A Model of Dynamic Compensation and Capital Structure

Zhiguo He*

April 2010

Abstract

This paper studies the optimal compensation problem between shareholders and the agent in the Leland (1994) capital structure model, and finds that the debt-overhang effect on the endogenous managerial incentives lowers the optimal leverage. Consistent with data, our model delivers a negative relation between pay-performance sensitivity and firm size, and the interaction between debt-overhang and agency issue leads smaller firms to take less leverage relative to their larger peers. During financial distress, a firm's cash-flow becomes more sensitive to underlying performance shocks due to debt-overhang. The implications on credit spreads and debt covenants are also considered.

Key Words: Continuous-time Contracting, Capital Structure, CARA (Exponential) Preference, Firm Growth, Size-Heterogeneity, Pay-Performance Sensitivity.

*University of Chicago, Booth School of Business, zhiguo.he@chicagobooth.edu. An earlier version of this paper was circulated under the title "Agency Problems, Firm Valuation, and Capital Structure." I thank an anonymous referee, Mike Fishman, Arvind Krishnamurthy, Patrick Bolton, Chris Hennessy, Debbie Lucas, Ilya Strabeleuv, and Yuri Tserlukevich (AFA 2009 discussant) for valuable suggestions. All errors are mine.

1 Introduction

This paper embeds optimal contracting between the agent (manager) and shareholders into the cash-flow framework commonly used in the literature of structural models of capital structure (Leland (1994)). By connecting these two literatures, we provide a general framework to study the impacts of agency characteristics on firm valuation and capital structure. Moreover, the dynamic nature of this framework allows us to calibrate our model, and in turn quantitatively assess the agency impact on the firm's leverage decision.

We characterize the optimal contract between shareholders and the agent explicitly. In determining the leverage level, debt bears an additional "*debt-overhang*" cost relative to the bankruptcy cost in standard models (a la Leland (1994)): by interpreting the agent's effort as a form of investment, shareholders implement diminishing effort (as cut back investment) during financial distress. As a result, the agency problem reduces the optimal leverage from 63.2% based on Leland (1994) to as low as 39.5% in our calibration. Consistent with the data, our model predicts that small firms take less leverage relative to their large peers. The debt-overhang problem also implies that the firm's cash-flow is more sensitive to underlying shocks, reinforcing the standard leverage effect.

In Section 2, we start by offering a general analysis in solving the optimal contracting problem. The analysis hinges on the agent's constant-absolute-risk-aversion (CARA, or exponential) preference. In contrast to Holmstrom and Milgrom (1987) where the lump-sum compensation is considered, the agent in our model has intermediate consumption flows, and can privately save. To solve for the optimal contract, we employ the approach in Sannikov (2008) and take the agent's continuation value (*continuation payoff*, or *promised utility*) as one of our state variables. The absence of wealth effect under CARA preference allows us to characterize the optimal contract by an Ordinary Differential Equation (ODE) in Section 2.3. We derive the (second-best) firm value and the agent's pay-performance sensitivity (the dollar-to-dollar measure as in Jensen and Murphy (1990)) based on the optimal contracting. We also characterize the condition that ensures the empirical regularity of an inverse relation between the agent's pay-performance sensitivity and firm size.

Section 3 applies the optimal contracting result to the cash-flows framework in Leland (1994).

There, the firm growth is endogenously affected by the agent's effort, and in the optimal contract both the pay-performance sensitivity and the firm growth are decreasing in firm size. To investigate the impact of agency issues on capital structure, Section 3.3 introduces debt into the baseline model. For better comparison to Leland (1994) and other related work, we leave the debt contract to take the same form as in Leland (1994)—specifically, only consol bond is considered, and shareholders have the option to default when the firm profitability deteriorates. On the contracting side, we bond the agent and shareholders together through an optimal contract solved in Section 2,[1] and shareholders and the agent can revise the compensation contract dynamically, so that the compensation contract is a best response to the capital structure.[2] Essentially, these simplifying assumptions capture the key notion that, in US corporations, managers are only responsible to shareholders (e.g., Allen, Brealey, and Myers (2006)).

We then solve for the optimal capital structure and the optimal employment contract in Section 3.3. Compared to Leland (1994), our model features a debt-overhang problem. Specifically, in our endogenous firm growth framework where the firm growth is controlled by the manager and/or shareholders, when close to bankruptcy shareholders will assign diminishing incentives to the agent. This corresponds to "debt-overhang," i.e., reducing the "positive-NPV" effort investment in financially distressed firms. In other words, beyond the standard bankruptcy cost, debt bears another form of cost, as debt-overhang interferes with agency issues. As a result, our model produces lower optimal leverage ratios relative to the Leland (1994) benchmark.

The debt-overhang generates a negative relation between leverage and agent's working incentives, a prediction opposite to Cadenillas, Cvitanic, and Zapatero (2004) where the debt level and agent's incentives are positively related. In that paper, they study a dynamic compensation/capital-structure model where the agent controls both the drift (effort) and the volatility (project selection) of the firm value. There, the agent's compensation space is restricted to equity shares, and shareholders commit to this static compensation scheme. Setting a higher leverage directly reduces the

[1] In our model, the agent—once bonded with shareholders by an optimal contract—will have perfectly aligned interest with shareholders when dealing with debt holders. As a result, the default policy will be independent of whether shareholders or the agent control the bankruptcy decision. This is different from Morellec (2004) where the agent tends to keep the firm alive longer for more private benefit.

[2] This assumption can be justified by the fact that, under this CARA framework, the long-term contract is renegotiation-proof and can be implemented by a sequence of short-term contracts (see Fudenburg, Holmstrom, and Milgrom (1990)).

value of the agent's equity compensation. Under their assumption of the agent's log utility, this induces a higher sensitivity of the agent's value to his performance, therefore stronger working incentives. In contrast, we show that in a dynamic model, if shareholders and the agent can revise the compensation contract ex-post,[3] then there is an opposite effect in addition to these channels, and it is an empirical question that which force indeed prevails under various economic circumstances.

Further, the interaction between agency issue and debt-overhang predicts that smaller firms take less leverage, which is consistent with empirical regularity. In our model, motivating the agent is less costly in small firms; this is a result of matching the empirically observed negative relationship between pay-performance sensitivities and firm sizes. Therefore, shareholders in small firms will implement a higher effort (or, higher investment) without debt. Since the presence of debt cuts back effort investment, debt-overhang becomes more severe in small firms. Taking this higher debt-overhang cost into account, small firms will choose a lower optimal leverage. In our calibrations, for small firms, the predicted optimal leverage ratios—with or without the agency issue—can have a sizeable difference (63.2% vs. 39.5%).

In the literature, several attempts have been made to incorporate other important agency issues into the corporate security pricing setting. For instance, Leland (1998) studies the "risk-shifting" issue due to the endogenous choice of firm's volatility; there, the agent and shareholders are treated as one party. In this paper, we focus on debt-overhang.[4] Our paper distinguishes itself from the above-mentioned literature, in that we study the agency impact based on the optimal dynamic contracting approach. Even though it seems appealing to restrict the compensation contract space within commonly observed forms as in Cadenillas, Cvitanic, and Zapatero (2004), one might wonder whether the derived impact of agency problems is sensitive to specific contract forms.[5] The

[3] This is consistent with both the practice of resetting the strike price of stock options, and the empirical results in Bryan, Hwang, and Lilien (2000) who investigate the stock-based compensations in a panel of firms.

[4] Based on the free cash-flow problem, Morellec (2004) introduces a tension between the agent and shareholders, and the empire-building agent tends to set a lower leverage ratio for rent-maximizing purposes. Cadenillas, Cvitanic, and Zapatero (2004) study a different version of agency problem, where they restrict the compensation contract space to be equity. Because the equity payoff ties to the debt face value, in their model the capital structure becomes a direct compensation scheme. In contrast, in our model the impact of leverage decision on the compensation contract is indirect. Subramanian (2007) studies a similar question as this paper, but only allows for single-period employment contracts in his analysis.

[5] Technically speaking, in the aforementioned papers, either the volatility choice in Leland (1998) which is observable in Brownian setting, or the over-investment (observable) decision in Morellec (2004), can be easily resolved by optimal contracting. These extreme examples illustrate the sensitivity of agency costs to the contracting space.

optimal contracting approach is free of this issue.

This paper is also related to the on-going continuous-time contracting literature. Sannikov (2008) studies a general dynamic agency problem without private savings. Williams (2006) focuses on the general hidden-state problem, and solves an example with CARA preference. In our paper, based on the continuation value approach advocated in Sannikov (2008), we analyze the general cash-flow setup, and focus on the applications to corporate finance. DeMarzo and Fishman (2006), DeMarzo and Sannikov (2006), and Biais et al. (2007) solve a dynamic contracting problem with a risk-neutral agent, where the limited liability restriction is imposed.[6] In contrast, this paper employes the Holmstrom and Milgrom (1987) framework, which not only allows for a risk-averse agent, but also easily accommodates a second state variable to capture the firm's time-varying profitability.[7] Compared to Holmstrom and Milgrom (1987), we allow for the agent's intermediate consumption, and therefore their approach is no longer applicable.

The rest of this paper is organized as follows. Section 2 derives an ODE which characterizes the optimal contracting, and Section 3 applies the contracting result to Leland (1994). Section 4 concludes. All proofs are in the Appendix.

2 General Model and Optimal Contracting

2.1 General Model

We study an infinite-horizon, continuous-time agency problem. The firm (shareholders) hires an agent to operate the business. The firm produces cash flows δ_t per unit of time, where δ_t follows the stochastic process

$$d\delta_t = \mu\left(\delta_t, a_t\right) dt + \sigma\left(\delta_t\right) dZ_t. \tag{1}$$

We will also interpret δ_t as firm size in this paper. Through unobservable effort $a_t \in [0, \overline{a}]$, the agent controls the cash-flow growth rate $\mu\left(\delta_t, a_t\right)$, where $\mu_a\left(\delta, a\right) \equiv \frac{\partial \mu(\delta,a)}{\partial a} > 0$ and $\mu_{aa}\left(\delta, a\right) \equiv$

[6]For extensions, e.g., He (2009) studies the optimal executive compensation in a geometric Brownian cash-flow setting, and Piskorski and Tchistyi (2007) study the optimal mortgage design by considering the exogenous regime switching in the investors' discount rate.

[7]For another example among various extensions of Holmstrom and Milgrom (1987), Schattler and Sung (1993) offer a general treatment for the continuous-time contracting with CARA preference, but under the original Holmstrom and Milgrom (1987) setting, i.e., a finite time horizon with a lump-sum compensation. Because the lump-sum compensation (consumption) occurs at the end of employment period, as opposed to flows studied in this paper, there is no issue of private savings in Schattler and Sung (1993).

$\frac{\partial^2 \mu(\delta,a)}{\partial a^2} \leq 0$. The performances $\{\delta_t\}$ are contractible.

Shareholders (the principal) are risk-neutral, and they discount future cash-flows at the constant market interest rate r. To focus on the optimal contracting, throughout this section the firm is unlevered; we introduce debt holders in Section 3.3.

The agent, with a CARA instantaneous utility and a time discounting factor r, maximizes his expected life-time utility

$$\mathbb{E}\left[\int_0^\infty -\frac{1}{\gamma} e^{-\gamma(c_t - g(\delta_t, a_t)) - rt} dt\right],$$

where $c_t \in \mathbb{R}$ is the agent's consumption rate, and $g(\delta, a)$ is the agent's monetary effort cost with $g_a(\delta, a) = \frac{\partial g(\delta,a)}{\partial a} > 0$ and $g_{aa}(\delta, a) = \frac{\partial^2 g(\delta,a)}{\partial a^2} > 0$. To ensure that pay-performance-sensitivity is falling with firm size in the optimal contract, later in Section 2.4.2 we will impose restrictions on the dependence of $\mu(\delta, a)$ and $g(\delta, a)$ on firm size δ.

We allow for the agent's private (unobservable) savings. The agent can borrow and save at the risk-free rate r in his personal savings account. The account balance, as well as the agent's actual consumption, is unobservable to shareholders. It is the agent's intermediate consumption, associated with the possibility of private savings, that distinguishes our analysis from the classic Holmstrom and Milgrom (1987).

2.2 Contracting Problem

We distinguish the policies *recommended* by the contract, from the agent's *actual* policies; the latter will be indicated by a "hat" on top of the relevant symbols.

The employment contract $\Pi = \{c; a\}$ specifies the agent's recommended consumption process c and the recommended effort process $\{a\}$. The process $\{c\}$ can also be interpreted as the wage process. Both elements are adapted to the filtration generated by $\{\delta\}$; in other words, they are functions of the agent's performance history. To simplify the analysis, unless otherwise stated, in this section we assume that the effort process $\{a\}$ takes interior solutions.

