Financing a Black Box: Dynamic Investment with Persistent Private Information

Felix Zhiyu Feng*

May 2021

Abstract

The increasing sophistication of modern production technology, globalization, and major market disruptions such as the outburst of a pandemic can impose strong challenges for general investors to directly observe and precisely evaluate the productivity, return, and growth potential of their investments. The annals of corporate management in recent years are rife with financial misconduct or value distortion as a result of the lack of transparency inside a business, implying a substantial degree of agency frictions that likely bear persistent, long-term effects. This paper studies the implication of persistent private information on a firm's optimal financing and investment policies. In a dynamic agency model, an investor supplies capital to an entrepreneur with an opaque production technology. The investor observes neither the true productivity of the technology nor the actual amount of the output produced. The entrepreneur can generate private benefit from misreporting productivity and diverting output, both of which bear a persistent negative effect on the long-term growth of the technology. Compared to standard agency-based investment models, the persistence of the agency friction rationalizes over-investment especially among firms with a strong history of cash flow but a low Tobin's q, and reconciles the optimal financing policy with the empirical observations of a strong investment-cash-flow sensitivity and a weak or even negative investment-q sensitivity. When the optimal contract is implemented with equity, bonds, and cash reserves, the model implies a substantial credit yield spread that widens when the persistent effect of the agency friction is stronger.

JEL Classification: G32, D86, D25, L26

Key Words: dynamic agency, persistent private information, investment, entrepreneurship, q-theory.

*Michael G. Foster School of Business, University of Washington, Seattle, WA, 98105, USA. Email: ffeng@uw.edu.
I am indebted to Curtis Taylor and Mark Westerfield for discussions at the early stage of this project. I thank Chaojun Wang, Colin Ward, Cole Williams, the seminar and conference participants at University of Notre Dame, Econometric Society Winter Meetings, DSE Winter School, and the Industrial Organization Society IIOC for helpful comments and suggestions. All errors are my own.

1 Introduction

In early 2020, Luckin Coffee, a NASDAQ-listed coffeehouse chain, announced that it has found internal financial misconduct, following allegations that it had fabricated revenues and exaggerated sales, as well as misreported marketing and operating costs to cover up its lack of earnings. This came as a surprise to general investors, as evidenced by the 80% drop in stock price on the day of the announcement. Meanwhile, GSX Techedu, a NYSE-listed online tutoring service provider, denied a similar allegation that it had fabricated the numbers of users and overstated its profitability, stating that the allegation "contains inaccurate and disorderly data sources, and shows a lack of understanding of GSX's business". The share price of GSX rebounded following the statement.

The incidents of Luckin Coffee, GSX, and others (e.g. Nikola, Theranos, WeWork, to name a few) highlight a problem prevalent in today's financial market for outside investors: the lack of transparency in the technology employed and the operations conducted inside a business. Despite the increasingly convenient access to information in general, gauging the true productivity, return, and growth potential of a business has arguably become more challenging to outside investors for several reasons: first and foremost, the technology adopted by modern businesses has become increasingly sophisticated. Second, globalization means investors may hold a portfolio of businesses with distinct operation models at distant locations, and the cost of direct monitoring and precise valuation of one specific business can be prohibitively high. Finally, uncertainties in the market, such as the COVID-19 pandemic, can exacerbate the difficulties in monitoring and valuation by injecting a substantial amount of noise into the performance measures, such as output, that investor normally rely on.

The opaqueness of modern businesses poses a challenge to investors and entrepreneurs, as well as to academic researchers. To investors, many investment opportunities, especially in high tech industries or startups, are like black boxes. Investors can control the input: how much capital to invest; and receive reports on some output, such as revenues, profits, or growth. However, investors likely have little knowledge of how the reported outputs were produced or whether they accurately reflect the true productivity and return of their investment. To entrepreneurs, the opaqueness of their technology may be an obstacle preventing them from receiving the financing and investment needed, and it may not be simply overcome with more communication to the investors, due to the level of expertise required in the production process.[1] To researchers, such opaqueness potentially represents a substantial degree of agency friction (e.g., financial misconduct, value distortion, etc.) that likely bears persistent, long-term

[1] For example, because a large number of students often enter a virtual classroom simultaneously, GSX was suspected of using bots in lieu of actual subscribers. GSX explained that the synchronized entry is the result of participants being placed into a "prep session" prior to the beginning of class to then be admitted into the classroom by the instructor collectively.

negative consequences if it is not properly addressed. Thus, it is natural to ask: what are the implications of the information asymmetry and agency friction on the financing and investment policies for businesses with opaque technologies?

This paper aims to answer this question by using an agency-based dynamic investment model with persistent private information. In the model, a risk-neutral investor (the principal, she) provides capital to an entrepreneur (the agent, he) who possesses a technology that produces cash flows out of capital. The technology is opaque to the investor: she only has control of the input – the amount of capital invested. She receives an output: the cash flow reported to her by the entrepreneur. Other than the input and the output, she does not observe the true productivity of the entrepreneur or whether the reported cash flow is accurate. The entrepreneur must exert costly effort in order to increase productivity over time, but without proper incentives, he may engage in the socially suboptimal actions of shirking and cash flow diversion, which produce private benefit for him. He can also misreport the realization of cash flows to conceal his actions. Diversion and misreporting have a persistent, long-term negative impact on the business, as they reduce the expected growth rate of productivity.

Using the analytical tools developed in Williams (2011, 2015), and Marinovic and Varas (2019), I established the incentive compatibility (IC) conditions that prevent shirking, cash-flow diversion, and misreporting. Due to the persistent effect of those actions, the IC conditions rely on two state variables: the usual *continuation utility* of the entrepreneur, which is his expected present value of all future compensation, and the *stock of future incentives*, which is the expected present value of all future pay-performance sensitivity. The investor's incentive compatible financing policy and the growth rate of productivity at each point in time both depend on the stock of future incentives accumulated up to that point.

The additional dimension of the agency problem complicates the investor's optimal financing policy, which in general is a function of several state variables. To reduce the complexity of the investor's problem and provide clear, empirically relevant implications, I consider a special setting that has been widely adopted in the existing literature: a risk-neutral entrepreneur with liquidity constraint. That is, the entrepreneur's flow utility must be non-negative and his reported cost of effort must be compensated by the investor at all times. This specification renders the model comparable to the standard agency-based dynamic q-theory investment models, such as Bolton, Chen, and Wang (2011) and DeMarzo, Fishman, He, and Wang (2012), as well as models of financially constrained entrepreneur with unobservable production costs, such as Krishna, Lopomo, and Taylor (2013) and Krasikov and Lamba (2020).

Under the assumption of a risk-neutral entrepreneur with liquidity constraint, the optimal contract can be characterized with an ordinary differential equation (ODE) of a single state variable representing the stock of future incentives *per unit of capital*. This allows me to characterize the optimal investment policy and equilibrium productivity growth rate as functions of the single state variable, and highlight the distinct implications of this model compared to the standard dynamic q-theory models without the persistent impact of the entrepreneur's actions. One distinct implication is that the model predicts that investment can be both lower or higher than the first-best level, whereas the standard q-theory models only predict underinvestment. In particular, the model predicts over-investment to be likely prominent among firms with strong cash flows but low Tobin's q, which is consistent with the empirical evidence in Blanchard, Lopez-de Silanes, and Shleifer (1994). Secondly, the model implies a potentially strong investment-to-cash-flow sensitivity and relatively weak investment-to-q sensitivity, a prediction that reconciles with the empirical observations summarized in Ai, Li, and Li (2017) and Cao, Lorenzoni, and Walentin (2019) but is not available in standard q-theory models in which average productivity is assumed to be exogenous and constant.

The optimal contract can be implemented with common securities such as stocks, bonds, and cash reserves. The implementation implies a substantial, state-dependent credit yield spread when the entrepreneur is low on cash reserves. In particular, all else equal, the credit yield spread widens when the persistent impact of the agency frictions is stronger, offering a potential avenue for future studies of the credit spread puzzle from the perspective of persistent private information.

The model can be extended to accommodate different settings while the main empirical implications of the model remain robust. In one extension, the investor maximizes the valuation of the business instead of the discounted cash flows. The implications of this extension under the assumption of a risk-neutral entrepreneur with liquidity constraint are qualitatively similar to those produced in the baseline model: in equilibrium, both underinvestment and overinvestment can arise, and the optimal investment policy has a strong correlation with the return-of-capital and a weak or even negative correlation with Tobin's q. In another extension, the entrepreneur has a constant-absolute-risk-averse (CARA) utility function and the access to a private savings technology. This is a common modeling technique that takes advantage of the lack of the wealth effect with the CARA utility and reduces the dimension of the investor's problem.[2] Following this technique, the investor's value function can be once again summarized as an ODE with a single state variable. In addition to producing similar implications as the baseline model, this extension implies that the optimal financing policy can be mon-monotonic, and the correlation between firm value and cash-flow history varies with the entrepreneur's degree of risk aversion: the more risk-averse the entrepreneur is, the weaker such correlation should be.

Related Literature This paper bridges several strands of research. One strand is the finance literature on agency-based investment theories, such as Bolton, Chen, and Wang (2011), De-

[2]The same technique is adopted in He (2011), Williams (2015), He, Wei, Yu, and Gao (2017), Marinovic and Varas (2019), Cetemen, Feng, and Urgun (2020) and a number of other studies for the same purpose.

Marzo, Fishman, He, and Wang (2012), Decamps, Gryglewicz, Morellec, and Villeneuve (2016), Cao, Lorenzoni, and Walentin (2019), Ai, Kiku, Li, and Tong (2020), etc. These studies typically model a risk-neutral agent with one or several sources of uncertainty. However, the output (e.g. cash flow) is usually publicly observable, and the agent's actions do not bear a persistent effect. Consequently, the incentive compatibility condition boils down to a static tradeoff between instantaneous private benefit and continuation utility.

Another strand of related research is the industrial organization literature on firm financing, such as Albuquerque and Hopenhayn (2004), Quadrini (2004), Clementi and Hopenhayn (2006), DeMarzo and Fishman (2007), Krishna, Lopomo, and Taylor (2013), Fu and Krishna (2019), Krasikov and Lamba (2020), etc. These studies are cast in discrete-time settings with one of two types of agency frictions: unobservable cash flows, or private production cost. The latter is also typically accompanied with a liquidity constraint on the agent's side. The focus of these studies is usually the time-series dynamics of financing decisions based on the realized paths of various shocks. In particular, Fu and Krishna (2019) and Krasikov and Lamba (2020) study a cash flow diversion model and a private production cost model, respectively, assuming that the shocks to cash flow or production cost are serially correlated. In comparison, this paper models both cash flow diversion and private cost of production with a liquidity constraint. It offers implications on not only the time-series paths of financing but also the cross-sectional predictions on the relationship between the investment policies and the empirical determinants of such policies, including, Tobin's q and cash flow.

In terms of the methodology, this paper belongs to the active literature of continuous-time dynamic agency models, including DeMarzo and Sannikov (2006), Biais, Mariotti, Plantin, and Rochet (2007), Sannikov (2008), and Zhu (2013) with transitory shocks, and Hoffmann and Pfeil (2010), Piskorski and Tchistyi (2010), Li (2017), and Feng (2020) with persistent but publicly observable shocks to model parameters. The studies most closely related to this paper in terms of methodology are Williams (2011, 2015), and Marinovic and Varas (2019). These studies assume that the principal observes a noisy signal and that the agent can take private actions with a persistent impact on the future generation of that signal. They utilize a stochastic maximum principle approach involving a change of the probability measures – a technique also adopted in this paper.[3] However, all of the aforementioned research focuses on the design and implementation of the optimal compensation contract, whereas this paper undertakes the analysis of optimal financing and investment behaviors in the presence of persistent private information.

More broadly speaking, this paper complements the studies of (discrete-time) mechanism

[3]A critical result of this technique is that the agent's IC condition involves at least two state variables: the usual continuation utility, and the stock of future incentives. This is different from dynamic agency models with project selection or capital budgeting such as Varas (2018) and Malenko (2019), etc. In those studies, although the agent's action produces a persistent effect, the incentive for such action can be pinned down at the time of the action, and a single state variable is still sufficient to characterize the optimal contract.

design problems with persistent information, such as Fernandes and Phelan (2000), Battaglini (2005), Zhang (2009), Kapička (2013), Pavan, Segal, and Toikka (2014), Sannikov (2014), Tchistyi (2016) and others.[4] In these studies, the persistence of the agency friction mainly arises from the serial correlation among the shocks to a noisy signal as opposed to the long-term impact of the agent's private actions, which is what this paper examines.

2 The Model

In this section, I introduce the basic model setup in 2.1, followed by the information structure and agency frictions in 2.2. I establish the general incentive compatibility conditions in 2.3 and define the optimal contract in 2.4. Discussions of different interpretations of the model assumptions are also offered.

2.1 The Basic Environment

Time is continuous. A risk-neutral investor (the principal, she) has access to the capital market, and an entrepreneur (the agent, he) has the technology that produces cash flows out of capital. They share the same discount factor $r \geq 0$ and join to form a business, in which the investor provides financing by raising capital, and the entrepreneur produces output using capital and his technology.

The production of cash flow follows a standard neoclassic linear investment model such as that in Bolton et al. (2011) and DeMarzo et al. (2012). Let K_t represent the level of capital stock and A_t the cumulative productivity process. The incremental (gross) cash flow dY_t is given by

$$dY_t = K_t dA_t \,. \tag{1}$$

The accumulation of capital follows

$$dK_t = (I_t - \delta K_t)dt \,, \tag{2}$$

where I_t is the capital investment and $\delta \geq 0$ is the rate of capital depreciation. The evolution of productivity A_t is stochastic and given by

$$dA_t = R_t dt + \sigma dZ_t \,, \tag{3}$$

where R_t is the expected growth rate of productivity, Z_t is a standard Brownian motion on a

[4] See Bergemann and Välimäki (2019) for a survey on dynamic mechanism design with persistent information.

complete probability space, and σ is the constant volatility of the productivity growth process.[5] The details of R_t (its dependence on the entrepreneur's actions) are presented in the next subsection after introducing the information and agency frictions.

