I Introduction

The neoclassical investment theory implies that firms' investment should not respond to any other variables after controlling for Tobin's Q, or the ratio of firms' market value to the replacement cost of their capital stock. Empirically, however, regressions of investment on Tobin's Q and cash flow typically have a large coefficient on cash flow and a small coefficient on Tobin's Q. The significant investment-cash-flow sensitivity in the data presents a serious difficulty to the frictionless neoclassical investment theory and is typically interpreted as the result of financing constraints. It is natural to ask what theory of financing constraints can quantitatively account for the observed investment-cash-flow sensitivity in the data. However, any such attempt faces the following challenge: Empirically, investment-cash-flow sensitivity is often *negatively* related to other measures of financing constraints (Kaplan and Zingales (1997)). In particular, larger and older firms, which are often thought to be *less* financially constrained, typically have *higher* investment-cash-flow sensitivity than smaller and younger ones.

In this paper, we embed moral hazard into an otherwise standard neoclassical model of investment to provide a resolution of the above puzzle. Financing constraints endogenously arise as the result of optimal contracting under agency frictions. In our model, even after controlling for Tobin's Q, investment responds to cash-flow shocks because the later carry valuable information useful for incentive provision. More importantly, under the optimal contract, larger and more mature firms are *less* financially constrained. Nevertheless, they exhibit significantly *stronger* investment-cash-flow sensitivity than do smaller and younger ones, as in the data.

In our model, entrepreneurs have access to a productive project in which cash flows depend on capital input but are also subject to shocks that are unobservable to outside investors. Because entrepreneurs do not have enough wealth, outside investors need to cover temporary losses due to the unexpected cash flow shocks in order to continue the project. Moral hazard arises because entrepreneurs can secretly steal output for private consumption. In this environment, we show that the optimal dynamic contract has the following three predictions. First, because of the presence of moral hazard, it is optimal for outside investors to finance the project on a scale lower than its first-best level. That is, moral hazard endogenously leads to financing constraints. In our setup, unobservable cash flow shocks enter multiplicatively into the production function. As a result, increases in capital input lower its marginal product and raise the level of noise in output. Because efficiency requires equalizing the marginal product of capital to its marginal cost, which includes both the physical cost and the agency cost of capital, moral hazard leads to a lower level of investment due to the endogenous financing constraint.

Second, the optimal contact requires investment to respond positively to shocks to firms' cash flow even after controlling for Tobin's Q. To prevent entrepreneurs from diverting firms' cash flow, the optimal contract rewards high output by providing more capital and punishes low output by tightening the financing constraint. While Tobin's Q is a strictly increasing function of the marginal (physical) product of capital, cash flow contains independent information that is informative about the entrepreneur's hidden action. Because under the optimal contract, investment and financing are used for incentive provision, they optimally responds to shocks in cash flow that are orthogonal to Tobin's Q.

Third, the magnitude of investment-cash-flow sensitivity *increases* with firm size and *decreases* with the tightness of firms' financing constraint. In the presence of moral hazard, investment-cash-flow sensitivity is used as a tool for incentive provision: Following a sequence of positive performances, the financing constraint is relaxed and the probability of inefficient liquidation diminishes. As a result, large and mature firms in our model are less financially constrained, and the marginal agency cost of investment is lower in these firms. The lower cost of agency frictions allows less constrained firms to use high investment-cash-flow sensitivity to provide stronger incentives. In contrast, young and small firms face tighter financing constraints and higher probabilities of inefficient liquidation. They optimally chose not to use high-powered incentives because they raise the probability of inefficient liquidation.

We calibrate our model to evaluate its ability to quantitatively account for the stylized patterns in investment-cash-flow sensitivity in the data. We show that our model matches well the empirical negative relationship between firm size and investment rate, and the negative relationship between size and Tobin's Q. In our model simulations, cash flow is a robust predictor of investment even after controlling for Tobin's Q. More importantly, we demonstrate that the investment-cash-flow sensitivity is robustly increasing firm size, as in the data, despite that large firms typically have more relaxed financing constraint in the sense that the Lagrangian multiplier of their incentive compatibility constraint and the marginal product of their capital are both lower than those in small firms.

Related literature The puzzling fact of the existence of investment-cash-flow sensitivity and the inconsistent and often negative relationship between investment-cash flow sensitivity and other measures of financing constraints are well documented in the empirical literature. Beginning from Fazzari, Hubbard, and Petersen (1988), a large literature documents a robust investment-cash-flow sensitivity in the data and many use it as a quantitative measure of the severity of financing constraints, for example, Hoshi, Kashyap, and Scharfstein (1991) and Almeida and Campello (2007), as well as the papers referenced in Hubbard (1998).[1]

At the same time, Kaplan and Zingales (1997) and Cleary (1999) find a negative relationship between the magnitude of investment-cash-flow sensitivity and the severity of firms' financing constraints. Subsequently, Vogt (1994) and Kadapakkam, Kumar, and Riddick (1998) document robust evidence that investment-cash-flow sensitivity increases with firm size. The more recent literature proposes a variety of empirical measures of financing constraints, for example, Lamont, Polk, and Saaá-Requejo (2001), Whited and Wu (2006), Hadlock and Pierce (2010), Hoberg and Maksimovic (2015), and Farre-Mensa and Ljungqvist (2015), many providing contradicting results with the investment-cash-flow sensitivity. Our purpose is to present a theory of endogenous financing constraints to provide a coherent interpretation of the above (seemingly) contradictory empirical findings.

Our theoretical model builds on the large literature on dynamic moral hazard with investment, especially continuous-time models, for example, DeMarzo and Sannikov (2006), Biais, Mariotti, Plantin, and Rochet (2007), and Biais, Mariotti, and Villeneuve (2010), He (2009), DeMarzo, Fishman, He, and Wang (2012), and Zhu (2013). Our model is an extension of the above papers that allows for neoclassical production and adjustment cost. All of the above models feature a single state variable, which implies that Q theory must hold locally, as we show later in the paper. In contrast, the marginal product of capital and the promised utility in our model are two independent state variables, and this feature allows us to replicate quantitatively the empirical failure of Q-theory.

Our paper is evidently related to the literature that studies the investment-cash-flow sensitivity and its dependence on firm characteristics. Gomes (2001) emphasizes the importance of the general equilibrium effect and aggregate shocks; Alti (2003) focuses

[1]Several authors, for example, Blundell, Bond, Devereux, and Schiantarelli (1992), Gilchrist and Himmelberg (1995), and Erickson and Whited (2000), emphasize the importance of measurement error in investment-cash-flow regressions.

on the role of learning; Moyen (2004) proposes a heterogenous firm model with financing constraints; Lorenzoni and Walentin (2007) study the impact of limited commitment; and Abel and Eberly (2011) develop a model of stochastic depreciation rates. None of the above models focus on moral hazard as a micro-foundation for financial constraints and quantitatively account for the positive relationship between investment-cash-flow sensitivity and firm size and age as we do in this paper.

The rest of the paper is organized as follows: In Section II, we layout the basic model; in Section III, we describe the optimal contract; Section IV is about the investment-cash-flow sensitivity under the optimal contract; in Section A we show the calibration of the model and its quantitative performance in matching the data; and in Section VI we conclude the paper.

II The model

Time is infinite and continuous. A unit measure of risk-neutral entrepreneurs arrives at the economy per unit of time. Entrepreneurs are endowed with a technology that produces consumption goods from capital. However, they do not have enough initial wealth and borrow from outside investors to finance capital. Given an initial input, K_0, and a sequence of investment, $\{I_t\}$, capital can be accumulated according to the following neoclassical capital accumulation technology:

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

We assume that investing I_t requires a flow cost of $I_t + \phi\left(\frac{I_t}{K_t}\right) K_t$, which has to be financed from outside investors. Here, we assume that ϕ is the standard quadratic adjustment cost function:

$$\phi\,(i) = \frac{1}{2}\phi_0\,(i - (\delta + \mu_Z))^2\,,$$

where $\delta + \mu_Z$ is the steady-state investment rate and μ_Z will be defined later.

Like in DeMarzo and Sannikov (2006) and DeMarzo, Fishman, He, and Wang (2012), we specify the production technology by a cumulative output process, $\{Y_t\}_{t=0}^{\infty}$. Given a sequence of input, $\{K_t\}$, the flow of output at time t, dY_t is given by

$$dY_t = K_t^{\alpha} Z_t^{1-\alpha} dt + K_t \sigma dB_t, \tag{1}$$

where Z_t is the productivity of the firm observable to both the entrepreneur and outside investors, and $\alpha \in (0,1)$ is a return to scale parameter. To allow for long-run growth, we assume that Z_t follows a geometric Brownian motion with growth rate μ_Z and volatility σ_Z:

$$Z_t = \exp\left(\mu_Z t + \sigma_Z B_{Z,t}\right), \text{ for all } t > 0. \tag{2}$$

In equation (1), although $K_t^\alpha Z_t^{1-\alpha}$ is perfectly observable to outside investors, Brownian motion, dB_t, is not. As a result, output is observed only with noise, and σ_Z determines the level of noise in output. Even though output is, on average, positive, $Z_t^{1-\alpha} K_t^\alpha > 0$, the entrepreneur can incur temporary losses due to the presence of productivity shock, dB_t. Large losses must be covered by deep-pocketed outside investors whenever the entrepreneur's own fund is insufficient. Here, we assume that the level of noise in output is proportional to the size of the firm. This assumption is made for two reasons. First, our model allows for long-run growth of firms. Proportional noise prevents firms from growing out of the agency frictions in the long-run. Second, as in Edmans, Gabaix, and Landier (2008) and Edmans and Gabaix (2016), our assumption guarantees that the agent's hidden action is multiplicative in firm size.

In our setup, moral hazard arises because output is not fully observable, and the entrepreneur can claim large losses, privately consume all output, and declare bankruptcy. Efficiency requires entrepreneurs to deploy capital and produce output, while moral hazard implies that lending contracts must provide appropriate incentives for entrepreneurs not to misreport output and hide cash flow for private consumption. Intuitively, the optimal lending contract must reward high output by providing more financing and payment, and punish low output by reducing subsequent lending and terminating the project. We assume that upon termination, the outside investor can fully recover the firm's capital stock, K_T, but all cash flows are lost.