For simplicity we assume that the agent's initial wealth is 0.[8] Given $\Pi = \{c; a\}$, the agent's

[8]This assumption is innocuous given the CARA preference. If the agent's initial wealth W_0 is observable, then the contract may ask the agent to hand his wealth over to the principal (shareholders) who can plan for the agent subsequently through the contract. Even if the initial wealth W_0 is the agent's private information, the absence of wealth effect implies that facing any contract, the agent will take the same effort policy as another hypothetical agent with 0 initial wealth (except consuming rW_0 more each period)—see the argument in Section 2.3.2 and Lemma 3.

problem is:

$$\begin{aligned} V_0(\Pi) = \max_{\{\widehat{c}_t, \widehat{a}_t\}} \quad & \mathbb{E}\left[\int_0^\infty -\frac{1}{\gamma} e^{-\gamma(\widehat{c}_t - g(\delta_t, \widehat{a}_t)) - rt} dt\right] \\ s.t. \quad & dS_t = rS_t dt + c_t dt - \widehat{c}_t dt,\ S_0 = 0, \\ & d\delta_t = \mu(\delta_t, \widehat{a}_t) dt + \sigma(\delta_t) dZ_t, \end{aligned} \tag{2}$$

where we impose transversality condition $\lim_{T\to\infty} \mathbb{E}\left[e^{-rT} S_T\right] = 0$ for all feasible policies, and $V_0(\Pi)$ denotes the agent's time-0 value derived from the contract Π. It is clear that both the consumption policy $\{c\}$ and effort policy $\{a\}$ are "recommended" only. For instance, the first constraint states that the change of the agent's savings dS_t, is the interest accrual $rS_t dt$, plus the wage deposit $c_t dt$, and minus the consumption withdrawal $\widehat{c}_t dt$. To save, the agent can set his consumption $\widehat{c}_t$ strictly below the wage c_t.

Suppose that the agent has a time-0 outside option v_0. Shareholders solve the following problem:

$$\begin{aligned} \max_{\Pi} \quad & \mathbb{E}^{\delta(\Pi)}\left[\int_0^\infty e^{-rt}(\delta_t - c_t) dt\right] \\ s.t. \quad & V_0(\Pi) \geq v_0, \end{aligned}$$

where $\mathbb{E}^{\delta(\Pi)}[\cdot]$ indicates the dependence of probability measure (over $\{\delta\}$) on the employment contract Π when the agent solves his problem (2). The second line is the agent's participation constraint. As in Holmstrom and Milgrom (1987), under this CARA framework without limited liability, the participation constraint always binds, and the outside option v_0 only affects the optimal contract by a constant transfer between these two parties.

We define the class of incentive-compatible and no-savings contracts as follows.

Definition 1 *A contract $\Pi = \{c; a\}$ is incentive-compatible and no-savings if the solution to the agent's problem (2) is $\{c; a\}$.*

In words, a contract Π is incentive-compatible and no-savings, if the agent, once facing the contract Π, finds it optimal to exert the recommended effort (i.e., incentive-compatible), and follow the recommended consumption plans (i.e., no-savings).

Therefore the principal can easily design an optimal scheme to induce truth-telling, in that the contract promises the agent rW_0 more per period if at $t = 0$ the agent hands W_0 over to the principal.

Because shareholders have the same saving technology (with rate r) as the agent, the following lemma allows us to focus on incentive-compatible and no-savings contracts only, a result similar to the *Revelation Principle*. Essentially, when shareholders can fully commit,[9] they can save for the agent on his behalf.

Lemma 2 *Without loss of generality, we focus on the incentive-compatible and no-savings contracts.*

2.3 Model Solution

2.3.1 Agent's Continuation Value

Following the literature, in solving the optimal contract, we take the agent's *continuation value* (*continuation payoff*, or *promised utility*) as the state variable. Formally, given the contract $\Pi = \{c; a\}$, the agent's continuation value is defined as:

$$V_t\left(\Pi, \delta_t\right) = \mathbb{E}_t\left[\int_t^{\infty} -\frac{1}{\gamma} e^{-\gamma(c_s - g(\delta_s, a_s)) - r(s-t)} ds\right]. \tag{3}$$

This payoff is a function of the compensation contract Π and the current cash flow level δ_t. To be specific, it is the agent's payoff (given δ_t) obtained under the policies specified by Π: The agent exerts effort policy $\{a_s : s \geq t\}$ recommended by Π, and consumes exactly his future wages $\{c_s : s \geq t\}$. In Section 2.3.3, we will invoke the important fact that, these recommended policies have to be indeed optimal among all policies in the agent's problem stated in (2).

By the martingale representation theorem (e.g., Sannikov (2008)), Eq. (3) implies that the agent's continuation value evolves as:

$$dV_t = rV_t dt - u\left(c_t, a_t\right) dt + \beta_t\left(-\gamma r V_t\right)\left[d\delta_t - \mu\left(\delta_t, a_t\right) dt\right], \tag{4}$$

where the agent's instantaneous utility $u\left(c_t, a_t\right)$ is $u\left(c_t, a_t\right) = -e^{-\gamma(c_t - g(\delta_t, a_t))}/\gamma$, and $\{\beta\}$ is a progressively measurable process. Here, $-\gamma r V_t > 0$ (note that $V_t < 0$ in (3)) is a scaling factor that facilitates the economic interpretation of β_t in Section 2.3.3.

[9] When we introduce bankruptcy later into this model, the full-commitment ability also requires that, even in the event of bankruptcy, shareholders can fulfill any promise that they have made to the agent before bankruptcy. For more discussion on this issue, see Section 3.4.2.

To read the evolution of the agent's continuation value in (4), his expected total value change is

$$\mathbb{E}_t \left[dV_t + u\left(c_t, a_t\right) dt\right] = rV_t dt,$$

which is the required return for the agent. The key element in (4) lies in the volatility part. It is the volatility of the agent's continuation value that controls the agent's working incentives. Intuitively, as clear from reading (4), the volatility part $\beta_t \left(-\gamma r V_t\right) \left[d\delta_t - \mu\left(\delta_t, a_t\right) dt\right]$ directly links to the observable performance $d\delta_t$, and $\beta_t \left(-\gamma r V_t\right)$ measures the punishing/rewarding extent in the employment contract. As a result, in Section 2.3.3 we will connect $\beta_t \left(-\gamma r V_t\right)$—which is the incentive imposed by the contract—to the implemented effort a_t.

2.3.2 Absence of Wealth Effect

The CARA preference plays a key role in solving for the optimal contract. In essence, the absence of wealth effect allows us to derive the agent's deviation value (when he deviates to other off-equilibrium non-zero savings) only based on the agent's equilibrium value V without savings.

Lemma 3 *At any time $t \geq 0$, consider a deviating agent with savings S who faces the contract Π, and denote by $V_t\left(S; \Pi, \delta_t\right)$ his deviation continuation value. We have*

$$V_t\left(S; \Pi, \delta_t\right) = V_t\left(0; \Pi, \delta_t\right) \cdot e^{-\gamma r S} = V_t \cdot e^{-\gamma r S}, \tag{5}$$

where we have used the fact that $V_t\left(0; \Pi, \delta_t\right)$ is the agent's continuation value V_t along the no-savings path defined in (3).

The intuition is rather simple. For a CARA agent without wealth effect, given the extra savings S, his new optimal policy is to take the optimal consumption-effort policy without savings, but consume an extra rS more for all future dates $s \geq t$. Since $u\left(c_s + rS, a_s\right) = e^{-\gamma r S} u\left(c_s, a_s\right)$, this explains the factor $e^{-\gamma r S}$ in (5). Essentially, for CARA preference, the agent's problem is translation-invariant to his underlying wealth level. Without CARA preference, the agent's working incentives will be wealth-dependent, and the deviation value representations—as simple as (5)—are no longer available.

2.3.3 Equilibrium Evolution of V

For incentive-compatible and no-savings contracts, the recommended consumption-effort policies specified in Π have to be optimal among all policies. Based on this requirement, we now derive the necessary and sufficient conditions for the equilibrium evolution of V in (4).

No savings. Fix the effort policy first. By the optimality of the agent's consumption-savings policy in problem (2), his marginal utility from consumption must equal his marginal value of wealth:

$$u_c(c_t, a_t) = \frac{\partial}{\partial S} V_t(0; \Pi, \delta_t).$$

Therefore, the necessary condition for Π to rule out private savings is:

$$u_c(c_t, a_t) = e^{-\gamma(c_t - g(\delta_t, a_t))} = \frac{\partial}{\partial S} V_t(0; \Pi, \delta_t) = -\gamma r V_t \Rightarrow u(c_t, a_t) = rV_t, \tag{6}$$

where the third equation uses the functional form of $V_t(S; \Pi, \delta_t)$ in (5). Plugging this result into (4), we observe that $u(c_t, a_t)$ just offsets rV_t, and Eq. (4) becomes

$$dV_t = \beta_t(-\gamma r V_t)[d\delta_t - \mu(\delta_t, a_t)\, dt]. \tag{7}$$

Therefore, the agent's continuation value V_t follows a martingale.

Two points are note-worthy. First, because $u(c_t, a_t) = -e^{-\gamma(c_t - g(\delta_t, a_t))}/\gamma$, the relation $u(c_t, a_t) = rV_t$ implies that the equilibrium consumption (or the agent's wage) is

$$c_t = g(\delta_t, a_t) - \frac{\ln \gamma r}{\gamma} - \frac{1}{\gamma} \ln(-V_t). \tag{8}$$

Second, because $u_c(c_t, a_t) = -\gamma r V_t$ as shown in (6), the agent's marginal utility also follows a martingale. Naturally, this is a consequence of the agent's optimal consumption-savings policy, which is in direct contrast to the optimal contracting with observable savings as studied in Rogerson (1985) and Sannikov (2008). There, the principal can dictate the agent's consumption plan that are suboptimal from the agent's view.

Incentive compatibility. Now we turn to incentive provision to pinpoint the diffusion loading β_t in (7). As discussed before, $\beta_t(-\gamma r V_t)$ measures the agent's continuation *utility* sensitivity with respect to the unexpected performance $d\delta_t - \mu(\delta_t, a_t)\,dt$. Now the role of the scaling factor $-\gamma r V_t$ becomes clear: since $-\gamma r V_t$ is the agent's marginal utility u_c as shown in (6), by transforming *utilities* to *dollars*, β_t directly measures the (monetary) compensation sensitivity with respect to his performance.

Consider the agent's effort decision. On one hand, choosing $\widehat{a}_t$ affects the agent's instantaneous utility $u(c_t, \widehat{a}_t)$. On the other hand, $\widehat{a}_t$ sets the drift of his performance $d\delta_t$ which affects his expected continuation payoff $\mathbb{E}_t[dV_t(\widehat{a}_t)]$ in (7) via $\beta_t \cdot u_c \cdot \mu(\delta_t, \widehat{a}_t)$. By balancing the impacts on his instantaneous utility and continuation payoff, the agent is solving:

$$\max_{\widehat{a}_t} u(c_t, \widehat{a}_t) + \beta_t \cdot u_c \cdot \mu(\delta_t, \widehat{a}_t).$$

Because $u(c_t, \widehat{a}_t) = u(c_t - g(\widehat{a}_t))$, we have $u_a = u_c \cdot (-g_a(\delta_t, a_t))$. Therefore, implementing $\widehat{a}_t = a_t$ requires that[10]

$$-g_a(\delta_t, a_t) + \beta_t \mu_a(\delta_t, a_t) = 0 \Rightarrow \beta_t = \frac{g_a(\delta_t, a_t)}{\mu_a(\delta_t, a_t)}, \tag{9}$$

and it is easy to check that this first-order condition is also sufficient.

Eq. (9) gives an equilibrium relation between the recommended effort a_t and the incentive loading β_t. Intuitively, $\mu_a(\delta_t, a_t)$ is the agent's effort impact on the instantaneous performance, and $\beta_t \mu_a(\delta_t, a_t)$ gives the agent's *monetary* marginal benefit of his effort. To be incentive compatible, the marginal benefit must equal the agent's monetary marginal effort cost $g_a(\delta_t, a_t)$. And, because g (or μ) is convex (or concave) in a, one can show that the required incentive loading β_t is increasing in a_t. In other words, implementing a higher level of effort needs greater incentives.

In sum, for Π to be incentive-compatible and no-savings, we must have (recall (7)):

$$dV_t = \frac{g_a(\delta_t, a_t)}{\mu_a(\delta_t, a_t)} \gamma r(-V_t)\,\sigma(\delta_t)\,dZ_t, \tag{10}$$

where we have replaced the innovation term in (7) by $\sigma(\delta_t)\,dZ_t$ due to Eq. (1).

We have used the agent's First-Order-Conditions (FOCs) regarding the recommended consumption and effort policies to derive necessary conditions for the dynamics of V. It is well known

[10] Note that the main driving force underlying (9) is the monetary effort cost specification, i.e., $u(c_t, a_t) = u(c_t - g(a_t))$, rather than the CARA preference. To see this, if we write $dV_t = rV_t dt - u(c_t, a_t)\,dt + \beta_t \cdot u_c \cdot [d\delta_t - \mu(\delta_t, a_t)\,dt]$ in (4), then β_t is still the monetary incentive loading measured in dollars, and the exact same argument gives us the result in (9). However, as we show in (6), the CARA preference implies a convenient result that $u_c = -\gamma r V_t$, which makes the evolution of the agent's continuation value V dependent on V itself.

that with private-savings, FOCs cannot guarantee the global optimality of the recommended policies (e.g., Kocherlakota (2004) and He (2010)). But for the case of CARA preference without wealth effect, FOCs are both necessary and sufficient, a result which we show in Appendix A.3.

2.3.4 Optimal Contracting

Shareholders' Value Function Given the state variables δ and V, the shareholders' value function is

$$\begin{aligned} J(\delta, V) = \max \quad & \mathbb{E}\left[\int_t^\infty e^{-r(s-t)} (\delta_s - c_s)\, ds \middle| \delta_t = \delta\right] \\ s.t. \quad & V_t(\Pi, \delta) = V. \end{aligned} \tag{11}$$

The absence of wealth effect, thanks to the CARA preference, leads us to guess that

$$J(\delta, V) = f(\delta) - \frac{-1}{\gamma r} \ln(-\gamma r V),$$

where $\frac{-1}{\gamma r} \ln(-\gamma r V) > 0$ is the agent's certainty-equivalent given his continuation value V. We will refer to the agent's certainty-equivalent as the agent's inside stake in later discussions.