Capital investment must be financed by the investor, who bears an adjustment cost $G(I_t, K_t)$. I assume G is the standard neoclassic quadratic adjustment function. That is, $G = g(i)K$, where $i \equiv I/K$, and

$$g(i) = i + \frac{\theta}{2} i^2 , \tag{4}$$

for some $\theta > 0$. The investor also makes instantaneous compensation $C_t \geq 0$ to the entrepreneur. Therefore, the flow profit (net cash flow) of the business is

$$K_t dA_t - G_t dt - C_t dt . \tag{5}$$

2.2 Information and Agency Frictions

There are two sources of agency frictions in this model. The first is a standard (two-dimensional) moral hazard problem: on the one hand, the entrepreneur can divert cash flows to produce private benefit for himself. For every dollar of cash flow lost due to diversion, the entrepreneur receives $\lambda \in (0,1)$ dollars in private benefit. On the other hand, the entrepreneur bears a private cost $h(\mu, K)$ for generating the expected productivity growth rate μ. I assume $h(\mu, K) = \alpha\mu^2 K/2$ for some $\alpha > 0$. That is, the cost of productivity growth is quadratic in μ and homogeneous of degree one in K. In the absence of proper incentives and compensation for such cost, the entrepreneur may choose to shirk, exerting lower effort than the investor desires.

The flow utility of the entrepreneur u_t is a function of the compensation from the investor C_t, the amount of capital diverted B_t, and the level of effort μ_t. That is,

$$u_t \equiv u\left(C_t + \lambda B_t - \frac{\alpha}{2}\mu_t^2 K_t\right) , \tag{6}$$

where u is a continuous, twice-differentiable, monotonically increasing ($u' > 0$), and weakly concave ($u'' \leq 0$) utility function.

The second source of the agency friction, and the reason the entrepreneur can engage in cash-flow diversion and shirking without being detected, stems from the opaqueness of the entrepreneur's technology. In practice, the production process inside a firm, especially for startup in high-tech industries, is rarely transparent to outside investors, To capture this opaqueness, I assume that the investor can observe neither the actual cash flow Y_t nor the true level of pro-

[5] Alternatively, one can assume that productivity grows deterministically while capital growth is subject to Brownian shocks. The results are largely unchanged.

ductivity A_t. She only receives the entrepreneur's report on the value and the evolution paths of those variables. Specially, the *reported* cash flow, denoted by $d\hat{Y}_t$, follows

$$d\hat{Y}_t = K_t d\hat{A}_t \,, \tag{7}$$

and $\hat{A}_t$, the *reported* productivity, evolves according to

$$d\hat{A}_t = \hat{\mu}_t dt + \sigma d\hat{Z}_t \,, \tag{8}$$

where $\hat{\mu}_t$ is the entrepreneur's reported effort in productivity growth, and $\hat{Z}_t$ is the reported realization of the Brownian motion.[6] Both variables can potentially differ from their true values μ_t and Z_t, which represent the entrepreneur's true effort and the actual realization of the Brownian motion, respectively.

Both cash-flow diversion and misreporting hurt the long-term growth of the business. I define the entrepreneur's "stock of misconduct" M_t as the difference between the reported cash flow (7) and the true cash flow (1) plus the amount of cash flow diverted, that is

$$dM_t = d\hat{Y}_t - dY_t + B_t dt \,. \tag{9}$$

To capture the persistent effect of the entrepreneur's misconduct, I assume that the true productivity growth rate $R_t = \mu_t - \rho M_t$, where $\rho > 0$ is the marginal effect of misconduct on future productivity growth. Thus, the true law of motion for productivity A_t is

$$dA_t = \mu_t dt - \rho M_t dt + \sigma dZ_t \,. \tag{10}$$

The model setup, the information and agency frictions can be interpreted in different ways:

- The setup can be interpreted as a model of startup financing (as in Clementi and Hopenhayn (2006), Krishna et al. (2013), Fu and Krishna (2019), and Krasikov and Lamba (2020)), with the investor being a venture capitalist or a hedge fund. The investor's decision i_t represents the amount of capital injected in a startup with an opaque production technology. Alternatively, the setup can be interpreted as a model of firm investment (as in Bolton et al. (2011), DeMarzo et al. (2012), Ai et al. (2017), and Cao et al. (2019)), with the investor being the shareholders, and the entrepreneur being the CEO or managers with superior information of the technology and organization of the production process inside the firm. The investor's decision i_t represents the amount of capital allocated to the managers for production, but the managers' interests are not always aligned with those of

[6] Equation (7) and (8) assume that the investor does not wish to implement any misconduct in equilibrium (see Definition 2), and the entrepreneur's report must be consistent with some possible realizations of a process driven by Brownian shocks (e.g., no discrete jumps).

the shareholders without proper incentives. Because of these alternative interpretations, I use the terms *investment* and *financing* interchangeably in the subsequent analyses.[7]

- Cash-flow diversion and costly effort are among the most commonly studied agency frictions in both the finance and industrial organization literature.[8] Moreover, instead of the managerial effort spent on productivity growth, μ can alternatively be interpreted as working capital (e.g. cash) that must be privately deployed by the entrepreneur to maintain a certain level of the productivity growth. In Section 3 I study a version of the model in which the entrepreneur faces a liquidity constraint (e.g., as in Krishna et al. (2013) and Krasikov and Lamba (2020)). In that case, the cost of effort (or the working capital deployed) $h(\mu, K)$ must be fully covered by the investor, and misreporting effort is consequently equivalent to the entrepreneur misreporting his production cost.

- The output Y_t is interpreted as the cash flow for the most part of the paper, as discounted cash flow is one of the most popular methods of business or project valuation for investors. The other popular method is discounted earnings, in which case the agency friction can be naturally understood as earnings manipulation.[9] Finally, the output can also be modeled as the firm's market value. In Section 4 I demonstrate how the baseline model can be modified if maximizing market value is the objective of the investor and how it produces very similar results in terms of the dynamics of financing and investments.

- The persistent effect of misconduct in (10) follows the examples in Williams (2011, 2015) and Marinovic and Varas (2019). $\rho > 0$ captures the idea that misconduct in the past hurts firm growth in the future. Examples of this long-term negative impact include the time and energy the entrepreneur spends to conceal his past misconduct, litigation risks, and production disruptions during an external investigation. As noted in Williams (2011), equation (10) implies that the true productivity growth is an Ornstein-Uhlenbeck process, which is a continuous-time version of a stationary Gausssian autoregressive process, with λ governing the rate of mean reversion and hence the level of persistence of the process.[10]

[7] An important strand of the literature on firm financing focuses on the *structure* of the capital raised (e.g., debt versus equity). In this paper, the investor's financing decision mainly refers to the *quantity* of capital invested.

[8] Models in which cash-flow diversion is the main agency friction include Quadrini (2004), Clementi and Hopenhayn (2006), Ai et al. (2017), Fu and Krishna (2019), etc. Models in which costly effort is the main agency friction include Krishna et al. (2013), Krasikov and Lamba (2020), etc. Studies such as DeMarzo and Sannikov (2006) and DeMarzo et al. (2012) assume the cost of effort is linear, which results in shirking being equivalent to cash-flow diversion. In this model, the cost of effort is strictly convex and hence not equivalent to cash-flow diversion.

[9] There is broad literature on the prevalence and long-term effect of earnings manipulation, such as Stein (1989), Edmans, Gabaix, Sadzik, and Sannikov (2012), Varas (2018), Zhu (2018), etc. See Marinovic and Varas (2019) and the references within.

[10] Note that the accumulation of M_t is dynamic. Thus, this version of modeling the persistent effect of an agent's private actions is different from studies in which the agent takes a one-time action with a persistent effect, such as Hoffmann, Inderst, and Opp (2020) and Georgiadis and Szentes (2020).

In short, the investor of this model faces a "black box" investment opportunity. She controls the input: capital investment I_t, and observes an output: the reported cash flow $\hat{Y}_t$. However, she does not observe the cash flow production process. The entrepreneur may shirk or divert cash flows for private benefit. The investor cannot directly detect shirking or cash-flow diversion because she also does not observe the actual realization of the cash flow.[11]

I adopt the following definitions for the remaining of the paper:

Definition 1 (Contract) *A contract is the investor's investment and compensation policies* $\{I_t, C_t\}_{\{t\geq 0\}}$ *and her recommended production growth rate* $\{\mu_t\}_{\{t\geq 0\}}$ *based on the entrepreneur's reported paths of cash flows and productivity.*

Definition 2 (Incentive Compatibility) *A contract is incentive compatible if it implements no cash-flow diversion (*$B_t = 0$*) and no misreporting (*$\hat{\mu}_t = \mu_t$*,* $\hat{Y}_t = Y_t$*, and* $M_t = 0$*) for all* $t \geq 0$*.*

I also assume that the investor can fully commit to the contract once it is signed. The analysis of an investor with limited commitment ability may be an interesting extension, which I briefly discuss in the concluding section.

2.3 The Entrepreneur's Problem

With the information structure and agency frictions specified, I now define the entrepreneur's optimization problem. Let $\mathscr{F}_t$ denote the filtration generated by the true paths of the Brownian motion Z_t, $\hat{\Pi}$ denote the set of the entrepreneur's reports $(\hat{Y}_t, \hat{A}_t, \hat{\mu}_t, \hat{Z}_t)$, T denote some stopping time (can be endogenous) and W_T denote the entrepreneur's continuation payoff at T. The entrepreneur's optimization problem given any contract $\{I_t, C_t, \mu_t\}_{t\geq 0}$ is

$$\max_{B,\,\mu,\,\hat{\Pi}} \quad \mathrm{E}\left[\int_0^T e^{-rt} u\left(C_t + \lambda B_t - \frac{\alpha}{2}\mu_t^2 K_t\right) dt + e^{-rT} W_T \middle| \hat{\mathscr{F}}_t\right], \tag{11}$$

subject to the reported paths of cash flows (7) and productivity (8) as well as the true paths (1), (2), and (10). To solve this problem, I adopt the technique used in Williams (2011, 2015) and Marinovic and Varas (2019) and perform a change of the probability measure from that generated by the true Brownian motion Z_t to that generated by the reported Brownian motion $\hat{Z}_t$. Consequently, the entrepreneur's problem can be expressed in two state variables. One is his continuation utility, W_t, defined as:

$$W_t = \mathrm{E}\left[\int_t^T e^{-r(s-t)} u\left(C_s + \lambda B_s - \frac{\alpha}{2}\mu_s^2 K_s\right) ds + e^{-r(T-t)} W_T \middle| \hat{\mathscr{F}}_t\right], \tag{12}$$

[11] Section 4 offers discussions on the implications of the agency frictions when the investor can directly observe the true cash flow history.

where, compared to (11), the expectation is taken under $\hat{\mathscr{F}}_t$, the filtration generated by $\hat{Z}_t$. The other state variable, denoted P_t, arises from the persistent effect of the entrepreneur's stock of misconduct (M_t) and is given by

$$P_t = -\mathrm{E}\left[\int_t^T e^{-r(s-t)}\rho P_s ds + e^{-r(T-t)}P_T \middle| \hat{\mathscr{F}}_t\right] < 0\,. \tag{13}$$

Broadly speaking, P_t captures the tradeoff between his misconduct and its impact on future productivity growth. It is an important variable to keep track of in models with persistent private information. More detailed explanation on this variable will be given after the next proposition, which states the necessary conditions for an incentive compatible contract:

Proposition 1 *Under the incentive compatible contract there exist $\mathscr{F}$-adapted processes $\{\phi_t, \beta_t\}$ such that*

$$dP_t = (r-\rho)P_t dt - \phi_t P_t dZ_t\,, \tag{14}$$

$$dW_t = rW_t dt - u_t dt - \beta_t W_t dZ_t\,, \tag{15}$$

Let u_B and u_μ denote the partial derivatives of the entrepreneur's utility function with respect to B_t and μ_t, respectively. The necessary conditions for the contract to be incentive compatible are

$$u_B \le -P_t \tag{16}$$

$$u_\mu = P_t \tag{17}$$

$$\beta_t = \sigma\frac{P_t}{W_t}\,. \tag{18}$$

Equations (14) and (15) characterize the laws of motion of the two state variables P_t and W_t, where ϕ_t and β_t capture the sensitivity of each state variable to the reported paths of the Brownian motion. In particular, β, as in standard dynamic moral hazard models, represents the entrepreneur's pay-performance sensitivity or his "skin-in-the-game". Equations (16) and (17) are the incentive-compatibility (IC) conditions for no cash-flow diversion and no shirking. Intuitively, both cash-flow diversion and shirking generate immediate utility. However, they also increase the stock of misconduct, which negatively affects the productivity growth rate in the future. The marginal value of such a negative long-run effect is captured by P_t.[12]

Equation (18) is the IC condition for no misreporting the path of the Brownian motion dZ_t.

[12]Condition (16) can be binding or slack depending on whether cash-flow diversion is allowed to be negative, i.e., whether the entrepreneur is allowed to privately save the cash flows diverted and reinvest them in the future. Section 3 explores an example in which private savings are not allowed. In that case, (16) implies a lower bound on the state variable. Section 4.2 explores an example in which private savings are allowed. In that case, (16) binds with equality. In comparison, (17) always binds, because the entrepreneur's effort can be higher or lower than the recommended level at any time.

