Optimal Short-Termism*

Dirk Hackbarth† Alejandro Rivera‡ Tak-Yuen Wong§

October 28, 2017

Abstract

This paper studies incentives in a dynamic contracting framework of a levered firm. In particular, the manager selects long-term and short-term efforts, while shareholders choose initially optimal leverage and ex-post optimal default policies. There are three results. First, shareholders trade off the benefits of short-termism (current cash flows) against benefits of higher growth from long-term effort (future cash flows), but because shareholders only split the latter with bondholders, they find short-termism ex-post optimal. Second, bright (grim) growth prospects imply lower (higher) optimal levels of short-termism. Third, the endogenous default threshold rises with the substitutability of tasks and, for a positive correlation of shocks, the endogenous default threshold is hump-shaped in the volatility of permanent shocks, but increases monotonically with the volatility of transitory shocks. Finally, we quantify agency costs of short-term and long-term effort, cost of short-termism, effects of investor time horizons, credit spreads, and risk-shifting.

*JEL Classification Numbers: D86, G13, G32, G33, J33.
Keywords: Capital structure, Contracting, Multi-tasking.

---

*We are grateful to Zhiguo He, Ron Kaniel, Leonid Kogan, Robert Marquez, Erwan Morellec, Stephen Terry, Valery Polkovnichenko, Antoinette Schoar, and participants at the 2017 Asian Meeting of Econometric Society (Hong Kong) for helpful comments.

†Questrom School of Business, Boston University. Email: dirkhackbarth@gmail.com.
‡Naveen Jindal School of Management, University of Texas at Dallas. Email: alejorivera1@gmail.com
§The School of Finance, Shanghai University of Finance and Economics. Email: etywong110@gmail.com.
1 Introduction

Corporate short-termism has been criticized often, because it could harm long-term performance. That is, managers arguably take actions that are favorable for them in the short-term at the expense of shareholders’ interest in increasing stock prices. Are corporate managers myopic when they do not invest sufficiently for the long term and hence short-termism is sub-optimal for shareholders? Or rather, can the behavior of managers be a result of equity value-maximization that crucially depends on firm characteristics, such as debt-equity ratio and growth prospects?

In this paper, we study optimal contracting between the agent (manager) and shareholders in a dynamic framework of a company funded with equity and risky debt. Shareholders choose optimal debt and default policies (Leland, 1994). In addition, shareholders design incentives based upon which the manager selects long-term effort (growth) and short-term effort (profit) in a multi-tasking environment (Holstrom and Milgrom, 1991). Following He (2011), capital structure, contracting, and provision of efforts are jointly optimized in our setting.

On the one hand, recoveries in bankruptcy transfer cash flows from equity to debt, so shareholders do not internalize all benefits from long-term effort (underinvestment). On the other hand, shareholders receive all benefits from short-term effort immediately, because it generates higher contemporaneous cash flows that are not transferred to debt in bankruptcy. More specifically, debt has two implications in our model. First, it decreases long-term effort, i.e., increases underinvestment (Myers, 1977). Second, it does not reduce the gains from short-term effort. However, because it decreases long-term effort, the manager has more resources for exerting short-term effort. If the marginal benefit of short-term effort exceeds its marginal cost, it is optimal for shareholders to induce short-term effort through an optimal contract. In equilibrium, shareholders trade off benefits of short-termism (current cash flows) against benefits of higher growth from long-term effort (future cash flows). In essence, there is an optimal level of short-termism for shareholders, because short-term effort reduces default risk.

We model two stochastic processes: the firm’s size and the firm’s profitability. The per period...
firm cash flows are given by the product of these two processes. Shocks to firm size have a persistent effect on the firm’s future cash flows (permanent shocks), while shocks to the firm profitability only have a contemporaneous effect on cash flows (transitory shocks). In our multi-tasking framework, managerial effort can be allocated towards increasing the baseline growth rate of the firm (long-term effort) or increasing the baseline profitability of the firm (short-term effort). The timing is as follows: the initial owners of the firm issue infinite maturity debt. Optimal leverage trades off the tax advantage of debt with the costs of bankruptcy. Once debt is in place, shareholders design an incentive compatible contract to implement the effort and default policies that maximize equity value. Since effort is not observable, incentive compatibility requires exposing the manager to the permanent and transitory shocks. Upon default, bondholders collect the firm assets net of costs of bankruptcy.

While the paper’s main result is that short-termism can be optimal for shareholders, it is also the first to highlight an asymmetric interaction between underinvestment and short-termism. When the levered firm moves towards financial distress, the underinvestment problem increases and hence it becomes increasingly desirable for shareholders to implement higher levels of short-term effort. Less long-term effort reduces the risk of investment benefits being largely reaped by bondholders and also make it cheaper for shareholders to incentivize more short-term effort (short-termism). Such higher levels of short-termism are optimal for shareholders, but detrimental to bondholders. Therefore, short-termism is another dimensions of the agency cost of debt. Propositions 1 and 2 formalize that short-termism is an indirect — in addition to underinvestment as a direct — agency cost of debt, and show, perhaps surprisingly, that if shareholders commit to no underinvest, then this is also a commitment to no short-termism, but the converse is not true.

There are a number of additional results. First, our model highlights potential endogeneity concerns in certain empirical specifications. We find that firms with bright growth prospects optimally choose to focus on long-term growth, while firms with grim growth prospects optimally focus on the short-term. Hence, in equilibrium one would observe that high growth firms are those which invest in the long-term. This does not mean that low growth firms should mimic the long-term approach of their high growth counterparts, because it would be value destroying for low growth firms to implement higher levels of long-term effort.

---

3 See Summers (2017) for an analogy of golfers with long swings, such as Phil Mickelson, who hit the ball further and more accurately than amateur golfers with short swings. An index of swing length would be correlated with most measures of golf performance. However, this does not mean that amateur golfers should lengthen their swings. Mimicking the long swing of professional golfers like Phil Mickelson would be detrimental for most amateur golfers’ performance.

4 Indeed, Kaplan (2017) finds that there is little long-term evidence in favor of the so-called short-termism critique.
Second, the continuous-time, dynamic framework also permits analytic comparative statics, which are especially helpful for understanding the effect of parameters on equity value and default boundary. To begin, note that because default is equity value-maximizing, a higher equity value implies a lower default boundary and vice versa. For example, higher costs of short-term or long-term efforts lower equity value and hence raise the default threshold. Higher baseline profitability or higher baseline growth increase equity value and hence decrease the default threshold. Moreover, the default threshold is inverted U-shaped in the volatility of the permanent shocks. To see this, consider an increase in the volatility of the permanent shocks. On the one hand, the information filtering effect accelerates default time. With a higher volatility, the firm size is a more noisy signal about the agent’s long-term action. By the informativeness principle (Holmstrom, 1979), the shareholders reduce the long-term incentives, because they are costlier (incentive effect). This results in lower cash flow growth rate and hence higher default threshold. On the other hand, a more volatile permanent shock makes the option to default (or wait) more valuable (real option effect). In sum, for low values of permanent shock volatility, the incentive effect dominates, while for high values the real option effect dominates. In contrast, the default threshold rises with the volatility of the transitory shocks because of the incentive effect. As profitability becomes a noisier signal about short-term effort, shareholders diminish short-term incentives. If the firm is in financial distress, shareholders absorb more losses and endogenously default sooner. So, equity value decreases in the volatility of the transitory shocks. Finally, a higher correlation of transitory and permanent shocks increases the risk borne by the manager, which increases compensation costs. However, a higher correlation does not affect the option to default. Therefore, correlation unambiguously reduce equity value.

Third, we adopt a baseline calibration and quantify the agency cost of debt, which we decompose it into underinvestment and excessive short-termism. We compute equity and debt values if shareholders can commit to unlevered effort policies, and compare them to values with levered effort policies. The reduction in total firm value due to debt overhang is about one percent, where up to one half of this reduction is due to excessive short-termism. However, contrary to standard intuition, managerial short-termism is not detrimental to equity value, but in fact desirable. Short-termism is an indirect cost of debt overhang and hence there are two related commitment problems, one for underinvestment and one for short-termism, which bondholders have to recognize at the outset.

5In a dynamic liquidity management model, Décamps et al. (2017) find correlation decreases default risk, because their state variable, scaled cash holdings, drifts away faster from the default boundary when the correlation is higher.

6He (2011) assume implicitly commitment to no short-termism and hence obtains higher ex-ante firm values.
Fourth, we extend the model to the case in which a subset of shareholders with a shorter time horizon (higher discount rate) takes control of the firm. Impatient shareholders will find optimal to implement higher short-term effort and lower long-term effort, which in turn will reduce equity value for the patient, regular shareholders and bondholders, who both employ the baseline discount rate. In our baseline calibration, a one percentage point increase in the discount rate of impatient shareholders leads approximately to a reduction of one percent in equity value, five percent of debt value, and two and a half percent of total firm value. Thus, our model predicts that a transfer of control to investors with shorter time horizons induces a sizable reduction in debt, equity, and firm values, which is consistent with the conventional critique of short-termism (e.g., Stein, 1989).7

Fifth, our model reveals opposite effects for increases in the volatility of permanent shocks (growth shocks) versus and increases in the volatility of transitory shocks (profit shocks). In particular, higher volatility of permanent shocks increases the value of the shareholders option to default thereby increasing equity value at the expense of bondholders. However, higher volatility of transitory shocks is value reducing for both shareholders and bondholders, since higher volatility increases the cost of providing managerial incentives. Thus, our model predicts that shareholders have an incentive to risk-shift with regards to permanent shocks but not to transitory shocks.

Our paper relates to the literature on financial markets and managerial myopia. Early works by Stein (1988, 1989), Shleifer and Vishny (1990), and Von Thadden (1995) argue that short-termism arises when a manager faces takeover threats or arbitrageurs with short-horizon; and also when a firm is financed by short-term debts. Holmstrom and Tirole (1993) show that the liquidity of the stock market affects the efficiency of equity-based compensation in disciplining managerial myopia. Froot, Perold and Stein (1992) discuss a potential link between the short-term horizon of shareholders and short-term managerial behavior. Compared to these works, our full-fledged dynamic framework allows us to quantify the impact of short-termism on firm valuations and decisions.

More recently, Bolton et al. (2006) argue that overly optimistic investors choose an equity-based compensation that weights the short-run stock performance more heavily and thus induce myopia. Edmans et al. (2012) and Marinovic and Varas (2017) show that equity investing implements the manipulation-minimizing optimal dynamic contract. While these works focus on short-term actions which are value-destroying (e.g., earnings manipulation), our paper focuses in the case in

7However, ex-post equity values are higher in our setting, because short-termism enhances equity value.

7To the extent that activist investors, hedge funds, and vulture funds have a shorter time horizon, they may not necessarily add value, but of course they also influence incentives of managers in other ways.
which short-term effort can be value-enhancing (e.g., cost cutting, streamlining). Thus, we do not seek for contracts that induce no short-termism. Instead, we characterize the optimal amount of short-termism, leverage, and default in a joint optimization problem.

Our paper builds also on recent research about corporate policies under permanent and transitory shocks. Gorbenko and Strebulaev (2010) study the effect of temporary Poisson shocks in the Leland framework. Décamps et al. (2016) study permanent and transitory shocks in an equity financed firm with liquidity concerns. Byun et al. (2017) and Gryglewicz et al. (2017) empirically test the predictions of the two-shock models. In contrast, we assume costless external financing and focus on the moral hazard dimensions of persistent and transitory shocks. These allow us to capture the endogenous variations in the distribution of returns of the long-term and short-term and relate them to the debt overhang problem. Thus, we view our contribution as complementary to theirs.

Numerous empirical studies examine the determinants and effects of short-termism, especially in accounting. In a notable example, Edmans et al. (2017) use stock option vesting periods to corroborate the link between managerial incentives and short-termistic investment. Brochet et al. (2015) create an index of short-termism from the language used by executives during conference calls and document that short-term oriented companies have lower accounting performance in the future. Terry (2017) finds short-termism matters at the aggregate level in a quantitative macro model. Our result on optimal short-termism complements one of his specifications in which short-term targets ameliorate empire-building (i.e., overinvestment in R&D) problems.

The paper proceeds as follows. Section 2 describes the model and Section 3 solves it. To study the properties of the solution numerically in Section 5, we further specify the model in Section 4. Section 6 concludes and the Appendices contain mathematical developments.

## 2 Model Setup

Consider a continuous-time principal-agent model with infinite time horizon. At time 0, the firm (principal) hires a manager (agent) to operate a project, which produces a stream of cash flows subject to both permanent and transitory shocks. The principal is risk neutral and discounts cash flows at rate \( r > 0 \). Permanent shocks affect the long-term prospects of the project. In particular, denote \( \delta_t \) as the firm size, and it follows the stochastic process

\[
d\delta_t = \mu(a_t, \delta_t)dt + \sigma(\delta_t)dZ_t^P,
\]  

(1)
where $\mu(a, \delta_t)$ is the growth rate of firm size, $a_t \geq 0$ is an unobservable long-term effort exerted by the agent, $\sigma(\delta_t) > 0$ is the volatility, and $Z^P_t$ is a standard Brownian motion. The long-term effort increases the growth rate at a decreasing rate, so we assume the partial derivatives are $\mu_a(a, \delta) > 0$ and $\mu_{aa}(a, \delta) \leq 0$. In addition to permanent shocks, cash flows are subject to transitory shocks.