Using (1) and (10), the HJB equation for the shareholders' problem in (11) is:

$$rJ(\delta, V) = \max_{a\in[0,\overline{a}]} \left\{ \begin{array}{c} \delta - c(a,\delta,V) + J_\delta \cdot \mu(\delta, a) + \frac{1}{2} J_{\delta\delta} \sigma(\delta)^2 \\ J_{\delta V} \frac{g_a(\delta_t, a_t)}{\mu_a(\delta_t, a_t)} \gamma r (-V_t) \sigma(\delta_t)^2 + \frac{1}{2} J_{VV} \gamma^2 r^2 V^2 \left[\frac{g_a(\delta_t,a_t)\sigma(\delta_t)}{\mu_a(\delta_t,a_t)}\right]^2 \end{array} \right\},$$

where $c(a, \delta, V)$ takes the form in (8). Plugging (8) into the above equation, and noting that $J_\delta = f'(\delta)$, $J_{\delta\delta} = f''(\delta)$, $J_{VV} = -\frac{1}{\gamma r}\frac{1}{V^2}$, and $J_{\delta V} = 0$, we obtain the following ODE for $f(\cdot)$:

$$rf(\delta) = \max_{a\in[0,\overline{a}]} \left\{ \delta + f'(\delta)\mu(\delta, a) + \frac{1}{2} f''(\delta)\sigma(\delta)^2 - g(\delta, a) - \frac{1}{2}\gamma r \left[\frac{g_a(\delta, a)\sigma(\delta)}{\mu_a(\delta, a)}\right]^2 \right\}. \tag{12}$$

Here, δ is cash inflow, the second and third terms capture the expected instantaneous change of $f(\delta)$ due to δ, and the last two terms are the effort-related costs. The optimal effort a^* is characterized by:

$$\arg\max_{a\in[0,\overline{a}]} \left\{ f'(\delta)\mu(\delta, a) - g(\delta, a) - \frac{1}{2}\gamma r \left[\frac{g_a(\delta, a)\sigma(\delta)}{\mu_a(\delta, a)}\right]^2 \right\} \tag{13}$$

Similar to Holmstrom and Milgrom (1987), in (12) there are two kinds of costs in implementing effort a. The first is the direct monetary effort cost $g(\delta, a)$, and the second is the risk-compensation term for the risk-averse agent to bear incentives:

$$\frac{1}{2}\gamma r\left[\frac{g_a(\delta_t,a_t)\,\sigma(\delta_t)}{\mu_a(\delta_t,a_t)}\right]^2. \tag{14}$$

This additional agency-related cost, as in Holmstrom and Milgrom (1987), captures the key trade-off between incentive provision and risk-sharing in the optimal contract.

The solution to (12), combined with (8), (10), and certain problem-specific boundary conditions, characterizes the optimal contracting. In Appendix A.3 we give an example of square root process where the optimal contract admits a closed form solution. More importantly, with certain technical conditions, in Appendix A.3 we provide detailed verification argument to show rigorously that the derived contract is indeed optimal. The proof techniques also apply to other more general cases.

2.4 Model Implications

2.4.1 Firm Value and Agent's Deferred Compensation

We interpret $f(\delta)$ as the firm value. From the derivation in Section 2.3.4, we see that in our CARA setting, maximizing shareholders' value $J(\delta,V)$ is equivalent to maximizing the firm value $f(\delta)$ in this model, as both aim to minimize the agency cost.

The agency cost under the CARA setup has one particular feature. Given the promised continuation value V to the agent, the cost of delivering V, from the shareholders' view, is its certainty equivalent $\frac{-1}{\gamma r}\ln(-\gamma r V)$, plus some additional agency cost due to inefficient incentive-risk allocation. Interestingly, under the CARA setup, this additional agency cost is independent of the agent's continuation value V.[11] In other words, the severity of agency problems are only reflected in the functional form of f as a solution to the ODE (12), not the agent's continuation value per se.

From the view of implementation, the firm value $f(\delta)$ is the sum of the (common) shareholders' value $J(\delta,V)$, plus the agent's inside stake which is measured by his certainty equivalent $-\frac{1}{\gamma r}\ln(-\gamma r V_t)$. Specifically, in implementing the optimal contract, shareholders set up a deferred-

[11]This is different from other dynamic agency models with inefficient termination (e.g., DeMarzo and Fishman (2008)) where the agency cost is linked to the agent's continuation value directly. There, the lower the agent's continuation value, the more likely the agent will be terminated, and therefore the higher the agency cost.

compensation fund inside the firm with a balance

$$W_t = -\frac{1}{\gamma r} \ln \left(-\gamma r V_t \right); \tag{15}$$

and shareholders adjust this balance continuously according to the evolution of V_t in (10).[12] By keeping the agent's stake inside the firm, the firm (market) value becomes the total value enjoyed by both the agent and shareholders. To the extent that in practice the agent's non-marketable rent (e.g., future wages) is small relative to the firm value, this treatment is a close approximation.[13]

Under the optimal employment contract, the deferred compensation fund W follows:

$$dW_t = \frac{g_a \left(\delta_t, a_t \right)}{\mu_a \left(\delta_t, a_t \right)} \sigma \left(\delta_t \right) dZ_t + \frac{1}{2} \gamma r \left[\frac{g_a \left(\delta_t, a_t \right) \sigma \left(\delta_t \right)}{\mu_a \left(\delta_t, a_t \right)} \right]^2 dt. \tag{16}$$

Here, the first diffusion term provides incentives, and the second drift term captures risk compensation. Interestingly, under the optimal contract the agent's consumption c_t cancels with the interest rW_t earned by the deferred-compensation fund and the effort cost reimbursement g_t (check Eq. (8)).

2.4.2 Pay-Performance Sensitivity and Size Dependence

By interpreting the deferred-compensation balance W_t as the agent's financial wealth, we can derive the agent's pay-performance sensitivity in this model. In the literature, the executive's (dollar-to-dollar) pay-permeance sensitivity (later called PPS) has received great attention since Jensen and Murphy (1992). The central question, whose answer is just PPS, is that: "*how much does the executive's wealth change when the firm value changes by one dollar?*"

In the current continuous-time framework, the agent's pay-performance sensitivity can be measured by the response of the balance of W, to a unit shock of the firm value.[14] Specifically, by

[12] When we introduce bankruptcy in Section 3, it is important to ensure that this deferred compensation has priority over debt in the event of default. As in Westerfield (2006), this balance can also be interpreted as the committed separation payment if either party wants to renege in the future. Theoretically, the CARA framework cannot rule out the possibility of $W_t < 0$. We interpret this case as the agent to take a personal debt, and the debt is netted out in calculating the total firm value.

[13] In this implementation shareholders conduct all the savings for the agent, as his wealth is kept inside the firm. Another equivalent implementation is to put W_t into the agent's personal account, but allow for two-way transfers between the agent and shareholders according to (10). This corresponds to the case where the agent's entire rent is non-marketable, and the firm's market value becomes $J(\delta, V)$.

[14] Strictly speaking, in the executive compensation literature, the pay-performance sensitivity is with respect to the shareholders' value, which should exclude the agent's non-marketable stake. There are two reasons why this treatment

neglecting all drift terms, we have (recall that $\beta_t(\delta_t, a_t) = \frac{g_a(\delta_t, a_t)}{\mu_a(\delta_t, a_t)}$ in (9)):

$$PPS = \frac{dW_t}{df(\delta_t)} = \frac{\beta_t(\delta_t, a_t^*)\,\sigma(\delta_t)\,dZ}{f'(\delta_t)\,\sigma(\delta_t)\,dZ} = \frac{\beta_t(\delta_t, a_t^*)}{f'(\delta_t)} = \frac{g_a(\delta_t, a_t^*)}{\mu_a(\delta_t, a_t^*)}\frac{1}{f'(\delta_t)}. \tag{17}$$

Intuitively, PPS is the ratio between β_t which is the agent's dollar incentive, and $f'(\delta_t)$ which captures the value change (in dollars) of the firm. Note that the optimal effort a_t^* in (17) is endogenously determined by the optimization problem in (13).

The result in (17) implies that the agent's pay-performance-sensitivity PPS depends on firm size δ_t. Our later calibration aims to replicate the well-known empirical regularity that PPS is negatively related to firm size (e.g., Murphy (1999)).[15] To this end, we now impose some structure on our model to study the general pattern of relationship between PPS and firm size.

When does PPS decrease with firm size? Suppose that

$$\mu(\delta_t, a_t) = \mu_0(\delta_t) + a_t\delta_t^{\mu_1},\ \sigma(\delta_t) = \sigma\delta_t^{\sigma_1},\ \text{and}\ g(\delta_t, a_t) = g_0(\delta_t) + \frac{\theta}{2}a_t^2\delta_t^{g_1},$$

which imply that

$$\mu_a(\delta_t, a_t) = \delta_t^{\mu_1},\ \text{and}\ g_a(\delta_t, a_t) = \theta a_t\delta_t^{g_1}, \tag{18}$$

Here, μ_1, σ_1, θ, and g_1 are constants. Note that this general specification encompasses Baker and Hall (2004) who argue that the effort impact on the firm growth (which we will refer to as effort benefit) might be size-dependent, i.e., $\mu_1 > 0$. We focus on μ_1, g_1 and σ_1; they characterize the dependence of the agent's effort benefit, direct monetary effort cost, and indirect risk-compensation cost on firm size, respectively.

Given this structure, the first-order-condition for (13) (assuming an interior solution a_t^*) is:

$$f'(\delta)\,\delta_t^{\mu_1} - \theta a_t^*\delta_t^{g_1} - \gamma r\sigma^2\theta^2 a_t^*\delta_t^{2(g_1+\sigma_1-\mu_1)} = 0,$$

is inessential: 1) the magnitude of PPS is small ($1 \sim 5\%$); 2) empirically, the executive's PPS mainly comes from his/her inside holdings which are marketable. For other definitions of pay-permeance sensitivities (e.g., pay-performance elasticities) and an agency model distinct from the Holmstrom and Milgrom (1987) framework, see the recent paper by Edmans, Gabaix, and Landier (2007).

[15] The PPS in executive compensation literature only considers the CEO's incentive holdings. Our model takes this interpretation as well, so that the agent is the single top manager of the firm. Readers can also interpret the manager here as a team of top managers, and the relevant PPS measure becomes the inside holdings of the firm's officers and directors. Even though it is theoretically possible that a larger firm might have more top managers who, as a team, hold more inside shares, empirically the opposite holds. For instance, Holderness, Krosner, and Sheehan (1999) document a negative relationship between the total ownership of officers and directors and firm size.

which implies the optimal effort as:

$$a_t^* = \frac{f'(\delta)\,\delta_t^{\mu_1}}{\theta\delta_t^{g_1} + \gamma r\sigma^2\theta^2\delta_t^{2(g_1+\sigma_1-\mu_1)}}. \tag{19}$$

Plugging (18) and (19) into (17), we find that $f'(\delta)$ cancels, and

$$PPS = \frac{1}{1 + \theta\gamma r\sigma^2\delta_t^{g_1+2\sigma_1-2\mu_1}}. \tag{20}$$

Therefore, the necessary and sufficient condition for a negative relation between PPS and firm size δ_t is

$$g_1 + 2\sigma_1 - 2\mu_1 > 0. \tag{21}$$

In words, when the size-dependence of the effort cost (either direct cost part g_1 or indirect cost part σ_1) is sufficiently large, or the size-dependence of effort benefit is sufficiently small, the firm should design an incentive contract whose power is decreasing in firm size.

When we apply our optimal contracting results to the Leland framework in Section 3, we will set

$$\mu(\delta, a) = (\phi + a)\,\delta,\ \sigma(\delta) = \sigma\delta, \text{ and } g(\delta, a) = \frac{\theta}{2}a^2\delta.$$

Here, $g_1 = \sigma_1 = \mu_1 = 1$, and $g_1 + 2\sigma_1 - 2\mu_1 = 1$. Therefore the PPS in the optimal contract is

$$PPS = \frac{1}{1 + \theta\gamma r\sigma^2\delta_t},$$

which is falling with firm size.[16]

There are some attempts in the literature to estimate these parameters. It is well-known (as the "leverage" effect) that the large firm has a greater dollar variance but a smaller return variance, i.e., $\sigma_1 \in (0, 1)$. Cheung and Ng (1992) fit an EGARCH model with CEV (constant-elasticity-variance) specification to a large sample of individual stocks (as opposed to certain stock index which is common in this literature), and find that σ_1 falls in the range of 0.84 (in the 1960's) and 0.94 (in the 1980's). This estimation is subject to the caveat that we are approximating δ_t by the firm's stock price. For μ_1 and g_1, Baker and Hall (2004) assume that the agent's effort cost is independent of firm size (i.e., $g_1 = 0$), and find that $\mu_1 = 0.4$. If we instead set $g_1 = 1$, then one

[16] On the other hand, if we consider another specification $g_1 = \mu_1 = 1$ but $\sigma_1 = 0.5$ as in Appendix A.3, then $g_1 + 2\sigma_1 - 2\mu_1 = 0$, and $PPS = \frac{1}{1+\theta\gamma r\sigma^2}$ is a constant independent of firm size.

can show that the effort benefit measure that Baker and Hall (2004) are estimating is effectively $\mu_1 - 0.5$ (see Baker and Hall (2004) for details). Therefore, the estimate for μ_1 in our model (with $g_1 = 1$) is approximately 0.9 (close to the choice of $\mu_1 = 1$ in Section 3).[17] The bottom line is, the condition (21) that guarantees a negative relation between PPS and firm size holds for these estimates, which extends indirect support to our model.