This condition highlights a critical difference that persistent private information makes compared to the IC condition in standard dynamic moral hazard problems without such persistence. In the standard problems, a strong pay-performance sensitivity is usually required to prevent the agent from adopting actions that are suboptimal for the principal. The more "skin-in-the-game" the agent has, the less motivated he is to engage in cash-flow diversion or shirking. In this model with persistent private information, strong pay-performance sensitivity potentially has a negative effect on the investor because it gives the entrepreneur incentives to falsify strong cash flows in order to boost his continuation utility. Instead, the IC condition restricts the level of pay-performance sensitivity according to P_t. Substituting (18) into (13) yields

$$P_t = -\mathrm{E}_t\left[\int_t^T e^{-r(s-t)}\left(\frac{\rho}{\sigma}\right)\beta_s W_s ds + e^{-r(T-t)}P_T\right]. \tag{19}$$

That is, P_t can be interpreted as the (negative) expected present value of all future pay-performance sensitivity or, more simply, the *stock of future incentives*. If the entrepreneur falsely reports a high realization of the cash flow, his continuation utility increases.[13] Meanwhile, the stock of his future pay-performance sensitivity also increases. As a result, future reports of low realizations of the cash flow will result in a more severe punishment. The degree of such punishment is captured by P_t and restricts the level of pay-performance sensitivity the contract can implement at each moment.

According to (14), P_t is a geometric Brownian motion (GBM) with expected growth rate $r-\rho$. The larger ρ is, the faster P_t decreases over time. Given that $P_t < 0$, (14) implies that the larger ρ is, the faster the entrepreneur's future incentives accumulate on average over time. This is an important intuition that plays a vital role in understanding the implications of the agency frictions on the dynamics of financing and investment discussed later.

In the interest of space, detailed derivations of the IC conditions in Proposition 1 are left to Appendix A. These conditions are also necessary, and their sufficiency is discussed in Appendix B. From now on, all subsequent analyses assume that the necessary conditions are met and the contracts are always incentive compatible. Consequently, the ˆ notation on all variables will be dropped, as all reported values are equivalent to their true values in equilibrium.

2.4 The Investor's Problem

The investor's objective is to maximize her valuation of the business, which is the expected present value of all the future (net) cash flows (gross output minus production cost minus compensation). Let F denote the investor's valuation, and F_T represent her valuation after contract

[13] Analyses in the next two sections show that $\beta < 0$ when $W > 0$ and $\beta > 0$ when $W < 0$. Consequently, W always increases with a positive realization/report of the cash flow.

termination (possibly endogenous). Under an incentive compatible contract, F is defined as

$$F_t = \max_{\{I_t, C_t, \mu_t, \phi_t, \beta_t\}} \mathrm{E}\left[\int_t^T e^{-r(s-t)}(dY_s - G_s ds - C_s ds) + e^{-r(T-t)}F_T \middle| \mathscr{F}_t\right], \tag{20}$$

subject to the laws of motion dK_t, dP_t, and dW_t as specified in (2), (14), and (15). The investor has quintuple controls: investment I_t, compensation C_t, recommended productivity growth rate μ_t, plus ϕ_t and β_t, the sensitivity of P_t and W_t to the reported Brownian shocks. Her value function F satisfies the following general Hamilton-Jacobi-Bellman (HJB) equation:

$$\begin{aligned} rF(K,P,W) = \max_{I,C,\mu,\phi,\beta} \; & \mu K - G(I,K) - C + (I-\delta K)F_K + (r-\rho)PF_P + \frac{1}{2}\phi^2 P^2 F_{PP} \\ & + (rW - u)F_W + \frac{1}{2}\beta^2 W^2 F_{WW} + \phi\beta PW F_{PW}, \end{aligned} \tag{21}$$

subject to the IC conditions (16), (17), (18), and some boundary conditions that are problem-specific and will be established later.

I adopt the following definition of an *optimal contract*:

Definition 3 (Optimal Contract) *A contract is optimal if it maximizes the investor's objective function (20) over the set of contracts with the following properties:*
i) Are incentive compatible.[14]
ii) Assign the entrepreneur some initial level of continuation utility W_0 and the initial stock of incentives P_0.[15]
iii) Respect the other problem-specific constraints that the entrepreneur has.[16]

In general, solving the investor's HJB equation (21) and characterizing the optimal contract is an analytically challenging task involving a partial differential equation (PDE) with multiple state and control variables. In the next sections, I explore a special setting commonly used in the literature with broad applications: a risk-neutral entrepreneur with liquidity constraint. This example reduces the investor's HJB equation (21) to a tractable ordinary differential equation (ODE), allowing the delivery of clear analytical predictions.

[14] This definition restricts the analysis to contracts that involve no misconduct. Zhu (2013) shows that in a simple dynamic moral hazard problem such as that in DeMarzo and Sannikov (2006), allowing actions such as shirking may not always be suboptimal for the principal. To my knowledge, such analysis is difficult in models with persistent private information and not conducted in existing studies such as Williams (2011, 2015) or Marinovic and Varas (2019). I thus abstract from such analysis and focus on contracts that are incentive compatible.

[15] I leave the determination of W_0 and P_0 outside the model. If the investor's value function has an interior maximum, then W_0 and P_0 are typically chosen at such maximum.

[16] E.g. the entrepreneur's "liquidity constraint" in Section 3 and "no-hidden-savings" constraint in Section 4.2.

3 Risk-Neutral Entrepreneur with Liquidity Constraint

In this section I solve the optimal financing and investment contract under the assumption of a risk-neutral entrepreneur with liquidity constraint. 3.1 introduces the assumptions used in this section. The resulting optimal contract is solved in 3.2, followed by the discussion of its empirical implications in 3.3. A security implementation of the optimal contract is presented in 3.4 along with its implications in asset pricing, followed by a short summary in 3.5.

3.1 Assumptions

The following assumptions are made throughout the section:

Assumption 1 (Risk Neutrality) *The entrepreneur's flow utility u_t is given by*

$$u_t = \left[c_t + \lambda b_t - \frac{\alpha}{2}\mu_t^2\right] K_t \ . \tag{22}$$

where $c_t \equiv C_t / K_t$ and $b_t \equiv B_t / K_t$ represent the rate of consumption and cash flow diversions per unit of capital, respectively.

Assumption 2 (Liquidity Constraint) *The entrepreneur cannot bear negative utility at any point in time. That is, $u_t \geq 0$ for all $t \geq 0$.*

Liquidity constraint of this kind has been extensively analyzed especially in the industrial organization literature, such as Clementi and Hopenhayn (2006), Biais et al. (2007), Krishna et al. (2013) and Krasikov and Lamba (2020). In this model, the costly effort μ can be interpreted as the working capital (e.g. cash) needed for maintaining a certain level of productivity growth. However, the entrepreneur is cash-strapped and must be provided the cash needed for investing in productivity growth. I also assume that the entrepreneur cannot save by himself. That is, he cannot relax the liquidity constraint himself through saving and reinvesting, which implies that b_t cannot be negative. This no-saving constraint is standard and made in all aforementioned studies featuring the liquidity constraint. It is also innocuous, because the investor and the entrepreneur share the same discount factor r, and any saving that is beneficial to the business can be done by the investor.

The liquidity constraint is not the same as "limited liability", the assumption usually made in models with a risk-neutral agent. Limited liability restricts the agent to be paid with non-negative consumption, or $c_t \geq 0$. In comparison, the liquidity constraint in this model states that the entrepreneur must be allocated the requisite working capital prior to production, or $u_t \geq 0$. The liquidity constraint also prevents the investor from selling capital to the entrepreneur, in which case the investor acts as an intermediary and the resulting contract would achieve the first-best outcomes trivially under risk-neutrality and equal discounting.

Finally, the optimal contract in this section involves possible termination. To maintain sufficient tractability, I assume the following termination technology:

Assumption 3 (Termination Technology) *If the contract is terminated at time T, the remaining capital is liquidated at $l < 1$ dollar per unit. The investor receives the proceeds from the liquidation and the entrepreneur receives his continuation utility W_T in expectation, subject to a noisy signal that respects the definition of P_T as the marginal value of misconduct in (13).*

Simplify put, this assumption ensures that when the contract ends, the terminal value of the state variables W_T and P_T are still properly defined.[17] As the ensuing analysis shows, contract termination in this model is not due to the usual limited liability constraint in most models with a risk-neutral agent. Instead, it is due to an endogenous incentive constraint resulting from the persistence of the entrepreneur's private actions.

3.2 Incentive Compatibility and the Optimal Contract

Given the entrepreneur's utility function (22), the IC conditions (16) and (17) become

$$-P_t \geq \lambda K_t \tag{23}$$

$$\mu_t = \frac{-P_t}{\alpha K_t} \ . \tag{24}$$

The two IC conditions can be summarized using a single state variable X_t, defined as

$$X_t \equiv -\frac{P_t}{K_t} > 0 \ . \tag{25}$$

X_t represents the *stock of future incentives per unit of capital.*[18] Equation (23), the IC condition that prevents cash-flow diversion, implies $X_t \geq \lambda$. On the one hand, cash-flow diversion generates private benefit for the entrepreneur with λ capturing its marginal benefit. On the other hand, cash-flow diversion increases the stock of misconduct, which reduces future productivity growth. Therefore, X measures the marginal cost of cash-flow diversion, which is the loss of future continuation utility due to lower productivity. Linearity in both the marginal benefit and the marginal cost of cash-flow diversion implies the lower bound (23) for X_t. The same intuition applies to (24), the IC condition for productivity growth. However, the quadratic cost in productivity growth pins μ_t down as a function of X_t.

[17]The precise information structure of the signal and the associated compensation policies are provided in Appendix B but the basic idea runs as follows: when the contract ends at time T, the investor receives a binary signal and pays the entrepreneur either $\zeta_h W_T$ or $\zeta_l W_T$ based on the signal. The amount of misconduct accumulated at T determines the probability that the investor receives the positive signal. ζ_h and ζ_l are set such that the entrepreneur's expected payment is exactly W_T and his marginal value of misconduct is exactly P_T.

[18]Note here X is defined as $-P/X$ for easy exposition because $P < 0$. Therefore, a higher X corresponds to a larger absolute value of P, or a larger stock of future incentives, holding K constant.

Besides the IC conditions, the assumptions made in Section 3.1 combined with various linearities built into the model also simplify the investor's problem, which can now be characterized as an ODE of X_t summarized as follows:

Proposition 2 *Under Assumptions 1, 2, and 3, the investor's value function* $F(K,P,W) = f(X)K - W$*, where* $f(X)$ *solves the following HJB equation:*

$$r f(X) = \mu(X) - g(i(X)) - c(X) + (i(X) - \delta)\left[f(X) - X f'(X)\right] + (\rho - r) X f'(X)\,, \tag{26}$$

with boundary conditions $f(\lambda) = l$*. The investor's optimal policies* $\{i_t, \mu_t, c_t\}$ *are all functions of* X_t*, given by*

$$i(X) = \frac{f(X) - X f'(X) - 1}{\theta} \tag{27}$$

$$\mu(X) = \frac{X}{\alpha} \tag{28}$$

$$c(X) = \frac{X^2}{2\alpha}\,. \tag{29}$$

Proposition 2 reveals some unique features of this model. First, the optimal contract involves endogenous *incentive-driven* contract termination at $X_t = \lambda$, because there is not a sufficient amount of incentives to prevent the entrepreneur from diverting cash flows. The business is dissolved and the investor receives l dollar per unit of the remaining capital according to the assumption made in Section 3.1. This is in contrast to the usual *limited-liability-driven* termination in most dynamic models with a risk-neutral agent (e.g., DeMarzo and Sannikov (2006), Bolton et al. (2011), DeMarzo et al. (2012), etc). In those modes, the agent's continuation utility decreases after negative shocks. The contract must be terminated and the agent fired if his continuation utility falls below his outside option (typically 0). In this model, limited liability is irrelevant, because the entrepreneur's liquidity constraint implies that both his flow utility and continuation utility are always non-negative.[19]

Another feature of this model different from existing models with a risk-neutral agent is the underlying source of inefficiency in the optimal contract. In the existing models, the agent's continuation utility is the key state variable and the likelihood of contract termination due to the agent's limited liability is the primary source of inefficiency. In this model, the primary source of inefficiency stems from the likelihood of contract termination due to insufficient future incentives (i.e., $X_t \geq \lambda$), while limited liability is irrelevant. Moreover, conditional on that the liquidity constraint is met, the investor can always make extra lump-sum payments to the entrepreneur and reduce his continuation utility anytime. Consequently, the marginal value of

[19]Readers can also infer from (15) that W_t never goes to zero because it follows a geometric Brownian motion when $u_t = 0$, which is indeed the case in equilibrium and is discussed in the next paragraph.

the entrepreneur's continuation utility to the investor is a constant (i.e. $F_W = -1$). This immediately implies that the entrepreneur's liquidity constraint is always binding under the optimal contract. That is, $u_t = 0$ and $c_t = \alpha\mu_t^2/2$ before the contract terminates. Nevertheless, the entrepreneur's continuation utility still varies with the reported cash flows. That is, $\beta_t > 0$, based on the IC condition (18). As discussed in Section 2.3, this is needed to prevent the entrepreneur from misreporting the realization of the Brownian motion. Because the investor and the entrepreneur share the same discount factor, any continuation utility the entrepreneur has accumulated is paid out in a lump-sum at the time of contract termination.[20]

3.3 Empirical Implications

The optimal contract characterized in Proposition 2 generates a number of empirically relevant implications on the dynamic behaviors of firm financing and investment. This section discusses two sets of those implications: this first pertains to the value of the business and the optimal investment policy as functions of the state variable X; the second pertains to model-implied correlation between investment, Tobin's q, and expected cash flow. To highlight what differences the persistent private information and other features of this model make, the discussion focuses on comparing those implications with those of three widely referenced benchmarks: Hayashi (1982), Bolton et al. (2011), and DeMarzo et al. (2012). The first is a standard neoclassic investment model without agency frictions. The latter two are agency-based dynamic investment models with a risk-neutral agent but without the persistence in the agent's private actions.