The contemporaneous profitability follows the stochastic process

$$dA_t = \alpha(e_t)dt + \sigma_A dZ^T_t,$$

where $\alpha(e_t)$ is the drift rate of the profitability, $e_t \geq 0$ is an unobservable short-term effort exerted by the agent, $\sigma_A > 0$ is the volatility, and $Z^A_t$ is a standard Brownian motion. Here, the short-term effort is productive and increases the drift at a decreasing rate: $\alpha_e(e) > 0$ and $\alpha_{ee}(e) \leq 0$. Over a small time interval $(t, t + dt)$, the project generates cash flows

$$dY_t = \delta_t A_t = \delta_t \left( \alpha(e_t)dt + \sigma_A dZ^T_t \right). \quad (3)$$

Permanent shocks $Z^P_t$ and transitory shocks $Z^T_t$ have a correlation coefficient $\rho \in [-1, 1]$, so that $\mathbb{E}_t [dZ^P_t dZ^T_t] = \rho dt$. The principal can observe the paths of the cash flows $(Y_t)_{t \geq 0}$ and firm size $\delta = (\delta_t)_{t \geq 0}$. This implies the incremental profitability $dA_t$ is observable to the principal as well.

Our formulation is in fact a multi-tasking agency problem. While the agent’s long-term effort $a_t$ directly increases the growth rate of firm size, the short-term action $e_t$ increases the contemporaneous profitability which only affects the current cash flows. To see this more closely, suppose $\mu(a, \delta) = \mu(a)\delta$ and $\sigma(\delta) = \sigma_{\delta}\delta$, then $\delta_t$ follows a geometric Brownian motion with the drift rate controlled by the agent: $d\delta_t = \delta_t \left( \mu(a_t)dt + \sigma_{\delta} dZ^P_t \right)$. The time-$t$ project value under a fixed effort pair $(\bar{a}, \bar{e})$, for example, the first-best effort (see Section 4), is

$$V(\delta_t) = \mathbb{E}^{\bar{a}, \bar{e}} \left[ \int_t^{\infty} e^{-r(s-t)} dY_s \right] = \frac{\alpha(\bar{e})\delta_t}{r - \mu(\bar{a})}.$$

From this expression, a positive permanent shock $dZ^P_t > 0$ increases the firm size $\delta_t$ and thus the project value. In contrast, a positive profitability shock $dZ^A_t > 0$ has no effect on the project value.

The technology specification admits two important special cases. First, when $\sigma_A = 0$, $e_t$ is perfectly observable and the principal can always implement optimal short-term effort. The cash flow
dynamics will be driven solely by the firm size, as in many structural models of corporate finance.\textsuperscript{10} Second, when $\sigma(\delta_t) = 0$, $a_t$ is perfectly observable and the principal can always implement optimal long-term effort. The profitability will drive the cash flows\textsuperscript{11}

The agent is risk averse and her instantaneous utility takes the form of exponential preferences

$$u(c_t, a_t, e_t) = \frac{-1}{\gamma} e^{-\gamma(c_t - g(a_t, e_t; \delta_t))},$$

where $c_t \in \mathbb{R}$ is the consumption rate and $g(a_t, e_t; \delta_t)$ is the monetary cost of effort. We assume the effort cost is increasing and convex in each task: $g_a(a, e; \delta) > 0$, $g_{ao}(a, e; \delta) \geq 0$, $g_e(a, e; \delta) > 0$, and $g_{ee}(a, e; \delta) \geq 0$. We allow for interdependency of tasks. When $g_{ae}(a, e; \delta) \leq 0$, tasks are complement; and when $g_{ae}(a, e; \delta) \geq 0$, tasks are substitute. Lastly, $\gamma$ is the coefficient of absolute risk aversion under CARA utility. Given a stream of consumption $\langle c_t \rangle_{t \geq 0}$, her expected discounted utility is

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

In addition, the agent has access to a private saving account, for which she can borrow and save at the interest rate $r$. We denote $S_t$ as the account balance at time $t$, and for simplicity, we assume that the agent has no initial saving $S_0 = 0$.

### 2.1 The Contracting Problem

At time 0, the principal designs a contract to maximize her expected discounted profits. Both principal and agent can fully commit to the contract, $\Gamma = \langle c, a, e \rangle$, which specifies the agent’s wage process $\langle c_t \rangle_{t \geq 0}$, and the recommended effort pair $\langle a_t, e_t \rangle_{t \geq 0}$; all processes are adapted to the filtration generated by $\langle \delta_t, A_t \rangle_{t \geq 0}$. Given a contract $\Gamma$, the agent solves the following problem

$$W_0(\delta_0, \Gamma) = \max_{\langle \hat{c}_t, \hat{a}_t, \hat{e}_t \rangle_{t \geq 0}} \mathbb{E}^{\hat{a}, \hat{e}} \left[ \int_0^\infty -\frac{1}{\gamma} e^{-\gamma(\hat{c}_t - g(\hat{a}_t, \hat{e}_t; \delta_t))} e^{-rt} dt \right]$$

s.t. $dS_t = rS_t dt + c_t dt - \hat{c}_t dt$, $S_0 = 0$, $S_t \geq 0$, \hfill (4)

and also subject to (1) to (3). In (4), $W_0(\delta_0, \Gamma)$ is the agent’s time-0 value under the contract $\Gamma$ and $\mathbb{E}^{(a,e)}[\cdot]$ is the expectation under the probability measure induced by $(a, e)$. The inter-temporal budgeting constraint specifies the evolution of the agent’s savings account: Over a small time

\textsuperscript{10}See, e.g., Leland (1994), Hackbarth et al. (2006), Goldstein et al. (2001), Strebulaev (2007), and, for a recent survey, Strebulaev and Whited (2012). Note that He (2011) is a special case of our model when $\sigma_A = 0$ and $a(e) \equiv 1$.

\textsuperscript{11}The technology resembles liquidity management models (Bolton, Chen, and Wang, 2011; Décamps et al., 2012) and models with dynamic agency (DeMarzo and Sannikov, 2006; DeMarzo et al., 2012; Miao and Rivera, 2016).
interval \((t, t + dt)\) the change in saving \(dS_t\) is the accumulated interest \(rS_t dt\) plus his wage \(c_t dt\) minus his actual consumption \(\hat{c}_t dt\).

We define a contract as incentive-compatible and no-savings (ICNS) if the solution to the agent’s problem is \((c, a, e)\). That is, the agent follows the recommended effort pair obediently (i.e., \(\hat{a}_t = a_t\) and \(\hat{e}_t = e_t\)); and does not save or withdraw from the bank account (i.e., \(\hat{c}_t = c_t\))\(^{12}\). Formally, the principal’s optimal contracting problem is

\[
\max_{\Gamma} \mathbb{E}^{(a,e)} \left[ \int_{0}^{\infty} e^{-rt}(dY_t - c_t dt) \right] \\
\text{s.t. \(\Gamma\) is ICNS and \(W_0(\Gamma) \geq w_0\)}
\]

where the first constraint requires the contract is incentive-compatible and no-savings, and the second constraint is the agent’s participation constraint for an initial outside option with value \(w_0\).

3 Model Solution

3.1 Necessary and Sufficient Conditions for ICNS Contracts

To characterize the optimal contract, we need to characterize the necessary and sufficient conditions for the contract to be incentive-compatible and no-savings. We first derive the dynamics of the agent’s continuation value.

3.1.1 Dynamics of continuation value

Following the dynamic contracting theory (see Sannikov (2008)), we take the agent’s continuation value as the state variable. Given a contract \(\Gamma\), the agent’s continuation value is defined as

\[
W_t(\delta_t, \Gamma) \equiv \mathbb{E}^{(a,e)} \left[ \int_{t}^{\infty} -\frac{1}{\gamma} e^{-\gamma (c_s - g(a_s, e_s; \delta_s)) - r(s-t)} ds \right]
\] (5)

Here, the agent’s continuation value depends on the firm size \(\delta_t\), which serves as a natural state variable, and the contract \(\Gamma\) that induces continuation consumption \((c_s)_{s \geq t}\) and effort choices \((a_s, e_s)_{s \geq t}\).

By the martingale representation theorem, the dynamics of the continuation value is

\[
dW_t = rW_t dt - u(c_t, a_t, e_t)dt + (-\gamma r W_t) \left( \beta_t^P \sigma(dZ_t^P) + \beta_T^T \delta_t \sigma_A dZ_t^T \right)
\] (6)

\(^{12}\)The focus on incentive-compatible and no saving contracts is without loss of generality. This is because the principal has full commitment and she can save on behalf of the agent.
where \((\beta_t^P, \beta_t^T)_{t \geq 0}\) are progressively measurable processes that capture the agent’s exposure to permanent and transitory shocks. In equation (6), the scaling factor \(-\gamma r W_t\) translates value in dollars to value in utility. This is because \(-\gamma r W_t > 0\) equals the agent’s marginal utility of consumption as we will see in the next subsection. On the equilibrium path, the last two terms on the right-hand side of (6) are the increment of the standard Brownian motions. Hence, \(E_{t}^{(a,e)} [dW_t + u(c_t, a_t, e_t)dt] = rW_t dt\), which states that the current flow utility and the expected changes in the continuation value equals the required return, and it reflects the promise-keeping constraint.

The novelty of our model is that the agent’s effort choice affect both the firm size and the contemporaneous profitability. Thus, there are both short-term incentives and long-term incentives that control the effort choices of the multi-tasking agent. The incentive provisions are captured by the volatility terms in (6). In a more detailed way, substitute (1) and (2) into (6) to obtain

\[
dW_t = (rW_t - u(c_t, a_t, e_t)) dt + (-\gamma r W_t) (\beta_t^P (d\delta_t - \mu(a_t, \delta_t)dt) + \beta_t^T \delta_t (dA_t - \alpha(e_t) dt))
\]

where the “transitory” (“permanent”) part captures the compensation promised to the agent due to the exposure to the transitory (permanent) shocks.

### 3.1.2 Conditions for no-savings

It is well-known that the wealth effect is absent with CARA preferences. This implies that the agent’s continuation value will scale with savings \(S_t\) in a convenient way. Specifically, consider the optimal continuation policy \((c_s, a_s, e_s)_{s \geq t}\) given a contract \(\Gamma\); and suppose the agent is given extra savings \(S\) at time \(t\). In the absence of wealth effect, the new optimal continuation policy is to follow the original continuation effort choices and consume an extra amount \(rS\) at every time in the future \(s \geq t\), that is \((c_s + rS, a_s, e_s)_{s \geq t}\) is the new optimal policy. It follows from the exponential preferences that \(u(c_s + rS, a_s, e_s) = e^{-\gamma r S} u(c_s, a_s, e_s)\) for all \(s \geq t\). Then, in terms of utility, an agent with savings \(S\) at time \(t\) must have continuation value

\[
W_t(\delta_t, \Gamma; S) = e^{-\gamma r S} W_t(\delta_t, \Gamma; 0),
\]

where \(W_t(\delta_t, \Gamma; 0)\) is the agent’s continuation value without savings as defined in (5).

Given an effort policy, the agent’s problem (4) implies a necessary condition for consumption on the no-savings path: The agent’s marginal utility from consumption must equal her marginal value.

\[^{13}\]For a formal proof, see Lemma 3 in He (2011).
of wealth. That is, \( u_c(c_t, a_t, e_t) = \frac{\partial}{\partial S} W_t(\delta_t, \Gamma; 0) \). Then by condition (7), we have \( \frac{\partial}{\partial S} W_t(\delta_t, \Gamma; S) = -\gamma r e^{-\gamma r S} W_t(\delta_t, \Gamma; 0) \). Evaluating this expression at \( S = 0 \) we obtain

\[
 u_c(c_t, a_t, e_t) = \frac{\partial}{\partial S} W_t(\delta_t, \Gamma; 0) = -\gamma r W_t(\delta_t, \Gamma; 0) \tag{8}
\]

as a necessary condition for the contract \( \Gamma \) to induce no savings. Condition \( \ref{eq:8} \) has a few implications. First, \( u_c = -\gamma r W_t \) implies the scaling factor in (6) is the marginal utility of consumption. Therefore, this factor translates dollar values to units of utility and allows us to interpret the incentive loading \( \beta^P_t \) and \( \beta^T_t \) as monetary incentives. Second, given our CARA assumption it must be true that \( u(c_t, a_t, e_t) = r W_t \) and the drift of (6) vanishes. The continuation of the agent’s continuation value satisfies

\[
d W_t = (-\gamma r W_t) \left( \beta^T_t \delta_t (dA_t - \alpha(e_t)dt) + \beta^P_t (d\delta_t - \mu(a_t, \delta_t)dt) \right)
\]

in equilibrium. As a result, the continuation \( W_t \) evolves as a martingale under the no-savings contract. This also implies that the marginal utility of consumption \( u_c = -\gamma r W_t \) follows a martingale.