CRRA (power) preference Our entire analysis hinges on the assumption of CARA (constant-absolute-risk-aversion, exponential) preference, which implies that the agent's risk absorbing capacity is independent of his wealth level. As an important ingredient for PPS, the risk absorbing capacity directly determines the risk compensation cost in (14), which in turn pins down the size-dependence of PPS in (20). What can we say if instead the agent has a CRRA (constant-relative-risk-aversion, power) preference?

Unfortunately the wealth effect in the CRRA preference complicates the optimal contracting significantly, and we do not know much about the solution to that problem.[18] In the literature there are several attempts to accommodate this question. For example, Baker and Hall (2004) solve an (static) optimal contracting problem as if the agent has an exponential utility; however, they specify the agent's absolute-risk-aversion parameter $\gamma(W_t) = \frac{\gamma_0}{W_t}$ (where γ_0 is a positive constant) to be proportional to the inverse of his wealth $1/W_t$, as if the agent has a power utility. Here we will take this simple approach as well.

It is important to note that the agent's wealth W_t is not necessarily proportional to firm size δ_t. Ideally we would like to derive the path of W_t endogenously from the model, but it is not available in CRRA setting.[19] Because the question at hand is a calibration question, we resort help from data. We know empirically that managers in larger firms—although get higher pay in terms of dollar amounts—have lower inside stakes in their firms. Most of literature (e.g., Baker and

[17] Condition (21) is easier to be satisfied if we take the Baker and Hall (2004) assumption of $g_1 = 0$ and therefore $\mu_1 = 0.4$.

[18] Edmans et al. (2009) consider a multiplicative effort cost model, and impose some modification on the timing structure to make the linear contract optimal in every instant. They solve a long-term optimal contracting problem in closed-form if the optimal contract aims to implement a maximum target effort (exogenously given). In the optimal contract, the implemented effort is the constant target effort level, and the return PPS (i.e., the log change of manager's compensation to the log change of firm value) is also a constant independent of firm size.

[19] In our theretical result with CARA preference, W_t in (16) can be negative which is inappropriate to define $\gamma(W_t) = \frac{\gamma_0}{W_t}$.

Hall (2004)) use the manager's total compensation $Comp_t$ to approximate his wealth W_t. In fact, Gabaix and Landier (2008) calibrate that $Comp_t \propto \delta_t^{1/3}$, i.e., the elasticity between pay level and firm size is $1/3$. Given this result, we estimate that

$$\gamma(W_t) = \frac{\gamma_0}{W_t} = \widehat{\gamma}_0 \delta_t^{-1/3},$$

where $\widehat{\gamma}_0$ is another positive constant potentially different from γ_0. Plugging this result into (20), we have

$$PPS = \frac{1}{1 + \theta \widehat{\gamma}_0 r \sigma^2 \delta_t^{g_1 + 2\sigma_1 - 2\mu_1 - 1/3}}.$$

Therefore, under the CRRA preference, the necessary and sufficient condition that ensures a negative relation between PPS and firm size δ_t becomes:

$$g_1 + 2\sigma_1 - 2\mu_1 > \frac{1}{3}. \tag{22}$$

This condition holds for the case $g_1 = \sigma_1 = \mu_1 = 1$ that we are going to study in Section 3, as well as for the empirical estimates of $\{g_1, \sigma_1, \mu_1\}$ discussed in the end of Section 2.4.2. Thus, even taking into account the fact that the agent might have a risk-aversion decreasing with his wealth (as implied by CRRA preferences), our model has qualitatively similar predictions under reasonable parameterizations.

3 Revisiting Leland (1994): Optimal Capital Structure

3.1 Model Specification

Following Leland (1994), we consider the case that

$$d\delta_t = (\phi + a_t)\,\delta_t dt + \sigma \delta_t dZ_t,$$

where ϕ and σ are constants. In the language of Eq. (1), we have $\mu(\delta, a) = (\phi + a)\,\delta$, and $\sigma(\delta) = \sigma\delta$. Here, ϕ is the baseline growth level, and by exerting effort the agent can accelerate the firm growth. The agent's effort cost takes the form $g(\delta, a) = \frac{\theta}{2} a^2 \delta$ which is quadratic in a and linear in size δ.

Recall that in Section 2.1 we restrict the agent's action space to a bounded interval $[0, \overline{a}]$, and the calibration in the unlevered firm might call for a binding effort $a_t = \overline{a}$ in the optimal contract. In fact, under the parametrization considered later, the first-best solution has $a_t^{FB} = \overline{a}$. To characterize the first-best solution, we can simply set $\gamma = 0$ in (12) (so there is no agency problem), and as a result

$$rf^{FB}(\delta) = \max_{a\in[0,\overline{a}]} \left\{ \delta + f^{FB\prime}(\delta)(\phi + a)\delta + \frac{1}{2}f^{FB\prime\prime}(\delta)\sigma^2\delta^2 - \frac{\theta}{2}a^2\delta \right\}, \tag{23}$$

where $f^{FB}(\delta)$ is the first-best firm value without agency problems. Because all model elements are proportional to δ, we guess that $f^{FB}(\delta) = A^{FB}\delta$, where A^{FB} is a constant to be solved. Plugging into (23), we have

$$rA^{FB} = \max_{a\in[0,\overline{a}]} \left\{ 1 + A^{FB}(\phi + a) - \frac{\theta}{2}a^2 \right\},$$

which jointly determines a^{FB} and A^{FB}. In Appendix A.4 we give the exact condition under which a^{FB} binds at $\overline{a}$.

Independent of whether a^{FB} binds at $\overline{a}$ or not, the scale-invariance of this model implies that, in the first-best case, the firm's cash-flow—as well as the firm value—follows a geometric Brownian motion. Due to its analytical convenience, this setup has become the workhorse in the literature of structural models of capital structure (e.g., Leland (1994), Goldstein, Ju, and Leland (2001), and Strebulaev (2006)).

3.2 Optimal Contracting in an Unlevered Firm

Before we introduces debt into this framework, we apply the optimal contracting results obtained in Section 2 to an unlevered firm. To implement effort a_t, Eq. (9) implies that the agent's incentive slope $\beta_t = \theta a_t$. Then Eq. (12) becomes:

$$rf(\delta) = \max_{a\in[0,\overline{a}]} \left\{ \delta + f'(\delta_t)\cdot(\phi + a)\delta + \frac{1}{2}f''(\delta_t)\sigma^2\delta^2 - \frac{\theta}{2}a^2\delta - \frac{1}{2}\gamma r\theta^2 a^2\sigma^2\delta^2 \right\}. \tag{24}$$

Simple calculation yields the optimal effort as (the optimal effort might bind at $\overline{a}$ along the optimal path):[20]

$$a_t^* = \min\left(\frac{f'(\delta_t)}{\theta(1 + \theta\gamma r\sigma^2\delta_t)}, \overline{a} \right). \tag{25}$$

[20]When a binds at $\overline{a}$, the same incentive loading $\beta_t = \theta\overline{a}$ applies—investors can set a higher incentive loading, but it is costly to do so because the agent is risk-averse. And, because firm value is increasing in the cash-flow level δ, one can formally show that in this model $f'(\delta)$ is always positive, therefore a^* never binds at zero. For formal proofs, see Appendix A.5.

As discussed in Section 2.4.2, when the optimal effort level is interior, we have

$$PPS = \frac{1}{1 + \theta\gamma r\sigma^2\delta_t}, \tag{26}$$

which is decreasing in firm size δ_t. Fundamentally, this result is due to the fact that as the firm grows, the quadratic risk-compensation cost $\frac{1}{2}\gamma r\theta^2 a^2\sigma^2\delta_t^2$ is in the order of δ_t^2, while the incentive benefit is in the order of δ_t (check (24); in Appendix A.6 we show that when $\delta_t \to \infty$, $f'(\delta_t) \to \frac{1}{r-\phi}$). This exactly reflects the common wisdom that managers in larger firms have lower-powered incentive schemes due to risk considerations.[21,22]

Another appealing feature, which is closely related to the pattern of PPS falling in firm size, is that the endogenous firm growth rate $\phi + a_t^*$ is decreasing in firm size δ_t as well (see Figure 2 in Section 3.4.3 for a numerical example). The negative relationship between firm size and growth is studied in, for instance, Cooley and Quadrini (2001). In this model, because incentivizing the agent is more costly in larger firms, the optimal contract implements a lower effort in larger firms, leading to a lower growth.

3.3 Optimal Employment Contract and Leverage

3.3.1 Additional Assumptions

Now suppose that the firm issues debt to take advantage of tax shields. Relative to the standard bilateral contracting framework between investors (the principal) and the agent, now we have heterogenous investors—shareholders and debt holders. To abstract from complicated contracting issues among three parties, we take the following treatment. First, the debt contract takes the form studied in Leland (1994), i.e., only the consol bond (with a constant coupon rate C) is considered, and shareholders (with their perfectly-aligned agent when they are dealing with debt holders)

[21]For instance, Murphy (1999) states that, "*The inverse relation between company size and pay performance sensitivities is not surprising, since risk-averse and wealth-constrained CEOs of large firms can feasibly own only a tiny fraction of the company...the result merely underscores that increased agency problems are a cost of company size that must weighed against the benefits of expanded scale and scope.*" Notice that under an optimal contracting framework, wealth-constrainedness is not an appealing justification, because empirically CEOs receive a fair amount of fixed salaries in their compensation packages.

[22]This loose reasoning is precise when the manager's risk-absorbing capacity is independent of firm size, which holds only for CARA preference. However, this statement is probably better interpreted as following: Even though the managers in larger firms might have greater risk-absorbing capacity (presumably because they recieve higher pay), their greater risk-absorbing capacity cannot offset the higher total risk in larger firms. For a related discussion of CRRA preference, see the end of Section 2.4.2.

can endogenously default when the firm's financial status deteriorates. Second, we assume that shareholders can fulfill the promise of the agent's continuation value at bankruptcy as a part of employment contract. In other words, when bankruptcy occurs, the agent is guaranteed with the deferred-compensation fund W defined in (15). We will comment on this assumption in Section 3.4.2.

Another important timing assumption is that in this model, shareholders design the employment contract as an optimal response to the leverage decision. Theoretically, this is consistent with the fact that, in this CARA framework a long-term optimal contract can be implemented by a sequence of short-term contracts (Fudenburg, Holmstrom, and Milgrom (1990)). Essentially, in the CARA framework studied here, shareholders and the agent can revise the contract (as long as both parties agree to do so) once the debt is issued,[23] which generates the debt-overhang problem analyzed in Section 3.4.2.

These assumptions represent a minimum, but essential, departure from Leland (1994). They reflect the key economic rationale regarding the manager's objective in US corporations: managers are supposed to be responsible to shareholders only (Allen, Brealey, and Myers (2006)).

3.3.2 Equity Value and Endogenous Default

Similar to the case of unlevered firms, the shareholders' value function is $J^E(\delta, V) = f^E(\delta) + \frac{1}{\gamma r}\ln(-\gamma r V)$. The separability between δ and V hinges on the assumption that in the leveraged firm the shareholders can always keep the promise of delivering the continuation payoff V to the agent, even when the firm goes bankrupt at $\delta = \delta_B$. In the implementation, the promise is guaranteed by the deferred-compensation fund which has priority over debt in the event of bankruptcy.

As in Section 2.3.4, by writing down the HJB equation for $J^E(\delta, V)$, we reach the following ODE for the equity value $f^E(\delta)$, where the control is over a_t and the bankruptcy boundary δ_B:

$$rf^E(\delta) = \max_{a\in[0,\overline{a}],\delta_B} \left\{ \delta - C(1-\tau) + f^{E\prime}(\delta)\cdot(\phi + a)\delta + \frac{1}{2}f^{E\prime\prime}(\delta)\sigma^2\delta^2 - \frac{1}{2}\theta a^2\delta - \frac{1}{2}\gamma r\theta^2 a^2\sigma^2\delta^2 \right\}, \tag{27}$$

where C is the coupon rate, and τ is the corporate tax rate. The equity value $f^E(\delta)$ is the

[23] In this CARA setup we allow for renegotiations in deriving the optimal contract. To see this, the resulting optimal contract is renegotiation-proof, as the Pareto boundary is always downward-sloping, i.e., $J_V = \frac{1}{\gamma r V} < 0$ (note that $V < 0$ in this model).

sum of (common) shareholders' value $J^E(\delta, V)$, and the agent's deferred-compensation fund $-\frac{1}{\gamma r}\ln(-\gamma r V)$. Compared to (24) without debt, (27) has an additional cash outflow $C(1-\tau)$ as the after-tax coupon payment. Similar to (25), the optimal effort is

$$a_t^* = \min\left(\frac{f^{E\prime}(\delta_t)}{\theta(1+\theta\gamma r\sigma^2\delta_t)}, \overline{a}\right). \tag{28}$$

Plugging it into (27), we have an ODE to characterize the optimal contracting.