A. Business Value and Optimal Investment

The left panel of Figure 1 plots the scaled value function $f(X)$, that is, the value of the business or the investment opportunity per unit of capital.[21] When X is low, the value of the business is low, due to the inefficient contract termination when X reaches the incentive termination boundary $X = \lambda$. Higher X reduces the risk of inefficient termination. However, when X is too high, the value of the business starts to decrease, mainly due to an inefficiently high level of investment, which is discussed next.

The right panel of Figure 1 plots the optimal investment policy i as a function of X. For comparison, the first-best level of investment in the absence of agency frictions is also plotted. The first-best level is achieved when both A_t and Y_t are directly observable to the investor. In that case, the first-best level of investment i^{FB} coincides with the solution in Hayashi (1982)

[20]In contrast, most existing dynamic models with a risk-neutral agent assume that the agent is more impatient than the principal. The agent receives cash payments if his continuation utility is sufficiently high.

[21]Recall that the investor's value function $F(K,P,W) = f(X)K - W$. Therefore, $f(X) = (F+W)/K$.

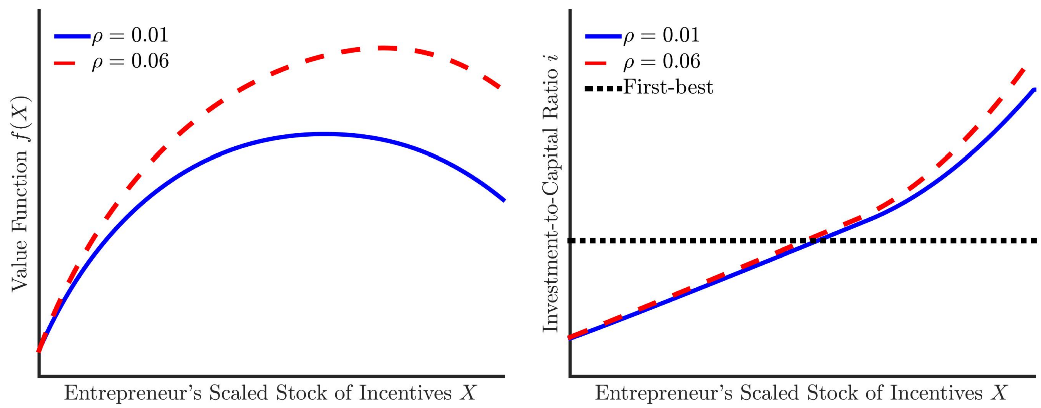

Figure 1: **Optimal Contract: Risk-Neutral Entrepreneur with Liquidity Constraint**
This figure presents the scaled value function $f(X)$ (left panel) and the investment-to-capital ratio $i(X)$ (right panel) under the optimal contract characterized in Proposition 2. The parameter values are $r = 0.05$, $\sigma = 0.25$, $\delta = 0.2$, $\lambda = 0.1$, $\theta = 1$, $\alpha = 1$, and $l = 0.9$. The blue solid line corresponds to $\rho = 0.01$, and the red dashed line corresponds to $\rho = 0.06$.

and is a constant given by

$$g'(i^{FB}) = \max_i \frac{\mu^{FB} - g(i)}{r + \delta - i}, \tag{30}$$

where $\mu^{FB} = 1/\alpha$. The first-best is independent of λ and ρ, because those parameters are only related to the agency frictions of cash-flow diversion and misreporting; it is time-invariant, because the investor is risk-neutral; and it is independent of σ and the history of productivity and cash flow realizations. It is, however, related to α. Because the cost of productivity growth is assumed to be convex, the first-best level of productivity μ^{FB} is a function α.[22]

In the presence of the agency frictions introduced in this model, the optimal investment decision becomes a function of the state variable X. When X, the stock of future incentives, is low, the optimal investment intensity is kept low, because the likelihood of the incentive termination is high. At the termination boundary ($X = \lambda$), the marginal value of capital is the liquidation value $l < 1$. When both l and X are low, termination is inefficient and imminent. The optimal contract therefore reduces the investment intensity and may even scale down the size of the firm, resulting in $i(X) < 0$. In other words, the optimal contract implies *underinvestment* by the investor when X is low.

The optimal investment intensity is a monotonically increasing function of X. This can also

[22] Note that the finite first-best μ^{FB} is not the result of the entrepreneur's liquidity constraint. Even if the entrepreneur can withstand negative flow utility, the cost of μ still needs to be paid in the first-best scenario.

been seen from differentiating (27) with respect to X, which yields

$$i'(X) = -\frac{Xf''(X)}{\theta} > 0\,, \tag{31}$$

due to the concavity of $f(X)$. Interestingly, $i(X)$ eventually rises above the first-best level, resulting in *overinvestment* from the investor. This is in sharp contrast to Bolton et al. (2011) and DeMarzo et al. (2012), in which the agency friction implies underinvestment only. In those studies, the impact of the agent's private actions is transitory, and the principal's cost of providing incentives vanishes when the agent's continuation utility is at a sufficiently high level. Because continuation utility only approaches such level from below, investment increases gradually with continuation utility and never surpasses the first-best.[23]

In this model, overinvestment is a unique result of the persistent impact of the agency frictions. Observe from (30) that i^{FB} is derived under a finite, first-best level of productivity μ^{FB}. In contrast, in the presence of the agency frictions, $\mu(X)$, given by (28), is an increasing function in X. A sufficiently high X may lead to an inefficiently high μ, resulting in over-investment.

To illustrate the underlying mechanism in more detail, recall that (28) is the IC condition for μ (equation 17) under risk-neutrality, which prevents the entrepreneur from deviating from the productivity growth rate the investor desires. When X, the stock of future incentives, is low, the investor prefers a low growth rate because she is primarily concerned with the entrepreneur shirking and does not have a sufficient stock of incentives to sustain a high growth rate. When X increases, the exact opposite concern arises: the entrepreneur is tempted to take advantage of the large stock of incentives (pay-performance sensitivity) accumulated by accelerating the productivity growth rate, boosting future cash flows, and consequently receiving higher continuation utility. To prevent such deviation, the investor is compelled to implement a growth rate higher than the first-best level, because that increases the marginal cost of additional growth in productivity for the entrepreneur. In short, over-investment is the result of the investor implementing an inefficiently high growth rate to prevent the entrepreneur from exploiting the large stock of future pay-performance sensitivity he has amassed through the contract.

In practice, overinvestment is observed in many scenarios, such as aggressive but inefficient mergers and acquisitions.[24] Blanchard, Lopez-de Silanes, and Shleifer (1994) find that firms with a strong series of cash flows may over-invest even when their investment opportunities are poor, as measured by a low Tobin's q. As Part B of this section will show, strong cash flows

[23] In fact, in Bolton et al. (2011) and DeMarzo et al. (2012), the equilibrium investment-to-capital ratio is *always lower* than the first-best level even when the cost of incentives vanishes. This is because the agents in those models are assumed to be more impatient than the principal. Thus, there is an efficiency loss for any level of continuation utility the principal promises to the agent.

[24] Wang (2018) summarizes the empirical findings on market reactions to M&A news. The reactions appear to be mostly negative for the acquirers, potentially reflecting the concerns of the general investors of the acquirers regarding the value of such expansions.

in this model lead to a higher X but not necessarily a higher q. Consequently, the prediction of this model provides a possible explanation for the over-investment problem in practice from the perspective of agency frictions with a persistent effect.[25]

Figure 1 also highlights how the value of the business and the optimal investment intensity varies with the degree of persistence of the agency frictions, measured by ρ. The left panel shows that the value function increases in ρ. This can be explained by examining the expected growth rate of the state variable X_t, given by

$$E(dX_t) = \rho - r + \delta - i \ . \tag{32}$$

That is, on average, X_t increases faster with a larger ρ. The intuition is very similar to that highlighted in equation (14), which shows that the stock of future incentives $-P_t$ on average increases faster when ρ is higher. Put differently, the stronger the negative effect of the entrepreneur's private actions, the faster the stock of future incentives accumulates. Because of the dynamic nature of the model, a larger stock of future incentives mitigates the agency friction, resulting in a higher valuation of the business or the investment opportunity. Meanwhile, a higher ρ uniformly elevates the optimal investment intensity for all X. That is, a strong negative long-term effect of the entrepreneur's private actions alleviates the underinvestment problem while exacerbating overinvestment.

B. Investment Sensitivity to Tobin's q and Cash Flows

The q-theory is among the most fundamental and widely adopted theories of firm investment. Hayashi (1982), Bolton et al. (2011), and DeMarzo et al. (2012) all examine the dynamic behaviors of investment within the q-theory framework. In the literature, two measures of q are commonly studied: the marginal q (q_m) and the average (Tobin's) q (q_a). They correspond to the marginal and average value of capital, respectively. In this model, the total value of capital, or the value of the business, is $F(K,P,W)+W$. Thus, Proposition 2 implies that both q_m and q_a can be expressed as simple functions of the state variable X:

$$q_m \equiv \frac{\partial(F+W)}{\partial K} = f(X) - Xf'(X) \tag{33}$$

$$q_a \equiv \frac{F+W}{K} = f(X) \ . \tag{34}$$

[25]This theory thus complements the recent studies that also feature over-investment as the result of dynamic agency frictions with different mechanisms, e.g., Bolton, Wang, and Yang (2019) with human capital, Gryglewicz, Mayer, and Morellec (2020) with correlated short-run and long-run effort, Szydlowski (2019) with multi-tasking, Ai et al. (2020) with limited commitment, and Gryglewicz and Hartman-Glaser (2020) with real options.

Comparing (33) with the optimal investment rule in (27) shows that optimal investment is fully determined by q_m. This result is also found in Bolton et al. (2011) and DeMarzo et al. (2012) and is not surprising. For the purpose of analytical tractability, both studies as well as this paper make various assumptions to ensure that the value function is homogeneous of degree one in capital. Consequently, the first-order condition for $i(X)$ (27) relates optimal investment with the marginal value of capital, which is fully captured by q_m.[26]

Despite the predictive power of the marginal q, Tobin's (average) q, or q_a, has been much more commonly used in empirical studies examining investment behaviors, mainly because it is easier to measure. It is well-known from standard neoclassical models such as Hayashi (1982) that q_a is identical to q_m in a frictionless environment. However, models with agency frictions predict a wedge between q_a and q_m, implying potential measurement error when using q_a as a proxy for q_m.[27] Moreover, empirical studies have consistently documented a large sensitivity of investment to cash flow and a small sensitivity of investment to Tobin's q. The large investment-cash-flow sensitivity is also more prominent among larger and older firms.[28] These findings cannot be reconciled with Bolton et al. (2011) and DeMarzo et al. (2012) for two reasons. First, those studies assume that the expected cash flow μ is exogenous and constant. Secondly, in those studies, both q_a and investment are monotonically increasing in the state variable. Consequently, those studies would predict a large coefficient of investment on q_a and a small coefficient on cash flow.

The prediction of this model is different from that of Bolton et al. (2011) and DeMarzo et al. (2012) but consistent with the empirical observations. To illustrate this prediction, the left panel of Figure 2 plots investment i along with q_a and the expected gross cash flow μ. All three variables are functions of X, but investment and μ are both increasing in X, while q_a is hump-shaped in X. This implies a potentially strong correlation between investment and cash flow and a weak or even negative correlation between investment and Tobin's q. Indeed, after combining (27), (28), and (33) and applying Ito's lemma, the variations in investment can be written

[26] The analysis in Erickson and Whited (2000) also corroborates, both theoretically and empirically, that marginal q being a major determinant of investment (when measured correctly).

[27] Unlike Bolton et al. (2011) and DeMarzo et al. (2012), in this paper, the wedge between q_a and q_m does not disappear for either low or high X. It does not disappear for low X because of a technical reason: while those studies assume contract termination at $W = 0$, in this model, contract termination occurs at $X = \lambda > 0$. The wedge does not disappear for high X because of the persistence of the agency friction: while in those studies the agency friction is fully resolved when W is sufficiently high, in this model, the agency friction is never fully resolved even with arbitrarily high levels of incentives.

[28] Ai et al. (2017) and Cao et al. (2019) summarize the empirical evidence for these observations and offer their explanations. In addition to cash flow shocks, Ai et al. (2017) introduce a separate productivity shock and a liquidity constraint, while Cao et al. (2019) introduce a "news shock" that allows agents to observe the realization of future productivity in advance. In a recent study, Ward (2020) proposes an alternative explanation for the investment-q sensitivity (but not the investment-cash-flow sensitivity) based on separating the agency frictions associated with intangible capital from the frictions associated with physical capital.

as the following function of the variations in q_a and μ:

$$di_t = \eta_1 dq_{a,t} + \eta_2 dY_t\,, \tag{35}$$

where

$$\eta_1 = \frac{1}{\theta} \tag{36}$$

$$\eta_2 = -\frac{\alpha}{\theta}\left[f'(X_t) + X_t f''(X_t)\right]\ . \tag{37}$$

The investment-q sensitivity is captured by η_1, which is a constant. The investment-cash-flow sensitivity is captured by η_2, which is a function of X and therefore depends on the realized cash-flow history. The relative sizes of the two coefficients are illustrated in the right panel of Figure 2. Compared to η_1, η_2 can be very large if one or more of the following is true: α, the marginal cost of productivity growth, is large; $f'(X)$ is very negative; or $Xf''(X)$ is large in absolute value (recall that $f''(X) < 0$). In particular, both a negative $f'(X)$ and a large (absolute value of) $Xf''(X)$ occur when X is high. As shown in (32), a larger ρ implies that X on average increases over time. Therefore, if the long-run negative effect of the agency frictions is sufficiently strong, the model predicts that larger and older firms, which are often thought to be less financially constrained, will typically have a higher investment-cash-flow sensitivity than would smaller and younger firms, as documented in Ai et al. (2017). Combined with the prediction in Part A that firms with a larger X also over-invest relative to the first-best, the implications of this model are also consistent with Blanchard et al. (1994), who find that firms with a strong cash flow history may overinvest despite having a low Tobin's q.