Finally, the no private saving condition \( u(c_t, a_t, e_t) = r W_t \) allows us to pin down the consumption process for a given continuation value and effort levels:

\[
c_t = g(a_t, e_t; \delta_t) - \frac{1}{\gamma} \ln(-\gamma r W_t). \tag{9}
\]

### 3.1.3 Incentive compatibility

We now characterize the necessary and sufficient conditions for incentive compatibility. Given the monetary incentives \( (\beta^P_t, \beta^T_t) \), the agent chooses long-term action \( \hat{a}_t \) and short-term action \( \hat{e}_t \) to maximize her continuation value at each point in time. Specifically, the payoff consists of the current flow utility \( u(c_t, \hat{a}_t, \hat{e}_t) \) and the expected change in continuation value \( \mathbb{E}_t [d W_t(\hat{a}_t, \hat{e}_t)] \). Effort choice affect the current utility because they are costly, but they also affect the evolution of continuation value because effort changes the growth rate of the firm size \( d\delta_t \) as well as profitability shocks \( dA_t \), which are in turn connected to the continuation value through the monetary incentives \( (\beta^P_t, \beta^T_t) \). Therefore, the agent solves:

\[
\max_{(a_t, e_t)} \left\{ u(c_t, a_t, e_t) + \beta^P_t u_c(a_t, \delta_t) + \beta^T_t u_c(e_t) \delta_t \right\}
\]
The first-order conditions are
\[ -g_e(a_t, e_t; \delta_t) + \beta_t^P \mu_e(a_t, \delta_t) = 0 \Rightarrow \beta_t^P = \frac{g_e(a_t, e_t; \delta_t)}{\mu_e(a_t, \delta_t)} \tag{10} \]
\[ -g_e(a_t, e_t; \delta_t) + \beta_t^T \alpha_e(e_t)\delta_t = 0 \Rightarrow \beta_t^T = \frac{g_e(a_t, e_t; \delta_t)}{\alpha_e(e_t)\delta_t} \tag{11} \]

Intuitively, given incentives \((\beta_t^P, \beta_t^T)\), the agent balances the marginal monetary benefit of effort (given by the “\(\beta_t^t\)” term times the marginal impact of the drift of the respective process) and the marginal monetary cost of effort. As the agent is multi-tasking, the marginal cost of effort on one task depends also on the effort exerted on the other task. To implement \(\hat{a}_t = a_t\) and \(\hat{e}_t = e_t\), that is, for the long-term effort \(a_t\) and the short-term effort \(e_t\) to be incentive-compatible, the monetary incentives \((\beta_t^P, \beta_t^T)\) must satisfy conditions (10) and (11) simultaneously. \(^{14}\)

### 3.2 Optimal Contract

Given the dynamics of continuation value and the conditions for incentive-compatible and no-savings contracts, we can now write the principal’s problem recursively. At each point in time, given the state variables \(\delta_t\) and \(W_t\), the principal’s problem is
\[
P(\delta_t, W_t) = \max_{(\hat{a}_t, \hat{e}_t, \hat{\delta}_t)} \mathbb{E} \left[ \int_{t}^{\infty} e^{-r(s-t)} (dY_s - c_sds) \right]
\]
subject to \(dW_t = (-\gamma r W_t) \left( \beta_t^T \delta_t \sigma_A dZ_t^T + \beta_t^P \sigma(\delta_t) dZ_t^P \right)\)

and conditions (9), (10), (11)

where \(P(\delta_t, W_t)\) denotes the principal’s value function. Following He (2011), we guess that the principal’s value takes the form:
\[
P(\delta_t, W_t) = \underbrace{f(\delta_t)}_{\text{firm value}} - \underbrace{-\frac{1}{\gamma} \ln(-\gamma r W_t)}_{\text{agent’s certainty equivalent}} \tag{12}
\]

The Hamilton-Jacobi-Bellman (HJB) equation for the principal’s problem is
\[
rP(\delta, W) = \max_{(a, e)} \left\{ \delta \alpha(\delta) - c(a, e, \delta, W) + P_\delta \mu(a, \delta) + \frac{1}{2} P_{\delta\delta} \sigma(\delta)^2 + P_{WW} \beta_t^P (-\gamma r W) \sigma(\delta)^2 \right\}
\]
\[
+ \frac{1}{2} P_{WW} (-\gamma r W)^2 \left( (\beta_t^T)^2 \sigma_A^2 \delta_t^2 + (\beta_t^P)^2 \sigma(\delta_t)^2 + 2 \rho \beta_t^P \beta_t^T \delta \sigma_A \sigma(\delta) \right) \}
\]

where \(c(a, e, \delta, W)\) satisfies (9), \(\beta_t^P\) and \(\beta_t^T\) satisfy (10) and (11) respectively. From the conjectured value function (12), \(P_\delta = f'(\delta), P_{\delta\delta} = f''(\delta), P_{WW} = -\frac{1}{\gamma r W^2}\), and \(P_{\delta W} = 0\). Plugging these

\(^{14}\)In Appendix B, we provide a verification theorem regarding the global optimality of the obedient and no-saving policy. The result implies that the first-order conditions (10) and (11) together with the no-saving conditions are sufficient for contracts to be incentive-compatible and no-savings.
expressions into the HJB equation, the firm value \( f(\delta) \) must satisfy the ODE

\[
rf(\delta) = \max_{(a,e)} \left\{ \delta \alpha(e) - g(a, e; \delta) + \mu(a, \delta) f'(\delta) + \frac{1}{2} \sigma(\delta)^2 f''(\delta) \right\}
\]

(13)

The first line of (13) is the expected cash flows net of the direct monetary cost of effort plus the expected capital gain due to changes in the firm size. The second line of (13) represents the incentive cost for the principal to induce both the long-term and short-term effort. Incentive costs arise because the risk-averse agent is exposed to both the permanent and transitory shocks for her to work hard, and the risk compensation is required.

### 3.3 Deferred Compensation

Following the implementation of the optimal contract in He (2011), we interpret the agent’s certainty equivalent as a deferred compensation balance \( B_t = -\frac{1}{\gamma r} \ln(-\gamma r W_t) \). Under the optimal contract, the dynamics of the balance are given by

\[
dB_t = \frac{1}{2} \gamma r \left( (\beta T^t)^2 \delta^2 t^2 + (\beta P_t)^2 \delta (\delta t)^2 + 2 \rho \beta T^t \beta P_t \delta t \sigma_A(A \sigma(\delta)) \right) dt + \beta T^t \delta t \sigma_A dZ^T_t + \beta P_t \sigma(\delta t) dZ^P_t.
\]

(14)

In words, the agent’s stake inside the firm is given by \( B_t \) at any point in time; and the shareholders adjust the balance continuously according to (14) in order to provide the appropriate incentives to the agent. The adjustment in (14) include the drift term which reflects the incentive cost in (13); and the volatility terms which reflects the agent’s exposure to both the transitory and permanent shocks. In the optimal consumption (9), the agent’s wage compensates her for the effort cost and the interest \( rB_t \) earned from the deferred compensation balance.

### 4 Dynamic Agency and Capital Structure

This section embeds the dynamic agency problem into a Leland-style model. Following He (2011), the embedding allows us to endogenize the firm’s cash flows process. For tractability, we assume

\[
\begin{align*}
&dA_t = (\psi + e_t) dt + \sigma_A dZ^A_t \\
d\delta_t = (\phi + a_t) \delta_t dt + \sigma_\delta \delta_t dZ^P_t
\end{align*}
\]

where \( \psi, \phi, \) and \( \sigma_\delta \) are constant. In other words, we have \( \mu(a, \delta) = (\phi + a) \delta \) and \( \sigma(\delta) = \sigma_\delta \delta \) in equation (1), and \( \alpha(e) = \psi + e \) in equation (2), where \( \phi \) is the baseline growth rate and \( \psi \) is the baseline expected per-period profitability. Notice that \( \delta_t \) evolves as a geometric Brownian motion with the agent controlling the drift rate. Moreover, the agent’s effort cost takes the quadratic form
\( g(a, e, \delta) = \frac{1}{2} (\theta_a a^2 + \theta_e e^2 + 2\theta_{ae} ae) \delta \). The tasks are asymmetric in cost, the cost is proportional to firm size \( \delta \), and the two efforts can be either complements (\( \theta_{ae} \leq 0 \)) or substitutes (\( \theta_{ae} \geq 0 \)).

To derive the first-best contract, suppose there is no agency problem, for example, when actions are perfectly observable \( \sigma_A = \sigma_\delta = 0 \) or the agent is risk-neutral \( \gamma = 0 \), so the risk-compensation in HJB-equation (13) vanishes. Here, we assume \( \gamma = 0 \) and the first-best firm value \( f_{FB}(\delta) \) satisfies

\[
rf_{FB}(\delta) = \max_{(a, e)} \left\{ \delta(\psi + e) - \frac{1}{2} (\theta_a a^2 + \theta_e e^2 + 2\theta_{ae} ae) \delta + (\phi + a)\delta f'_{FB}(\delta) + \frac{1}{2} \sigma_\delta^2 \delta^2 f''_{FB}(\delta) \right\}
\]

(15)

Since the cash flows and expected capital gains are proportional to \( \delta \), homogeneity implies \( f_{FB}(\delta) = q\delta \), where \( q \) is a constant that reflects the marginal value per unit of firm size. Substituting the conjecture into (15), we have an equation that determines the unknown coefficient \( q \):

\[
rq = \max_{(a, e)} \left\{ (\psi + e) - \frac{1}{2} (\theta_a a^2 + \theta_e e^2 + 2\theta_{ae} ae) + (\phi + a)q \right\}
\]

The equation implies the first-order conditions \( q = (\theta_a a^{FB} + \theta_{ae} e^{FB}) \) and \( 1 = \theta_e e^{FB} + \theta_{ae} a^{FB} \) for long-term and short-term efforts. These conditions determine the coefficient \( q \) (valuation multiple).

There are a few implications. First, the marginal value per unit of firm size is constant. This implies, together with the scaling in the marginal effort costs, that first-best efforts are independent of firm size. Second, permanent shocks, transitory shocks, and their correlation do not affect \( q \), and hence efforts are independent of volatility. However, the firm value \( q\delta_t \) is still volatile, it evolves as a geometric Brownian motion. Third, a positive permanent shock \( dZ_t^{P} > 0 \) increases firm value through its effect on \( \delta_t \), but a positive transitory shocks \( dA_t > 0 \) has no effect on firm value.

4.1 Optimal Contract in an Unlevered Firm

Now we characterize the optimal contract in an unlevered firm. Let \( f_u(\delta) \) be the value of an unlevered firm. Then using the incentive compatibility conditions (10), \( \beta_t^P = \theta_a a_t + \theta_{ae} e_t \), and (11), \( \beta_t^T = \theta_e e_t + \theta_{ae} a_t \), the HJB-equation (13) for \( f_u(\delta) \) becomes

\[
r f_u(\delta) = \max_{\delta \in [0, \tilde{\delta}]} \left\{ \delta(\psi + e) - \frac{1}{2} (\theta_a a^2 + 2\theta_{ae} ae + \theta_e e^2) \delta + (\phi + a)\delta f'_u(\delta) + \frac{1}{2} \sigma_\delta^2 \delta^2 f''_u(\delta) \right\}
\]

(16)

For an interior solution, we use the first order conditions for the maximization of (16) to obtain

\[
a_u^*(\delta) = f'_u(\delta) \tilde{D} - \tilde{B} \quad \text{and} \quad e_u^*(\delta) = \frac{\hat{A} - f'_u(\delta) \hat{B}}{\hat{A} - \hat{B}^2},
\]

(17)
where $\hat{A}$, $\hat{B}$, and $\hat{D}$ are given Appendix B.\footnote{In Appendix B we show how to deal with the cases in which the constraints on the effort policies bind.} Using equation (17), one can show that the marginal unlevered firm value $f'_u(\delta) \to \frac{\psi}{r-\sigma}$ as $\delta \to \infty$. Together with the boundary condition $f_u(0) = 0$, the HJB-equation (16) can be solved numerically. For the special case where $\rho = 0$ and $\theta_{ae} = 0$, tasks are independent and the optimal effort policies are:

$$a^*_u(\delta) = \frac{f'_u(\delta)}{\theta_a(1 + \gamma r \theta_a^2 \sigma^2 \delta)} \quad \text{and} \quad e^*_u(\delta) = \frac{1}{\theta_e(1 + \gamma r \theta_e^2 \sigma^2 A \delta)}.$$  

(18)

We can make a few observations. First, since long-term effort $a^*$ affects the firm size, it depends on the marginal value of firm size $f'_u(\delta)$. However, $f'_u(\delta)$ is irrelevant to the short-term action $e^*$ because it only affects the current profitability. Second, the optimal effort depends on the volatility: $\sigma_\delta$ affects $a^*$ and $\sigma_A$ affects $e^*$. This is because the volatility of the processes $A_t$ and $\delta_t$ measure how informative these processes are about $e$ and $a$ respectively.

## 4.2 Optimal Contract in a Levered Firm

As in Leland (1994), the firm issues a perpetual debt with a constant coupon rate $C$. With a marginal corporate tax rate $\tau \in (0,1)$, the tax shield per unit of time is $\tau C$. And we interpret the quantity $dY_t = \delta_t dA_t$ as the after-tax cash flows. Once debt is in place shareholders implement the effort and default policies that are optimal for them. We refer to this specification as the base case model. We denote $f_E(\delta)$ as the value of equity and $D(\delta)$ as the value of corporate debt. From (12), $f_E(\delta) = P_E(\delta,W) + \frac{1}{\gamma r} \ln(-\gamma r W_t)$. That is, equity value consists of the shareholders value $P_E(\delta,W)$ (outside equity) and the agent’s value $\frac{1}{\gamma r} \ln(-\gamma r W_t)$ (inside equity).