Formally, a contract must specify $\left\{ \left(K_t \left(\hat{Y}^t, Z^t \right), C_t \left(\hat{Y}^t, Z^t \right) \right)_{t=0}^{\tau \wedge T}, T \right\}$, where K_t is the capital input; C_t is the cumulative payment to the entrepreneur; and T is the liquidation time of the project. We assume that entrepreneurs exit the economy at rate κ per unit of time and let τ denote the stopping time of exit. We use $\hat{Y}^t$ to denote the history of the entrepreneur's report of output up to time t and Z^t to denote the history of observable productivity shocks up to time t.

Given a contract, the entrepreneur's utility from reporting $\left\{\hat{Y}_t\right\}$ is given by:

$$E_0\left[\int_0^{\tau\wedge T} e^{-\beta t}\left[dC_t\left(\hat{Y}^t, Z^t\right) + dY_t - d\hat{Y}_t\right]\right].$$

That is, by (mis)reporting Y_t as $\hat{Y}_t$, the entrepreneur receives consumption $dC_t\left(\hat{Y}^t, Z^t\right)$ and consumes the difference, $dY_t - d\hat{Y}_t$. Incentive compatibility requires that misreporting never can be optimal, that is,

$$E_0\left[\int_0^{\tau\wedge T} e^{-\beta t} dC_t\left(Y^t, Z^t\right)\right] \geq E_0\left[\int_0^{\tau\wedge T} e^{-\beta t}\left[dC_t\left(\hat{Y}^t, Z^t\right) + dY_t - d\hat{Y}_t\right]\right], \tag{3}$$

for all $\left\{\hat{Y}_t\right\}$. If the entrepreneur truthfully reports, the present value of the firm's cash flow to the outside investor is

$$E_0\left[\int_0^{\tau\wedge T} e^{-rt}\left[dY_t - dC_t\left(Y^t, Z^t\right) - \left(I_t + \phi\left(\frac{I_t}{K_t}\right)K_t\right)dt\right] - K_0\right]. \tag{4}$$

The optimal contracting problem maximizes outside investor's payoff, (4), subject to incentive compatibility, (3), and a participation constraint:

$$E_0\left[\int_0^{\tau} e^{-\beta t} dC_t\left(Y^t, Z^t\right)\right] \geq U_0.$$

III Optimal contracting

As is standard in the dynamic contracting literature, we use promised utility as a state variable. Given a contract $\left\{(K_t, C_t)_{t=0}^{\tau\wedge T}, T\right\}$, the entrepreneur's time-t continuation utility can be calculated as:

$$U_t = E_t\left[\int_t^{\tau\wedge T} e^{-\beta(s-t)} dC_s\right] \text{ for } t \in [0, \tau\wedge T].$$

The present value of firm's cash flow is a function of current productivity, Z_t, current capital stock, K_t, and the promised utility to the entrepreneur:

$$V\left(Z_t, K_t, U_t\right) = E_t\left[\int_t^{\tau\wedge T} e^{-rs}\left[dY_s - dC_s - \left(I_s + \phi\left(\frac{I_s}{K_s}\right)K_t\right)ds\right]\right].$$

Given the homogeneity of the problem, it is straightforward to show that the value function satisfies

$$V(Z,K,U) = Zv\left(\frac{K}{Z},\frac{U}{Z}\right),$$

for some function $v(k,u)$, where we define $k = \frac{K}{Z}$ and $u = \frac{U}{Z}$. We also normalize the compensation policy and define $c = \frac{C}{Z}$. The definition of promised utility implies that if the entrepreneur does not divert cash flow, then normalized continuation utility must follow

$$du_t = u_t(\beta + \kappa - \mu_Z)\,dt - dc_t + g_t\sigma dB_t + k_t h_t \sigma_Z dB_{Z,t}, \tag{5}$$

where $\{g_t\}$ and $\{h_t\}$ are two predictable processes such that $\{g_t Z_t\}$ and $\{k_t h_t Z_t\}$ are square integrable.

Here, g_t and h_t are the sensitivities of the entrepreneur's normalized continuation utility with respect to the unobservable shocks in firm's cash flow, dB_t, and the observable productivity shocks, dB_{Zt}, respectively. Standard results imply that the incentive compatibility constraint (3) can be written as a restriction on g_t, which we formally state as follows.

Lemma 1 *A contract $(\{C_t\},\{K_t\},T)$ is incentive compatible, if and only if*

$$g_t \geq k_t \textit{ for all } t \in [0, \tau \wedge T]. \tag{6}$$

Proof. See Appendix A. ■

Condition (6) requires the level of the sensitivity g_t to be proportional to the capital stock of the firm. This is because the level of noise in firm's output, and therefore the amount of cash flow that the entrepreneur can divert for personal consumption, is proportional to K_t. Under the optimal contract, this incentive constraint must be binding because the firm is effectively risk averse due to the presence of agency frictions (the value function v is concave).

By using Lemma 1, we can show that the value function must satisfy the following HJB equation in the interior of its domain:

$$(r + \kappa - \mu_Z)\,v(k,u) = \max_{i,h}\left\{\begin{array}{c} k^\alpha - ik - k\phi(i) + (\beta + \kappa - \mu_Z)\,uv_u(k,u) \\ +\,(i - \delta - \mu_Z)\,kv_k(k,u) \\ +\frac{1}{2}\left(\sigma^2 + \left(h - \frac{u}{k}\right)\sigma_Z^2\right)k^2 v_{uu}(k,u) \\ +\frac{1}{2}k^2\sigma_Z^2 v_{kk}(k,u) - k^2\left(h - \frac{u}{k}\right)\sigma_Z^2 v_{ku}(k,u). \end{array}\right\}, \tag{7}$$

where $i = \frac{I}{K}$ is the investment-to-capital ratio.

In addition, $\forall k$, $v(k,u)$ must satisfy the following conditions on the boundaries:

$$v(k,0) = k, \tag{8}$$

$$v(k,u) \geq \max_{x\geq 0} \{v(k,u-x) - x\}. \tag{9}$$

Equation (8) sets the value function at the liquidation boundary: upon liquidation, the entrepreneur receives zero payoff, while the outside investor can recover the capital stock, but loses all potential future cash flows. Equation (9) also has an intuitive interpretation. Promised utility, u, can be delivered in two ways, immediate cash reward or promised payment in the future. Paying x amount of cash immediately costs the outside investor x and reduces the obligation of promised utility by the same amount. Because cash payment is always an option, we must have $v(k,u) \geq v(k,u-x) - x$ for all $x > 0$. This observation motivates the definition of the payoff boundary:

$$\hat{u}(k) = \inf \{u : \forall x > 0,\ v(k,u) = v(k,u-x) - x\}.$$

Clearly, in the region in which $u > \hat{u}(k)$, $v_u(k,u) = -1$. Here, because the entrepreneur is less patient, immediate payoff is always optimal, and under the optimal contract, u instantaneously moves to $\hat{u}(k)$. For $u < \hat{u}(k)$, concavity implies that $\frac{\partial}{\partial u} v(k,u) > -1$, and the dynamics of u follows equation (5).

To illustrate the dynamics of the state variables under the optimal contract, in Figure 1, we plot the domain of the state variables (left) and the expected direction of their movement under the optimal contract (right).[2] The solid line in both figures is the payoff boundary $\hat{u}(k)$. The dashed line represents the trajectory of (k,u) such that the marginal Q equals one: $v_k(k_0,u_0) = 1$. Without adjustment cost, the state variables stay on this line with probability one. With adjustment cost, the system tends to move toward the dashed line, but may be pushed away due to the unexpected Brownian motion shocks in productivity. Finally, note that the optimal contract starts at (k_0,u_0), which is the initial condition that solves the optimization problem (4) and is marked as a star in the figure.

As shown in the right panel of Figure 1, over time, both state variables u and k tend to increase toward the payoff boundary, where $v_u(k,u) = -1$, and concavity diminishes. The increases in k imply that firms grow in size over time, and size and

[2]The plots in Figure 1 are based on the calibrated parameter values introduced in Section V.

Figure 1: The domain of the state variables.

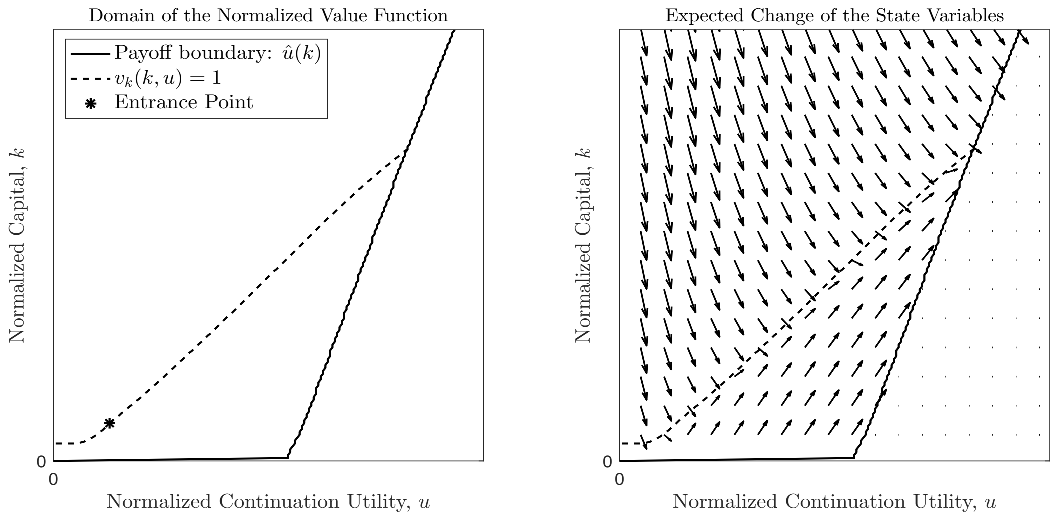


Figure 1 plots the domain of the state variables, (k, u) (left panel) and their expected changes under the optimal contract (right panel). In both panels, the solid line is the payoff boundary $\hat{u}\,(k)$. Under the optimal contract, the state variables, (k, u), stay between the liquidation boundary, $u = 0$ and the payoff boundary $\hat{u}(k)$. The dashed line is the trajectory of (k, u) where marginal Q equal one.

age are positively correlated in the equilibrium distribution. The positive trend in u implies that agency frictions diminish over time: Larger and more mature firms are less financially constrained.