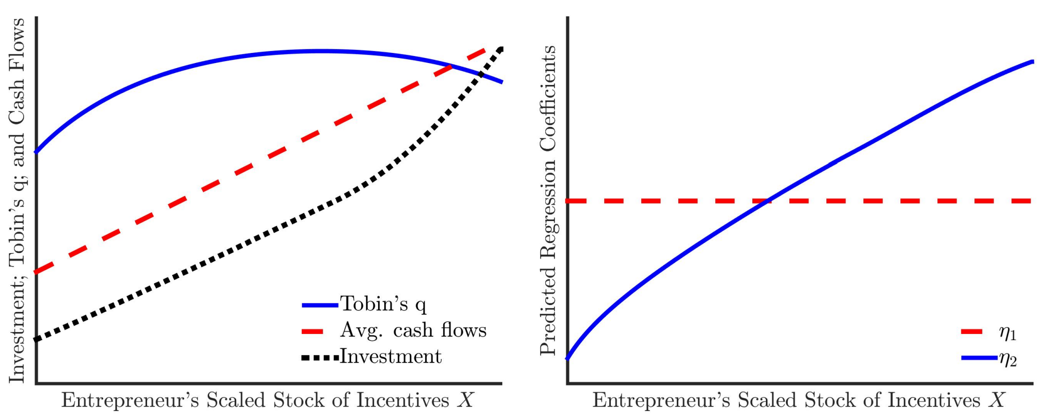

Figure 2: **Investment, q, & Cash Flow: Risk-Neutral Entrepreneur with Liquidity Constraint**
This figure presents the investment i, Tobin's q (q_a), and expected cash flows μ in the left panel, and the investment-to-q sensitivity ($\eta_1 = 1/\theta$) and the investment-to-cash-flows sensitivity (η_2) in the right panel. Parameters are the same as those in Figure 1 with $\rho = 0.01$.

3.4 Security Implementation and Asset Pricing Implications

The optimal contract derived in Proposition 2 can be implemented with common securities. This section considers one particular implementation that has received broad interests in the literature and yields unique asset pricing implications from the persistent effect of the agency frictions.

Specifically, the implementation involves three types of securities: stocks, bonds, and cash reserves. Cash reserves, denoted by J_t, accumulates (or depreciates) at the rate $\rho - r$.[29] Cash reserves must be maintained above the minimum level λK_t. Once $J_t \leq \lambda K_t$, the entrepreneur's contract is terminated and capital liquidated.Stocks make continuous dividend payout. The minimal dividend rate D_t is given by

$$dD_t = J_t \left[d\hat{Y}_t - \frac{J_t}{\alpha} - g(i_t)K_t - \left(\frac{J_t^2}{2\alpha K_t} \right) \right] + (1 - J_t)d\hat{Y}_t \,. \tag{38}$$

Conditional on meeting the minimal dividend payout, the entrepreneur has the discretion to choose the productivity growth effort μ_t (subject to the effort cost $h(\mu_t, K_t)$), investment i_t, and the amount of reported cash flow $d\hat{Y}_t$. Bonds make zero coupon payment except for their face value at the time of liquidation. The face value of the bonds held by the entrepreneur is denoted by L_t and follows

$$dL_t = rL_t + d\hat{Y}_t - dD_t \,. \tag{39}$$

Together, these securities implement the optimal contract in Proposition 2. Intuitively, both the investor and the entrepreneur share the same discount factor, so the compensation to the entrepreneur can be costlessly deferred to the time of liquidation. The IC conditions for μ_t is binding in equilibrium, and thus can be regulated with the balance of the cash reserves J_t, which is observable to the investor. Any excess cash flow reported by the entrepreneur must be accompanied by an appropriate amount of increase in dividend payout to the investor, which can only be produced if the entrepreneur exerts excess effort in addition to the amount he is compensated for, thus offsetting his private benefit from mis-reporting.

The persistent effect of the agency frictions in this model has unique implications on the pricing of the securities in the above implementation. In particular, following Biais et al. (2007), I define the *credit yield spread* Δ_t as a measure of the risk of liquidation at any time t. Δ_t solves

$$\int_t^\infty e^{-(r+\Delta_t)(s-t)} ds = \mathrm{E}\left[\int_t^T e^{-r(s-t)} ds \right] \,. \tag{40}$$

[29] Theoretically, the model can accommodate both $\rho - r > 0$ and $\rho - r < 0$. The former can be interpreted as an interest, as in DeMarzo et al. (2012), and the latter can be interpreted as a carrying cost, as in Bolton et al. (2011).

Because liquidation occurs when J_t reaches λK_t, Δ_t is state-dependent, and can be written as a function of J_t, or $\Delta(J_t)$, where

$$\Delta(J) = \frac{rle^{\left(\frac{\rho-r}{r}\right)(J-\lambda)}}{1 - le^{\left(\frac{\rho-r}{r}\right)(J-\lambda)}}. \tag{41}$$

Figure 3 plots $\Delta(J)$ for different levels of ρ. As it shows, for any fixed J, $\Delta(J)$ is higher when ρ is larger. In other words, the credit yield spread implied by the security implementation of this model is wider when the long-run marginal effect of the agency frictions is stronger. Admittedly, similar to DeMarzo et al. (2012), DeMarzo and Sannikov (2006) and Biais et al. (2007), the implementation is not unique. Moreover, the current model is built on the dynamic agency theories of investment in corporate finance and is not well-posed to be directly taken for asset pricing tests. Nevertheless, the model offers a potential rationale for the long-debated credit spread puzzle from a the new perspective: the persistence of the agency friction between insider managers and outside investors.

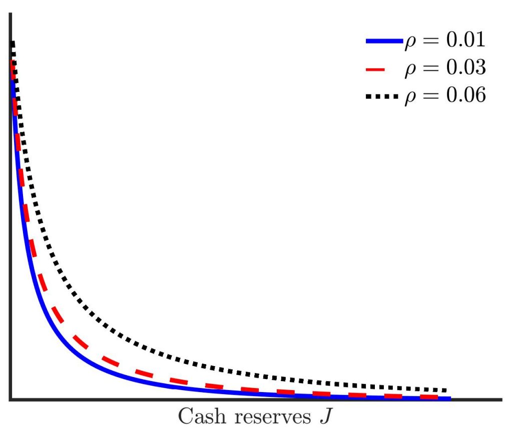

Figure 3: **Credit Yield Spread**

This figure presents the credit yield spread $\Delta(J)$. Parameter values (r, λ, l) are the same as those in Figure 1. The blue solid line, red dashed line, and black dotted line correspond to $\rho = 0.01$, $\rho = 0.03$, and $\rho = 0.06$, respectively.

3.5 Section Summary

When the entrepreneur is risk-neutral and faces a liquidity constraint, the optimal contract can be characterized with an ODE of a single state variable $X = -P/K$, which represents the stock of future incentives for the entrepreneur per unit of capital. The optimal investment policy and the expected growth rate of productivity are both increasing functions of X. The model generates a number of empirically relevant implications: first, over-investment relative to the first-best is possible, especially among firms with a strong history of cash flow but a low Tobin's q. Second, there is a potentially high investment-cash-flow sensitivity and relatively low investment-q sensitivity. Both implications are consistent with empirical observations but are absent from from existing agency-based dynamic investment models (e.g., Bolton et al. (2011) and DeMarzo et al. (2012)) without the persistent impact of agency frictions. If the optimal contract is implemented with stocks, bonds, and cash reserves, the implied credit yield spread can be substantially wide especially with a strong persistent effect of the agency frictions.

4 Extensions and Discussions

This section offers some extensions of the baseline model. I demonstrate that the model can be easily modified to accommodate the setting in which the investor's objective is to maximize the valuation of the business, and the setting in which the entrepreneur is risk averse. The latter also offers some intuitions on the implications of the unobservable cash flow assumption.

4.1 A Valuation Model

In the baseline model, the output is assumed to be cash flows produced by the entrepreneur's technology. As discussed in Section 2.2, this is both a common choice in the literature and a popular objective of the investors in practice. However, the model can be easily modified to study the optimal investment policies when the investor's object is to maximize the valuation of the business.

The valuation model requires slightly different assumptions than those made for the cash-flow model. Let V_t be the total value of the business given by

$$V_t = A_t K_t \,, \tag{42}$$

where A_t, instead of representing the productivity, now measures the (gross) *return of capital* (ROC). Because returns are most naturally modeled with a geometric Brownian motion, A_t is assumed to evolve according to

$$dA_t = A_t \left(\mu_t - \rho M_t dt + \sigma dZ_t\right) \,. \tag{43}$$

The exact same agency frictions in the baseline model still exit for the valuation model: the investor does not observe the true return A_t or the total valuation V_t. The entrepreneur reports the values $\hat{A}_t$ and $\hat{V}_t$, while he privately controls the expected return μ_t and can divert firm value (B_t) to generate private benefit.[30] The "stock of misconduct" M_t follows the same definition in (39) as the difference between the reported firm value $\hat{V}_t$ and its true value V_t plus the amount

[30]For example, using depository shares of Luckin Coffee as collateral, the CEO and the chairman of the board of Luckin both borrowed hundreds of millions dollars from Goldman Sacks and other financial institutions for their personal investments elsewhere (e.g., in Car Inc.). Following Luckin's accounting scandal and the subsequent collapse of its share prices, they failed to pay back those loans.

of diversion:

$$dM_t = d\hat{V}_t - dV_t + B_t dt\,. \tag{44}$$

A contract is defined according to Definition 1: the investor decides the investment policy I_t and compensation to the entrepreneur C_t based on the entrepreneur's reports. Incentive compatibility follows Definition 2, involving no shirking, misreporting, or diversion of firm value. The entrepreneur's problem is similar to (11), subject to the true and reported paths of K, A, and V. The investor's problem is similar to (20), and the optimal contract follows Definition 3.

In the subsequent analysis, I assume that the entrepreneur is risk-neutral with a liquidity constraint as in Section 3.1. This setting delivers the most interesting and empirically relevant implications. To ensure tractability, some modifications of the assumptions made in the previous sections are necessary:

Assumption 4 *The entrepreneur is risk-neutral with a liquidity constraint. His utility function is given by $u_t = u(c_t, b_t, \mu_t)$ where*

$$u(c, b, \mu) = \left[c + \lambda b - \frac{\alpha}{2}\mu^2\right] V\ . \tag{45}$$

The adjustment cost of capital is given by $G(I, V) = g(i)V$, where $g(i) = i + \frac{\theta}{2}i^2$. The lower case letters c, b, and i, all represent the value of their corresponding capital letters scaled by firm value V. When the contract terminates, the business is liquidated at a discounted value lV with $l < 1$.

These assumptions are made such that the value function, or Tobin's average q of the business, can still be characterized with a single state variable $X \equiv P/V$. The following proposition, which is similar to Proposition 2, summarizes the optimal contract and the resulting policies:

Proposition 3 *Under Assumption 4, the investor's value function $F(V, A, P, W) = f(X)V - W$, where $f(X)$ solves the following HJB equation:*

$$rf(X) = \mu(X) + i - \delta - g(i(X)) - c(X) + (i(X) - \delta)\left[f(X) - Xf'(X)\right] + (\rho - r)Xf'(X)\ , \tag{46}$$

with boundaries condition $f(\lambda) = l$. The investor's optimal policies $\{i_t, \mu_t, c_t\}$ are all functions of X_t given by

$$i(X) = \frac{f(X) - Xf'(X)}{\theta}\ , \tag{47}$$

$$\mu(X) = \frac{X}{\alpha}\ , \tag{48}$$

$$c(X) = \frac{X^2}{2\alpha}\ . \tag{49}$$

The implications of Proposition 3 are illustrated in Figure 4. The left panel plots the scaled value function $f(X)$, which also equals the Tobin's (average) q of the business defined as $q_a \equiv (F+W)/V = f(X)$. The right panel plots the equilibrium investment $i(X)$, expected return μ, and the first-best level of investment. Similar to the implications in Section 3.3, investment can be higher or lower than the first-best level, which explains the hump-shape of the value function. Both the investment and expected return of capital are increasing in the state variable X, while q_a is non-monotonic in X. Therefore, the model predicts a strong correlation between investment and the expected return of capital and a weak or even negative correlation between investment and Tobin's q.