When δ falls to a certain level, say δ_B, shareholders refuse to serve the coupon payment by simply declaring bankruptcy. This is captured by the value-matching boundary condition

$$f^E(\delta_B) = 0, \tag{29}$$

and the smooth-pasting condition

$$f^{E\prime}(\delta_B) = 0. \tag{30}$$

Both conditions are standard in this literature (e.g., Leland (1994)). Note that these conditions are a result of maximizing the shareholders' value $J^E(\delta, V)$. But since these policies are toward debt holders, it is equivalent to maximizing $f^E(\delta)$, i.e., the joint (ex-post) surplus enjoyed by shareholders and the agent.[24]

For the boundary condition on the other end, when δ takes a sufficiently large value $\overline{\delta} \to \infty$, the bankruptcy event is negligible. In the Appendix A.6 we show that,

$$f^E(\overline{\delta}) \simeq \overline{f}(\overline{\delta}) - \frac{C(1-\tau)}{r}, \tag{31}$$

where $\overline{f}(\cdot)$ captures the firm value under a Gordon growth model with a growth rate ϕ (see Eq. (37) in the Appendix A.6), and $\frac{C(1-\tau)}{r}$ is the value for a perpetual after-tax coupon payment. Then we can numerically solve for $f^E(\cdot)$ and δ_B, based on conditions (29), (30), and (31). For detailed numerical methods, see Appendix A.6.

3.3.3 Debt Value and Capital Structure

Given the implemented effort policy $a^*(\delta)$ in (28), we can evaluate the consol bond with a promised coupon rate C. Since debt holders anticipate the optimal contracting between shareholders and the

[24]This result implies that despite the agency conflicts between the agent and shareholders, under the optimal contract they have perfectly aligned interests with respect to the policy toward debt holders. In other words, the default policy will not depend on whether shareholders or the agent is in charge of the bankruptcy decision. This differs from Morellec (2004) where the agent tends to keep the firm alive longer for more private benefits.

agent, the value of the corporate debt, $D(\delta)$, satisfies:

$$rD(\delta) = C + D_\delta(\delta) \cdot (\phi + a^*(\delta))\,\delta + \frac{1}{2} D_{\delta\delta}(\delta)\,\sigma^2\delta^2,$$

with $D(\delta_B) = (1-\alpha) f(\delta_B)$ where $\alpha < 1$ is the percentage bankruptcy cost, and $D\left(\overline{\delta}\right) \to \frac{C}{r}$ as $\overline{\delta} \to \infty$. Here we simply assume that, once bankruptcy occurs, debt holders pay the bankruptcy cost $\alpha f(\delta_B)$, and then keep running the project as an unlevered firm.[25]

Given the time-0 cash-flow level δ_0, shareholders choose coupon C to maximize the total levered firm value $f^E(\delta_0; C) + D(\delta_0; C)$ before the debt issuance; they then design the optimal contract with an agent who has an outside option v_0. As discussed in Section 3.3.1, this timing assumption is equivalent to allowing shareholders and the agent to revise the employment contract ex-post after the debt issuance.

We define the firm's optimal leverage ratio as

$$LR(\delta_0) \equiv \frac{D(\delta_0; C^*(\delta_0))}{f^E(\delta_0; C^*(\delta_0)) + D(\delta_0; C^*(\delta_0))}.$$

In Leland (1994), the scale invariance implies that the optimal leverage ratio LR is independent of firm size δ_0. However, we have seen that in our model the quadratic risk-compensation eliminates the scale invariance. In fact, in the following calibration exercises, we will mainly investigate the divergent leverage decisions for different-sized firms.

3.3.4 Parameterization

Table 1 tabulates our baseline parametrization. Interest rate $r = 5\%$, bankruptcy cost $\alpha = 25\%$, and tax rate $\tau = 20\%$ (considering personal tax effect), are typical in the literature (e.g., Leland (1998)).

We also record the average growth rate in the 50-year simulation, and this measure helps us pin down ϕ and $\overline{a}$. In the literature with constant coefficients, Goldstein, Ju, and Leland (2001) calibrate a slightly negative μ, and Leland (1998) chooses the growth rate $\mu = 1\%$. Under the choice of $\phi = -0.5\%$ and $\overline{a} = 5\%$ ($\overline{a}$ is irrelevant for levered firms as the optimal effort $\{a^*\}$ never binds at 5%; see Figure 2), the simulated average growth rates fit these numbers squarely across various firm sizes (see Table 2).

[25] Also, the new agent's outside option is $v_0 = \frac{-1}{\gamma r}$, so $W_0 = 0$. Our result is insensitive to the treatment of unlevered firm after the bankruptcy.

Table 1: Baseline Parameters

Agency		Leland (1998)	
Risk Aversion	$\gamma = 5$	Volatility	$\sigma = 25\%$
Effort Cost	$\theta = 35$	Interest Rate	$r = 5\%$
Lower Bound Growth	$\phi = -0.5\%$	Bankruptcy Cost	$\alpha = 25\%$
Upper Bound Effort	$\overline{a} = 5\%$	Marginal Tax Rate	$\tau = 20\%$

Because in our calibration the optimal effort never binds at $\overline{a}$ in levered firms, we have the pay-performance sensitivity as in (26):[26]

$$PPS = \frac{1}{1 + \theta\gamma r\sigma^2\delta_t}.$$

Based on this result, we choose the agency-related parameters (risk aversion $\gamma = 5$ which is the median value used in Haubrich (1994), and effort cost $\theta = 35$) and the starting firm size δ_0 to match the PPS documented in the empirical literature. Jensen and Murphy (1990) report a PPS of 0.3% in their sample (1969-1983), while Hall and Liebman (1998) document a higher PPS with mean 2.5%. Aggarwal and Samwick (1998) control for the firm risk, and report a mean PPS of 6.94% from the OLS regression. For the size-dependence pattern of PPS, Murphy (1999) finds that for large S&P500 firms, the PPS is approximately 1%; for Midcap firms, it is 1.5%; and for small firms, it is 3%. Hall and Liebman (1998) document a mean PPS around 2.5%, and note that in their sample, "*the largest firms (with market value over $10 billion) have a median PPS that is more than an order of magnitude smaller than the smallest firms (market value less than $500 million).*"

3.4 Results and Discussions

3.4.1 Optimal Leverage Ratio

Because debt-overhang adversely affects the firm's *endogenous* growth, relative to Leland (1994) firms will take less leverage for their optimal capital structure. Figure 1 shows the optimal leverage

[26] The presence of debt does not affect the expression of PPS in (26). To see this, the agent's performance is measured as the change of equity value $f^E(\delta)$. However, $f^{E\prime}(\delta)$ cancels in the expression (26) when the optimal effort $a^*(\delta)$ takes an interior solution (check the derivation in Eq. (20)).

ratio (the solid line) for firms with different sizes. For better comparison, in Figure 1 we also provide two benchmark optimal leverage ratios predicted by the Leland (1994) model, with exogenous constant growth. The dashed line with circles (63.21%) is the first-best case, where the cash-flow growth rate is $\overline{a} + \phi = 4.5\%$. However, as we have seen in Section 3.2, the agency problem alone reduces the firm growth, which would lead to a lower leverage ratio even without debt-overhang. To address this issue, we take the results of unlevered firms from Section 3.2, simulate the model, and obtain time-series averages of the cash-flow growth and volatility.[27] We then use these estimates as inputs to calculate the Leland leverage ratio, which is graphed in the dotted line with asterisks (61.59%) in Figure 1.

The optimal leverage ratios are reported in Table 2 along with other measures. For each initial cash-flow level δ_0, we calculate the sample mean of growth and volatility of $d\delta/\delta$ over 100 years in 500 simulations (see Panel C in Table 1). We also report the sample average of pay-performance sensitivity based on (26); these numbers fit the empirical estimates discussed in the last section squarely. We find that for small firms ($\delta_0 = 50$), the optimal leverage ratio falls from 61.59% (or 63.21% in the first-best case) to 39.35%.[28] In contrast to small firms, we find that the optimal leverage ratio for large firms is close to the result under Leland (1994), a cross-sectional result that we will come back shortly.

3.4.2 Debt-Overhang

In Figure 1 we observe that the optimal leverage ratio is lower relative to Leland (1994). The reason is debt-overhang, where we interpret the agent's effort as a form of "investment;" see Hen-

[27] Specifically, to match the relevant range for δ in levered firms, we set the initial $\delta_0 = 250$, and stop the process once δ reaches 7.97 (which is the lowest default boundary in our calibration). Also, the simulation length is 50 years to mitigate the impact of initial condition. We then average the time-series mean of growth rate and volatility across 500 simulations, which gives an average growth rate (volatility) as 3.31% (25.02%). Other treatment gives similar results.

[28] This reduction is comparable to other modifications of the Leland (1994) benchmark. For instance, by combining both the "callable" feature of the debt and upward capital restructuring together, Goldstein, Ju, and Leland (2001) reduce the optimal leverage from 49.8% to 37.1% in their baseline case.

It is not easy to accommodate the dynamic capital restructuring into our framework, as we do not have scale-invariance in this model. Certainly, as in Goldstein, Ju, and Leland (2001), the possibility of raising leverage in the future should reduce the firm's initial leverage. In fact, dynamic capital restructuring would have an interesting impact to our model if the restructuring was *downward*, i.e., reducing debt when the firm's cash flow level drops. This is because in this model almost all the action is on the downside where debt overhang cuts down the efficient effort. If we only allow for upward restructuring as in Goldstein, Ju, and Leland (2001), the interaction effect should be small.

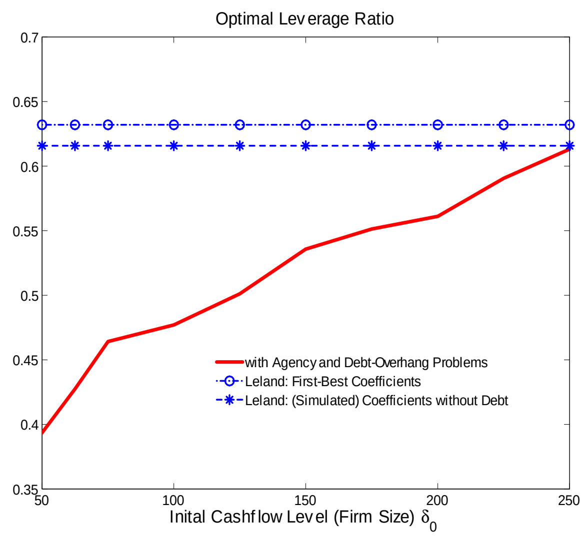

Figure 1: Optimal leverage ratio as a function of initial cash-flow level (firm size). We plot the two benchmark leverage ratios under Leland (1994). The first one is based on the first-best coefficients ($\mu = 4.5\%$ and $\sigma = 25\%$), which gives a leverage ratio 63.21% plotted in the dashed line with circles. The second one is based on the time-series averages in simulating the unlevered firm in Section 3.2 ($\mu = 3.31\%$ and $\sigma = 25.02\%$; for simulation details, see footnote 27); this case yields a leverage ratio 61.59% plotted in the dotted line with asterisks. The parameters are $r = 5\%$, $\sigma = 25\%$, $\theta = 35$, $\gamma = 5$, $\phi = -0.5\%$, $\overline{a} = 5\%$, $\alpha = 25\%$, $\tau = 20\%$.

nessy (2004) for a similar mechanism. In our model, shareholders design an employment contract to maximize the ex-post equity value, and the smooth-pasting condition (30) implies that $f^{E\prime}(\delta)$ goes to zero as δ approaches the default boundary δ_B. It implies that, once the firm is close to bankruptcy, shareholders gain almost nothing by improving the firm's performance. Then, according to (28) which says that the optimal effort a^* is proportional to $f^{E\prime}(\delta)$, shareholders implement diminishing effort (through providing diminishing incentives) during financial distress. As a result, in addition to the traditional bankruptcy cost, in our model the debt bears another form of cost due to debt-overhang.

We illustrate the above mechanism in Figure 2. The left panel plots the implemented effort investment a^* in solid line as a function of firm's financial status δ for small firms ($\delta_0 = 50$). As we will explain shortly, the debt-overhang problem is more severe in small firms. We also plot the optimal effort policy without debt-overhang (the dashed line), which corresponds to the case of unlevered firms studied in Section 3.2.

Relative to the effort policy without debt-overhang plotted in the dotted line, we observe an

Table 2: Optimal Capital Structure for Firms with Different Sizes

	Initial Cash-flow Level (Firm Size) δ_0				
	50	100	150	200	250
	Panel A: Optimal Debt Policies				
Optimal Coupon C^*	24.50	61.35	106.00	147.62	203.15
Default Boundary δ_B	7.97	22.79	40.58	57.28	79.98
Scaled Def. Boundary $\frac{\delta_B}{\delta_0}\%$	15.95	22.79	27.05	28.64	31.99
	Panel B: Valuation and Leverage				
Debt Value $D(\delta_0)$	446.05	997.44	1640.02	2267.70	3090.52
Lev. Ratio $\frac{D(\delta_0)}{D(\delta_0)+f^E(\delta_0)}\%$	39.35	47.70	53.57	56.10	61.32
Credit Spreads bps	49.22	115.12	146.34	150.98	157.34
	Panel C: Simulation Results				
Average Growth $\%$	1.95	0.50	0.12	-0.05	-0.17
Average Volatility $\%$	25.04	24.98	25.04	25.03	25.01
Pay-Perf. Sensitivity $\%$	5.33	2.52	1.69	1.29	0.99

The parameters are $r = 5\%$, $\sigma = 25\%$, $\theta = 35$, $\gamma = 5$, $\phi = -0.5\%$, $\overline{a} = 5\%$, $\alpha = 25\%$, $\tau = 20\%$. Credit spreads are calculated as $(C/D - r) \times 10000$. We simulate the model for 50 years to obtain the average growth rate and volatility for $d\delta/\delta$, given the initial δ_0. We also calculate the agent's average pay-performance sensitivity based on (26).

abrupt drop of implemented effort when the firm is in the verge of bankruptcy ($\delta \to \delta_B = 7.97$). From the view of social welfare, in this situation a higher effort is desirable, because it helps avoid the costly bankruptcy once δ hits δ_B. However, it is not in the shareholders' interest to ask the agent to work hard. When the firm approaches bankruptcy, shareholders obtain zero marginal value from improving δ. Consequently, they implement diminishing effort, a typical symptom of debt-overhang.