Having obtained the solution of the optimal contract, in the next two sections, we analyze its implications on financing constraints and investment-cash-flow sensitivity, both qualitatively and quantitatively.

IV Investment-cash-flow sensitivity

A Endogenous financing constraint

To better illustrate the implication of our model on financing constraints, we consider a special case without adjustment cost: $\phi\,(i) = 0$. Under this assumption, the shadow value of capital in the firm is the same as its market price. Therefore, the value function must be of the form $v\,(k, u) = \bar{v}\,(u) + k$. This allows us to reduce one state variable and represent the value function and policy functions as functions of u. For every u, there is a most efficient level of capital, $k\,(u)$, where, in the absence of adjustment cost, capital instantaneously moves to $k(u)$. Using (7), it is not difficult to show that $\bar{v}(u)$ and $k(u)$ must satisfy:

$$(r+\kappa-\mu_Z)\,\bar{v}(u)=\max_k\left\{k^{\alpha}-(r+\delta+\kappa)\,k+(\beta+\kappa-\mu_Z)\,u\bar{v}'(u)+\frac{1}{2}\bar{v}''(u)\,k^2\sigma^2\right\}, \tag{10}$$

with the boundary condition

$$\bar{v}(0)=0;\quad \bar{v}'(\hat{u})=-1;\quad \bar{v}''(\hat{u})=0.$$

Note that without moral hazard, the optimal level of capital must equalize the marginal product of capital to the user cost of capital: $\alpha k^{\alpha-1}=(r+\delta+\kappa)$. Let k^* denote the first-best level of capital, that is, $k^*=\left[\frac{1}{\alpha}(r+\delta+\kappa)\right]^{\frac{1}{\alpha-1}}$. The following proposition formalizes the notion of financing constraint induced by moral hazard:

Proposition 1 (Financing Constraint) *In the model without adjustment cost, the optimal level of capital input, $k(u)$ is given by:*

$$\alpha k(u)^{\alpha-1}=(r+\delta+\kappa)-k(u)\,\bar{v}''(u)\,\sigma^2. \tag{11}$$

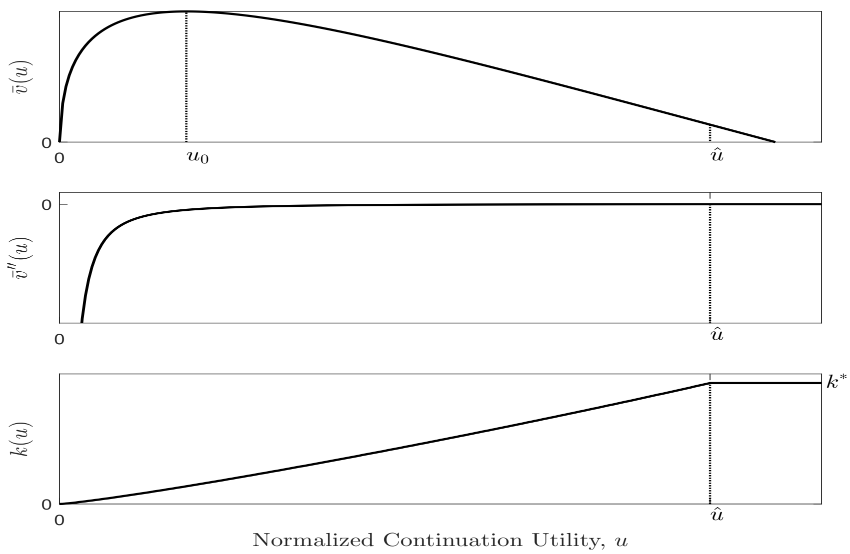
In particular, $k(u)<k^*$, where k^* is the first-best level of capital.

Proof. Equation (11) is straightforward from (10), and, concavity of $\bar{v}$ (Lemma 2 in Appendix C) implies $k(u)<k^*$. ■

Equation (11) has an intuitive interpretation. The term $\alpha k^{\alpha-1}$ is the marginal product of capital, and $r+\delta+\kappa$ is the marginal physical cost of capital, or the user cost of capital. The additional term $k\bar{v}''(u)\,\sigma^2$ can be interpreted as the marginal agency cost of capital. Due to the presence of agency frictions, the firm is effectively risk averse: $v''(u)<0$. Investing an additional unit of capital is costly because it allows the entrepreneur to divert more cash flow. As a result, the presence of moral hazard constraints the level of financing below its first best level.

Figure 2 illustrates the shape of the value function $\bar{v}(u)$ and its implications on financing constraints. Clearly, the value function $\bar{v}(u)$ in the top panel is strictly concave on $(0,\hat{u})$. This reflects moral-hazard-induced risk aversion: Firms dislike variations in entrepreneur's continuation utility because they increase the likelihood of inefficient liquidation. The middle panel plots the absolute value of the second-order derivative of the value function, $\bar{v}''(u)$. Note that risk aversion is maximized at the liquidation

Figure 2: The value function and the optimal policy.



The top panel of Figure 2 plots the normalized value function $\bar{v}(u)$. The middle panel plots the absolute value of the second order derivative of the value function, and the bottom panel plots the optimal choice of capital as a function of normalized utility, $k(u)$.

boundary, $u = 0$, and decreases gradually to zero as u approaches the payoff boundary, $\hat{u}$. Intuitively, the incentive provision requires tying the entrepreneur's continuation utility to performance. This is especially costly when u is close to zero, because small shocks in output may trigger inefficient liquidation. As u moves away from its left boundary, the probability of liquation diminishes and the firm becomes close to risk neutral. Because $k\bar{v}''(u)\,\sigma^2$ is a measure of the marginal agency cost of investment, the optimal level of financing starts at zero and converges to the first-best level, k^*, as risk aversion vanishes at $\hat{u}$.

B Investment-cash-flow sensitivity

The above version of our model without adjustment cost generates financing constraints due to the agency cost of investment. However, Tobin's Q perfectly predicts investment, at least locally, because all policies are functions of a single state variable, u. In fact, any monotone function of u should perfectly predict investment.

Our full model breaks the link between investment and Tobin's Q because it has two state variables, k and u. In addition, the dynamics of k_t and u_t are driven by two independent shocks, the observable productivity shock, $dB_{Z,t}$, and the unobservable cash flow shock, dB_t. The presence of multiple shocks and multiple state variables implies that investment and Tobin's Q will not be (locally) perfectly correlated. That is, Q-theory fails to perfectly predict investment even if the econometrician has infinite an amount of data.

To understand the investment-cash-flow sensitivity regression in our model, we first write investment and Tobin's Q as functions of state variables. Using equation (7), the optimal investment policy in the full model, $i(k,u)$, must satisfy

$$i(k,u) = (\delta + \mu_Z) + \frac{1}{\phi_0}\left(v_k(k,u) - 1\right).$$

Also, we can define $q(k,u) = \frac{v(k,u)}{k}$ as the average Q for the firm.

Next, for any diffusion process x_t, we define $d\hat{x}_t = dx_t - E_t[dx_t]$. Using Ito's formula and the law of motion of k_t and u_t, we can write the innovations in investment as

$$d\hat{\imath}(k_t,u_t) = \left[i_u(k_t,u_t)\,h(k_t,u_t) - i_k(k_t,u_t)\right]k_t\sigma_Z dB_{Z,t} + i_u(k_t,u_t)\,k_t\sigma dB_t, \tag{12}$$

where $h(k,u)$ is the sensitivity of continuation utility with respect to the observable productivity shock dZ_t and satisfies the first-order condition:

$$h(k,u) = \frac{u}{k} + \frac{v_{ku}(k,u)}{v_{uu}(k,u)}.$$

We can similarly write the innovations of $d\hat{q}(k_t,u_t)$ as

$$d\hat{q}(k_t,u_t) = \left[q_u(k_t,u_t)\,h(k_t,u_t) - q_k(k_t,u_t)\right]k_t\sigma_Z dB_{Z,t} + q_u(k_t,u_t)\,k_t\sigma dB_t \tag{13}$$

Note that the innovations in i_t and q_t are driven by two sources of shocks, $dB_{Z,t}$ and dB_t, which are independent of each other. In addition, no single state variable can summarize the two sources of shocks. As a result, changes in q_t typically do not contain enough information to predict changes in i_t. Let $dy_t = \frac{dY_t}{K_t}$ be output normalized by capital stock and then $d\hat{y}_t = \sigma dB_t$. The following proposition derives the local regression coefficient of $d\hat{\imath}_t$ on changes in Tobin's Q, $d\hat{q}_t$, and cash flow, $d\hat{y}_t$:

Proposition 2 (Investment-Cash-Flow Sensitivity) *In the model with adjustment* *cost, under the optimal contract,*

$$d\hat{\imath}_t = \beta_q(k_t, u_t)\, d\hat{q}_t + \beta_y(k_t, u_t)\, d\hat{y}_t,$$

where

$$\beta_q(k_t, u_t) = \frac{[i_u(k_t, u_t)\, h(k_t, u_t) - i_k(k_t, u_t)]}{[q_u(k_t, u_t)\, h(k_t, u_t) - q_k(k_t, u_t)]}, \tag{14}$$

and

$$\beta_y(k_t, u_t) = k_t\left[i_u(k_t, u_t) - \beta_q(k_t, u_t)\, q_u(k_t, u_t)\right]. \tag{15}$$

Proof. See Appendix B. ■

We make several observations. First, by comparing equations (12), (13), and (14), it is clear that in a conditional regression of investment on Tobin's Q, the coefficient on q identifies the the relative sensitivity of investment with respect to the observable productivity shock, $dB_{Z,t}$. Just like in neoclassical models with adjustment cost, productivity shocks move investment and Tobin's Q in the same direction, and the regression coefficient of investment on Tobin's identifies the elasticity of investment and Tobin's Q with respect to these shocks.