The structural credit risk models have illustrated that endogenous default by the shareholders is an important mechanism in understanding credit risks. Let $\delta_B$ be the default threshold. We expect that when the firm’s fundamental (firm size) $\delta_t$ becomes sufficiently weak, especially after a sequence of negative permanent shocks, shareholders will default once $\delta_t < \delta_B$. Following He (2011), we assume that at default, the shareholders can fulfill the promise and pay the agent her remaining continuation value. In other words, the agent has higher seniority than bondholders. The latter receives the liquidation value of the firm and continues to run the firm as an unlevered firm. This implies the debt valued at bankruptcy is $D(\delta_B) = (1 - \alpha)f_u(\delta_B)$, where $\alpha \in (0,1)$ is a proportional bankruptcy cost parameter.
4.2.1 Shareholders value and endogenous default

The shareholders value function is given by $P_E(\delta, W) = f_E(\delta) - \frac{1}{\gamma r} \ln(-\gamma r W_t)$ under the optimal contract. With leverage, the equity value satisfies the following ODE:

$$rf_E(\delta) = \max_{e \in [0, \bar{e}], a \in [0, \bar{a}]} \left\{ \delta(\psi + e) - (1 - \tau)C - \frac{\theta_\alpha a^2 + 2\theta_\alpha a e + \theta_e e^2}{2}\delta + (\phi + a)\delta f'_E(\delta) + \frac{1}{2}\gamma r^2 \frac{\sigma^2}{\delta} f''_E(\delta) \right\}$$

subject to the value matching $f_E(\delta_B) = 0$, smooth-pasting $f'_E(\delta_B) = 0$, and transversality conditions $\lim_{\delta \to \infty} f''_E(\delta) \to -\frac{\psi}{r - \phi}$. The first two boundary conditions are the standard conditions in the case of endogenous (optimal) default. The transversality conditions states that as the firm grows arbitrarily large the growth rate is proportional to the cash flow per unit of capital $\psi$ capitalized by the baseline growth rate of the firm $\phi$. This is due to the fact that effort policies converge to zero as $\delta$ goes to infinity.

Compare (19) to the ODE (16) for the unlevered firm: equation (19) contains the promised coupon payment $C$ and the tax shield $\tau C$, as well as an additional control $\delta_B$ over the default threshold. Assuming an interior solution over the effort choices delivers solutions:

$$a^*_t(\delta) = \frac{f'_E(\delta)\bar{D} - \bar{B}}{\bar{A}D - \bar{B}^2} \quad \text{and} \quad e^*_t(\delta) = \frac{A - f'_E(\delta)\bar{B}}{A\bar{D} - \bar{B}^2}. \quad (20)$$

In our numerical simulations the constraints on $e$ never bind. Hence we focus on the cases when the constraints on $a$ bind. In particular in the appendix we show that when the constraint $a \geq 0$ binds the optimal policies are given by $e^*_t(\delta) = \frac{1}{\bar{D}}$ and $a^*_t(\delta) = 0$. If the constraint on long-term effort binds at $a \leq \bar{a}$ (upper bound), optimal policies are given by $e^*_t(\delta) = \frac{1 - \bar{a}B}{\bar{D}}$ and $a^*_t(\delta) = \bar{a}$. Because of the additional after-tax coupon payment $(1 - \tau)C$, shareholders absorb more losses when the firm’s fundamental (firm size) is weak: $\delta_t(\psi + e^*_t) < (1 - \tau)C$, net of the monetary compensation for the effort cost. Therefore, as $\delta_t$ falls to $\delta_B$, shareholders default optimally and refuse to fulfill the debt obligation. Standard conditions – value-matching $f_E(\delta_B) = 0$ and smooth-pasting $f'_E(\delta_B) = 0$ – characterize the endogenous default threshold and its optimality.

4.2.2 Debt valuation and optimal leverage

Once debt is in place, shareholders select the optimal long-term contract and default policy. Bondholders anticipate the effect of debt on shareholders’ future behavior. Hence, in pricing the per-
petual debt contract, creditors take the optimal effort policy \((a^*(\delta), e^*(\delta))\) in (20), and the default threshold \(\delta_B\) as given. For any coupon \(C\), the debt value \(D(\delta)\) satisfies the following ODE:

\[
r D(\delta) = C + (\phi + a^*(\delta)) \delta D'(\delta) + \frac{1}{2} \sigma^2 \delta^2 D''(\delta)
\]

(21)

with boundary conditions \(D(\delta) \to \frac{C}{r}\) as \(\delta \to \infty\), and \(D(\delta_B) = (1 - \alpha) f_u(\delta_B)\). Observe that in equation (21) long-term effort directly affects the expected capital gains of the debt contract. In contrast, the short-term action \(e^*(\delta)\) affects cash flows and so the ability of the shareholders to absorb losses. Thus, it implicitly affects the endogenous default time and the debt value through the boundary condition.

Given an initial firm size \(\delta_0\), initial shareholders choose coupon \(C\) to maximize levered firm value (ex-ante equity value) \(TV(\delta_0; C) = f_E(\delta_0; C) + D(\delta_0; C)\) at time 0. Then they design the optimal long-term contract with the agent that implements the effort policy \((a^*(\delta), e^*(\delta))\), and run the firm until they declare bankruptcy. We define the firm’s optimal initial market leverage ratio as

\[
ML(\delta_0) \equiv \frac{D(\delta_0; C^*(\delta_0))}{f_E(\delta_0; C^*(\delta_0)) + D(\delta_0; C^*(\delta_0))}.
\]

### 4.3 Short-termism as an Indirect Cost of Debt-overhang

In this section we highlight the asymmetry between underinvestment and short-termism with respect to the debt overhang problem. We argue that underinvestment is a direct consequence of debt, as it is well known in the literature. In contrast, short-termism is only affected by the presence of debt indirectly through the underinvestment problem. Moreover, the effect of debt on short-termism disappears when the costs of implementing a particular effort policy pair \((a, e)\) for the shareholders are independent. Independence occurs when the cross term in the cost function is zero \((\theta_{ae} = 0)\) and the shocks are uncorrelated \((\rho = 0)\). Proposition 1 formalizes this result.

Furthermore, we show that when shareholders can commit to the unlevered long-term effort policy (i.e. when they can commit to avoid underinvestment), this will suffice as a commitment device to avoid short-termism (i.e. they will automatically be committing to the unlevered short-term effort policy). Proposition 2 formalizes this result. Figure 1 summarizes the findings of this section.

**Proposition 1.** Suppose that \(\rho = \theta_{ae} = 0\) then the optimal short-term effort policy \(e(\delta)\) is independent of the coupon payment \(C\).
The intuition for this proposition is straightforward: debt generates underinvestment in long-term effort because shareholders pay up front for the cost of long-term effort, but do not fully internalize the benefit of this investment. The reason is because some of the cash flows generated by investment take place after default, thereby accruing to bondholders. In contrast, the benefit of short-term effort is immediately realized by shareholders, therefore they fully internalize the benefits of short-term effort. Hence, the only mechanism by which debt can distort the short-term effort policy is when the cost of short-term effort depends on the implemented level of long-term effort. When these costs are independent from each other (\( \theta_{ae} = 0 \) and \( \rho = 0 \)) there is no distortion, and the optimal \( e(\delta) \) is not affected by the presence of debt.

In Figure 2 we illustrate the optimal amount of long-term effort \( a(\delta_t) \) and short-term effort \( e(\delta_t) \) for the optimally levered firm and for an all equity financed firm (unlevered firm). First, consider the case in which the actions are independent (Panels A and B). Panel A corroborates the reduction in long-term effort due to debt. Panel B is an empirical illustration of Proposition 1 showing that short-term effort is identical in the levered and unlevered cases when the the two actions are independent. Second, consider the case in which the two actions are substitutes (Panels C and D). In this case optimal short-term effort is larger in the presence of debt. Short-term effort is not affected by debt overhang since it increases contemporaneous cash flows but has no permanent effect. Hence, shareholders find it optimal to incentivize the manager to focus in boosting short-term profitability as opposed to improving the long-term prospects of the firm. Our model reverses the common intuition that CEO short-termism is detrimental to shareholders. To the contrary we show that shareholders find it optimal to encourage managers to focus myopically in the short-term when there is a large probability the company will not survive in the long-term.

---

17 A similar mechanism is present in Manso (2008), who shows the agency cost of debt is proportional to the degree of irreversibility of the investment project. In our case the inefficiency depends on whether the effort policy will have a permanent effect (long-term effort) or a transitory effect (short-term effort) on the firm’s cash flows.

18 The previous intuition is straightforward in case of actions being substitutes, \( \theta_{ae} > 0 \). However, even if actions are independent, \( \theta_{ae} = 0 \), and the correlation between the transitory and the permanent shock \( \rho \) is positive, a similar result takes place. Intuitively, when \( \rho > 0 \) implementing both actions is very costly for shareholders because the manager will be exposed to two positively correlated shocks. Hence, the manager will need to be compensated for bearing this risk. Therefore, having correlated shocks (\( \rho > 0 \)) is “as if” the two actions were substitutes (i.e. \( \theta_{ae} > 0 \)).
we refer to the underinvestment problem as the fact that debt induces shareholders to implement
a lower long-term effort (i.e. $a(\delta) \geq a_u(\delta)$), and to the short-termism problem to the fact that
shareholders implement more short-term effort (i.e. $e(\delta) \geq e_u(\delta)$).

Figure 2. Long-term and short-term effort as a function of firm size.
This picture illustrates the effect of debt over-hang on the optimal amount of long-term effort
(Panel A) and short-term effort (Panel B) when the actions are independent $\theta_{ae} = 0$. Panel’s C
and D illustrate the effect when the actions are substitutes $\theta_{ae} = 1.5$. Other parameter values are
$\delta_0 = 100$, $r = 0.05$, $\theta_a = 30$, $\theta_e = 1$, $\gamma = 5$, $\sigma_\delta = 0.25$, $\sigma_A = 0.12$, $\rho = 0$, $\tau = 0.15$, $\alpha = 0.30$, $\phi = -0.005$, and $\psi = 1$. The dashed (solid) line represents for the unlevered (levered) firm.

Next, we show that commitment to the long-term effort policy $a_u(\delta)$ suffices as a commitment
device when shareholders would like to commit to the unlevered effort policies ($a_u, e_u$). To compare
and contrast the base case without commitment, we consider three important benchmarks:

1. The no debt-overhang case (denoted $NO$): In this case shareholders commit to implementing
the unlevered effort policies ($a_u, e_u$) for both short-term and long-term effort. We denote the
value function in this case by $f_{NO}(\delta)$.

2. The no short-termism case (denoted $NS$): In this case shareholders commit to implementing
the unlevered short-term effort policy $e_u$, but they can freely choose the long-term effort
policy. We denote the value function in this case by $f_{NS}(\delta)$.$^{19}$

$^{19}$Note that, in the $NS$ case, shareholders request some amount of short-term effort. This is different from assuming
away short-term effort by setting $\theta_e = \infty$, which, as indicated by Footnote 10, delivers He (2011) given that $\alpha(0) = 1$. 

18
3. The no underinvestment case (denoted $NU$): In this case shareholders commit to implementing the unlevered long-term effort policy $a_u$, but they can freely choose the short-term effort policy. We denote the value function in this case by $f_{NU}(\delta)$.

The rigorous statement of the problems that are solved by each of the cases above can be found in Appendix A. We can now state the main result of this section.

**Proposition 2.** $f(\delta)$ solves the no underinvestment case ($NU$) if and only if it solves the no overhang case ($NO$).

In other words, debt has an asymmetric effect on long-term and short-term effort policies. If shareholders commit to the unlevered long-term effort policy $a_u$, they find it optimal to (ex-post) implement the unlevered short-term effort policy $e_u$. Therefore, if they wanted to commit to the long-term effort pair $(a_u, e_u)$, committing to $a_u$ would suffice to achieve this goal.\footnote{This is useful, because committing to the unlevered firm policies is value enhancing (see Section 5.1).} However, the converse is not true. If shareholders can commit to the unlevered short-term effort policy $e_u$, they would ex-post choose a different long-term effort policy from $a_u$.

In summary, debt distorts the choice of long-term effort through the well known logic of the under-investment problem. When the cost of the actions are independent, the choice of short-term effort is unaffected by the presence of debt (Proposition 1). In the general case in which the actions are not independent, the choice of short-term effort is also distorted. However, this distortion is only indirect: debt distorts the choice of long-term effort, and then the distortion in long-term effort induces a distortion in short-term effort. Therefore, if one could prevent the distortion of debt on long-term effort there would be no distortion in short-term effort either (Proposition 2).

This result has two implications for debt covenant design. First, debt covenants that restrict shareholder payout can be interpreted as covenant that restrict the short-term effort policy. As shown in Section 5.1 such covenants would mitigate some of the cost of debt and increase total firm value. However, debt covenants that restrict the long-term effort policy of the firms are much more effective at maximizing firm value. Such covenants could potentially minimize the underinvestment problem, while at the same time discouraging shareholder to engage in excessive short-termism.
4.4 Comparative Statics

Now we provide the effects of the parameters on the equity value $f_E(\delta)$ and the endogenous default threshold $\delta_B$. Here, we focus on the ex-post value and default decision. That is, we treat the coupon as an exogenous parameter rather than taking into account the ex-ante capital structure choice.

$$\frac{\partial f_E(\delta)}{\partial \delta_B} = \frac{\partial \psi}{\partial \phi} \frac{\partial \psi}{\partial \gamma} \frac{\partial \psi}{\partial \rho} \frac{\partial \psi}{\partial C} \frac{\partial \psi}{\partial \tau} \frac{\partial \psi}{\partial \theta}$$

Table 1. Analytical comparative statics.