We emphasize that the driving forces of the debt-overhang result are the endogenous nature of firm growth, and the smooth-pasting condition of the shareholders' value; both ingredients are quite generic in practice. Therefore, if in reality the management gives place to existing shareholders (like the board) during financial distress, then debt-overhang is still present without the intermediate link of diminishing management incentives.

Leverage vs. management incentives. The debt-overhang generates a negative relation between leverage and agent's working incentives. This result contrasts to Cadenillas, Cvitanic, and Zapatero

(2004), where the log agent's compensation space is restricted to equity shares, and shareholders commit to this compensation scheme. In comparison, in our optimal dynamic contracting setup, we do not place any restriction on the contracting space, and allow for dynamically revising the employment contract between the agent and shareholders. Finally, as we discussed in Section 3.3.1, part of the implementation of the optimal contract requires the firm to set the agent's deferred compensation aside as cash; this way shareholders commit to fully insulate the agent's compensation from bankruptcy. We now discuss these assumptions by relating them to "inside debt" documented in Sundaram and Yermack (2007).

Inside debt. There are several interesting remarks regarding the above assumptions, which point to the robustness of our debt-overhang result. First, in reality, although we do observe certain revising activities such as resetting the strike price of executives' previously awarded options, modifying compensation contract is not frictionless. For instance, managers' pensions—as a form of deferred compensation—are calculated according to certain prespecified formulae. More importantly, these pensions represent unsecured, unfunded debt claims against firm assets ("inside debt" as advocated in Sundaram and Yermack (2007) and Edmans and Liu (2007)),[29] rather than a senior claim against a cash-based deferred compensation fund as in our implementation. As a result, this portion of compensation scheme can potentially alleviate the debt-overhang problem in this paper and the risk-shifting problem in Edmans and Liu (2009).[30]

Nevertheless, for "inside debt" to work, one also needs shareholders' and the agent's commitment on other compensation schemes to prohibit undoing the "inside debt." The important point is that, to rule out ex-post revising, shareholders and the agent need "commitment" with debt holders on *all* compensation schemes, rather than on *one* or *some* schemes. In other words, as long as shareholders can modify the residual compensation freely, they can still undo the "inside-debt",

[29]This means that when the firm becomes insolvent, pension beneficiaries have the same priority as other unsecured creditors. However, footnote 10 in Sundaram and Yermack (2007) gives an example of "secular" trust fund which secure an executive's pension in a bankruptcy-proof form.

[30]These papers are closely related to the early theoretical work by John and John (1993) who consider a risk-shifting problem. Essentially, for the agent to maximize the firm's value, the compensation should be less aligned with shareholders when the leverage is higher, which predicts a negative relation between leverage and PPS. There, commitment is also essential, even though it appears not so in a two-period model. In contrast, in our model, the agent is always perfectly aligned with shareholders in terms of incentives: when close to bankruptcy, even though the agent's incentives diminishes, the equity volatility vanishes in the same order as the agent's incentives, and PPS as a ratio—which is always (26)—does not go to zero.

and theoretically we come back to the optimal contract without commitment. This implies that our theoretical results are quite robust to the practice of "inside-debt." And, there is some indirect empirical evidence consistent with this view of "undoing" or "dynamically revising" the employment contract, which can be a potential topic for future empirical studies.[31]

3.4.3 Size-Heterogeneity

Our model offers another explanation why small firms take less leverage relative to their large peers, a stylized fact documented in the literature (e.g., Frank and Goyal (2005)). The mechanism here is rooted in divergent severities of the aforementioned debt-overhang problem for different sized firms. In this model, to be consistent with the inverse relation between PPS and firm size, larger firms implement lower effort. This implies that, for larger firms, debt-overhang—which reduces the profitable effort investment—is less of concern. Consequently, larger firms will issue more debt to maximize the ex-ante firm value.

This point is illustrated in the right panel in Figure 2, where we plot the same effort policies with and without debt-overhang for large firms ($\delta_0 = 250$). For better comparison, the right panel adopts the same scale as the left panel where small firms are considered. We find that debt-overhang becomes moderate for large firms. As shown, at their relatively high default boundary $\delta_B = 79.98$, the optimal effort even without debt-overhang is quite low (only about 1.14%). Therefore, the drop of a_t^* when larger firms approach bankruptcy—the exact force of debt-overhang—is less dramatic compared to smaller firms (the left panel). In sum, in smaller firms the debt-overhang cost is greater, leading to a lower predicted leverage ratio.

It is important to add that this result is not driven by CARA preference. Rather, we only use CARA preference as the analytical tool to match the empirical pattern of pay-performance sensitivity (PPS), and it is the negative relationship between PPS and firm size that implies

[31] Bryan, Hwang, and Lilien (2000) analyze the *Incentive-Intensity* (the change in the value of annual stock-based compensation per change in equity value) and *Mix* (ratio of the value of annual stock-based compensation to cash compensation) measures, which are based on the annual stock-based grants only (as opposed to cumulative inside holdings which relate to the agent's "wealth"). Interestingly, they find that both *Incentive-Intensity and Mix* decrease with firm leverage. Under the debt-overhang framework studied in this paper, to the extent that working incentives generated by "inside debt" is increasing with leverage, these empirical result is consistent with the dynamic revising activity. Also, decreasing *Mix* with leverage implies that financially troubled firms pay the agent more cash compensation, a result consistent with our model if the juniority of pensions force the firm to start paying out deferred compensation to the agent in the form of cash.

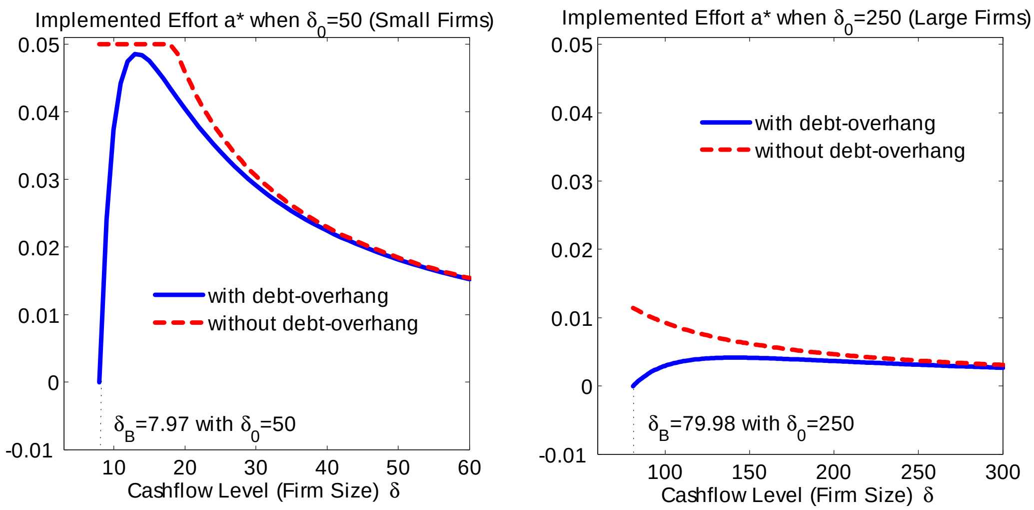

Figure 2: Implemented effort policies with debt-overhang for small ($\delta_0 = 50$) and large ($\delta_0 = 250$) firms under optimal leverage decisions. The optimal effort policy $a^*(\delta)$ is shown in the solid line; for better comparison, we also plot the optimal effort policy (dashed line) in an unlevered firm without debt-overhang. We mark the optimal endogenous default boundary, where shareholders optimally (ex-post) implement a zero effort. The parameters are $r = 5\%$, $\sigma = 25\%$, $\theta = 35$, $\gamma = 5$, $\phi = -0.5\%$, $\overline{a} = 5\%$, $\alpha = 25\%$, $\tau = 20\%$.

lower debt-overhang costs in large firms.

3.4.4 Default Policy and Credit Spreads

Table 2 also reports the endogenous default policies. The "scaled" default boundary δ_B/δ_0 are lower (so shareholders default later) than the Leland (1994) benchmark.[32] The intuition for shareholders to postpone bankruptcy is as follows. In our model, a firm with recent unsatisfactory performances has a lower cash-flow level, or smaller size. But given a smaller risk-compensation, shareholders find it cheaper to motivate the agent, which gives them more value to wait for future improvement. Because of the delayed default, our model produces similar, but slightly lower, credit spreads (Panel B in Table 2) than Leland (1994), *conditional on comparable leverages.*[33]

There are various theoretical models in which agency problems are countercyclical (e.g., Bernanke and Gertler (1989), Eisfeldt and Rampini (2007), etc.). Acknowledging that the agency issue be-

[32] For the Leland (1994) exogenous growth model, $\delta_B/\delta_0 = 39.73\%$ in the first-best parametrization, and $\delta_B/\delta_0 = 37.24\%$ in the unlevered firm parametrization.

[33] For instance, based on the growth and the volatility estimates (Panel C) of medium-sized firms ($\delta_0 = 150$), under Leland (1994) a 53.57% of leverage ratio translates into a credit spread of $175bps$, which is higher than $146bps$ in our model.

comes more severe in recessions (for instance, a higher cash-flow volatility, or a higher constant absolute risk-aversion when the agent's wealth is lower), it is interesting to further explore the impact of agency problems on credit spreads. Chen et al. (2006) argue that the key mechanism in explaining the high-credit-spread puzzle is that firms default more frequently in those higher risk premium states. This countercyclical default policy might be related to exacerbated agency issues in these bad states, because severe agency frictions can lead to a great reduction in shareholders' value of keeping the firm under their control. For instance, consider small firms $\delta_0 = 50$. By raising the agent's constant absolute risk aversion coefficient γ from 5 to 10, while fixing coupon level $C = 24.50$, the default boundary increases from 7.97 to 8.92, and as a result the credit spread goes up substantially (49 bps vs. 85 bps). We leave this for future research.

3.4.5 Asset Volatility and Leverage Effect

Various empirical studies find that equity returns become more volatile as the firm approaches bankruptcy. This phenomenon can be explained by the so-called *leverage* effect, even holding the volatility of the underlying asset constant.

In our model, the state variable for a firm's financial status is its cash-flow level, and we also assume a constant *instantaneous* (return) volatility in the continuous-time setup. However, when the sampling frequency cannot be arbitrarily high, the estimated variance will differ from the instantaneous volatility. Interestingly, because of the hump-shaped endogenous effort (see Figure 2), our model generates higher conditional variances (based on infrequent sampling) when firms are in financial distress.

The mechanism is as follows. Due to debt-overhang, the firm's endogenous growth is positively correlated with underlying performance shocks in financial distress. To see this, consider the implemented effort policy in the left panel in Figure 2. When δ is close to $\delta_B = 7.97$, $a^*(\delta)$ is increasing in δ. Then, a positive shock to δ, by reducing debt-overhang, gives rise to a higher effort $a^*(\delta)$. This further leads to a higher cash-flow growth rate $a^*(\delta) + \phi$, and in turn magnifies the positive shock. Therefore, when sampling is infrequent, the observed return volatility will be higher than the constant volatility $\sigma = 25\%$. For instance, when $\delta = 15$, over one year the annualized volatility based on monthly observed data is about 25.12%. On the other hand, when δ

is far away from the bankruptcy threshold, $a^*(\delta)$ is decreasing in δ, and the exact opposite force leads to a lower volatility estimate (when $\delta = 50$, the above mentioned annualized volatility is about 24.68%).

In sum, in our model, for a firm near bankruptcy, its financial status becomes more sensitive to underlying performance shocks. In fact, this general message does not depend on the discrete-sampling, as in our model the firm value (rather than the instantaneous cash flow rate) displays a higher instantaneous volatility during financial distress.[34]

3.4.6 Debt Covenants

A commonly observed debt covenant is that debt holders can force shareholders to go bankrupt when the firm's cash-flow δ hits a prespecified level δ_B^b. This covenant is along the same line as "*positive net-worth covenants for protected debt*" in Leland (1994) which stipulates that the firm defaults whenever the asset value drops to the debt face value. Under our cash-flow framework, this can also be interpreted as a hard covenant on the interest coverage ratio, which, combined with coupon C, gives the bankruptcy threshold.[35]

In the standard trade-off model as in Leland (1994), a forced earlier default always hurts firm value, because an earlier default reduces the tax benefit and raises the bankruptcy cost. In fact, if equity holders can fully commit in Leland (1994), then "no default" is the first-best outcome.

In contrast, due to the debt-overhang problem, in our model specifying $\delta_B^b > \delta_B$ might be welfare improving. The reason is that now around δ_B^b, we no longer have the smooth-pasting condition as in (30), and shareholders will benefit from improving the firm performance. Then, specifying $\delta_B^b > \delta_B$ is equivalent to committing to provide incentives (to the agent) even in the deep financial distress, and it might be socially optimal to do so. For instance, for small firms $\delta_0 = 50$, given the coupon $C = 24.50$, the optimal $\delta_B^b = 8.05 > \delta_B = 7.97$, and the firm value becomes $1133.81 > 1133.67$. Though small magnitude, this interesting finding distinguishes our

[34]The firm's asset value, generated by future cash-flows, can be defined as $X(\delta_t) = \mathbb{E}_t\left[\int_t^{\tau(\delta=\delta_B)} e^{-r(s-t)}\delta_s ds + f(\delta_B)\right]$, where $\tau(\delta=\delta_B)$ is the first passage time of δ hitting δ_B. Then, the instantaneous return volatility of $X(\delta_t)$ will be larger when δ_t is close to δ_B. The reason is that $X(\delta_t)$ takes into account the impact on the stochastic growth μ, which positively correlates with δ_t during distress.