Second, in the special case with no agency frictions, investment and Tobin's Q are independent of promised utilities, $i_u(k_t, u_t) = q_u(k_t, u_t) = 0$. In this case, our model reduces to the frictionless neoclasssical model of investment, and investment-cash-flow sensitivity is zero.

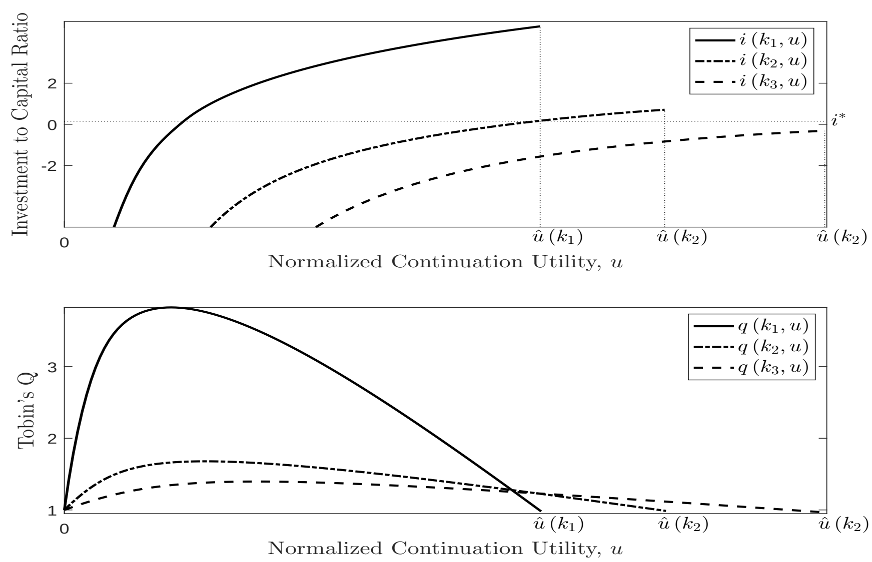
Third, a sufficient condition for a positive investment-cash-flow sensitivity is that $i_u(k_t, u_t) > 0$ and $q_u(k_t, u_t) < 0$. Although it is difficult to prove these results analytically, due to the complexity of the PDE (7), that investment increases with promised utility and Tobin's Q decreases with promised utility (in most part of its domain) is intuitive and is a robust feature of our model.

In our model, investment increases with promised utility, u, because incentives are provided by granting the entrepreneur more capital after positive performance. Intuitively, under moral hazard, financing is used as a tool for incentive provision. After positive surprises in output, the optimal contract rewards the entrepreneur by providing additional financing to allow him to invest at a level closer to the first-best level. This arrangement is optimal because, as promised utility increases, the probability of inefficient liquidation, and therefore the marginal agency cost of investment diminishes, allowing the project to operate at a scale closer to the first-best level. Similarly, the optimal contract tightens the financial constraint if the entrepreneur underperforms;

therefore, investment reduces whenever promised utility does so.

Second, that $q_u(k,u) < 0$ is also intuitive. To support more promised utility, the outside investor needs to allocate a higher fraction of firm's cash flow to the entrepreneur, and, as a result, the value of the firm to outside investors decreases in promised utility. In fact, as u increases, $q_u(k,u)$ converges to $-\frac{1}{k}$ for all k.

Figure 3: Investment-cash-flow sensitivity.



The top panel of Figure 3 plots the third order derivative of normalized value function. The bottom panel plots investment-cash-flow sensitivity, that is, the first order derivative of the policy function $k(u)$.

In Figure 3, we plot the investment policy $i(k,u)$ and Tobin's Q, $q(k,u)$, as functions of normalized promised utility for three levels of capital in the top and the bottom panels, respectively. In the top panel of the figure, the dotted line represents the first-best level of investment-to-capital ratio, $i^* = \delta + \mu_Z$, and $\hat{u}(k)$ represents the boundary of normalized utility at which the entrepreneur starts receive cash payment. We plot policy functions for three levels of normalized capital, where $k_1 < k_2 < k_3$, $k_1 < k^* < k_3$, and k_2 is close to k^*. Because the marginal product of capital, $\alpha k^{\alpha-1}$, is decreasing in k, investment rate and Tobin's Q both decrease in k. Note that k_1 is less than the efficient level of capital and k_3 is higher than k^*, and, as a result, $i(k_1,u)$ stays above i^* and $i(k_3,u)$ stays below as u increases toward $\hat{u}(k_1)$ and $u(k_3)$, respectively. Clearly, as shown in the Figure 3, in most of their domain, $i(k,u)$ is an increasing function of u and $q(k,u)$ is a decreasing function of u. Quantitatively, these patterns of our model account for the small coefficient on q and large coefficient on cash flow in investment-cash-flow sensitivity regressions.

C The role of firm size and age

Our model not only generates a positive investment-cash-flow sensitivity but also, and more importantly, the magnitude of investment-cash-flow sensitivity is inversely related to firm size and age, as in the data. To explain this implication of our model, in this section, we define investment-cash-flow sensitivity as the slope coefficient of a unit-variate regression of investment on cash flow and analyze its relationship with firm size and age in the simple model without adjustment cost.

The advantage of the simple model is that the absence of adjustment cost allows us to provide analytical results. As we will explain in Section C, in the model without adjustment cost, Tobin's Q perfectly predicts investment (locally) in conditional investment regressions. We quantitatively address this issue in Section V, where we show that relationship between investment-cash-flow sensitivity and firm size established in this section remains true even after controlling for Tobin's Q in our full model with adjustment cost.

In the model without adjustment cost, the rate of investment can be calculated from percentage changes in capital stock:

$$\frac{dK_t}{K_t} = \left[\mu_Z + \frac{k'(u_t)(\beta + \kappa - \mu_Z)u_t + \frac{1}{2}k''(u_t)\sigma_u^2 k(u_t)^2}{k(u_t)}\right]dt + k'(u_t)\sigma dB_t. \tag{16}$$

Therefore, the coefficient of an univariate regression of investment on cash flow ($d\hat{y}_t = \sigma dB_t$) is given by

Proposition 3 *In the model without adjustment cost, under the optimal contract, the instantaneous investment to cash flow sensitivity is*

$$\frac{Cov\left(\frac{dK_t}{K_t}, \frac{dY_t}{K_t}\right)}{Var\left(\frac{dY_t}{dK_t}\right)} = k'(u_t). \tag{17}$$

The above result is intuitive. Financing in our model is used as a tool for incentive provision. Positive cash flow shocks is rewarded by higher continuation utility and a relaxed financing constraint. Because the incentive compatibility constraint implies a unit elasticity of continuation utility with respect to cash flow shocks, investment-cash-flow sensitivity is proportional to the sensitivity of investment with respect to continuation utility.

By using the result of Proposition 3. We can now illustrate the monotonic relationship between investment-cash-flow sensitivity and firm size and age in our model.

First, under the optimal contract, investment-cash-flow sensitivity is an increasing function of u. To see this, differentiate both sides of equation (11), we can obtain the following expression for the investment-cash-flow sensitivity:

$$k'(u) = \frac{\sigma^2 k(u)\,\bar{v}'''(u)}{(1-\alpha)\,\alpha k^{\alpha-2}(u) - \sigma^2 \bar{v}''(u)}.$$

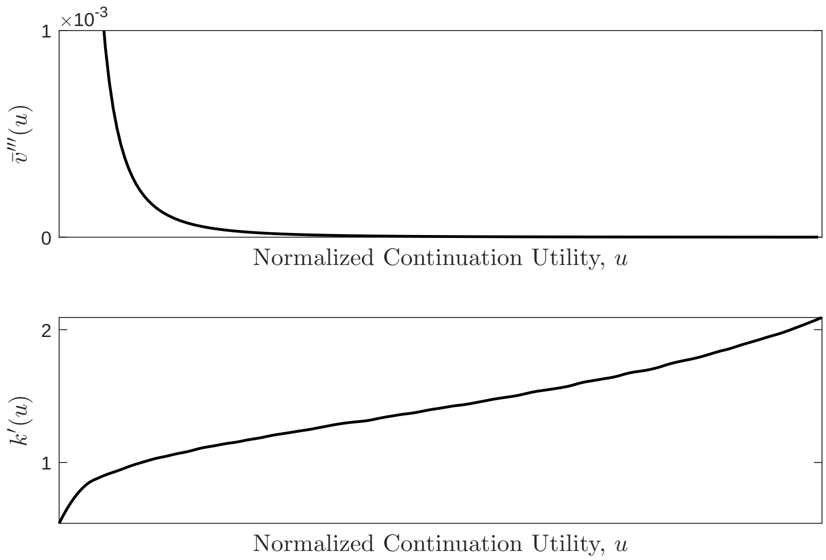
As we show in Figure 4, the third-order derivative, $v'''(u)$, is positive and therefore investment-cash-flow sensitivity is represented by $k'(u) > 0$. Intuitively, as promised utility, u, increases, financing constraint relaxes and the probability of inefficient liquidation diminishes. As a result, the concavity of value function gradually reduces and eventually disappears at the payoff boundary, $\hat{u}$. Furthermore, we show in Figure 4, the magnitude of investment-cash-flow volatility increases with u. As we move away from the inefficient liquidation boundary, $u = 0$, the marginal agency cost of investment diminishes, and it allows the investor to use a higher investment-cash-flow sensitivity to provide incentives. As a result, as the probability of inefficient liquidation diminishes, it is optimal not only to relax the entrepreneur's financing constraint but also to increase financing at a faster rate.

Second, as we explained in Figure 1 and Section III, in equilibrium promised utility, u, and firm size and age are positively correlated because both u and k are expected to increases with age under the optimal contract. As a result, investment-cash-flow sensitivity is positively correlated with size and age and negatively correlated with the severity of financing constraints. In the next section, we calibrate our model to evaluate the above implications of our model quantitatively.