Several results can be observed from Table 1. First, the increase in the drifts $\psi$ and $\phi$, a higher tax saving from the increase in the corporate tax rate $\tau$, the decrease in coupon $C$ and the agency parameter $\gamma$ will increase the equity value and delay default. The results are intuitive because these variations generate more cash flows to the shareholders.

Second, the increase in the effort parameters, $\theta_a$, $\theta_e$, and $\theta_{ae}$, implies a higher compensation the equity needs to make to the agent and hence the equity value reduces. However, this set of result only holds true when $\rho \geq 0$ because the tasks tend to be substitute in this case. Consider an increase in $\theta_a$. The direct effect implies a lower equity value because of the increased effort and incentive cost. In respond to the increase in $\theta_a$, the equity would like to implement a lower $a_t$. And a $\rho \geq 0$ implies $e_t$ will increase at the same time. This indirect effect from the information filtering channel implies a further increase in the effort and incentive compensation. As a result, equity value decreases and the shareholders default earlier. Panels C and D of Figure 3 below illustrate this comparative statics with $\rho = 0$: The default threshold is increasing in the cost of short-term effort and long-term effort respectively.

In a similar vein, the effect of the variation of the volatilities parameters, $\sigma_\delta$ and $\sigma_A$, on the default threshold can be understood from the information filtering channel. Consider an increase in the volatility $\sigma_\delta$, this effect makes the firm size $\delta_t$ a more noisy signal about the long-term effort $a_t$ and thus increases the incentive cost and accelerates default. However, the comparative statics of $\sigma_\delta$ is ambiguous. This is because an increase in the volatility $\sigma_\delta$ also generates a real option effect that delays the equity’s default decision. This can be observed from the terms involving $\sigma_\delta$.
on the HJB-equation \[ \text{(19)} \] (with $\rho = 0$):

$$
\frac{1}{2} \sigma^2 \delta^2 f''_E(\delta) - \frac{1}{2} \gamma r \delta^2 \left( \left( \theta_a + \theta_{ae} \right) \sigma^2 \right)
$$

While the second term captures the incentive cost effect, the first term captures the real option effect and it is increasing in $\sigma_\delta$ under the convexity of the shareholders value\textsuperscript{21}. The two countervailing forces generates an inverted U-shaped for the default threshold as a function of $\sigma_\delta$. Numerically, this result is illustrated on Panel A of Figure 3.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Comparative statics for default boundary.}
\end{figure}

The parameter values are $C = 47.83$, $\delta_0 = 100$, $r = 0.05$, $\theta_a = 30$, $\theta_e = 1$, $\theta_{ae} = 1.5$, $\gamma = 5$, $\sigma_\delta = 0.25$, $\sigma_A = 0.12$, $\rho = 0$, $\tau = 0.15$, $\alpha = 0.30$, $\phi = -0.005$, and $\psi = 1$.

In contrast, the increase in the volatility of transitory shocks $\sigma_A$ has no real option effect. While the variation makes the incentive cost of effort more expensive, the comparative statics is ambiguous. It is only definite when $\rho \geq 0$. The idea is similar to the comparative statics of the effort cost parameters: An increase in $\sigma_A$ will lower the short-term effort $e_t$, if the correlation was negative, then tasks tend to be complement and the shareholders would like to implement a lower long-term effort $a_t$ as well. The latter effect reduces the effort cost and counters the increase in incentive cost. Panel B of Figure 3 illustrates this comparative statics with $\rho = 0$.

Finally, we explore the effect of the correlation between the transitory and the permanent shock on the default decision of the firm. Panel A of Figure 3 shows the default boundary is increasing in

\textsuperscript{21} $f''_E(\delta) > 0$ in our numerical simulations.
the correlation between the transitory and the permanent shock. Higher correlation makes it more costly for shareholders to compensate a given effort policy \((a, e)\) since the (risk-averse) manager would necessarily bear more risk and thus require compensation for bearing such risk. This higher cost would induce the shareholders to default earlier. We can check this intuition by inspecting the HJB and collecting terms involving \(\rho\)

\[
\frac{-\gamma r \delta^2}{2} (2\rho(\theta_a a + \theta_{ae} e)(\theta_e e + \theta_{ae} a)\sigma \delta \sigma_A)
\]

and noticing that the incentive cost are increasing in \(\rho\). Panels B and C of Figure 4 depict effort levels for three different values of \(\rho\). As the correlation between shocks increases shareholders find it more costly to incentivize the manager to exert effort. Hence, they find it optimal to implement a lower level of short and long-term effort.

![Figure 4. Effects of correlation and substitutability.](image)

The parameter values are \(C = 47.83, \delta_0 = 100, \ r = 0.05, \ \theta_a = 30, \ \theta_e = 1, \ \theta_{ae} = 1.5, \ \gamma = 5, \ \sigma_\delta = 0.25, \ \sigma_A = 0.12, \ \tau = 0.15, \ \alpha = 0.30, \ \phi = -0.005, \ \psi = 1, \ \rho \text{ is either } -0.3, 0, \text{ or } 0.3, \text{ and } \theta_{ae} \text{ is either } 1.4, 1.5, \text{ or } 1.6.

Interestingly, the comparative statics of \(\theta_{ae}\) are similar to the ones of \(\rho\). Panels D, E, and F of Figure 4 compute comparatives statics for the default boundary and the effort policies with respect to the term governing the sustitability/complementarity of cost function of the manager \(\theta_{ae}\). As \(\theta_{ae}\) increase the two tasks become substitutes and it becomes more costly to incentivize a given effort policy \((a, e)\). Moreover, the marginal cost of increasing \(e\) for a given value of \(a\) is increasing.
in $\theta_{ae}$. The same is true for $\rho$, fix $a$ and suppose shareholders wants to increase $e$; the marginal cost of increasing $e$ is larger for a given $\rho$ since the agent will be exposed to more risk (due to the correlation between shocks). Thus, the equilibrium effects of $\rho$ and $\theta_{ae}$ are quite similar.

5 Quantitative Analysis

In this section we numerically explore the implications of our model. First, we compute in our calibrated model the various sources of firm value, and show that the total agency cost of debt is of the order of 1%. Of this 1% about half of the cost of debt comes from excessive short-termism. Second, we highlight potential endogeneity concerns when studying the relationship between excessive short-termism and firm’s growth rate. In particular, we show that firms with high (low) growth rates optimally choose to implement lower (higher) levels of short-term effort and higher (lower) levels of long-term effort. Third, we extend the model to the case in which a subset of investors with higher discount rates takes control of the firm. We find that impatient shareholders implement higher levels of short-term effort and lower levels of long-term effort, which in turn reduce equity value for the regular (patient) shareholders and bondholders. Finally, we show that an increase in the volatility of permanent shocks can be desirable for shareholders (risk-shifting type of intuition), but not an increase in volatility of transitory shocks.

For numerical solutions, we adopt the following baseline parameter values: interest rate $r = 5\%$, baseline growth rate of firm size $\phi = -0.5\%$, correlation $\rho = 0$, volatility of firm size $\sigma_\delta = 25\%$, profitability $\sigma_A = 12\%$, corporate tax rate $\tau = 15\%$, and bankruptcy cost $\alpha = 30\%$. Following He (2009, 2011), we set the baseline profitability $\psi = 1$, risk aversion $\gamma = 5$, and long-term effort costs to $\theta_a = 30$. In Appendix B we calibrate the cost of short-term effort $\theta_e = 1$ and the substitutability parameter $\theta_{ae} = 1.5$. For an initial firm size of, e.g., $\delta_0 = 100$ the firm’s optimal coupon is $C^* = 47.83$ and market leverage is $ML(\delta_0) = 39.42\%$.

5.1 Analysis of Sources of Firm Value

In this section we quantify the various sources of firm value. First, we define the tax advantage of debt $TB(\delta)$ and the cost of bankruptcy $BC(\delta)$ in the usual way

\[ TB(\delta) = E_t \left[ \int_t^\chi e^{-r(s-t)}Cds \right] \quad \text{and} \quad BC(\delta) = E_t \left[ e^{-r(\chi-t)}\alpha f_u(\delta_B) \right], \tag{22} \]

These baseline parameter values are consistent with, e.g., Bolton, Chen, and Wang (2011), He (2011), and Morellec, Nikolov, and Schürhoff (2012).
where $\chi = \inf\{t : \delta_t \leq \delta_B\}$ corresponds to the endogenously chosen default time.

Second, we define the total cost of managerial compensation $CC(\delta)$. The total expected cost of managerial compensation is the sum of the direct cost of compensating the manager for his effort $DC(\delta)$ plus the indirect cost of compensating the manager $IC(\delta)$ for his risk exposure. Formally,

$$DC(\delta) = Et \left[ \int_t^\chi e^{-r(s-t)}g(a_s, e_s; \delta_s)ds \right],$$

and

$$IC(\delta) = Et \left[ \int_t^\chi e^{-r(s-t)} \frac{1}{2} \gamma r \left( (\beta_s^T)^2 \sigma_A^2 \delta_s^2 + (\beta_s^e)^2 \sigma(\delta_s)^2 + 2 \rho \beta_s^P \beta_s^T \delta_s \sigma_A \sigma(\delta_s) \right) ds \right],$$

and

$$CC(\delta) = DC(\delta) + IC(\delta).$$

Table 2 shows the different components of firm value for our baseline calibration.

<table>
<thead>
<tr>
<th>Panel A: Optimal Leverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$</td>
</tr>
<tr>
<td>75</td>
</tr>
<tr>
<td>100</td>
</tr>
<tr>
<td>150</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Unlevered Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$</td>
</tr>
<tr>
<td>75</td>
</tr>
<tr>
<td>100</td>
</tr>
<tr>
<td>150</td>
</tr>
</tbody>
</table>

Table 2. Sources of firm value in base case.

This table shows the various sources of firm value for the base case (i.e. the case without commitment where effort policies are optimally chosen by shareholders after debt has been issued) for three different firm sizes $\delta$. For each firm size we calculate quantities at the optimal leverage $C^*$ and for the unlevered case $C = 0$. The parameter values are $r = 0.05$, $\theta_a = 30$, $\theta_e = 1$, $\theta_{ae} = 1.5$, $\gamma = 5$, $\sigma_\delta = 0.25$, $\sigma_A = 0.12$, $\rho = 0$, $\tau = 0.15$, $\alpha = 0.30$, $\phi = -0.005$, and $\psi = 1$.

5.1.1 Quantifying the cost of debt-overhang

In this section, we tease-out the total cost of debt-overhang (i.e. the cost of under-investment plus the cost of short-termism for firm value). In particular, we consider the case in which shareholders can commit ex-ante to implementing the optimal unlevered long-term and short-term effort policies (which are ex-post suboptimal for shareholders). This case corresponds to the no overhang case.
(NO) described in section 4.3. We denote the resulting equity, debt, total firm, and leverage values, respectively, by \( f_{NO}(\delta) \), \( D_{NO}(\delta) \), \( TV_{NO}(\delta; C) \), and \( ML_{NO}(\delta_0) \) (see Appendix A for details).

Table 3 computes firm values for three different initial firm sizes \( \delta_0 = 75, 100, 150 \). Column 1 computes quantities for the base case, i.e. the case when effort policies are optimally chosen by shareholders after debt has been issued (normalized relative to the unlevered unmanaged total firm value \( \delta_0/(\phi - r) \equiv 100 \) in parentheses). Column 2 computes quantities for the case in which there is commitment to the no-overhang policies as described at the beginning of this section. They correspond to the cases with the subscript \( NO \). Column 3 computes quantities for the base case but using the optimal coupon \( C^*_NO \) calculated under the no-overhang policies. Finally, column 4 computes the percentage change between columns 3 and 2.

A number of interesting observations can be made about Table 3. First, committing to the no-overhang effort policies increases normalized total firm value between 0.50 (large firms) to 0.92 (small firms). By alleviating short-termism and underinvestment (due to debt overhang) total firm value can be enhanced. Since small firms are more prone to the distortions of debt overhang it is expected that commitment to the no-overhang policies will induce a larger increment for small firms.

Second, committing to the no-overhang effort policies \( a_u \) and \( e_u \) reduces shareholder value. The reduction in shareholder value ranges from 0.53% (small firms) to 0.80% (large firms). Because shareholders cannot implement their optimal policies, for a given level of \( \delta \) their expected value is reduced (i.e. \( f_E(\delta) > f_{NO}(\delta) \)) and they choose to default earlier (i.e. \( \delta_B < \delta^*_B^{NO} \)). Since shareholders cannot reduce long-term investment when the firm is in financial distress and focus on implementing higher short-term profitability, it is more costly for them to keep on running their firm during financial distress. Thus, triggering earlier default.

Third, the increment in debt value as a result of committing to the no-overhang policies ranges from 1.99% (large firms) to 2.49% (small firms). On the one hand shareholders default earlier, but on the other hand shareholder are committed to a higher level of long-term effort which on expectation keeps \( \delta \) away from the default boundary. Our results indicate that in general the latter effect dominates (i.e. \( D_{NO}(\delta) > D(\delta) \)). Furthermore, the increment in debt value dominates the reduction in shareholder value and as discussed above total firm value goes up.