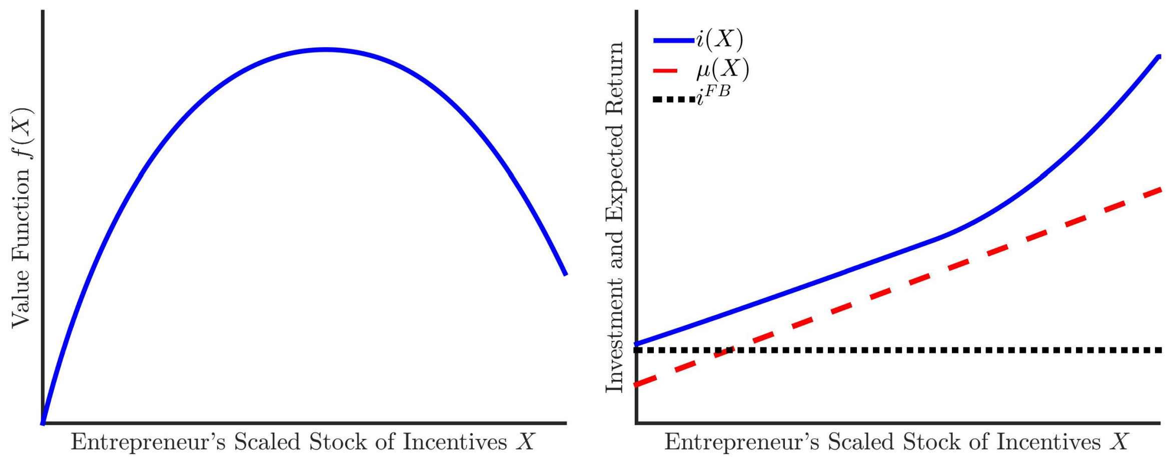

Figure 4: **Optimal Contract: the Valuation Model**

The left panel of this figure plots the scaled value function $f(X)$. The right panel plots the equilibrium investment $i(X)$, expected return of capital μ, and the first-best level of investment. Parameters are $r = 0.1$, $\sigma = 0.5$, $\delta = 0.2$, $\lambda = 0.1$, $\theta = 1$, $\alpha = 1$, $l = 0.9$, and $\rho = 0.01$.

4.2 Risk-averse Entrepreneur

In this section I explore the optimal financing and investment contract under the assumption of a risk-averse entrepreneur. Following the literature, I consider the combination of constant-absolute-risk-aversion (CARA) utility and hidden savings. The specific assumptions used in this section are presented in 4.2.1. The resulting optimal contract is solved in 4.2.2, followed by discussions in 4.2.3.

4.2.1 Assumptions

To highlight the implications of risk-aversion while maintaining tractability, I assume that the entrepreneur has CARA utility and hidden savings specified as follows:

Assumption 5 *The entrepreneur's flow utility u_t is given by*

$$u_t = -\frac{1}{\gamma} e^{-\gamma[c_t + \lambda b_t - \frac{\alpha}{2}\mu_t^2]K_t} \ , \tag{50}$$

where $\gamma > 0$ measures the degree of risk aversion. The entrepreneur also has a private savings account to which he can add funds or from which withdraw consumption. The account has a balance S_t and grows at a rate r, such that

$$dS_t = rS_t + (\hat{C}_t - C_t)dt \ , \tag{51}$$

where $\hat{C}_t$ is the amount of consumption allocated to the entrepreneur from the investor and C_t is the entrepreneur's actual consumption.[31]

Due to its lack of the wealth effect, CARA utility is a commonly adopted assumption in models with persistent private information (e.g. Williams (2015), He et al. (2017), and Marinovic and Varas (2019)) for generating tractable results. All of these studies also assume private savings, which allows the entrepreneur to smooth consumption intertemporally and restricts the ability of the investor to implement compensation and investment schemes that lead to steep expected consumption patterns. Another natural implication of private savings is that the entrepreneur can use his savings to boost the cash flow of the business without engaging in other misconduct. To save notation, I assume this is equivalent to negative cash-flow diversion, or $b_t \leq 0$. Finally, to provide a boundary condition, I assume the investor can "reset" the contract at the marginal cost of $\epsilon > 0$ when $P_t = 0$. That is, $F_P(K, P = 0, W) = \epsilon$. Note that this assumption is made in a heuristic form for illustration purpose only, although it can be more rigorously micro-founded, for example via costly refinancing (Bolton et al., 2011) or randomization.

Given Assumption 5, the entrepreneur's optimization problem in (11) entails an additional control variable C_t and an additional state variable S_t with the law of motion given by (51). The same additional control and state variables also apply to the investor's problem (20). An incentive compatible contract, as in Definition 2, entails the additional requirement for no private savings: that is, $\hat{C}_t = C_t$ and $S_t = 0$ for all $t \geq 0$. This additional requirement implies an IC constraint besides those presented in Proposition 1 and plays a vital role in simplifying the optimal contract, as discussed next.

[31]The notation $\hat{C}_t$ suggests that it also represents the amount of *reported* consumption by the entrepreneur to the investor, because it is assumed throughout the paper that the entrepreneur never reports any deviation from the investor's recommended strategies.

4.2.2 Incentive Compatibility and the Optimal Contract

Assumption 5 implies the following results from the entrepreneur's problem:

$$\gamma\lambda u_t K_t = P_t \tag{52}$$

$$\gamma\alpha u_t \mu_t K_t = P_t \tag{53}$$

$$u_t = rW_t \ . \tag{54}$$

Equations (52) and (53) are the IC constraints for no cash-flow diversion and no shirking, respectively. Their interpretations are identical to their counterparts (23) and (24) in the risk-neutral entrepreneur case: that is, the marginal cost of those suboptimal actions, captured by P_t, must equal their marginal benefit. Equation (54) comes from the IC condition for no private savings. It implies that the entrepreneur's current utility is proportional to his continuation utility, which is the critical simplifying result for adopting the assumptions of CARA utility and private savings. Consequently, the IC constraints (52) and (53) can both be written as linear functions of K_t and a new state variable X_t, now defined as

$$X_t \equiv \frac{P_t}{W_t} > 0 \ . \tag{55}$$

X_t represents the stock of future incentives per unit of continuation utility. Unlike the case explored in Section 3, the continuation utility and marginal utility of consumption of a entrepreneur with CARA utility play important roles in the characterization of the optimal contract. Nevertheless, because the investor is still risk-neutral, her problem turns out to be still homogeneous in K_t and can be summarized with X_t as the single state variable as follows:

Proposition 4 *Under Assumption 5, the investor's value function $F(K,P,W) = f(X)K$, where $f(X)$ solves the following HJB equation:*

$$rf(X) = \mu - g(i(X)) - c + (i-\delta)Xf'(X) \ , \tag{56}$$

with boundary conditions $f'(0) = \epsilon$. The investor's optimal policies $\{i_t, \mu_t, c_t\}$ are all functions of X_t, given by

$$i(X) = r - \rho + \delta + \sigma X - (\sigma X)^2 \tag{57}$$

$$\mu(X) = \frac{X}{\alpha\gamma r} \tag{58}$$

$$c(X) = \left(\frac{1}{2r\gamma\alpha} + \sigma^2\right)\frac{X^2}{r} \ . \tag{59}$$

Proposition 4 features several results unlike those discussed in Section 3. In particular, all of the policy variables (i, μ, c) are functions of the state variable X only; that is, they are independent of the investor's value function $f(X)$. In other words, all equilibrium policies are determined by the entrepreneur's IC conditions. Moreover, (57) indicates that investment is hump-shaped in X, while (58) and (59) combined imply that the expected net cash flow $\mu - c$ is also hump-shaped in X. The implications of these results are discussed in the next subsection.

It is worth noting that the equilibrium productivity growth rate μ is a linear function of X. This is in contrast to Marinovic and Varas (2019), in which the optimal compensation contract can achieve the constant, first-best level of effort when the contracting horizon is infinite. This is because unlike Marinovic and Varas (2019), the investor does not observe the cash flow and must rely on the entrepreneur's report of it. The necessary incentives for no misreporting determines a specific path of the pay-performance sensitivity, which in turn drives a specific relationship between effort and the state variable. Section 4.2.3 discuss in details the implications of the model when investor can directly observe the cash flow. In that case, the resulting contract resembles the contract in Marinovic and Varas (2019) much more closely.

The optimal financing policies obtained in Proposition 4 yields empirical implications comparable to those discussed in Section 3.3. In particular, numerical examples show that both overinvestment and underinvestment can occur in the equilibrium, which is in contrast to existing models with a risk-averse agent (e.g., Hirshleifer and Thakor (1992)) that predict only underinvestment as the result of the agency frictions. Moreover, the value of the business and the intensity of investment vary with the entrepreneur's degree of risk-aversion γ. γ negatively affects the equilibrium expected productivity growth rate, as shown in equation (58). When X is small and μ is too low compared to the first-best level, a higher γ exacerbates the underinvestment problem in productivity growth, resulting in a lower $f(X)$ for the investor. In contrast, when X is large and μ is too high compared to the first-best level, a higher γ alleviates the over-investment problem in productivity growth, resulting in a higher $f(X)$ for the investor. Because X varies according to the history of cash flows, these dynamics imply an empirically testable hypothesis: the correlation between a firm's value and its cash-flow history is stronger (weaker) when the entrepreneur is less (more) risk-averse.

4.2.3 The Role of Unobservable Cash Flows

The most critical assumption made in this paper is that of the entrepreneur's "black box" technology. The investor is assumed to be unable to observe the output (the cash flows) directly and must rely on the entrepreneur's report of it. In this section, maintaining the assumptions of CARA utility and hidden savings, I highlight the changes in the results of the model if cash flows become directly observable. The following proposition summarizes the IC conditions under this new setting:

Proposition 5 *Given Assumptions 5, and the assumption that the investor observes the realization of cash flows, there exists $\mathscr{F}$-adapted processes $\{\phi_t, \beta_t\}$ such that, under the incentive-compatible contract,*

$$dP_t = (r-\rho) P_t dt - \phi_t P_t dZ_t \ , \tag{60}$$

$$dW_t = -\beta_t W_t dZ_t \ . \tag{61}$$

Let $X_t \equiv P_t / W_t$. The necessary incentive compatibility conditions are

$$\alpha\gamma r \mu_t K_t = \frac{\beta_t}{\sigma} \ , \tag{62}$$

$$\lambda\gamma r K_t = X_t - \frac{\beta_t}{\sigma} \ . \tag{63}$$

The IC conditions in Proposition 5 are different from those in Proposition 4 but resemble those in Marinovic and Varas (2019), who also utilizes CARA utility and hidden savings. Because cash flow is observable, both shirking and cash flow diversion have the direct impact of lowering the current cash flow and therefore, lowering the entrepreneur's continuation utility. Consequently, the entrepreneur is more constrained in his private actions: if he diverts cash flows but does not want to suffer immediate punishment in his continuation utility, he must exert higher effort in raising the productivity growth rate. [32] In particular, equation (62) states that the entrepreneur's effort is fully determined by the *current* pay-performance sensitivity β_t, which is similar to the static tradeoff between effort and continuation utility in standard moral hazard models without persistent private information (e.g., DeMarzo and Sannikov (2006), Biais et al. (2007), Sannikov (2008), etc). Meanwhile, cash-flow diversion has the further cumulative impact on the long-term cash-flow growth. However, because the entrepreneur cannot misreport cash flows, the size of the investment K_t is bounded by not only the stock of future incentives but also the current incentive. When X_t is too low (e.g., because the entrepreneur's continuation utility W_t is too high), inefficient termination may be required, as the stock of incentives is not sufficient to support the investment of even the smallest business.

Despite the resemblance of the *agent's* problem between this version of the model and that studied in Marinovic and Varas (2019), the *principal's* problem remains meaningfully different. In addition to the design of managerial compensation in Marinovic and Varas (2019), the investor in this model also designs the optimal productivity growth and financing policies. The underlying mechanisms driving those policies are similar to those analyzed in Proposition 4 while taking into account the different IC conditions presented in Proposition 5.

[32] Feng and Westerfield (2021) adopt a similar technique to study the private choice of volatility in a continuous-time dynamic principal-agent model. In that model, the visibility of cash flow means that its total volatility can be inferred by the principal. However, the agent has private control over the *composition* of the cash-flow volatility and must be offered appropriate incentives to not deviate from the composition that the principal desires.

5 Conclusion

Financing and investment decisions are critical for the growth and prosperity of a modern economy. However, those decisions are faced with new challenges given the recent development of technology and globalization. On the one hand, the success of a modern business relies heavily on adopting cutting-edge technologies. On the other hand, the rapid development of modern technologies, often taking place at various locations in the world, has created an information barrier between investors and entrepreneurs. Certain businesses and investment opportunities may be black boxes to general investors: while investors control their investment and receive reports on the outputs, they have little knowledge of the internal production process of those reports and how accurately they reflect the true productivity and return to their investments. This information barrier provides a breeding ground for agency frictions, which can result in not only instantaneous social costs but also persistent impact on a firm's long-term growth.

This paper aims to understand the implications of opaque production technologies on firms' financing and investment behaviors. Based on a dynamic investment model with persistent private information, I demonstrate various implications that are absent in standard models without the persistence of private information but are consistent with empirical observations. Moreover, the model illustrates the critical roles played by certain economic factors that are usually omitted in standard investment and financing theories. Those factors include the investor's ability to directly observe the output and the degree of persistence of the impact of agency frictions on future firm value.

There are several directions in which this model might be fruitfully extended. The most natural is perhaps relaxing the investor's commitment requirement. To maintain tractability, the current model assumes that the investor can fully commit to her investment policies, even when under- or over-investment becomes excessively inefficient and the resulting firm value is extremely low. One may consider a model in which the investor has the ability to walk away from the business in those situations, thus connecting this study to the literature of limited commitment on the principal's side.[33] More generally, an explicit, detailed analysis of the investor's exiting strategy through the means of liquidation, merger and acquisition, IPO, etc., may yield interesting new insights. I leave such analyses for future research.

[33] As in e.g. Ai and Li (2015), Miao and Zhang (2015), Ai et al. (2020), Feng (2020), etc.