V Quantitative results

In this section, we formally evaluate the ability of our model to account quantitatively for the pattern of investment-cash-flow sensitivity in the data. Consistent with the

Figure 4: Investment-cash-flow sensitivity.



The top panel of 4 plots the third order derivative of normalized value function. The bottom panel plots investment-cash-flow sensitivity, that is, the first order derivative of the policy function $k(u)$.

previous literature, we use the data on manufacturing firms in the Compustat data set for the period of 1967 to 2015 to evaluate our model. The rest of this section is organized as follows. We first describe our calibration procedure, and then demonstrate that our model is consistent with the stylized empirical patterns of firm investment and Tobin's Q. Finally, we present investment-cash-flow sensitivity regressions in our model and compare them with the data.

A Model calibration

Our model parameters can be divided into two groups. The first group can be chosen independently of the other parameters of the model to target data moments. We choose the interest rate $r = 4\%$ to match the average return of risky and risk-free asset in the U.S. post-war period. We set the depreciate rate of capital to $\delta = 10\%$ per annum, consistent with the real business-cycle literature.[3] We set the return to scale parameter, α, to 0.8 to match the average cash-flow-to-capital ratio in the data and choose $\kappa = 5\%$ to match the average death rate of all firms in our data set.

[3] For example, see Kydland and Prescott (1982), Rouwenhorst (1995), and King and Rebelo (1999).

The remaining five parameters, listed Table 1, including the mean and the standard deviation of productivity growth, μ_Z and σ_Z, the volatility of cash flow shocks, σ, the entrepreneur discount rate, β, and the adjustment cost parameter, ϕ, are estimated through simulated method of moments. We target nine moments and list them in Table 2. Although all moments are jointly determined by parameter values, the ones we include in estimation are particularly informative about these parameters. In our model, the growth rate of unconstrained firms is determined by the productivity growth, μ_Z, while the same moment of constrained firms depends on μ_Z, as well as parameters that affect the magnitude of agency frictions, such as the entrepreneur discount rate β and the volatility of the cash-flow shock, σ. Therefore, we include the growth rate, investment-to-capital ratio for all firms as well as for mature firms in our estimation.[4] Next, Tobin's Q is informative about both the adjustment cost parameter ϕ and the magnitude of agency frictions, and we include this moment for all firms, as well as for mature firms in our estimation. Finally, volatility of cash flow is informative about σ, and the volatility of growth rate and investment rate are informative about both σ and σ_Z. Overall, our model matches these moments fairly well.

B Model fit

In this section, we show that our calibrated model is consistent with several stylized features of the data on firm investment and Tobin's Q, although our model is not calibrated to target these moments. First, it is well known that firm investment rate decreases with size and age; see, for example, Evans (1987) and Hall (1987). As shown in Table 3, our model is largely consistent with this feature of the data. Large and older firms tend to invest at a lower rate, both in the model and the data.

Second, Tobin's Q is also monotonically decreasing with firm size and age. We compare the model's implication on Tobin's Q with the data in Table 4. The monotonic pattern of Tobin's Q with respect to firm size and age implied by our model is largely consistent with the data.

Finally, our model is also able to replicate a power law in firm size distribution. As documented by Axtell (2001) and Luttmer (2007), the distribution of firm size is well approximated by a Pareto distribution with a tail slope close to 1.05. Our model produces a power law distribution of firm size with a slope of 1.04, which is close to its

[4] Mature firms are defined as all firms older than the median firm age. In the data, firm age is measured by founding year age, the construction of which is provided in Appendix E.

empirical counterpart.

C Investment-cash-flow sensitivity regressions

In this section, we conduct the standard investment-cash-flow sensitivity regressions from the model simulation and compare them with the data. We show that our model can quantitatively replicate the patterns of investment-cash-flow sensitivity regressions in the data and can account for the robust positive empirical relationship between investment-cash-flow sensitivity and firm size and age.

Following Fazzari, Hubbard, and Petersen (1988), we estimate investment-cash-flow sensitivity as follows:

$$\frac{I_{it}}{K_{it}} = \alpha_i + \alpha_t + \beta_Q Q_{it-1} + \beta_{CF} \frac{CF_{it}}{K_{it}} + \varepsilon_{it}. \tag{18}$$

Here, β_Q measures investment-Q sensitivity; β_{CF} measures the investment-cash-flow sensitivity; and α_i and α_t denote the firm and year fixed effects, respectively. The standard errors are heteroskedasticity consistent and clustered at the firm level. In our model, we aggregate continuous-time quantities to the annual level so that model output is directly comparable with that in the data. We report our regression results in Table 5. In the empirical literature, various measures of investment have been used. For robustness, we replicate the investment-cash-flow sensitivity regressions by using five different measures of investment that appear in the prior literature. Appendix E provides the definitions and constructions.

We make several observations from Table 5. First, in the data, even after controlling for Tobin's Q, cash flow carries significant explanatory power for investment. The investment-cash-flow sensitivity coefficients are significantly positive in all five specifications. The coefficient estimates range from 0.026 to 0.054; and are all significant at the 99% confidence level, consistent with the original study of Fazzari, Hubbard, and Petersen (1988). Second, our model replicates the pattern of these regressions in the data. Because our simulated model is immune to measurement errors and because investment in the data potentially can be affected by other unmodeled shocks, the estimates of the cash-flow sensitivity coefficient in our model are significantly larger than their data counterparts. However, the pattern of monotonicity is clear, both in the model and in the data.

More importantly, our model not only replicates the positive regression coefficient on cash flow but it also predicts that the magnitude of investment-cash-flow sensitivity increases with firm size and age. We compare this implication of our model with the data in Table 6, where we divide the samples into five groups according to their size and age with equal number of firms in each group. We compute the regression (18) results for each group and report the coefficient β_{CF} in the data and in the model in the above table.

As we can see from the table, the magnitude of the sensitivity monotonically increases with firm size (Panel A) and firm age (Panel B), both in the data and the model. Our empirical results are robust across all five definitions of investment and are consistent with the prior literature as shown in Table 7; see, for example, Vogt (1994) and Kadapakkam, Kumar, and Riddick (1998). Our model reproduces this pattern of investment-cash-flow sensitivity. As we explained Section IV, investment-cash-flow sensitivity is used as a tool for incentive provision in our model. Small and young firms are more constrained and closer to the liquidation boundary. As a result, they face strong agency frictions and are extremely risk averse. Optimality implies that these firms should not use higher powered incentives, and therefore their investment-cash-flow sensitivity is low compared with mature and large firms.

VI Conclusion

We present a dynamic model with neoclassical investment technology and with moral hazard to rationalize the stylized facts on investment-cash-flow sensitivity. In our model, optimal contracting under moral hazard problem induces an endogenous financing constraint, and investment-cash-flow sensitivity is used as an incentive provision device. Because cash flows carry information about entrepreneurs' (hidden) action that are not contained in Tobin's Q, firm investment optimally responds to cash flows even after controlling for Tobin's Q. In addition, our model provides a quantitative explanation of the positive relationship between investment-cash-flow sensitivity and firm size and age. In our model, large and mature firms are less financially constrained. Because the marginal agency cost of investment is lower in these firms, firms optimally use higher investment-cash-flow sensitivity to provide stronger incentives. In contrast, small and young firms cannot use higher powered incentives due to the higher moral-hazard-induced risk aversion.

Tables

Table 1: Estimated Parameters

Parameters	Descriptions	Values
μ_Z	Mean productivity growth rate	0.050
σ_Z	Volatility of productivity growth rate	0.050
β	Entrepreneur's discount rate	0.090
σ	Volatility of unobservable cash flow shock	0.320
ϕ_0	Adjustment cost parameter	0.02

Notes - This table lists the group of calibrated parameters that we estimated through simulated method of moments to minimize the Euclidean distance between the eleven targeted moments listed in Table 2 from the model and the data.

Table 2: Targeted Moments

Targeted Moments	Data	Model
Median growth rate		
all firms	0.088	0.084
mature firms	0.076	0.074
Median investment to capital ratio		
all firms	0.205	0.194
mature firms	0.186	0.173
Median Tobin's Q		
all firms	2.700	2.590
mature firms	2.225	2.170
Volatility of growth rate	0.244	0.255
Volatility of investment to capital ratio	0.286	0.277
Volatility of cash flow to capital ratio	0.447	0.421

Notes - This table lists the nine moments that we target through the simulated method of moments in order to pin down the calibrated parameters in Table 1. The empirical data moments are based on Compustat manufacturing firms annual database from 1967 to 2015. Mature firms include the firms with founding year age larger than the median age in the sample. The data construction is detailed in Appendix E.

Table 3: Investment to capital ratio by size and age

	Size Group		Age Group	
Group No.	Data	Model	Data	Model
1	0.214	0.201	0.209	0.197
2	0.198	0.220	0.206	0.192
3	0.188	0.184	0.194	0.183
4	0.185	0.181	0.187	0.173
5	0.179	0.141	0.183	0.169

Notes - This table presents the median value of investment to capital ratio ($\frac{I}{K}$) by size and founding year age groups, both in the data and model simulation. We follow the standard sorting procedure in the data by assigning firms into portfolios using breakpoints based on the NYSE traded firms. In the model, firms are sorted using breakpoints that are equally-numbered in sorting variables. Portfolios are re-balanced at the annual frequency. The details on sample and variable constrictions are described in Appendix E.

Table 4: Tobin's Q by size and age

	Size Group		Age Group	
Group No.	Data	Model	Data	Model
1	3.03	3.77	2.89	3.72
2	3.00	3.40	2.78	3.04
3	2.24	2.48	2.36	2.30
4	2.00	1.57	2.24	1.91
5	1.97	1.25	2.20	1.83

Notes - This table presents the median Tobin's Q by size and founding year age groups, both in the data and model simulation. We follow the standard sorting procedure by assigning firms into five portfolios with equal number of firms, and rebalance at the annual frequency. The details on sample and variable constrictions are described in Appendix E.