Finally, the effect of debt-overhang in terms of short-termism and under-investment is stronger for small firms. To mitigate the effect of debt overhang, optimal leverage is lower for small firms...
Table 3. Firm value without debt-overhang in effort policies.

This table calculates the changes in equity, debt, and total firm value when there is no debt-overhang over the effort policies. Column 1 corresponds to the base case without commitment where effort policies are optimally chosen by shareholders (after debt has been issued). Column 2 corresponds to the no overhang case \((NO)\). Column 3 recomputes the base case when the coupon is given by the no overhang case \(C_{NO}^*\). Column 4 computes percentage changes between Columns 2 and 3.

Normalized total firm values are in parentheses as the percentage value relative to the unlevered unmanaged total firm value \(\delta_0/(\phi - r) \equiv 100\). The parameter values are \(r = 0.05, \theta_a = 30, \theta_e = 1, \theta_{ae} = 1.5, \gamma = 5, \sigma_\delta = 0.25, \sigma_A = 0.12, \rho = 0, \tau = 0.15, \alpha = 0.30, \phi = -0.005, \) and \(\psi = 1\).
Optimal leverage in the base case ranges from 36.97% (small firms) to 43.57% (large firms). On the other hand, when firms can commit to the no overhang policies (Column 2) there is no need to adjust leverage to mitigate the effects of short-termism and underinvestment, thus optimal leverage is much less sensitive to initial firm size.

5.1.2 Quantifying the cost of short-termism

In this section we quantify the effect of short-termism on firm value. In particular, we consider first the case in which shareholders can commit ex-ante to implementing the unlevered short-term effort policy (but which does not maximize ex-post shareholder value). However, shareholders are free to choose the long-term effort policy that is optimal for them. This case corresponds to the no short-termism case ($NS$) described in section 4.3. We denote the resulting equity, debt, total firm, and leverage values by $f_{NS}(\delta)$, $D_{NS}(\delta)$, $TV_{NS}(\delta_0;C)$, and $ML_{NS}(\delta_0)$ respectively.

Table 4 performs a similar exercise as Table 3 but computes the changes in firm value resulting from committing to the unlevered short-term effort policy. As expected the ability to commit to $e_u$ increases total firm value, but the effect is more modest than when the firm can commit to both $e_u$ and $a_u$. In particular, normalized total firm value increases by 0.31 (large firms) or 0.44 (small firms), which is approximately half of the increment observed in the previous case. Similarly, the reduction in shareholder value ranges from 0.23% (small firms) to 0.49% (large firms), while the increment in debt value ranges from 1.21% (large firms) to 1.38% (small firms).

In sum, this section shows that commitment to the unlevered short-term policy has the expected qualitative effects on firm values: total firm value increases, debt values increases, and shareholder value decreases. The quantitative effects are about half of those observed in the no-overhang case. Economically, the experiments imply that outlawing short-termism destroys shareholder value once debt is in place, but increases debt value. Hence policy proposals should recognize ex-ante beneficial and the ex-post harmful implications and, in particular, the expected net result for firm value.

5.2 Endogeneity of Growth Rate

In this section we compute optimal policies for firms that have bright prospects and firms that have grim prospects. We do so by computing comparative statics with respect to the baseline growth rate of firm size $\phi$ and interpret a large (low) value of $\phi$ as having bright (grim) prospects, because this firm is expected to experience fast (slow) growth in the future.
<table>
<thead>
<tr>
<th></th>
<th>Base Case</th>
<th>No short-termism (NS)</th>
<th>Base Case with $C_{NS}^*$</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_0$</td>
<td>150</td>
<td>150</td>
<td>150</td>
<td></td>
</tr>
<tr>
<td>$\delta_0/(\phi - r)$</td>
<td>2727.27 (100)</td>
<td>3070.32 (112.57)</td>
<td>3070.31 (112.57)</td>
<td>0.27% (0.31)</td>
</tr>
<tr>
<td>$f_u(\delta_0)$</td>
<td>2959.19 (108.50)</td>
<td>3078.63 (112.88)</td>
<td>3078.63 (112.88)</td>
<td></td>
</tr>
<tr>
<td>$TV_{NS}(\delta_0)$</td>
<td>1732.44</td>
<td>1716.89</td>
<td>1725.51</td>
<td>-0.49%</td>
</tr>
<tr>
<td>$D_{NS}(\delta_0)$</td>
<td>1337.88</td>
<td>1361.74</td>
<td>1344.80</td>
<td>1.26%</td>
</tr>
<tr>
<td>$C_{NS}^*$</td>
<td>81.07</td>
<td>81.62</td>
<td>81.62</td>
<td></td>
</tr>
<tr>
<td>$\delta_B$</td>
<td>29.81</td>
<td>31.60</td>
<td>30.03</td>
<td></td>
</tr>
<tr>
<td>$ML_{NS}(\delta_0)$</td>
<td>43.57</td>
<td>44.23</td>
<td>43.80</td>
<td></td>
</tr>
<tr>
<td>$CS_{NS}(\delta_0)$</td>
<td>105.95</td>
<td>99.41</td>
<td>106.92</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Base Case</th>
<th>No short-termism (NS)</th>
<th>Base Case with $C_{NS}^*$</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_0$</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>$\delta_0/(\phi - r)$</td>
<td>1818.18 (100)</td>
<td>2124.98 (116.87)</td>
<td>2124.64 (116.85)</td>
<td>0.35% (0.42)</td>
</tr>
<tr>
<td>$f_u(\delta_0)$</td>
<td>2045.68 (112.51)</td>
<td>2132.24 (117.27)</td>
<td>2132.24 (117.27)</td>
<td></td>
</tr>
<tr>
<td>$TV_{NS}(\delta_0)$</td>
<td>1287.24</td>
<td>1235.31</td>
<td>1240.01</td>
<td>-0.37%</td>
</tr>
<tr>
<td>$f(\delta_0)$</td>
<td>837.74</td>
<td>896.92</td>
<td>884.63</td>
<td>1.38%</td>
</tr>
<tr>
<td>$C_{NS}^*$</td>
<td>47.83</td>
<td>51.32</td>
<td>51.32</td>
<td></td>
</tr>
<tr>
<td>$\delta_B$</td>
<td>16.59</td>
<td>19.00</td>
<td>17.97</td>
<td></td>
</tr>
<tr>
<td>$ML_{NS}(\delta_0)$</td>
<td>39.42</td>
<td>42.06</td>
<td>41.63</td>
<td></td>
</tr>
<tr>
<td>$CS_{NS}(\delta_0)$</td>
<td>70.93</td>
<td>72.24</td>
<td>80.12</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Base Case</th>
<th>No short-termism (NS)</th>
<th>Base Case with $C_{NS}^*$</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_0$</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>$\delta_0/(\phi - r)$</td>
<td>1363.63 (100)</td>
<td>1654.89 (121.35)</td>
<td>1654.12 (121.30)</td>
<td>0.36% (0.44)</td>
</tr>
<tr>
<td>$f_u(\delta_0)$</td>
<td>1587.70 (116.43)</td>
<td>1660.15 (121.74)</td>
<td>1660.15 (121.74)</td>
<td></td>
</tr>
<tr>
<td>$TV_{NS}(\delta_0)$</td>
<td>1043.06</td>
<td>969.44</td>
<td>971.67</td>
<td>-0.23%</td>
</tr>
<tr>
<td>$f(\delta_0)$</td>
<td>611.82</td>
<td>690.71</td>
<td>682.45</td>
<td>1.21%</td>
</tr>
<tr>
<td>$C_{NS}^*$</td>
<td>33.38</td>
<td>38.36</td>
<td>38.36</td>
<td></td>
</tr>
<tr>
<td>$\delta_B$</td>
<td>10.89</td>
<td>13.30</td>
<td>12.86</td>
<td></td>
</tr>
<tr>
<td>$ML_{NS}(\delta_0)$</td>
<td>36.97</td>
<td>41.60</td>
<td>41.25</td>
<td></td>
</tr>
<tr>
<td>$CS_{NS}(\delta_0)$</td>
<td>45.57</td>
<td>55.45</td>
<td>62.08</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Firm value without debt-overhang in short-term effort.
This table calculates the changes in equity, debt, and total firm value when there is no debt overhang over the short-term effort policies. Column 1 corresponds to the base case without commitment where effort policies are optimally chosen by shareholders (after debt has been issued). Column 2 corresponds to the no short-termism (NS) case. Column 3 recomputes the base case when the optimal coupon is given by the no short-termism case $C_{NS}^*$. Column 4 computes percentage changes between Columns 2 and 3. Normalized total firm values are in parentheses as the percentage value relative to the unlevered unmanaged total firm value $\delta_0/(\phi - r) \equiv 100$. The parameter values are $r = 0.05, \theta_d = 30, \theta_e = 1, \theta_{ae} = 1.5, \gamma = 5, \sigma_\delta = 0.25, \sigma_A = 0.12, \rho = 0, \tau = 0.15, \alpha = 0.30, \phi = -0.005$, and $\psi = 1$. 

28
Figure 5. Effect of growth rate.

The parameter values are $C^* = 47.83$, $\delta_0 = 100$, $r = 0.05$, $\theta_a = 30$, $\theta_e = 1$, $\theta_{ae} = 1.5$, $\gamma = 5$, $\sigma_\delta = 0.25$, $\sigma_A = 0.12$, $\rho = 0$, $\tau = 0.15$, $\alpha = 0.30$, and $\psi = 1$.

Figure 5 shows that firms with bright prospects focus a lot more in the long-term and put little effort in the short-term. Conversely, firms with grim prospects optimally focus on the short-term and put little effort in the long-term. Our results show that studies suggesting excessive short-termism need to be aware of simultaneous causality. One may be tempted to infer that success is a consequence of long-term investment. However, our model shows that firms that are more likely to succeed (high $\phi$) optimally focus on the long-term. Consistent with Summers (2017), success is thus not the result but also the cause of long-term investment (less short-termism).

5.3 Impact of Investor Horizon in Optimal Policies

In this section we study the impact of investor’s horizon in the optimal policies of the firm. In particular, we relax the assumption that shareholders, bondholders, and the manager have the same discount rate $r$. Instead, we will maintain the assumption that the manager’s discount rate is $r$, but we will allow the shareholders to have discount rate $r^S = r + \lambda$. As $\lambda$ increases shareholders become more impatient and we interpret this as having a shorter investment horizon.

\footnote{Consider the example of IBM. IBM's sales in 2017 are at the same level as they were in 1997. Over this period it has reduced costs and cut investment by half. Our model suggests that such decisions are not suboptimal and myopic. Instead, they are optimal responses to the new generation of technology firms, which took over the industry.}
Figure 6. Effect of investor horizon.

The parameter values are $C^* = 47.83$, $\delta_0 = 100$, $r = 0.05$, $\theta_a = 30$, $\theta_e = 1$, $\theta_{ae} = 1.5$, $\gamma = 5$, $\sigma_\delta = 0.25$, $\sigma_A = 0.12$, $\tau = 0.15$, $\alpha = 0.30$, $\phi = -0.005$, $\psi = 1$, and $r^S$ is either 0.05, 0.055, or 0.06.

Figure 6 depicts the optimal investment policies for three different values of $r^S$. Panel A shows that as shareholders investment horizon decreases, the optimal amount of long-term effort implemented goes down. This is intuitive because of the term-structure of the payoffs associated with long-term effort $a(\delta)$. Shareholders pay upfront for the cost of effort (both direct and indirect), but the payoff is realized over time via the increment in firm size $\delta$. A higher discount rate will reduce the value of future cash flows and thus reduce optimal long-term effort. Panel B shows that a shorter investment horizon increases short-term effort. The payoff from short-term effort is realized immediately for shareholders via higher contemporaneous cash flows. Under the assumption that the two tasks are substitutes (or that transitory and permanent shocks are positively correlated), the reduction in investment horizon will encourage shareholders to increase short-term effort at the expense of long-term effort. Furthermore, a shorter investment horizon reduces shareholder value for a given firm size, rendering it optimal to default earlier.