Appendix A. Proofs

Proof of Proposition 1

The proof follows the stochastic maximum principle technique with a change of measure developed in Williams (2011, 2015), and Marinovic and Varas (2019). Let $\mathbb{P}$ be the probability measure under the entrepreneur's actions and $\hat{\mathbb{P}}$ be the probability measure induced by his report, there exists a process η_t such that the Radon-Nikodym derivative between $\hat{\mathbb{P}}$ and $\mathbb{P}$ is given by

$$\xi_t \equiv \frac{d\hat{\mathbb{P}}}{d\mathbb{P}} = \exp\left(-\frac{1}{2}\int_0^t \eta_s^2 ds + \int_0^t \eta_s dZ_s\right) . \tag{A-1}$$

This implies

$$d\xi_t = \eta_t \xi_t d\hat{Z}_t \tag{A-2}$$
$$dZ_t = -\eta_t dt + d\hat{Z}_t . \tag{A-3}$$

This technique allows me to evaluate the entrepreneur's expected payoff from any deviation on the probability measure induced by $\hat{Z}_t$. The entrepreneur's problem is therefore

$$\max \quad \mathrm{E}^{\hat{Z}}\left[\int_0^\tau e^{-rt}\xi_t u\left(C_t + \lambda B_t - \frac{\alpha}{2}\hat{\mu}_t^2 K_t\right) dt + e^{-r\tau}\xi_\tau W_\tau\right] , \tag{A-4}$$

subject to

$$d\xi_t = \eta_t \xi_t d\hat{Z}_t \tag{A-5}$$
$$dM_t = \left(\hat{\mu}_t - \mu_t - \rho M_t + B_t - \eta_t \sigma\right) dt , \tag{A-6}$$

where $\{\mu_t, B_t, \eta_t\}$ are the entrepreneur's choice variables and $\hat{\mu}_t$ is the investor's report (which obviously must equal the recommended action). This optimization problem can be written as the following (first-order current value) Hamiltonian system:

$$\mathscr{H} = \xi u\left(C + \lambda B - \frac{\alpha}{2}\hat{\mu}^2 K\right) + q^\xi \eta \xi + p^M\left(\hat{\mu} - \mu - \rho M + B - \eta\sigma\right) , \tag{A-7}$$

with adjoint variables p_t^M and p_t^M following the laws of motion at the optimum ($\mu = \hat{\mu}$, $B = \eta = M = 0$, $\xi = 1$):

$$dp^M = rp^M dt - \left(\frac{\partial \mathscr{H}}{\partial M}\right) dt + q^X d\hat{Z} = rp^M dt - \rho p^M dt + q^M d\hat{Z} \tag{A-8}$$
$$dp^\xi = rp^\xi dt - \left(\frac{\partial \mathscr{H}}{\partial \xi}\right) dt + q^\xi d\hat{Z} = rp^\xi dt - u_t dt + q^\xi d\hat{Z} . \tag{A-9}$$

These adjoint variables can be interpreted by integrating their laws of motion to obtain

$$p_t^\xi = \mathrm{E}_t\left[\int_t^T e^{-r(s-t)} u_s ds + e^{-r(T-t)} W_T\right] , \tag{A-10}$$

which represents the entrepreneur's continuation utility, and can thus be denoted W_t, following the dynamic contracting literature conventions. W_t is a martingale and by the martingale representation theorem, there exists a $\mathscr{F}$-adapted process β_t such that

$$dW_t = rW_t dt - u_t dt - \beta_t W_t dZ_t \ . \tag{A-11}$$

Similarly, p_t^M is the solution to

$$p_t^M = \mathrm{E}_t \left[\int_t^T e^{-r(s-t)} \rho p_s^M ds + e^{-r(T-t)} p_T^M \right] < 0 \ , \tag{A-12}$$

which represents the discounted marginal value of the persistent impact of misconduct. Define $P_t \equiv p_t^M$ and by the martingale representation theorem, there exists a $\mathscr{F}$-adapted process ϕ_t such that

$$dP_t = (r-\rho)P_t dt - \phi_t P_t dZ_t \ . \tag{A-13}$$

Finally, the Hamiltonian system yields the following first-order conditions evaluated at the optimum:

$$\mathscr{H}_B: \quad u_B + P \le 0 \tag{A-14}$$
$$\mathscr{H}_\mu: \quad u_\mu - P = 0 \tag{A-15}$$
$$\mathscr{H}_\eta: \quad \beta_t = \sigma \frac{P_t}{W_t} \ , \tag{A-16}$$

where the last equation uses the fact that $q^\xi = -\beta W$. These first-order conditions correspond to (16), (17), and (18), respectively. ☐

Proof of Proposition 2

Given Assumptions 1 and 2, the entrepreneur's first-order conditions can be re-written as

$$\lambda K + P \le 0 \Rightarrow \lambda \le X \tag{A-17}$$
$$\alpha \mu K = P \Rightarrow \mu = \frac{X}{\alpha} \ , \tag{A-18}$$

where $X \equiv P/K$. The investor's HJB equation (21) becomes

$$\begin{aligned} rF(K,P,W) = \max_{\mu,c,\phi,\beta} & \big(\mu - g(i) - c\big) K + (i-\delta)KF_K + (r-\rho)PF_P + \frac{1}{2}\phi^2 P^2 F_{PP} \\ & + (r-u)WF_W + \frac{1}{2}\beta^2 W^2 F_{WW} - \phi\beta PWF_{PW} \ , \end{aligned} \tag{A-19}$$

with constraints (A-16), (A-17), (A-18), and $c \ge \alpha\mu^2 K/2$. Conjecture that $F(K,P,W) = f(X)K - W$, then $F_K = f(X) - Xf'(X)$, $F_P = -f'(X)$, $F_W = -1$, $F_{WW} = F_{PW} = 0$, which imply $\phi = 0$ and $c = \alpha\mu^2 K/2$. Substituting these terms into the investor's HJB equation implies $f(X)$ solves:

$$rf(X) = \max_i \ \big(\mu - g(i) - c\big) + (i-\delta)(f(X) - Xf'(X)) + (r-\rho)f'(X) \ , \tag{A-20}$$

with boundary condition $f(\lambda) = l$. The first-order condition for i yields:

$$g'(i(X)) = 1 + \theta i = f(X) - Xf'(X) \Rightarrow i(X) = \frac{f(X) - Xf'(X) - 1}{\theta} \ . \tag{A-21}$$

☐

Proof of Proposition 3

The proof is very similar to that for Proposition 2. Thus, only the differences will be highlighted below. Under Assumption 4, the entrepreneur's IC conditions (A-14) and (A-15) become

$$\lambda V \leq -P \tag{A-22}$$
$$\alpha\mu V = P \ , \tag{A-23}$$

while (A-16) stays the same. Consequently, let $X \equiv -P/V$, the investor's HJB equation (21) becomes

$$\begin{aligned} rF(V,A,P,W) = \max_{\mu,c,\phi,\beta} \left(\mu + i - \delta - g(i) - c\right) V + (i-\delta) V F_V + (r-\rho) P F_P + \frac{1}{2}\phi^2 P^2 F_{PP} + \\ + \mu A F_A + (r-u) W F_W + \frac{1}{2}\beta^2 W^2 F_{WW} - \phi\beta P W F_{PW} \ , \end{aligned} \tag{A-24}$$

with constraints (A-16), (A-22), (A-23), and $c \geq \alpha\mu^2 V/2$. Conjecture that $F(V,A,P,W) = f(X)V - W$, then $F_V = f(V) - Vf'(V)$, $F_P = -f'(V)$, $F_W = -1$, $F_{WW} = F_{PW} = 0$. Together they imply $\phi = 0$ and $c = \alpha\mu^2 V/2$. Substituting these terms into the investor's HJB equation implies that $f(X)$ solves

$$rf(X) = \max_i \ \left(\mu + i - \delta - g(i) - c\right) + (i-\delta)(f(X) - Xf'(X)) + (r-\rho) f'(X) \ , \tag{A-25}$$

with boundary condition $f(\lambda) = l$. The first-order condition for i yields:

$$g'(i) = 1 + \theta i = 1 + f(X) - Xf'(X) \Rightarrow i = \frac{f(X) - Xf'(X)}{\theta} \ . \tag{A-26}$$

□

Proof of Proposition 4

Under Assumption 5, the entrepreneur's problem (A-4) has an additional state variable S_t where

$$dS_t = rS_t dt + (\hat{C}_t - C_t) dt \ , \tag{A-27}$$

where $\hat{C}_t$ represents the entrepreneur's reported level of consumption (which obviously must equal the investor's recommended level). The Hamiltonian system (A-7) now becomes

$$\mathscr{H} = \xi u \left(C + \lambda B - \frac{\alpha}{2}\hat{\mu}^2 K\right) + q^{\xi}\eta\xi + p^M \left(\hat{\mu} - \mu - R(M) + B - \eta\sigma\right) + p^S (rS + \hat{C} - C) \ , \tag{A-28}$$

with an additional adjoint variable p_t^S following the law of motion at the optimum ($\mu = \hat{\mu}$, $C = \hat{C}$, $B = \eta = M = S = 0$, $\xi = 1$):

$$dp^S = rp^S dt - \left(\frac{\partial \mathscr{H}}{\partial S}\right) dt + q^S d\hat{Z} = q^S d\hat{Z} \ , \tag{A-29}$$

and an additional first-order condition for C:

$$\mathscr{H}_C : \quad -\gamma u + p^S = 0 \ . \tag{A-30}$$

He (2011), He et al. (2017), Williams (2015) and Marinovic and Varas (2019) show that (A-29) and (A-30) combined imply that $u = rW$. Consequently, the first order conditions (A-14), (A-15), and (A-16) become

$$\gamma\lambda r K_t = X_t \tag{A-31}$$

$$\gamma\alpha r \mu_t K_t = X_t \tag{A-32}$$

$$\beta_t = \sigma X_t \ , \tag{A-33}$$

where $X_t \equiv P_t / W_t$, and (15) becomes

$$dW_t = -\beta W_t dZ_t \ . \tag{A-34}$$

By Ito's lemma,

$$dX_t = \left[(r-\rho) + \beta_t - \phi_t \beta_t\right] X_t dt + \left(\beta_t - \phi_t\right) X_t dZ_t \ . \tag{A-35}$$

Therefore, (A-31) and (A-33) imply that $\phi_t = \sigma X_t$, and

$$i_t = r - \rho + \delta + \sigma X_t - (\sigma X_t)^2 \ . \tag{A-36}$$

Finally, following the same calculation in Williams (2015) and Marinovic and Varas (2019),

$$c_t = \frac{1}{r}\left(\frac{1}{2r\gamma\alpha} + \sigma^2\right) X_t^2 \ . \tag{A-37}$$

Substituting all results above into the investor's HJB equation (21) implies that the investor's value function $F(K, P, W) = f(X)K$, where $f(X)$ solves

$$r f(X) = \mu(X) - g(i(X)) - c(X) + (i(X) - \delta) X f'(X) \ , \tag{A-38}$$

and $\mu(X), i(X), c(X)$ are given by (A-32), (A-36), and (A-37), respectively. □

Proof of Proposition 5

When Y_t is observable, $d\hat{Y}_t = dY_t$, and the measure difference η_t defined in (A-1) must satisfy

$$\eta_t = -\frac{\hat{\mu}_t - \mu_t + b_t}{\sigma} \ . \tag{A-39}$$

η is no longer a choice variable for the entrepreneur's maximization problem (A-4). Under CARA utility and hidden savings, the Hamiltonian system (A-7) becomes

$$\mathcal{H} = \xi u(c, b, \mu) - q^{\xi}\left(\frac{\hat{\mu} - \mu + b}{\sigma}\right)\xi + p^M b + p^S (rS + (\hat{c} - c)K) \ . \tag{A-40}$$

The first-order conditions for $\{b, \mu\}$ evaluated at the optimum ($\mu = \hat{\mu}, c = \hat{c}, b = m = S = 0, \xi = 1$) are

$$\mathcal{H}_b: \quad \gamma\lambda u K = \frac{q^{\xi}}{\sigma} - p^M \tag{A-41}$$

$$\mathcal{H}_\mu: \quad \gamma\alpha\mu K = \frac{q^{\xi}}{\sigma} \ . \tag{A-42}$$

The same argument in the proof of Proposition 4 implies that $u = rW$ and $q^{\xi} = -\beta W$. Therefore, let $X_t \equiv p^M / W_t$, the first-order conditions above become

$$\mathcal{H}_b : \quad \gamma\lambda r K = X - \frac{\beta}{\sigma} \tag{A-43}$$

$$\mathcal{H}_\mu : \quad \gamma\alpha r \mu K = \frac{\beta}{\sigma} \ . \tag{A-44}$$

□

Appendix B. Sufficiency of the IC Conditions

The IC conditions in Propositions 1, 2 and 4 are given in the forms of necessary conditions. Williams (2011) provides a set of sufficient conditions for models involving persistent private information under relatively general assumptions. However, as noted in Williams (2011), in many cases those sufficient conditions are "overly stringent or difficult to verify", and the verification typically require re-solving the agent's problem via numerical methods given the contract derived under the necessary conditions.

In light of these observations, in this section I mainly discuss the sufficiency of Propositions 2. Although they are derived under different entrepreneur's utility functions, in terms of their sufficiency, the same basic idea applies: that is, taking the contract as given, using the entrepreneur's utility and marginal utility under the contract to bound the entrepreneur's utility gains from any deviation.