Table 5: Investment-cash flow sensitivity regressions

	Model	Data				
	$\frac{I}{K}$	$\frac{I_1}{K}$	$\frac{I_2}{K}$	$\frac{I_3}{K}$	$\frac{I_4}{K}$	$\frac{I_5}{K}$
β_Q	0.006	0.007***	0.007***	0.008***	0.010***	0.012***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
β_{CF}	0.661	0.026***	0.027***	0.039***	0.043***	0.054***
		(0.002)	(0.002)	(0.003)	(0.004)	(0.006)
Obs.		63517	48019	63516	63516	63517
$adj.R^2$		0.158	0.142	0.112	0.102	0.058

Notes - This table presents the coefficient estimates of the regression $\frac{I_{it}}{K_{it}} = \alpha_i + \alpha_t + \beta_Q Q_{it-1} + \beta_{CF} \frac{CF_{it}}{K_{it}} + \varepsilon_{it}$ from the model simulation and the data. The empirical sample is based on Compustat annual database from 1967 to 2015. Standard errors reported in the parentheses are heteroskedasticity consistent and clustered as the firm level. We denote p-values smaller than 1%, 5%, and 10% by ***, ** and *, respectively. For the robustness check, in the data, we use 5 definitions of investment used in the literature, I_1 to I_5, respectively, which are described in Appendix E.

Table 6: Investment-cash flow sensitivity regressions by size groups

Group	1	2	3	4	5
	β_{CF}	β_{CF}	β_{CF}	β_{CF}	β_{CF}
Panel A: Size Groups					
Model	0.368	0.548	0.622	0.673	0.660
Data	0.030***	0.064***	0.053***	0.094***	0.077***
	(0.003)	(0.012)	(0.012)	(0.015)	(0.021)
Panel B: Age Groups					
Model	0.473	0.574	0.607	0.618	0.619
Data	0.022***	0.027***	0.062***	0.050***	0.064***
	(0.003)	(0.005)	(0.011)	(0.011)	(0.019)

Notes - This table presents the coefficient estimates of the investment-cash flow sensitivity, β_{CF}, by size groups (Panel A) and founding year age groups (Panel B) from the model simulation and the data. The empirical sample is based on Compustat annual database from 1967 to 2015. We follow the standard sorting procedure by assigning firms into five portfolios with equal number of firms, and rebalance at the annual frequency. We then run investment-cash flow sensitivity regressions $\frac{I_{it}}{K_{it}} = \alpha_i + \alpha_t + \beta_Q Q_{it-1} + \beta_{CF}\frac{CF_{it}}{K_{it}} + \varepsilon_{it}$ within each size group. Standard errors reported in the parentheses are heteroskedasticity consistent and clustered as the firm level. We denote p-values smaller than 1%, 5%, and 10% by ***, ** and *, respectively.

Table 7: Investment-cash flow sensitivity regressions by size groups, robustness check

Size Group	1	2	3	4	5
	β_{CF}	β_{CF}	β_{CF}	β_{CF}	β_{CF}
Panel B: Size Groups					
$\frac{I_2}{K}$	0.032***	0.068***	0.058***	0.138***	0.086***
	(0.003)	(0.013)	(0.016)	(0.020)	(0.030)
$\frac{I_3}{K}$	0.043***	0.101***	0.095***	0.166***	0.163***
	(0.003)	(0.018)	(0.022)	(0.032)	(0.028)
$\frac{I_4}{K}$	0.049***	0.108***	0.109***	0.198***	0.201***
	(0.005)	(0.022)	(0.024)	(0.040)	(0.038)
$\frac{I_5}{K}$	0.052***	0.135***	0.065	0.180**	0.197***
	(0.008)	(0.035)	(0.064)	(0.085)	(0.032)
Panel A: Age Groups					
$\frac{I_2}{K}$	0.021***	0.029***	0.072***	0.059***	0.092***
	(0.004)	(0.005)	(0.013)	(0.012)	(0.029)
$\frac{I_3}{K}$	0.029***	0.036***	0.111***	0.120***	0.106***
	(0.004)	(0.006)	(0.017)	(0.031)	(0.033)
$\frac{I_4}{K}$	0.030***	0.038***	0.142***	0.157***	0.200***
	(0.007)	(0.010)	(0.022)	(0.039)	(0.024)
$\frac{I_5}{K}$	0.029**	0.036**	0.171***	0.239***	0.198***
	(0.012)	(0.017)	(0.044)	(0.073)	(0.036)

Notes - This table complements Table 6 and shows that in the data the monotonically increasing pattern of the investment-cash flow sensitivity, β_{CF}, with respect to firm size (Panel A) and firm founding year age (Panel B) are robust for another four alternative measures of investment, I_2 to I_5, respectively, which are described in Appendix E. We follow the standard sorting procedure by assigning firms into five portfolios with equal number of firms, and rebalance at the annual frequency. We then run investment-cash flow sensitivity regressions $\frac{I_{it}}{K_{it}} = \alpha_i + \alpha_t + \beta_Q Q_{it-1} + \beta_{CF} \frac{CF_{it}}{K_{it}} + \varepsilon_{it}$ within each size group. Standard errors reported in the parentheses are heteroskedasticity consistent and clustered as the firm level. We denote p-values smaller than 1%, 5%, and 10% by ***, ** and *, respectively.

References

Abel, A., and J. Eberly, 2011, How Q and cash flow affect investment without frictions: An analytic explanation, *Review of Economic Studies* 78, 1179–1200.

Almeida, H., and M. Campello, 2007, Financial constraints, asset tangibility, and corporate investment, *Review of Financial Studies* 20, 1429–1460.

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

Axtell, Robert L., 2001, Zipf Distribution of U.S. Firm Sizes, *Science* 293, 1818–1820.

Biais, Bruno, Thomas Mariotti, Cuillaume Plantin, and Jean-Charles Rochet, 2007, Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications, *The Review of Economic Studies* 74, 345–390.

Biais, Bruno, Thomas Mariotti, and Jean-Charles Rochet Stéphane Villeneuve, 2010, Large Risks, Limited Liability, and Dynamic Moral Hazard, *Econometrica* 78, 73–118.

Blundell, R., S. Bond, M. Devereux, and F. Schiantarelli, 1992, Investment and Tobin's Q: Evidence from company panel data, *Journal of Econometrics* 51, 233–257.

Cleary, S., 1999, The relationship between firm investment and financial status, *Journal of Finance* 54, 673–692.

DeMarzo, Peter, Michael Fishman, Zhiguo He, and Neng Wang, 2012, Dynamic Agency and q Theory of Investment, *Journal of Finance* 67, 2295–2340.

DeMarzo, Peter, and Yuliy Sannikov, 2006, Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model, *Journal of Finance* 61, 2681–2724.

Edmans, Alex, and Xavier Gabaix, 2016, Executive Compensation: A Modern Primer, *Journal of Economic Literature* forthcoming.

Edmans, A., X. Gabaix, and A. Landier, 2008, A Calibratable Model of Optimal CEO Incentives in Market Equilibrium, *NBER Working Paper No. 13372*.

Eisfeldt, A., and A. Rampini, 2006, Capital reallocation and liquidity, *Journal of Monetary Economics* 53, 369–399.

Erickson, T., and T. Whited, 2000, Measurement error and the relationship between investment and q, *Journal of Political Economy* 108, 1027–1057.

Evans, David, 1987, Tests of Alternative Theories of Firm Growth, *Journal of Political Economy* 95, 657–674.

Farre-Mensa, J., and A. Ljungqvist, 2015, Do Measures of Financial Constraints Measure Financial Constraints?, *Review of Financial Studies* 29, 271–308.

Fazzari, Steven, Glenn Hubbard, and Bruce Petersen, 1988, Financing Constraints and Corporate Investment, *Brookings Papers on Economic Activity* 1988, 141–206.

Gilchrist, S., and C. Himmelberg, 1995, Evidence on the role of cash flow for investment, *Journal of Monetary Economics* 36, 541–572.

Gomes, J., 2001, Financing investment, *American Economic Review* 91, 1263–1285.

Hadlock, C., and J. Pierce, 2010, New evidence on measuring financial constraints: Moving beyond the KZ Index, *Review of Financial Studies* 23, 1909–1940.

Hall, Bronwyn, 1987, The reslationship between firm size and firm growth in the US manufacturing sector, *The Journal of Industrial Economics* 35, 583–606.

He, Zhiguo, 2009, Optimal Executive Compensation when Firm Size Follows Geometric Brownian Motion, *Review of Financial Studies* 22, 859–892.

Hennessy, C., A. Levy, and T. Whited, 2007, Testing Q theory with financing frictions, *Journal of Financial Economics* 83, 691–717.

Hoberg, G., and V. Maksimovic, 2015, Redefining financial constraints: a text-based analysis, *Review of Financial Studies* 28, 1312–1352.

Hoshi, T., A. Kashyap, and D. Scharfstein, 1991, Corporate Structure, Liquidity, and Investment: Evidence from Japanese Industrial Groups, *The Quarterly Journal of Economics* 106, 33–60.

Hubbard, R., 1998, Capital market imperfections and investment, *Journal of Economic Literature* 36, 193–227.

Jovanovic, B., and P. Rousseau, 2001, Why Wait? A Century of Life before IPO, *American Economic Review* 91, 336–341.

Kadapakkam, P., P. Kumar, and L. Riddick, 1998, The impact of cash flows and firm size on investment: The international evidence, *Journal of Banking and Finance* 22, 293–320.

Kaplan, S., and L. Zingales, 1997, Do investment-cash flow sensitivities provides useful measures of financing constraints?, *Quarterly Journal of Economics* 112, 169–215.

King, Robert G., and Sergio T. Rebelo, 1999, Resuscitating real business cycles, in *Handbook of Macroeconomics* (Elsevier,).

Kydland, Finn, and Edward Prescott, 1982, Time to Build and Aggregate Fluctuations, *Econometrica* 50, 1345–1370.