[35]Because part of the coupon C can be interpreted as constant operating cost (in other words, the leverage derived here includes operating leverage too), the effective interest coverage ratio can be much higher than implied by δ_B^b/C.

model from Leland (1994) and its variations where the firm growth is exogenously specified.

4 Concluding Remarks

By generalizing the optimal contracting result to widely-used cash-flow setups in finance, this paper offers a more tractable framework to investigate the impact of agency problems in various economic contexts. The absence of wealth effect due to CARA preference simplifies the optimal contracting problem with private savings, and we characterize the optimal contract by an ODE. When we apply our results to the geometric Brownian cash-flow setup as a revisit to Leland (1994), the underinvestment of effort due to debt-overhang produces a lower optimal leverage ratio, and the interesting interaction between agency issue and debt-overhang leads smaller firms to take less debt in their leverage decisions.

The relatively simple structure in this paper leaves several directions for future research. For instance, incorporating investment decisions would be desirable, as one can explore the investment distortion and its interaction with financing decisions under agency problems. Also, incorporating time-varying risk premia which is correlated with agency frictions (e.g., time-varying risk-aversion) will be valuable to investigate the agency impacts on credit spreads.

References

Aggarwal, R., and A. Samwick, 1999, "The Other Side of the Tradeoff: The Impact of Risk on Executive Compensation," *Journal of Political Economy* 107, 65-105.

Allen, F., S. Myers, and R. Brealey, 2006, "Principles of Corporate Finance," 8th Edition, New York, McGraw-Hill/Irwin.

Baker, G., and B. Hall, 2004, "CEO incentives and Firm Size," *Journal of Labor Economics* 22, 767-798.

Bernanke, B., and M. Gertler, 1989, "Agency Costs, Net Worth, and Business Fluctuations," *American Economic Review* 79, 14-31.

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

Bryan, S., L. Hwang, and S. Lilien, 2000, "CEO Stock-Based Compensation: An Empirical Analysis of Incentive-Intensity, Relative Mix, and Economic Determinants," *Journal of Business* 73, 661-693.

Cadenillas, A., J. Cvitanic, and F. Zapatero, 2004, "Leverage Decision and Manager Compensation with Choice of Effort and Volatility," *Journal of Financial Economics* 73, 71-92.

Chen, L., P. Collin-Dufresne, and R. Goldstein, 2006, "On the Relation between the Credit Spread Puzzle and the Equity Premium Puzzle," Working Paper, U.C. Berkeley.

Cheung, Y., and L. Ng, 1992, "Stock Price Dynamics and Firm Size: An empirical Investigation," *Journal of Finance* 47, 1985-1997.

Cooley, T., and V. Quadrini, 2001, "Financial Markets and Firm Dynamics," *American Economic Review* 91, 1286-1310.

DeMarzo, P., and M. Fishman, 2006, "Optimal Long-Term Financial Contracting," forthcoming in *Review of Financial Studies*.

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, E. and Q. Liu, 2008, "Inside Debt," forthcoming in *Review of Finance*.

Edmans, A., X. Gabaix, and A. Landier, 2007, "A Calibratable Model of Optimal CEO incentives in Market Equilibrium," Working paper, Wharton and NYU.

Edmans, A., X. Gabaix, T. Sadzik, and Y. Sannikov, 2009, "Dynamic Incentive Accounts," Working paper, Wharton, NYU, and Princeton.

Eisfeldt, A. and A. Rampini, 2007, "Managerial Incentives, Capital Reallocation, and the Business Cycle," forthcoming in *Journal of Financial Economics*.

Frank, M. and V. Goyal, 2005, "Trade-off and Pecking Order Theories of Debt," Handbook of Corporate Finance: Empirical Corporate Finance, Chapter 7, edited by B, Eckbo, Elsevier/North-holland, 2005.

Fudenberg, D., B. Holmstrom, and P. Milgrom, 1990, "Short-Term Contracts and Long-Term Agency Relationships," *Journal of Economic Theory* 51, 1-31.

Gabaix, X., and A. Landier, 2008, "Why Has CEO Pay Increased So Much?" *Quarterly Journal of Economics*.

Goldstein R., N. Ju, and H. Leland, 2001, "An EBIT-Based Model of Dynamic Capital Structure," *Journal of Business* 74, 483-512.

Hall, B., and J. Liebman, 1998, "Are CEOs Really Paid Like Bureaucrats?" *Quarterly Journal of Economics* 113, 653-69.

Haubrich, J., 1994, "Risk Aversion, Performance Pay, and the Principal-Agent Problem," *Journal of Political Economy* 102, 258-276.

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

He, Z., 2010, "Dynamic Compensation Contracts with Private Savings," Working Paper, University of Chicago.

Hennessy, C., 2004, "Tobin's Q, Debt Overhang, and Investment," *Journal of Finance* 59, 1717-1742.

Holderness, C., R. Kroszner, and D. Sheehan, 1999, “Were the Gold Old Days That Good? Changes in Managerial Stock Ownership Since the Great Depression,” *Journal of Finance* 54, 435-469.

Holmstrom, B., and P. Milgrom, 1987, “Aggregation and Linearity in the Provision of Intertemporal Incentives,” *Econometrica* 55, 303-328.

Jensen, M., and K. Murphy, 1990, “Performance Pay and Top-management Incentives,” *Journal of Political Economy* 98, 225-264.

John, T., and K. John, 1993, “Top-Management Compensation and Capital Structure,” *Journal of Finance* 48, 949-974.

Leland, H., 1994, “Corporate Debt Value, Bond Covenants, and Optimal Capital Structure,” *Journal of Finance* 49, 1213-1252.

Leland, H., 1998, “Agency Costs, Risk Management, and Capital Structure,” *Journal of Finance* 53, 1213-1243.

Morellec, E., 2004, “Can Managerial Discretion Explain Observed Leverage Ratios?” *Review of Financial Studies* 17, 257-294.

Murphy, K., 1999, “Executive Compensation”, in Orley Ashenfelter and David Card (eds.), *Handbook of Labor Economics*, Vol. 3, North Holland.

Piskorski, T., and A. Tchistyi, 2007, “Optimal Mortgage Design”, Working paper, NYU.

Rogerson, W., 1985, “Repeated Moral Hazard,” *Econometrica* 53, 69-76.

Sannikov, Y., 2008, “A Continuous-Time Version of the Principal-Agent Problem,” *Review of Economic Studies* 75, 957-984.

Schattler, H., and J. Sung, 1993, “The First-Order Approach to the Continuous-Time Principal-Agent Problem with Exponential Utility,” *Journal of Economic Theory* 61, 331-371.

Strebulaev, I. A., 2007, “Do Tests of Capital Structure Theory Mean What They Say?” *Journal of Finance* 62, 1747-1787.

Subramanian, A., 2007, “Managerial Flexibility, Agency Costs and Optimal Capital Structure,” Working Paper, Georgia State University.

Sundaram, R., and D. Yermack, 2007, “Pay Me Later: Inside Debt and Its Role in Managerial Compensation,” *Journal of Finance* 62, 1551-1587.

Westerfield, M., 2006, “Dynamic Contracts with Hidden Actions in Continuous Time,” Working Paper, USC.

Williams, N., 2006, “On Dynamic Principal-Agent Problems in Continuous Time,” Working Paper, Princeton.

A Appendix

A.1 Proof of Lemma 2

Consider any contract $\Pi = \{c; a\}$ which induces an optimal policy $\{\widehat{c}; \widehat{a}\}$ from the agent with a value $\widehat{V}_0$. From the agent's budget equation, we have

$$S_t = \int_0^t e^{r(t-s)} \left(c_s - \widehat{c}_s\right) ds,$$

which gives the agent's optimal savings path, with the transversality condition $\lim_{T\to\infty} \mathbb{E}\left[e^{-rT} S_T\right] = 0$. Note that the transversality condition holds for all measures induced by any feasible effort policies.

By invoking the replication argument similar to revelation principle, we consider giving the agent a direct contract $\widehat{\Pi} = \{\widehat{c}; \widehat{a}\}$; notice that shareholders have the same cost to deliver this contract as $\Pi = \{c; a\}$. Clearly taking consumption-effort policy $\{\widehat{c}; \widehat{a}\}$ is feasible for the agent with no private-savings. We need to show that $\{\widehat{c}; \widehat{a}\}$ is indeed optimal for the agent. Suppose that given the contract $\widehat{\Pi}$, the agent finds that $\{c'; a'\}$ yields a strictly higher payoff $V_0' > \widehat{V}_0$ in problem (2), with associated savings path

$$S_t' = \int_0^t e^{r(t-s)} \left(\widehat{c}_s - c_s'\right) ds,$$

which satisfies transversality condition. However, consider the following saving path

$$S_t'' = S_t + S_t' = \int_0^t e^{r(t-s)} \left(c_s - c_s'\right) ds,$$

which also satisfies the transversality condition. But this implies that given the original contract Π, the saving rule S_t'' supports $\{c'; a'\}$ but delivering a strictly higher payoff V_0'. This contradicts with the optimality of $\{\widehat{c}; \widehat{a}\}$ under the contract Π.

A.2 Proof of Lemma 3

Given any savings $S_t = S$ and a contract $\Pi = \{c; a\}$, from time-t on the agent's problem is (recall problem (2))

$$\begin{aligned} \max_{\{\widehat{c}_s\},\{\widehat{a}_s\}} \quad & \mathbb{E}\left[\int_t^\infty -\frac{1}{\gamma} e^{-\gamma(\widehat{c}_s - g(\delta_s, \widehat{a}_s)) - r(s-t)} ds\right] \\ s.t. \quad & dS_s = rS_s ds + c_s ds - \widehat{c}_s ds,\ S_t = S,\ s > t \\ & d\delta_s = \mu\left(\delta_s, \widehat{a}_s\right) ds + \sigma\left(\delta_s\right) dZ_s. \end{aligned} \tag{32}$$

Denote by $\{\widehat{c}_s^*, \widehat{a}_s^*\}$ the solution to the above problem, and by $V_t\left(\Pi, S\right)$ the resulting agent's value.

Now consider the problem with $S = 0$, which is the continuation payoff along the equilibrium path:

$$\begin{aligned} \max_{\{\widehat{c}_s\},\{\widehat{a}_s\}} \quad & \mathbb{E}\left[\int_t^\infty -\frac{1}{\gamma} e^{-\gamma(\widehat{c}_s - g(\delta_s, \widehat{a}_s)) - r(s-t)} ds\right] \\ s.t. \quad & dS_s = rS_s ds + c_s ds - \widehat{c}_s ds,\ S_t = 0,\ s > t \\ & d\delta_s = \mu\left(\delta_s, \widehat{a}_s\right) ds + \sigma\left(\delta_s\right) dZ_s. \end{aligned}$$

We claim that the solution to this problem is $\{\widehat{c}_s^* - rS, \widehat{a}_s^*\}$, and therefore the value is $V_t\left(\Pi, 0\right) = e^{\gamma r S} V_t\left(\Pi, S\right)$. There are two steps to show this. First, this solution is feasible. Second, suppose that there exists another policy

$\{\widehat{c}'_s, \widehat{a}'_s\}$ that is superior to $\{\widehat{c}^*_s - rS, \widehat{a}^*_s\}$, so that the associated value $V'_t(\Pi, 0) > e^{-\gamma r S} V_t(\Pi, S)$. Consider $\{\widehat{c}'_s + rS, \widehat{a}'_s\}$, which is feasible to the problem (32). But under this plan the agent's objective is

$$e^{-\gamma r S} \cdot \max \mathbb{E}_t \left[\int_t^{\infty} -\frac{1}{\gamma} e^{-\gamma\left(\widehat{c}'_s - g\left(\delta_s, \widehat{a}'_s\right)\right) - r(s-t)} ds \right] = e^{-\gamma r S} V'_t(\Pi, 0) > V_t(\Pi, S).$$

which contradicts with the optimality of $\{\widehat{c}^*_s, \widehat{a}^*_s\}$. As a result, $V_t(\Pi, S) = e^{-\gamma r S} V_t(\Pi, 0)$.