For the risk-neutral entrepreneur, let $U(P,W,M)$ denote his value function from a deviation, under Assumptions 1 and 2 , $U(P,W,M)$ solves the following HJB equation:

$$\begin{aligned} 0 = \max_{b,\mu,\eta} \ & \lambda b K - \frac{\alpha}{2}\left(\mu^2 - \hat{\mu}^2\right) K - rU + (r - \sigma\eta\phi) P U_P + (rW - \sigma\eta P) U_W \\ & + \left(b + \hat{\mu} - \mu + \sigma\eta\right) U_M + \phi^2 P^2 U_{PP} + \beta^2 W^2 U_{WW} + \beta\phi P W U_{PW} \ . \end{aligned} \tag{A-45}$$

It is easy to verify that $U = W + PM$. Therefore, $F_{PP} = F_{WW} = 0$, $U_P = M$, $U_W = 1$, and $U_M = P$. Substituting these results with the HJB equation yields

$$0 = \max_{b,\mu,\eta} \ \lambda b K - \frac{\alpha}{2}\left(\mu^2 - \hat{\mu}^2\right) K - \sigma\eta\phi P M - \sigma\eta P + \left(b + \hat{\mu} - \mu + \sigma\eta\right) P \ . \tag{A-46}$$

The first order conditions for b, μ are

$$\lambda K + P \le 0 \tag{A-47}$$

$$\alpha\mu K + P = 0 \ . \tag{A-48}$$

which implies $b = 0$ and $\mu = \hat{\mu}$ if (23) and (24) are satisfied. Furthermore, $\eta = 0$ because $\phi = 0$ under the optimal contract (see the proof of Proposition 2).

The argument above ensures that the entrepreneur does engage in intertemporal misconduct for the duration of the contract. Ensuring that he does not engage in misconduct at the time of contract termination (e.g. by reporting an infinitely large amount of output right before termination) requires a certain information structure and compensation policy through which the stock of misconduct can be unwound at termination. For example, suppose the investor receives a contractible binary signal that is either positive or negative, and can design the terminal payment to the entrepreneur based on the signal received. The probability of the negative signal is $\Pi(M_T)$, where $\pi_M = \Pi_M > 0$, and pays the entrepreneur either $\zeta_h W_T$ or $\zeta_l W_T$ ($\zeta_h > \zeta_l$) based on the signal. Intuitively, the signal can be be interpreted as an external audit when the investor liquidate or exit the business. The auditing technology is imperfect, but the more misconduct the entrepreneur has accumulated at the time of termination, the more likely he generates the negative signal and receive the lower payment. Then, the necessary conditions in Proposition 2 are sufficient as long as

$$W_T = \zeta_l W_T \Pi_0 + \zeta_h W_T (1 - \Pi_0) \tag{A-49}$$

$$P_T = \frac{\partial}{\partial M} [\zeta_l W_T \Pi_0 + \zeta_h W_T (1 - \Pi_0)] \tag{A-50}$$

for any W_T, P_T. That is, the entrepreneur's expected payment is exactly W_T and his marginal value of misconduct is exactly P_T. Recall the definition of $X = -P/W > 0$, (A-49) and (A-50) imply that

$$\zeta_h = 1 + \left(\frac{\Pi_0}{\pi_0} \right) X_T \tag{A-51}$$

$$\zeta_l = 1 - \left(\frac{1 - \Pi_0}{\pi_0} \right) X_T \tag{A-52}$$

Verifying the sufficiency of Proposition 4 follows a very similar procedure above utilizing the same technique as in Williams (2015) and Marinovic and Varas (2019), both of which study continuous-time dynamic moral hazard models of persistent private information with CARA utility and hidden savings. In particular, the arguments in Williams (2015) Section 5.4. resemble (A-45), (A-46), and the ensuing first-order conditions for the entrepreneur's intertemporal deviation strategies. The details are thus omitted in the interest of space. Meanwhile, Marinovic and Varas (2019) Section V.A. introduces a feasible compensation policy at the time of contract termination that achieves the same purpose of the analysis around (A-49) to (A-52). They assume that the optimal contract deterministically relaxes P_T over a fixed vesting period. In the setting of Proposition 4, this is equivalent to assuming that after contract termination at T, the terminal payment to the entrepreneur is vested gradually over a period of length $\tau > 0$ via a stream of consumption that is conditional on the realization of the output. Critically, the stock of misconduct accumulated up to time T still bears a negative, persistent impact on the drift of the output in this vesting period. Under CARA utility the investor's optimization problem at T can be formulated as the following cost-minimizing problem:

$$\max_{\beta_t} \int_T^{T+\tau} e^{-r(t-T)} \frac{\beta_t^2}{2r\gamma} dt \tag{A-53}$$

$$\text{s.t. } P_T = \int_T^{T+\tau} e^{-r(t-T)} \frac{\beta_t}{\sigma} W_T dt \tag{A-54}$$

where β_t is the entrepreneur's pay-performance-sensitivity in (A-34) and the constraint (A-54) comes from the of the interpretation of P_t as the stock of future as given in (13). Intuitively, if the entrepreneur carries any stock of misconduct to the time of termination, the average output during the vesting period is lower due to the persistent negative effect of the misconduct, which in turn reduces his average consumption in that period. Combining this termination strategy with the entrepreneur's intertemporal value function for the CARA utility ensures that the conditions in Proposition 4 are also sufficient.

To the best of my knowledge, the deterministic relaxation of the stock of incentives in Marinovic and Varas (2019) and the binary signal and compensation structure illustrated in equations (A-49) to (A-52) are the only two *analytically tractable* methods available for ensuring the sufficiency of the necessary conditions for dynamic moral hazard problems with persistent private information. Despite their rather parsimonious nature, as noted in Marinovic and Varas (2019), stochastic treatment of this terminal value problem has not been systematically studied in the literature yet and is thus left for future research.

References

Ai, H., D. Kiku, R. Li, and J. Tong (2020). A Unified Model of Firm Dynamics with Limited Commitment and Assortative Matching. *Journal of Finance*. Forthcoming.

Ai, H., K. Li, and R. Li (2017). Moral Hazard and Investment-Cash-Flow Sensitivity. Working paper.

Ai, H. and R. Li (2015). Investment and CEO Compensation Under Limited Commitment. *Journal of Financial Economics 116*(3), 452–472.

Albuquerque, R. and H. Hopenhayn (2004). Optimal Lending Contracts and Firm Dynamics. *Review of Economic Studies 71*, 285–315.

Battaglini, M. (2005). Long-term Contracting with Markovian Consumers. *American Economic Review 95*(3), 637–658.

Bergemann, D. and J. Välimäki (2019). Dynamic Mechanism Design: An Introduction. *Journal of Economic Literature 57*(2), 235–74.

Biais, B., T. Mariotti, G. Plantin, and J.-C. Rochet (2007). Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications. *Review of Economic Studies 74*, 345–390.

Blanchard, O. J., F. Lopez-de Silanes, and A. Shleifer (1994). What Do Firms Do with Cash Windfalls? *Journal of Financial Economics 36*(3), 337–360.

Bolton, P., H. Chen, and N. Wang (2011). A Unified Theory of Tobin's q, Corporate Investment, Financing, and Risk Management. *Journal of Finance 66*, 1545–1578.

Bolton, P., N. Wang, and J. Yang (2019). Optimal Contracting, Corporate Finance, and Valuation with Inalienable Human Capital. *Journal of Finance 74*(3), 1363–1429.

Cao, D., G. Lorenzoni, and K. Walentin (2019). Financial Frictions, Investment, and Tobin's q. *Journal of Monetary Economics 103*, 105–122.

Cetemen, D., F. Z. Feng, and C. Urgun (2020). Renegotiation and Dynamic Inconsistency: Contracting with Non-Exponential Discounting. Working paper.

Clementi, G. L. and H. Hopenhayn (2006). A Theory of Financing Constraints and Firm Dynamics. *Quarterly Journal of Economics 121*, 229–265.

Decamps, J.-P., S. Gryglewicz, E. Morellec, and S. Villeneuve (2016). Corporate Policies with Permanent and Transitory Shocks. *Review of Financial Studies 30*(1), 162–210.

DeMarzo, P. and M. Fishman (2007). Optimal Long-term Financial Contracting. *Review of Financial Studies 20*, 2079–2128.

DeMarzo, P., M. Fishman, Z. He, and N. Wang (2012). Dynamic Agency and the q Theory of Investment. *Journal of Finance 67*.

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

Edmans, A., X. Gabaix, T. Sadzik, and Y. Sannikov (2012). Dynamic CEO Compensation. *Journal of Finance 67*(5), 1603–1647.

Erickson, T. and T. M. Whited (2000). Measurement Error and the Relationship Between Investment and *q*. *Journal of Political Economy 108*(5), 1027–1057.

Feng, F. Z. (2020). Dynamic Compensation under Uncertainty Shock and Limited Commitment. *Journal of Financial and Quantitative Analysis*. Forthcoming.

Feng, F. Z. and M. Westerfield (2021). Dynamic Resource Allocation and Hidden Volatility. *Journal of Financial Economics 140*(2), 560–581.

Fernandes, A. and C. J. Phelan (2000). A Recursive Formulation for Repeated Agency with History Dependence. *Journal of Economic Theory 91*(2), 223–247.

Fu, S. and R. V. Krishna (2019). Dynamic Financial Contracting with Persistent Private Information. *RAND Journal of Economics 50*(2), 418–452.

Georgiadis, G. and B. Szentes (2020). Optimal Monitoring Design. *Econometrica*. Forthcoming.

Gryglewicz, S. and B. Hartman-Glaser (2020). Investment Timing and Incentive Costs. *Review of Financial Studies 33*(1), 309–357.

Gryglewicz, S., S. Mayer, and E. Morellec (2020). Agency Conflicts and Short-versus Long-termism in Corporate Policies. *Journal of Financial Economics 136*(3), 718–742.

Hayashi, F. (1982). Tobin's Marginal *q* and Average *q*: A Neoclassical Interpretation. *Econometrica 50*(1), 213–224.

He, Z. (2011). A Model of Dynamic Compensation and Capital Structure. *Journal of Financial Economics 100*(2), 351–366.

He, Z., B. Wei, J. Yu, and F. Gao (2017). Optimal Long-term Contracting with Learning. *Review of Financial Studies 30*(6), 2006–2065.

Hirshleifer, D. and A. V. Thakor (1992). Managerial Conservatism, Project Choice, and Debt. *Review of Financial Studies 5*(3), 437–470.

Hoffmann, F., R. Inderst, and M. Opp (2020). Only Time Will Tell: A Theory of Deferred Compensation. *Review of Economic Studies*. Forthcoming.

Hoffmann, F. and S. Pfeil (2010). Reward for Luck in a Dynamic Agency Model. *Review of Financial Studies 23*, 3329–3345.

Kapička, M. (2013). Efficient Allocations in Dynamic Private Information Economies with Persistent Shocks: A First-order Approach. *Review of Economic Studies 80*(3), 1027–1054.

Krasikov, I. and R. Lamba (2020). A Theory of Dynamic Contracting with Financial Constraints. *Journal of Economic Theory*. forthcoming.

Krishna, R. V., G. Lopomo, and C. R. Taylor (2013). Stairway to Heaven or Highway to Hell: Liquidity, Sweat Equity, and the Uncertain Path to Ownership. *RAND Journal of Economics 44*(1), 104–127.

Li, R. (2017). Dynamic Agency with Persistent Observable Shocks. *Journal of Mathematical Economics 71*, 74–91.

Malenko, A. (2019). Optimal Dynamic Capital Budgeting. *Review of Economic Studies 86*(4), 1747–1778.

Marinovic, I. and F. Varas (2019). CEO Horizon, Optimal Pay Duration, and the Escalation of Short-Termism. *Journal of Finance 74*(4), 2011–2053.

Miao, J. and Y. Zhang (2015). A Duality Spproach to Continuous-time Contracting Problems with Limited Commitment. *Journal of Economic Theory 159*, 929–988.

Pavan, A., I. Segal, and J. Toikka (2014). Dynamic Mechanism Design: A Myersonian Approach. *Econometrica 82*(2), 601–653.

Piskorski, T. and A. Tchistyi (2010). Optimal Mortgage Design. *Review of Financial Studies 23*, 3098–3140.

Quadrini, V. (2004). Investment and Liquidation in Renegotiation-proof Contracts with Moral Hazard. *Journal of Monetary Economics 51*(4), 713–751.

Sannikov, Y. (2008). A Continuous-time Version of the Principal-agent Problem. *Review of Economic Studies 75*, 957–984.

Sannikov, Y. (2014). Moral Hazard and Long-run Incentives. Working paper.

Stein, J. C. (1989). Efficient Capital Markets, Inefficient Firms: A Model of Myopic Corporate Behavior. *Quarterly Journal of Economics 104*(4), 655–669.

Szydlowski, M. (2019). Incentives, Project choice, and Dynamic Multitasking. *Theoretical Economics 14*(3), 813–847.

Tchistyi, A. (2016). Security Design with Correlated Hidden Cash Flows: The Optimality of Performance Pricing. *Working paper*.

Varas, F. (2018). Managerial Short-Termism, Turnover Policy, and the Dynamics of Incentives. *Review of Financial Studies 31*(9), 3409–3451.

Wang, W. (2018). Bid Anticipation, Information Revelation, and Merger Gains. *Journal of Financial Economics 128*(2), 320–343.

Ward, C. (2020). Agency in Intangibles. Working paper.

Williams, N. (2011). Persistent Private Information. *Econometrica 79*(4), 1233–1275.

Williams, N. (2015). A Solvable Continuous-time Dynamic Principal–agent Model. *Journal of Economic Theory 159*, 989–1015.

Zhang, Y. (2009). Dynamic Contracting with Persistent Shocks. *Journal of Economic Theory 144*(2), 635–675.

Zhu, J. Y. (2013). Optimal Contracts with Shirking. *Review of Economic Studies 80*, 812–839.

Zhu, J. Y. (2018). Myopic Agency. *Review of Economic Studies 85*(2), 1352–1388.