Lamont, O., C. Polk, and J. Saaá-Requejo, 2001, Financial Constraints and Stock Returns, *Review of Financial Studies* 14, 529–554.

Lorenzoni, G., and K. Walentin, 2007, Financial Frictions, Investment and Tobin's q, *NBER Working Paper No. 13092*.

Luttmer, Erzo, 2007, Selection, Growth, and the Size Distribution of Firms, *Quarterly Journal of Economics* 122, 1103–1144.

Moyen, N, 2004, Investment-Cash flow sensitivities: constrained versus unconstrained firms, *Journal of Finance* 59, 2061–2092.

Ritter, J., and T. Loughran, 2004, Why has IPO underpricing changed over time?, *Financial Management* 33, 5–37.

Rouwenhorst, K. Geert, 1995, Asset pricing implications of equilibrium business cycle models, in *Frontiers of Business Cycle Research* (Princeton University Press,).

Vogt, S., 1994, The cash flow/investment relationship: evidence from US manufacturing Firm, *Financial Management* 23, 3–20.

Whited, T., and G. Wu, 2006, Financial constraints risk, *Review of Financial Studies* 19, 531–559.

Zhu, John, 2013, Optimal Contracts with Shirking, *Review of Economic Studies* 80, 812–839.

Appendix

A Proof of Lemma 1

The Martingale representation theorem implies

$$dU_t = (\beta + \kappa) U_t dt - dC_t + G_t \sigma dB_t + H_t \sigma_Z dB_{Z,t} \text{ for all } t \in [0, \tau \wedge T] \tag{19}$$

with $\{G_t\}$ and $\{H_t\}$ being two predictable and square integrable processes. We the following incentive compatibility condition, which is equivalent to (6).

$$G_t \geq K_t \text{ for all } t \in [0, \tau \wedge T]. \tag{20}$$

Under the contract, given any diversion policy of the entrepreneur, $\{D_t\}$, define $\left\{B_t^D\right\}$ such that

$$dB_{u,t}^D = \frac{dY_t - dB_{o,t} - \left(Z_t^{1-\alpha} K_t^\alpha - D_t\right) dt}{\sigma K_t} \text{ for all } t \in [\tau \wedge T].$$

Given the realizations of $\{Y_t\}$ and $\{K_t\}$, we have

$$\sigma dB_t = \sigma dB_t^D + \frac{D_t}{K_t} dt. \tag{21}$$

For any $t < \tau \wedge T$, we define

$$\mathcal{G}_t^D = \int_0^t e^{-(\beta+\kappa)s} \left[D_s + dC_s\right] ds + e^{-(\beta+\kappa)t} U_t,$$

which is the time-t conditional expectation of the entrepreneur's utility if he diverts cash flows according to $\{D_t\}$ and stops diverting at time t. Equation (19) and (21) imply

$$e^{-(\beta+\kappa)t} d\mathcal{G}_t^D = D_t \left(1 - \frac{G_t}{K_t}\right) dt + G_t \sigma dB_t^D.$$

Under the diversion policy $\{D_t\}$, the second term on the right hand side is a martingale. Therefore $\left\{\mathcal{G}_t^D\right\}$ is a super martingale if and only if (20) is satisfied. Therefore, if and only if (20) is satisfied, it is optimal to not divert from the beginning of the contract and we have the desired result.

B Proof of Proposition 2

Equations (12) and (13) imply

$$d\hat{\imath}\left(k_t,u_t\right)=\left[i_u\left(k_t,u_t\right)h\left(k_t,u_t\right)-i_k\left(k_t,u_t\right)\right]k_t\sigma_Z dB_{Z,t}+i_u\left(k_t,u_t\right)k_t d\hat{y}_t$$

and

$$d\hat{q}\left(k_t,u_t\right)=\left[q_u\left(k_t,u_t\right)h\left(k_t,u_t\right)-q_k\left(k_t,u_t\right)\right]k_t\sigma_Z dB_{Z,t}+q_u\left(k_t,u_t\right)k_t d\hat{y}_t.$$

By solving $d\hat{\imath}$ as a linear function of $d\hat{q}$ and $d\hat{y}_t$, we have the desired result.

C Optimal contracting without adjustment cost

In this section, we discuss the value function and the optimal contract without adjustment cost in investment. The characterization of the optimal contract is extended from that presented DeMarzo and Sannikov (2006) and He (2009). The laws of motion of u_t, (5), implies that $\bar{v}(u)$ satisfies the following HJB differential equation.

$$\begin{aligned}0=\max_{k,g\geq k,h,dc}\quad & k^\alpha-(r+\delta+\kappa)k-(r+\kappa-\mu_Z)\bar{v}(u)+(\beta+\kappa-\mu_Z)u\bar{v}'(u)\\ & +\frac{1}{2}\bar{v}''(u)\left(g^2\sigma^2+(hk-u)^2\sigma_Z^2\right)-(1+\bar{v}'(u))\,dc \end{aligned}\tag{22}$$

Since the investor can always payoff part of the payments that promised to the entrepreneur immediately, we have $\bar{v}(u)\geq\bar{v}(u-dc)-dc$ for all $dc>0$ and thus $\bar{v}'(u)\geq-1$. Moreover, the optimality on the right hand side of (22) implies that $dc>0$ only if $\bar{v}'(u)=-1$. So, concavity of v, which is shown in Lemma 2, implies that there exists a level $\hat{u}>0$ such that $\bar{v}'(\hat{u})=-1$, $\bar{v}'(u)>-1$ for $u\in[0,\hat{u})$. Therefore $dc=0$ if $u\in[0,\hat{u}]$ and $dc=u-\hat{u}>0$ otherwise. As a result, we have (10).[5]

The smooth pasting condition implies $\bar{v}''(\hat{u})=0$ and, according to (10), we have

$$\bar{v}(\hat{u})=\frac{\pi^*}{r+\kappa-\mu_Z}-\frac{\beta+\kappa-\mu_Z}{r+\kappa-\mu_Z}\hat{u}.\tag{23}$$

Here π^* is the level of the normalized operating profit implied by the first-best level of

[5] Notice that, the incentive compatibility condition implies $g=k$ and the optimality condition implies that $hk-u=0$ for all u.

working capital k^* discussed in Appendix D. On the other boundary, we have $\bar{v}(0) = 0$. Now we have the following characterization of v and the optimal contract.

Proposition 4 *The optimal lending contract delivering the entrepreneur initial expected utility $U_0 = Z_0 u_0$ takes the following form: for $t \leq \tau \wedge T$, u_t evolves according to* (5) *when $u_t \in [0, \hat{u})$, $dc_t = 0$, $k_t = k(u_t)$, which is the maximizer of the right hand side of* (10)*, and $h(u_t) = \frac{u_t}{k(u_t)}$; when $u_t > \hat{u}$, $dc_t = u_t - \hat{u}$ reflecting u_t back to $\hat{u}$. The firm is liquidated at time T, the stopping time when u_t hits zero. The normalized value function, $\bar{v}$, satisfies* (10) *over $[0, \hat{u}]$ with boundary conditions, $\bar{v}(0) = 0$, $\bar{v}'(\hat{u}) = -1$, $v''(\hat{u}) = 0$, and* (23)*. For $u \geq \hat{u}$, $\bar{v}(u) = \bar{v}(\hat{u}) - (u - \hat{u})$. $\bar{v}$ is strictly concave over $[0, \hat{u}]$.*

The proof is divided into three lemmas. In the first one we show the concavity of the normalized value function.

Lemma 2 *The normalized value function v satisfying HJB* (7) *and the boundary conditions $v'(\hat{u}) = -1$,* (23)*, and $v''(\hat{u}) = 0$ is concave over $[0, \hat{u}]$.*

Proof. According to the Envelope theorem, by taking derivative with respect to u on both hand side of (10), we have

$$(\beta - r)\,\bar{v}'(u) + (\beta + \kappa - \mu_Z)\,u\bar{v}''(u) + \frac{1}{2}\bar{v}'''k(u)^2\sigma^2 = 0. \tag{24}$$

On the boundary $\hat{u}$, $\bar{v}'(\hat{u}) = -1$ and $\bar{v}''(\hat{u}) = 0$. Therefore, (24) implies

$$\bar{v}'''(\hat{u}) = \frac{2(\beta - r)}{k(\hat{u})^2\sigma^2} > 0$$

and, for a sufficient small real number $\epsilon > 0$, $\bar{v}''(\hat{u} - \epsilon) < 0$. Now, suppose that $\bar{v}$ is not concave and $\tilde{u}$ is the largest real number such that $\bar{v}''(u) = 0$. Then (10) implies that $k(\tilde{u}) = k^*$, the first-best level, and then

$$\bar{v}(\tilde{u}) = \frac{\pi^*}{\beta + \kappa - \mu_Z} + \frac{\beta + \kappa - \mu}{r + \kappa - \mu_Z}\tilde{u}\bar{v}'(\tilde{u}). \tag{25}$$

Notice that, for all $u \in (\tilde{u}, \hat{u})$, $\bar{v}''(u) < 0$ and $\bar{v}'(u) > -1$. Therefore, (25) is contradict to the boundary condition (23) and we have the desired result. ■

The next two lemmas verify the optimality of the normalized value function $\bar{v}$ and the contract.