A.3 Appendix for Section 2.3.4

A.3.1 An Example: Square-Root Mean-Reverting Cash-Flows

Suppose that $\mu(\delta_t, a_t) = (\Delta - \kappa\delta_t) + a_t\delta_t$, and $\sigma(\delta_t) = \sigma\sqrt{\delta_t}$, where Δ, κ, and σ are positive constants. The agent's effort cost $g(\delta, a) = \frac{\theta}{2} a^2 \delta$ which is quadratic in effort and linear in the cash-flow level. We have $\beta_t = \frac{g_a(\delta_t, a_t)}{\mu_a(\delta_t, a_t)} = \theta a$, and the risk-compensation is linear in firm size δ:

$$\frac{1}{2}\gamma r \left[\frac{g_a(\delta_t, a_t)\,\sigma(\delta_t)}{\mu_a(\delta_t, a_t)} \right]^2 = \frac{1}{2}\gamma r \theta^2 a^2 \sigma^2 \delta.$$

Since all elements in (12) are affine in δ, we guess that the firm value $f(\delta) = A + B\delta$, where A and B are constants to be determined. Focusing on the interior solution, we have $A = \frac{B\Delta}{r}$, and

$$B = \frac{2}{r + \kappa + \sqrt{(r+\kappa)^2 - \frac{2}{(1+\theta\gamma r\sigma^2)\theta}}}.$$

In the optimal contract, the constant optimal effort a^* is

$$a^* = \frac{B}{\theta\left(1 + \theta\gamma r\sigma^2\right)}.$$

The evolutions of the agent's wage (consumption) policy follows (8), and the agent's continuation value is governed by

$$dV_t = \gamma r(-V_t) \frac{B\sigma\sqrt{\delta_t}}{1 + \theta\gamma r\sigma^2} dZ_t. \tag{33}$$

This example gives a counterpart of Holmstrom and Milgrom (1987) who derive a constant optimal effort in a Gaussian framework.

We can solve Eq. (33) to obtain:

$$V_t = V_0 \exp\left(-\int_0^t \frac{1}{2}\gamma^2 r^2 \theta^2 a^{*2} \sigma^2 \delta_s ds - \int_0^t \gamma r \theta a^* \sigma \sqrt{\delta_t} dZ_s \right) < 0.$$

One can show that if $(\kappa - a^*)^2 \geq 2M^2\sigma^2$ where $M^2 = \frac{1}{2}\gamma^2 r^2 \theta^2 a^{*2} \sigma^2$, then $\{V\}$ is a indeed martingale. The proof is available upon request. Whenever the expression of V_t is needed, we will focus on this case in the following analysis.

A.3.2 Verification of Optimality of Contract

Now we verify that the optimal contract derived above is indeed optimal. As usual, we denote by c (a) the recommended consumption (effort), and $\widehat{c}\,(\widehat{a})$ is the agent's actual policies. The main argument applies to more general cases, with case-by-case technical requirement to ensure the transversality conditions.

Step 1. We show that the agent's policy stated in Section 2.3.3 is optimal, i.e., the contract is indeed incentive compatible and induce no savings. Given arbitrary policies $\{\widehat{c}, \widehat{a}\}$, consider the following auxiliary gain process

$$G_t^A = \int_0^t \frac{e^{-\gamma(\widehat{c}_s - g(\widehat{a}_s)) - rs}}{\gamma} ds + e^{-rt} e^{-\gamma r S_t} V_t,$$

where $S_t = \int_0^t e^{r(t-s)} \left(c_s - \widehat{c}_s\right) ds$ is the agent's cumulative (deviation) wealth, and

$$dV_t = \theta a^* \left(-\gamma r V_t\right) \left[d\delta_t - \mu\left(\delta_t, a_t^*\right) dt\right]$$

according to the optimal contract. Then

$$\begin{aligned} e^{\gamma r S_t} e^{rt} dG_t^A &= \frac{e^{-\gamma(\widehat{c}_t - g(\widehat{a}_t))}}{\gamma} e^{\gamma r S_t} dt - r V_t dt + \gamma r \left(r S_t + c_t - \widehat{c}_t\right) V_t dt \\ &\quad + \theta a^* \left(-\gamma r V_t\right) \left[\left(\Delta - \left(\kappa - \widehat{a}_t\right) \delta_t\right) dt + \sigma \sqrt{\delta_t} dZ_t - \mu\left(\delta_t, a_t^*\right) dt\right]. \end{aligned} \tag{34}$$

We aim to show that $\mathbb{E}_t \left[e^{\gamma r S_t} e^{rt} dG_t^A\right]$ is zero under optimal contract, while nonpositive for other policies. First, it is easy to see that the RHS of Eq. (34) (with control variables $\widehat{c}_t$ and $\widehat{a}_t$) is a well-defined concave problem, with FOC as

$$\begin{aligned} \widehat{c}_t &: \quad -\frac{e^{-\gamma(\widehat{c}_t - g(\widehat{a}_t))}}{\gamma} e^{\gamma r S_t} - r V_t = 0, \\ \widehat{a}_t &: \quad -e^{-\gamma(\widehat{c}_t - g(\widehat{a}_t))} e^{\gamma r S_t} \theta \widehat{a}_t \delta_t + \theta a^* \left(-\gamma r V_t\right) \delta_t = 0. \end{aligned}$$

These two conditions imply that the RHS of Eq. (34) takes its maximum value when $\widehat{a}_t = a' = a^*$, and $\widehat{c}_t = c'$ (here we denote by c' and a' as the agent's optimal policies) such that

$$-\frac{e^{-\gamma\left(c' - g(a^*)\right)}}{\gamma} e^{\gamma r S_t} = r V_t.$$

Since in the optimal contract $r V_t = u\left(c_t, a^*\right) = \frac{e^{-\gamma(c_t - g(a^*))}}{\gamma}$, we have

$$c' = r S_t + c_t,$$

which says that from time-t on it is optimal to consume future wages c and permanent interest income earned from today's private wealth $r S_t$. Plugging these results into Eq. (34), we have

$$e^{\gamma r S_t} e^{rt} dG_t^A = \text{non-positive drift} + \theta a^* \left(-\gamma r V_t\right) \sigma \sqrt{\delta_t} dZ_t, \tag{35}$$

and when the agent takes the recommended effort and consumption policy the drift is zero.

Eq. (35) implies that $\mathbb{E}\left[G_T^A\right] \le G_0^A + \int_0^T e^{-rt} e^{-\gamma r S_t} \theta a^* \left(-\gamma r V_t\right) \sigma \sqrt{\delta_t} dZ_t$, for $\forall T$. As usual, we impose extra conditions on feasible S_t so that the second term is indeed a martingale (when $S_t = 0$ along the equilibrium path, it is $\int_0^T e^{-rt}dV_t$, which is a martingale since $\{V\}$ is a martingale.) Because $G_0^A = V_0$, then the value from an arbitrary feasible policy delivers a value below V_0, i.e.,

$$\mathbb{E}\left[\int_0^\infty \frac{e^{-\gamma(\widehat{c}_s - g(\widehat{a}_s)) - rs}}{\gamma} ds\right] \leq V_0,$$

provided that the transversality condition $\mathbb{E}\left[\lim_{T\to\infty} e^{-rT}e^{-\gamma r S_T}V_T\right] = 0$ holds. When $S_t = 0$ along the equilibrium path, V is a martingale and therefore this transversality condition follows trivially.

Step 2. Now we verify the optimality of shareholders' policy. Define the shareholders' gain process as

$$G_t^I = \int_0^t e^{-rs}(\delta_s - c_s)\, ds + e^{-rt}J(\delta_t, V_t)$$

where $J(\delta_t, V_t) = \frac{B\Delta}{r} + B\delta_t + \frac{\ln(-\gamma r V_t)}{\gamma r}$. Due to construction, under any employment contract we have

$$dG_t^I = a_{G^{\mathrm{I}}}dt + e^{-rt}(B - \beta_t)\sigma\sqrt{\delta}dZ_t.$$

It is easy to show that due to construction, $a_{G^{\mathrm{I}}}$ is nonpositive, and zero under the optimal contract. we impose a usual square-integrable condition on feasible employment contracts:

$$\mathbb{E}\left[\int_0^T e^{-rt}\beta_t^2\sigma^2\delta_t dt\right] < \infty \text{ for } \forall T \tag{36}$$

Notice that in the optimal contract, β_t is the constant θa^*; therefore the optimal contract satisfies condition (36), and $\mathbb{E}\left[\int_0^T e^{-rt}G_t^I\right] = 0$ for all T.

Finally we check the shareholders' transversality condition $\mathbb{E}\left[\lim_{T\to\infty} e^{-rT}J(\delta_T, V_T)\right] = 0$. Under the optimal contract, since δ is stationary, we only have to check the term associated with V_T. But

$$\frac{\ln(-V_T)}{\gamma r} = \text{Constant} - \int_0^T \frac{1}{2}\gamma^2 r^2\theta^2 a^{*2}\sigma^2\delta_t dt + \int_0^T \gamma r\theta a^*\sigma\sqrt{\delta_t}dZ_t,$$

and again the second diffusion term is a martingale given the condition $(\kappa - a^*)^2 \geq 2M^2\sigma^2$ (the proof is available upon request.) Because $\mathbb{E}\left[\int_0^T \delta_t dt\right]$ is in the order of T when $T \to \infty$ as the long-run mean of δ_t is $\frac{\Delta}{\kappa - a^*}$, transversality condition holds for the optimal contract. For general policies, notice that $J(\delta_T, V_T) \leq \frac{B^{FB}\Delta}{r} + B^{FB}\delta_T + \frac{\ln(-\gamma r V_T)}{\gamma r}$ which is the first-best result. Because δ is mean-reverting, the sufficient condition is $\mathbb{E}\left[\lim_{T\to\infty} e^{-rT}\frac{\ln(-\gamma r V_T)}{\gamma r}\right] \leq 0$, which is the condition that we impose in addition to (36) for any feasible contract Π.

A.4 Appendix for Section 3.1

We have $rA^{FB} = \max_{a\in[0,\overline{a}]}\left\{1 + A^{FB}(\phi + a) - \frac{\theta}{2}a^2\right\}$, and are interested in characterizing the condition for the first-best effort a^{FB} to bind at $\overline{a}$. When a takes an interior solution, then $a^* = \frac{A^{FB}}{\theta}$, and $A^{FB} = \frac{2}{r-\phi+\sqrt{(r-\phi)^2 - \frac{2}{\theta}}}$. First, we need to ensure that $(r-\phi)^2 > \frac{2}{\theta}$ so that A^{FB} is real; otherwise the firm value is unbounded, resulting in an unbounded a^{FB}. Therefore, when $(r-\phi)^2 \leq \frac{2}{\theta}$, $a_t^{FB} = \overline{a}$. Second, even when $(r-\phi)^2 > \frac{2}{\theta}$, it is possible that the implied the first-best effort $\frac{1}{\theta}\frac{2}{r-\phi+\sqrt{(r-\phi)^2-\frac{2}{\theta}}} > \overline{a}$. In this case, the first-best effort also binds at $a_t^{FB} = \overline{a}$. Under both scenarios, $A^{FB} = \frac{1-\frac{\theta}{2}\overline{a}^2}{r-\phi-\overline{a}}$.

A.5 Appendix for Section 3.2

Under the boundary conditions specified in Section 3.2, f' is always positive in (24), therefore a^* never binds at zero in this problem. To see this, clearly $f'(0) > 0$ (notice $f(0) = 0$, and for $\delta > 0$, even with zero effort (so without the agent) the value is positive). When when $\delta \to \infty$, using (24) and (25) one can check that

$$f(\delta) \simeq \overline{f}(\delta) \equiv \frac{1}{r-\phi}\delta + \frac{1}{2(r-\phi)^2 \theta^2 \gamma r \sigma^2}. \tag{37}$$

which is increasing in δ. Now suppose that there exists $\delta_1 > 0$ such that $f'(\delta_1) = 0$; take the smallest one so that $f''(\delta_1) < 0$. Therefore we have $rf(\delta_1) = \delta_1 + \frac{1}{2}f''(\delta_1)\sigma^2\delta_1^2 < \delta_1$. But there must exists another point $\delta_2 > \delta_1$, such that f is deceasing in $[\delta_1, \delta_2]$, and f becomes increasing again after δ_2 (as f is increasing when δ is large enough). This implies $f(\delta_2) < f(\delta_1)$, $f'(\delta_2) = 0$ and $f''(\delta_2) > 0$. But (24) implies that $rf(\delta_2) > \delta_2 > \delta_1 > rf(\delta_1)$, contradiction with $f(\delta_2) < f(\delta_1)$.

A.6 Appendix for Section 3.3.2

When $\delta \to \infty$, the probability of bankruptcy is negligible, and the firm value can be viewed as the unlevered firm value minus the present value of after-tax coupon payment. Eq. (37) gives the unlevered firm value f when $\delta \to \infty$; notice that f' approaches to $\frac{1}{r-\phi}$ similar to the standard Gordon growth model. Therefore $f^E(\delta) \simeq \overline{f}(\delta) - \frac{(1-\tau)C}{r}$ when δ is sufficiently large.

There is one technical issue under this geometric Brownian motion setup. As illustrated in Appendix A.3, we require that the agent's continuation value V defined in (10) is indeed a martingale. Unfortunately, if δ follows an unbounded geometric Brownian motion, we are unable to verify that $\{V\}$ is martingale. To circumvent this technical issue, we simply assume that when the firm's δ reaches a sufficiently large $\overline{\delta}$, shareholders simply sell the firm at a price $\overline{f}(\overline{\delta})$, and the employment contract terminates once this upper threshold is reached. This simplifying assumption can be interpreted as an successful exit option (e.g., M&A) in reality, and has no impact on the static capital structure decision at date 0. Finally, once the state space becomes bounded, following the argument in Section A.3 one can easily show the optimality of the contract.

We use $Matlab$ built-in solver *bvp4c* to solve the model, setting the tolerance level to be 10^{-8}. We set $\overline{\delta} = 500$. In solving for the bankruptcy boundary δ_B, we first choose one candidate δ_B^1, and solve f^E based on (29), (30), and (31) using *bvp4c*. Then we evaluate $f^{E\prime}(\delta_B^1)$; if $f^{E\prime}(\delta_B^1)$ >(<)0 which means that δ_B^1 is too low (high), we adjust δ_B^1 upward (downward). We use a bisection method to search for δ_B, which converges rather quickly.