Finally, we quantify the changes in firm value associated with implementing policies dictated by shareholders with shorter investment horizons. In Table 5 we compute the net present value of cash flows that would accrue to shareholder and bondholders for the three different sets of effort policies depicted in Figure 6. That is, we compute the effort policies that shareholders with discount rates $r^S = 5\%, 5.5\%$, and $6\%$ would deem optimal, and then calculate bond and equity values under the
discount rate $r = 5\%$. The point of this exercise is to isolate the distortion in firm value associated with implementing the policies desired by shareholders with shorter investment horizons from the mechanical reduction in value resulting from a higher discount rate.\footnote{One can think of this exercise as the following thought experiment: a majority shareholder with discount rate $r^S$ takes control of the firm, and implements the effort policies that she finds optimal. We then compute the value of equity from the perspective of the minority shareholder with discount rate $r$. Similarly, for bond prices.}

| Panel A: $\delta_0 = 75, C^* = 33.38$ |
|-----------------|-----------------|-----------------|-----------------|
| $r^S$           | $TV(\delta)$   | $% \text{ change } TV(\delta)$ | $f_E(\delta)$  | $% \text{ change } f(\delta)$ | $D(\delta)$ | $% \text{ change } D(\delta)$ |
| 5%              | 1,654.88        | -                | 1,043.06        | -                | 611.82      | -                |
| 5.5%            | 1,633.90        | -0.73            | 1,038.73        | -0.41            | 595.16      | -2.72            |
| 6%              | 1,611.13        | -2.11            | 1,030.42        | -1.21            | 580.70      | -5.08            |

| Panel B: $\delta_0 = 100, C^* = 47.83$ |
|-----------------|-----------------|-----------------|
| $r^S$           | $TV(\delta)$   | $f(\delta)$    | $D(\delta)$    |
| 5%              | 2,124.98        | 1,287.25        | 837.74          |
| 5.5%            | 2,098.26        | 1,283.03        | 815.23          |
| 6%              | 2,072.73        | 1,275.90        | 796.82          |

Table 5. Firm Value for different investor horizons.
The parameter values are $r = 0.05, \theta_a = 30, \theta_e = 1, \theta_{ae} = 1.5, \gamma = 5, \sigma_\delta = 0.25, \sigma_A = 0.12, \rho = 0, \tau = 0.15, \alpha = 0.30, \text{ and } \phi = -0.005, \text{ and } \psi = 1$. Table 5 shows that implementing policies dictated by investors with shorter investment horizons reduces both equity and debt prices. It is intuitive that more under investment will be detrimental for bondholders. Moreover, shareholders (with a discount rate of $r$) are also made worse off because the implemented policies are no longer optimal for them. The implemented policies feature too little long-term effort and too much short-term effort. Importantly, the associated changes induced by an increase in the discount rate from 5% to 6% lead approximately to a reduction of 1% in equity value, 2% in total firm value, and 5% in debt value. Thus, the potential welfare losses resulting from investors with shorter investment horizons are potentially significant.\footnote{The model restricts attention to the effect of short-term investors in our multi-tasking problem, which does not suggest there are no benefits from shorter investor time horizons in practice (see, e.g., Giannetti and Yu (2017). However, investors with shorter time horizons may, in practice, be precisely those who are better at identifying firms with higher baseline growth rates $\phi$ that we analyzed in the previous section (see, e.g., Edmans (2017)).}

5.4 Credit Spreads and Risk-Shifting

In this section we show that equity value responds very differently to an increase in the value of volatility of transitory versus permanent shocks. Panels A and B of Figure 7 depict comparative statics of equity value and credit spreads for different values of $\sigma_\delta$. Equity value can be increasing in
Panels A and B (C and D) chart the comparative statics of equity value and credit spreads for different values of the volatility of transitory (permanent) shocks. The parameter values are $C = 47.83$, $\delta_0 = 100$, $r = 0.05$, $\theta_a = 30$, $\theta_e = 1$, $\theta_{ae} = 0$, $\gamma = 5$, $\sigma_\delta = 0.25$, $\sigma_A = 0.12$, $\tau = 0.15$, $\alpha = 0.30$, $\phi = -0.005$, and $\psi = 1$.

the volatility of permanent shocks. As we discussed in the previous section the real options effect and the incentive cost effect affect equity value. For the parameter values we considered the real option effect dominates and shareholders would benefit from an increase in $\sigma_\delta$. Intuitively, shareholders face an asymmetric payoff from higher volatility. On the upside they can benefit if higher $\sigma_\delta$ leads to a large value of $\delta$, but can exercise their option to default if it leads to a low value of $\delta$. The higher probability of default negatively impacts debt value and leads to higher credit spreads on debt. This conflicting views between shareholders and bondholders with regards to the desirability of increasing volatility is known in the literature as asset substitution or risk-shifting problem.

Interestingly such conflict of interests is not present when it comes to volatility of transitory shocks $\sigma_A$. Panels C and D of Figure 7 depict comparative statics of shareholder value and credit spreads for different values of $\sigma_A$. In the case of volatility of transitory shocks shareholders are made worse off as a result of the higher cost of incentive provision, but there is no real-option effect. This higher cost is born by the shareholders and renders their claim less valuable, as seen in the picture. Moreover, as we discussed in the previous section (Panel B in Figure 3) higher $\sigma_A$ leads to earlier liquidation. Since bondholders bear the costs of bankruptcy they are made worse off by the increase in transitory volatility (credit spreads increase). In Figure 7 we set $\theta_{ae} = 0$ to make
the results more visible. However they hold for our baseline calibration with $\theta_{ae} = 1.5$, but the changes are quantitatively smaller and harder to see in a figure.

Our results rationalize the finding that managers do not rank concerns about risk-shifting amongst their most pressing concerns. In fact, Graham and Harvey (2001) find that managers are not concerned about minimizing risk-shifting concerns when deciding on their debt policy (for example by issuing short-term debt). Our results show this is rational when these concerns are about increasing volatility of transitory shocks. Such volatility would hurt both shareholders and bondholders, and thus bondholders would know that it is also in shareholders interest to minimize it.

6 Conclusion

This paper studies the tension corporations face between incentivizing long-term growth versus maximizing their short-term profitability. We show that short-termism is not necessarily the result of myopic managerial behavior but in fact optimal for shareholders. Short-termism is particularly desirable for shareholders of a financially distressed firm financed with debt. Furthermore, we characterize excessive short-termism as an indirect cost of debt and show that it is quantitatively as important as the cost of underinvestment. Notably, the conventional wisdom to ban short-termism (e.g., by the Securities and Exchange Commission) destroys shareholder value and policy proposals need to be more nuanced to recognize agency costs of debt and dynamic effects of commitment.

Two additional results have important implications for empirical work and policy making. First, high growth firms endogenously choose to implement higher levels of long-term effort. Hence, empirical results connecting firm performance and a proxy for long-term focus need to carefully account for endogeneity concerns. Second, firms controlled by investors with sub-optimally short horizons focus excessively in short-term profitability at the expense of long-term growth. Our results indicate such distortions significantly reduce firm value. Thus, designing capital gains tax policies aimed at lengthening investors horizon can lead to an improvement in social welfare.

Finally, our analysis suggests that the excessive short-termism problem can be mitigated by considering a richer set of debt contracts. That is, performance-sensitive debt or finite maturity debt can discipline shareholders to strike the right balance between long and short-term effort. Also, compensation contracts with inside debt can serve as a commitment device for shareholders to incentivize long-term effort during financial distress. These questions are fruitful for future research.
References


Miao, Jianjun, and Alejandro Rivera, 2016, Robust Contracts in Continuous Time, *Econometrica* 84, 1405-1440.


Appendix A

In this appendix we formally state the three adjoint problems that we use through the body of the paper: the case with no overhang \((NO)\), the case with no short-termism \((NS)\), and the case with no underinvestment \((NU)\).

Case 1: No Overhang

Consider the case in which shareholders can commit ex-ante to implementing long-term and short-term effort policies that maximize total firm value (but which do not maximize ex-post shareholder value). We denote the resulting equity and debt values by \(f_{NO}(\delta)\) and \(D_{NO}(\delta)\) respectively, where the subscript \(NO\) stands for no debt-overhang. Formally, \(f_{NO}(\delta)\) solves the following ODE

\[
r f_{NO}(\delta) = \max_{\delta_B^{NO}} \left\{ \delta (\psi + \epsilon_u) - (1 - \tau)C - \frac{\theta_a a_u^2 + 2 \theta_a a_u \epsilon_u + \theta_a \epsilon_u^2}{2} \delta + (\phi + a_u) \delta f'_{NO}(\delta) + \frac{1}{2} \sigma_s^2 \delta^2 f''_{NO}(\delta) \right\}
\]

subject to the value matching, smooth-pasting, and transversality conditions:

\[
  f_{NO}(\delta_B^{NO}) = 0, \quad f'_{NO}(\delta_B^{NO}) = 0, \quad \lim_{\delta \to \infty} f'_{NO}(\delta) = \frac{\psi}{r - \phi}.
\]

where \(a_u\) and \(\epsilon_u\) are given by \([17]\) which correspond to the effort policies obtained in the case of an unlevered firm. We also compute optimal leverage for the case in which shareholders can commit to the no overhang policies \(a_u\) and \(\epsilon_u\). Initial shareholders choose coupon \(C\) to maximize the levered total firm value \(TV_{NO}(\delta_0; C) = f_{NO}(\delta_0; C) + D_{NO}(\delta_0; C)\). The resulting initial market leverage ratio in this case is given by

\[
  ML_{NO}(\delta_0) = \frac{D_{NO}(\delta_0; C_{NO}^*(\delta_0))}{f_{NO}(\delta_0; C_{NO}^*(\delta_0)) + D_{NO}(\delta_0; C_{NO}^*(\delta_0))},
\]

where \(C_{NO}\) denotes the optimal coupon.

Case 2: No short-termism

Consider the case in which shareholders can commit ex-ante to implementing the short-term effort policy that maximizes total firm value (but which does not maximize ex-post shareholder value). However, shareholders are free to choose the long-term effort policy that is optimal for them. We denote the resulting equity and debt values by \(f_{NS}(\delta)\) and \(D_{NS}(\delta)\) respectively, where the subscript \(NS\) stands for no short-termism. Formally, \(f_{NS}(\delta)\) solves the following ODE

\[
r f_{NS}(\delta) = \max_{\delta_B^{NS}} \left\{ \delta (\psi + \epsilon_u) - (1 - \tau)C - \frac{\theta_a a_u^2 + 2 \theta_a a_u \epsilon_u + \theta_a \epsilon_u^2}{2} \delta + (\phi + a_u) \delta f'_{NS}(\delta) + \frac{1}{2} \sigma_s^2 \delta^2 f''_{NS}(\delta) \right\}
\]

subject to the value matching, smooth-pasting, and transversality conditions:

\[
  f_{NS}(\delta_B^{NS}) = 0, \quad f'_{NS}(\delta_B^{NS}) = 0, \quad \lim_{\delta \to \infty} f'_{NS}(\delta) = \frac{\psi}{r - \phi}.
\]

where \(\epsilon_u\) is given by the solution of \([17]\) which correspond to the effort policies obtained in the case of an unlevered firm. We also compute optimal leverage for the case in which shareholders can commit to the
short-term effort policy \(e_a\). Initial shareholders choose coupon \(C\) to maximize the levered total firm value 
\[TV_{NS}(\delta_0; C) = f_{NS}(\delta_0; C) + D_{NS}(\delta_0; C).\] The resulting initial market leverage ratio in this case is given by
\[ML_{NS}(\delta_0) = \frac{D_{NS}(\delta_0; C_{NS}^{*}(\delta_0))}{f_{NS}(\delta_0; C_{NS}^{*}(\delta_0)) + D_{NS}(\delta_0; C_{NS}^{*}(\delta_0))},\]
where \(C_{NS}\) denotes the optimal coupon.

Case 2: No Underinvestment

Consider the case in which shareholders can commit ex-ante to implementing the long-term effort policy that maximizes total firm value (but which does not maximize ex-post shareholder value). However, shareholders are free to choose the short-term effort policy that is optimal for them. We denote the resulting equity and maximizes total firm value (but which does not maximize ex-post shareholder value). However, shareholders are free to choose the short-term effort policy that is optimal for them. We denote the resulting equity and debt values by \(f_{NU}(\delta)\) and \(D_{NU}(\delta)\) respectively, where the subscript \(NU\) stands for no underinvestment. Formally, \(f_{NU}(\delta)\) solves the following ODE

\[
r f_{NU}(\delta) = \max_{\tilde{\eta}^{(c,a)}} \left\{ \delta(e + A) - (1 - \tau)C - \left(\theta_a^2 + 2\theta_a a + \theta_a^2 \epsilon^2\right)\frac{\delta}{2} + (\phi + a_\epsilon)\delta f_{NU}(\delta) + \frac{1}{2}\sigma^2 \delta^2 f_{NU}(\delta) \right\}
\]

subject to the value matching, smooth-pasting, and transversality conditions:
\[f_{NU}(\delta_B^{NU}) = 0, \quad f_{NU}'(\delta_B^{NU}) = 0, \quad \lim_{\delta \to \infty} f_{NU}'(\delta) = \frac{\psi}{r - \phi}.
\]

where \(a_\circ(\delta)\) is given by the solution of \([17]\), and the subscript \(NU\) stands for “no underinvestment”.