Lemma 3 *Under the contract described in Proposition 4, the normalized firm value is characterized by* $\bar{v}$*.*

Proof. It is easy to check that the normalized continuation utility follows (5) under the described contract. Now suppose that $U_0 = u_0 \in [0, \hat{u})$. We define

$$\Psi_t = \int_0^t e^{-(r+\kappa)s} \left[Z_s^\alpha K_s^{1-\alpha} - (r+\delta+\kappa) K_s - dC_s \right] + e^{-(r+\kappa)t} Z_t v\left(u_t\right) \text{ for } t \in [0, \tau \wedge T].$$

In fact, Ψ_t is the time-t conditional expectation of the total expected payoff given that the contract is implemented from time zero to t and then the net present value of the future payoff from time t on is summarized by $Z_t \bar{v}(u_t)$. Obviously, $\Psi_0 = Z_0 \bar{v}(u_0) = \bar{v}(u_0)$. According to Ito's lemma.

$$e^{-(r+\kappa)t} d\Psi_t = Z_t \begin{bmatrix} k_t^\alpha - (r+\delta+\kappa) k_t - (r+\kappa-\mu_Z) \bar{v}(u_t) + (\beta+\kappa-\mu_Z) u_t \bar{v}'(u_t) \\ +\frac{1}{2} \bar{v}''(u_t) \left(g_t^2 \sigma^2 + (h_t k_t - u_t)^2 \sigma_Z^2 \right) - (1 + \bar{v}'(u_t)) dc_t \\ +\bar{v}'(u_t) k_t \sigma dB_t + (\bar{v}(u_t) - u_t \bar{v}'(u_t)) \sigma_Z dB_{Z,t} \end{bmatrix}. \tag{26}$$

Under the contract, $dc_t \neq 0$ if and only if $v'(u_t) = -1$. Moreover, k_t and h_t maximize the right hand side of (22). So (26) is simplified to

$$e^{-(r+\kappa)t} d\Psi_t = Z_t \left[\bar{v}'(u_t) k_t \sigma dB_t + (\bar{v}(u_t) - u_t \bar{v}'(u_t)) \sigma_Z dB_{Z,t} \right]$$

and $\{\Psi_t\}$ is a martingale. Therefore, the expected payoff under the contract is $E_0[\Psi_\infty] = \Psi_0 = v(u_0)$ and we have the desired result. ■

Lemma 4 *Any incentive compatible contract promising the entrepreneur expected utility* $U_0 = Z_0 u_0 = u_0$ *cannot generate an expected payoff larger than* $Z_0 \bar{v}(u_0) = \bar{v}(u_0)$*.*

Proof. Obviously, the only way to deliver a zero expected utility to the entrepreneur is by liquidating the firm. So, we assume that the firm is liquidated when the normalized continuation utility of the entrepreneur hits zero. Denote the stopping time by T_0 which could be infinite.[6] Let $\left(\left\{ \hat{C}_t \right\}, \left\{ \hat{K}_t \right\}, \hat{T} \right)$ be an alternative incentive compatible contract promising the entrepreneur initial expected utility u_0 and we use hatted letters denote corresponding terms under this contract. For any $t \in \left[0, \tau \wedge \hat{T} \wedge T_0\right]$, define

$$\hat{\Psi}_t = \int_0^t e^{-(r+\kappa)s} \left[\left(Z_s^\alpha \hat{K}_s^{1-\alpha} - (r+\delta+\kappa) \hat{K}_s \right) ds - d\hat{C}_s \right] + e^{-(r+\kappa)t} Z_t \bar{v}(\hat{u}_t).$$

[6] Under the described contract, $T = T_0$.

In fact, $\hat{\Psi}_t$ is the time-t conditional expectation of the investor's total payoff given that she implements the alternative contract from time zero to t and then switches to the described contract with the continuation utility of the entrepreneur being kept. Notice that $\hat{\Psi}_0 = Z_0 \bar{v}(u_0) = \bar{v}(u_0)$. Then

$$e^{-(r+\kappa)t} d\hat{\Psi}_t = Z_t \begin{bmatrix} \hat{k}_t^{\alpha} - (r+\delta+\kappa)\hat{k}_t - (r+\kappa-\mu_Z)\bar{v}(\hat{u}_t) + (\beta+\kappa-\mu_Z)\hat{u}_t \bar{v}'(\hat{u}_t) \\ +\frac{1}{2}\bar{v}''(\hat{u}_t)\left(\hat{g}_t^2\sigma^2 + \left(\hat{h}_t\hat{k}_t - \hat{u}_t\right)^2\sigma^2\right) - (1+\bar{v}'(\hat{u}_t))\, d\hat{c}_t \\ +\bar{v}'(\hat{u}_t)\hat{g}_t\sigma dB_t + (\bar{v}(\hat{u}_t) - \hat{u}_t\bar{v}'(\hat{u}_t))\sigma_Z dB_{Z,t} \end{bmatrix}.$$

Concavity of $\bar{v}$, the fact that $v' > -1$ and the HJB (22) imply that $\left\{\hat{\Psi}_t\right\}$ is a super martingale. Therefore the expected payoff of the investor under the alternative contract

$$E_0\left[\hat{\Psi}_{\tau\wedge\hat{T}\wedge T_0}\right] \le E_0\left[\hat{\Psi}_0\right] = Z_0\bar{v}(u_0) = \bar{v}(u_0).$$

■

D The first-best case

In this section, we consider the first-best case in which there is no information asymmetry or adjustment cost. Obviously, liquidation never takes place and $K_t = k^* Z_t$ for $t \le \tau$ with

$$k^* = \arg\max_{\tilde{k}} \tilde{k}^{\alpha} - (r+\delta+\kappa)\tilde{k} = \left(\frac{\alpha}{r+\delta+\kappa}\right)^{\frac{1}{1-\alpha}},$$

and the rate of the operating profit is $\pi^* Z_t$ with

$$\pi^* = (1-\alpha)\left(\frac{\alpha}{r+\delta+\kappa}\right)^{\frac{\alpha}{1-\alpha}}.$$

Therefore, the expected NPV of the total cash flows created by the firm is $Z_0\frac{\pi^*}{r+\kappa-\mu_Z} = \frac{\pi^*}{r+\kappa-\mu_Z}$. Since the investor is more patient than the entrepreneur, it is optimal to pay off the promised utility as a lump-sum payment at the beginning of the contract. Specifically, $dC_0 = U_0$ and $dC_t = 0$ for all $t > 0$ and then the investor's expected payoff is

$$Z_0\frac{\pi^*}{r+\kappa-\mu_Z} - U_0.$$

Let $\bar{v}^{FB}(u)$ be the first-best normalized value function. Then

$$\bar{v}^{FB}(u) = \frac{\pi^*}{r+\kappa-\mu_Z} - u \text{ for } u \geq 0.$$

Under the first-best contract, it is easy to see that all the firms grow at the same growth rate μ_Z; they have the same investment to capital ratio, $\mu_Z + \delta$, and the investment to cash flow sensitivity is zero; they have the same Tobin's Q which is

$$\frac{\bar{v}^{FB}(0)}{k^*} = \frac{1-\alpha}{\alpha}\left(\frac{r+\delta+\kappa}{r+\kappa-\mu_Z}\right).$$

E Data appendix

Data sources:

We construct our sample by using the annual Compustat data from 1967 to 2015. We exclude foreign firms (those with a foreign incorporation code). To be consistent with previous literature, our main tests include only manufacturing firms (with the standard industry classification (SIC) code between 2000 and 3999), but we also study a border set of sample including the nonmanufacturing firms (except for financial firms (SIC code 6000-6999) and utility firms (SIC 4000-4999)) as a robustness check.

Variable definitions:

Investment is the capital expenditure, (*capx*, Compustat data item 128). Capital stock is computed as one year lag of net property, plant, and equipment *(ppent, Item 8).* Consistent with the model, we use capital stock as a measure for firm size. Cash flow is the sum of income before extraordinary items *(ib, Item 18)* and depreciation and amortization *(dp, Item 14).* The market value of assets is equal to the book value of asset *(at, Item 120)*, plus market value common stocks *(prcc_f* $\times$ *csho, Item 30* $\times$ *Item 25)*, and minus the book value of common stocks *(ceq, Item 60)* and the deferred taxes *(txdb, Item 74).* And the Tobin's Q is calculated as the market value of assets minus the difference between the book value of asset and the capital, divided by the capital. We calculate firm's age based on the founding years from Ritter and Loughran (2004) and Jovanovic and Rousseau (2001).

We also test the robustness of our cash-flow sensitivity regressions with respect to different definitions of investment. Besides the standard definition, we use the following: (1) Compustat Item 30 (capxv), which includes increases in property, plant, and equipment from acquisitions that use purchase accounting, minus the sale of property *(sppe, Item 107)*; (2) change in net property, plant, and equipment; (3) change in net property, plant and equipment adding back depreciation; (4) the sum of capital expenditures and acquisitions *(acq, Item 129).* For these different variations of investment definitions, the first one is used by Hennessy, Levy, and Whited (2007); definitions (2) and (3) are from Kaplan and Zingales (1997); and the last definition is used by Eisfeldt and Rampini (2006). The standard definition of investment is denoted as I_1, while the remaining four alternative definitions of investment correspond to I_2 to I_5, respectively in Table 6 and Table 7.

Sample screening:

As standard in the literature, we use the following criteria to screen the sample. We require firms to have valid observations for all variables in the investment-cash-flow regression equation. To mitigate outliers, we first drop those observations with either capital, or book assets or sales smaller than 1 million U.S. dollars, and then we delete the outliers with investment, capital, Tobin's Q and cash flow to capital ratio beyond the 1st and 99th percentiles. In the investment-to-cash flow sensitivity regressions, we further conduct the following screenings. To alleviate Computstat's backfill bias, exclude firms for which we cannot compute the lagged cash flow to capital ration, $\frac{CF_{it-1}}{K_{it-1}}$. Following Almeida, Campello and Weisbach [2004], we exclude firms with asset or sales growth exceeding 100% to avoid potential business discontinuities caused by mergers and acquisitions. And we winsorize all the variables needed in the regressions at the 1st and 99th percentiles.

Empirical model:

Following Fazzari, Hubbard, and Petersen (1988), we estimate investment-cash-flow sensitivity as follows:

$$\frac{I_{it}}{K_{it}} = \alpha_i + \alpha_t + \beta_Q Q_{it-1} + \beta_{CF} \frac{CF_{it}}{K_{it}} + \varepsilon_{it}$$

in which β_Q measures investment-Q sensitivity; β_{CF} measures the investment-cash-flow sensitivity. α_i and α_t denote the firm and year fixed effects. The standard errors are heteroskedasticity-consistent and clustered at the firm level.

We run the regressions both for the whole sample and by size and age groups. The regression results are reported in Table 5 and Table 6, respectively.