Appendix B

Appendix for Section 3

Dynamics of continuation value: Given any contract \(\Gamma = \langle c, a, e\rangle\), define the value process of the agent’s expected discounted utility 
\[V_t = \mathbb{E}_t \left[ \int_0^\infty e^{-r_s}u(c_s, a_s, e_s)ds \right]\] for any \(t\). As \((V_t)_{t \geq 0}\) is a martingale, and under the obedient choices \(\hat{a}_t = a_t\) and \(\hat{e}_t = e_t\), \((Z_t^\rho)_{t \geq 0}\) and \((Z_t^\gamma)_{t \geq 0}\) are two independent standard Brownian motions from the principal’s perspective, the Martingale Representation Theorem implies that there exists a progressively measurable process \((\beta_t^\rho, \beta_t^\gamma)_{t \geq 0}\) such that for any \(t > 0\),

\[V_t = V_0 + \int_0^t (-\gamma r W_s) e^{-rs} \beta_s^\rho ds + \int_0^t (-\gamma r W_s) \beta_s^\gamma dZ_s^\gamma + \int_0^t (-\gamma r W_s) \beta_s^\rho \sigma_d(dZ_s^\gamma)
\]

By \([5]\), \(V_t = \int_0^t e^{-r_s} u(c_s, a_s, e_s)ds + e^{-rt} W_t\). Differentiation of the two expressions for \(V_t\) with respect to \(t\) gives the dynamics \([6]\) ■

Verification of Optimality of the Contract:

We verify that the contract derived “heuristically” is indeed optimal. We denote by \{\hat{c}, \hat{a}, \hat{e}\} the recommended consumption and effort, and by \{\hat{c}, \hat{a}, \hat{e}\} the agents best response to the contract. We proceed in two steps:

Step 1. We show that the necessary conditions derived in section 3 are indeed optimal for the contract. That is, we show that the contract is incentive compatible and induces no private savings. Fix an arbitrary policy \{\hat{c}, \hat{a}, \hat{e}\} and consider the gain process

\[G_t^{\hat{c}, \hat{a}, \hat{e}} = \int_0^t -\frac{e^{-\gamma(s-t)} - g(\hat{a}_s, \hat{e}_s)}{\gamma} e^{-r_s ds} + e^{-rt} e^{-\gamma r S_t} V_t.
\]

38
where the agent’s cumulative private savings $S_t$ evolves according to
\[ dS_t = rS_t dt + (c_t - \hat{c}_t) dt, \]
and the agent’s continuation value under the recommended policies according to
\[ dW_t = (-\gamma r W_t) \left( \beta_t^T \delta_t (dA_t - \alpha(e_t) dt) + \beta_t^P (d\delta_t - \mu(a_t, \delta_t) dt) \right). \]

We now consider the dynamics of $G_t^{(\hat{c}, \hat{a}, \hat{e})}$ under the probability measure induced by $\{\hat{a}, \hat{e}\}$:
\[
e^{rt} e^{\gamma r S_t} dG_t^{(\hat{c}, \hat{a}, \hat{e})} = -\frac{e^{-\gamma (\hat{c}_t - g(\hat{a}_t, \hat{e}_t))}}{\gamma} e^{\gamma r S_t} dt - rW_t dt - \gamma r (rS_t + c_t - \hat{c}_t) W_t dt \]
\[ - \gamma r W_t (\beta_t^T \delta_t (\alpha(e_t) - \alpha(e_t) dt) + \beta_t^P (\mu(\hat{a}_t, \delta_t) - \mu(a_t, \delta_t) dt)) \]
\[ + (-\gamma r W_t) \left[ \beta_t^P \sigma(\delta_t) dZ_t^P + \beta_t^T \delta_t \sigma A dZ_t^T \right]. \]

Maximizing the drift with respect to $\{\hat{c}, \hat{a}, \hat{e}\}$ show that it is a concave problem in which the FOC guarantee optimality. Moreover, the necessary conditions derived in section 3 ensure that $E_t \left[ e^{rt} e^{\gamma r S_t} dG_t^{(\hat{c}, \hat{a}, \hat{e})} \right] \leq 0$ with equality for the optimal policies $\{c + rS, a, e\}$. Hence, we obtain that
\[
e^{rt} e^{\gamma r S_t} dG_t^{(\hat{c}, \hat{a}, \hat{e})} = \text{non-positive drift} + (-\gamma r W_t) \left[ \beta_t^P \sigma(\delta_t) dZ_t^P + \beta_t^T \delta_t \sigma A dZ_t^T \right]. \tag{29} \]

Integrating equation (29) yields:
\[
G_T^{(\hat{c}, \hat{a}, \hat{e})} \leq G_0^{(\hat{c}, \hat{a}, \hat{e})} + \int_0^T e^{-rt} e^{-\gamma r S_t} (-\gamma r W_t) \left[ \beta_t^P \sigma(\delta_t) dZ_t^P + \beta_t^T \delta_t \sigma A dZ_t^T \right],
\]
for all $T$. We impose extra conditions on the feasible paths of $S_t$ to ensure that the second term is a martingale. Importantly, under the recommended policies the second term becomes $\int_0^T e^{-rt} dV_t$ which is a martingale (since $V_t$ is a martingale). Letting $T \rightarrow \infty$, taking expectations, and recalling that $G_0^{(\hat{c}, \hat{a}, \hat{e})} = W_0$ imply that for any feasible policy $\{\hat{c}, \hat{a}, \hat{e}\}$
\[
E \left[ \int_0^\infty -\frac{e^{-\gamma (\hat{c}_t - g(\hat{a}_t, \hat{e}_t))}}{\gamma} e^{-rs} ds \right] \leq V_0
\]
provided that the transversality condition $\lim T \rightarrow \infty E \left[ e^{-rt} e^{-\gamma r S_T} V_T \right] = 0$ holds. Under the optimal policy the transversality conditions holds since $S_T = 0$ and $V_T$ is a martingale. For a general policy the transversality condition precludes Ponzi schemes with regard to the agent’s private savings. This completes the proof, since the agent has no incentive to deviate from the recommended policies.

**Step 2**. We now verify the optimality of the shareholder’s policy. Define the gain process for an arbitrary feasible effort policy $I$
\[
G_t^I = \int_0^t e^{-rs} \left( \delta_s (\psi + e_s) - (1 - \tau)C - c_s \right) ds + e^{-rt} P(\delta_t, W_t),
\]
where $P(\delta, W) = f_E(\delta) - \frac{1}{2\gamma} \ln(-\gamma r W_1)$. The gain process satisfies the following SDE
\[
dG_t^I = a_{Gt} dt + e^{-rt} \left( \sigma_2 \delta_t \left( f_E(\delta_t) - \beta_t^P \right) dZ_t^P - \beta_t^T \delta_t \sigma A dZ_t^T \right).
\]
Since $f_E(\delta)$ satisfies the HJB (19) the drift $a_{Gt}$ is non-positive for an arbitrary $I$ and equal to zero for the proposed optimal policy. We impose integrability conditions on feasible contracts to ensure the second term is a martingale.

Finally, integrating yields:
\[
\int_0^\infty e^{-rs} \left( \delta_s (\psi + e_s) - (1 - \tau)C - c_s \right) ds \leq P(\delta_0, W_0)
\]
provided that the transversality condition
\[
E \left[ \lim_{T \to \infty} e^{-rT} P(\delta_T, W_T) \right] = E \left[ \lim_{T \to \infty} e^{-rT} \left( f_E(\delta) - \frac{-1}{\gamma r} \ln(-\gamma r W_t) \right) \right] = 0
\]
is satisfied. For the optimal contract this condition is satisfied since \( f_E(\delta) \) is bounded by a linear function (the first best) and the growth rate of \( \delta \) is bounded above by \( r \), and since the compensation fund evolves as a semi-martingale with bounded drift. For an arbitrary policy we impose this transversality condition for feasible contracts, which essentially ensure a no Ponzi scheme condition for the compensation fund of the manager.

**Calibration** We calibrate the parameter values \( \theta_{ae} \) and \( \theta_e \) such that the direct compensation for the manager’s effort of an unlevered firm for size \( \delta = 80 \) are such that 40% of the compensation comes from long-term effort \( a \), 40% comes from short-term effort \( 2/5 \) and the remaining 20% comes from the cross-term between the two effort choices. To be precise this means that:

\[
\frac{\theta_{ae} \sigma_a^2(\delta) \delta}{2} = \frac{\theta_e \sigma_e^2(\delta) \delta}{2} = 2 \theta_a e_u(\delta) a_u(\delta) \delta
\]

for \( \delta = 80 \) which is right in the middle of the various values of \( \delta_0 \) that we use in our numerical examples.

**Solution for optimal effort policies in the HJB:** The optimization on the RHS of the HJB \([19]\) can be rewritten as

\[
\max_{a \in [\bar{a}, \underline{a}], e \in [0, e]} \left\{ a f'(\delta) + e - (1/2) a^2 \left( \theta_a + \gamma r \delta (\theta_a^2 + \theta_e^2 \sigma_A^2 + 2 \theta_{ae} \theta_a \sigma_{\delta A}) \right) - (1/2) e^2 \left( \theta_e + \gamma r e (\theta_a e_A^2 + \theta_e e_e^2 \sigma_A^2 + 2 \theta_{ae} \theta_e \sigma_{e A}) \right) \right\}
\]

\[
- a e \left( \theta_{ae} + \gamma r \delta (\theta_a e_{ae} \sigma_{\delta e}^2 + \theta_e \theta_{ae} e_{ae} \sigma_{\delta e}^2 + (\theta_a \theta_e + \theta_{ae}^2) \sigma_{\delta e A}) \right)
\]

We’ve made parametric assumptions on the coefficients such that the objective function is concave in the controls \((a, e)\). Since the constraints are linear this renders the Kuhn-Tucker conditions necessary and sufficient for optimality. Since the constraints on \( e \) never bind we consider only the constraints on \( a \), namely \( a \geq 0 \) and \( a \leq \bar{a} \). We denote by \( \lambda_1 \) and \( \lambda_2 \) the respective multipliers. We first calculate the unconstrained optimum by taking the FOC with respect to \( a \) and \( e \) to obtain

\[
a^* = \frac{f'(\delta) \bar{D} - \bar{B}}{AD - B^2}, \quad e^* = \frac{\bar{A} - f'(\delta) \bar{B}}{AD - B^2}.
\]

Three cases are possible:

1. When \( 0 < a^* < \bar{a} \) we claim the tuple \((a^*, e^*, 0, 0)\) satisfies the Kuhn-Tucker conditions: Stationarity is satisfied by construction, feasibility is satisfied by assumption, complimentary slackness is satisfied (since both multipliers are zero), and positivity is satisfied (again since both multipliers are zero).

2. When \( a^* \leq 0 \) we claim the the tuple \((0, e, \lambda_1, 0)\) satisfies the Kuhn-Tucker conditions (where \( e = 1/ \bar{D} \) and \( \lambda_1 = \frac{\bar{D} f'(\delta) - \bar{B}}{\bar{D}} \)): Stationarity is satisfied by construction, feasibility is satisfied (since \( a = 0 \), complimentary slackness is satisfied (since \( a = 0 \)), and it is straightforward to check that positivity of \( \lambda_1 \) is satisfied, and \( \lambda_2 \) is zero.
3. When \( a^* \geq \bar{a} \) we claim the tuple \((\bar{a}, \hat{e}, 0, \lambda_2)\) satisfies the Kuhn-Tucker conditions (where \( \hat{e} = \frac{1-\bar{a}B}{D} \) and \( \lambda_2 = \frac{\bar{a}(\bar{D} - \bar{B}^2) - (f'()D - B)}{\bar{D}} \)): Stationarity is satisfied by construction; feasibility is satisfied (since \( a = \bar{a} \)), complimentary slackness is satisfied (since \( a = \bar{a} \)), and it is straightforward to check that positivity of \( \lambda_2 \) is satisfied, and \( \lambda_1 \) is zero.

**Proof of Proposition 1**

We will show that the optimal level of short-term effort in the case when \( \rho = \theta_{ac} = 0 \) is given by \( e^* = \frac{1}{D} \), where in this case \( \dot{D} = \theta_a(1 + \gamma r \delta \theta e \sigma^2) \). Therefore, the optimal level of short-term effort would be independent of the coupon payment \( C \), thereby proving our proposition.

Consider the case in which the constraints on \( a \) do not bind. First, notice that \( \rho = \theta_{ac} = 0 \) implies \( \dot{B} = 0 \). Second, substituting \( \dot{B} = 0 \) in (30) yields \( e^* = \frac{1}{D} \), as expected. A similar argument shows that \( e^* = \frac{1}{D} \) for the case in which the constraints on \( a \) bind.

**Proof of Proposition 2**

Suppose that \( f(\delta) \) solves (26), which using our condensed notation for \( \dot{A}, \dot{B} \) and \( \dot{D} \) can be rewritten as:

\[
rf(\delta) = \left\{ \begin{array}{l}
\alpha_u(\delta) f'(\delta) + e_u(\delta) - a^2(\delta) \dot{A} - e^2(\delta) \dot{D} - a(\delta) e(\delta) \dot{B} \\
+ \delta\psi - (1 - \tau)C - \phi \delta f'(\delta) + \frac{1}{2}\sigma^2\delta^2 f''(\delta)
\end{array} \right.
\]

(subject to)

\[
f(\delta_{BO}^N) = 0, \quad f'(\delta_{BO}^N) = 0, \quad \lim_{\delta \to \infty} f'(\delta) \to \frac{\psi}{r - \phi}.
\]

Recall that \( a_u(\delta) \) and \( e_u(\delta) \) satisfy:

\[
a_u(\delta), e_u(\delta) \in \arg \max_{a,e} \left\{ \alpha_u(\delta) f'(\delta) + e_u(\delta) - \frac{a^2(\delta) \dot{A}}{2} - e^2(\delta) \dot{D} - a(\delta) e(\delta) \dot{B} \right\}
\]

which implies that

\[
e_u(\delta) \in \arg \max_e \left\{ a_u(\delta) f'(\delta) + e(\delta) - \frac{a^2(\delta) \dot{A}}{2} - e^2(\delta) \dot{D} - a(\delta) e(\delta) \dot{B} \right\}
\]

Therefore \( f(\delta) \) satisfies

\[
rf(\delta) = \max_{e \in [0,e]} \left\{ \begin{array}{l}
a_u(\delta) f'(\delta) + e - \frac{a^2(\delta) \dot{A}}{2} - e^2(\delta) \dot{D} - a(\delta) e(\delta) \dot{B} \\
+ \delta\psi - (1 - \tau)C - \phi \delta f'(\delta) + \frac{1}{2}\sigma^2\delta^2 f''(\delta)
\end{array} \right.
\]

and \( \delta_{BO}^N = \delta_{BO}^N \) satisfies the boundary conditions by assumption. Therefore \( f(\delta) \) satisfies (28). The proof of the converse is essentially identical.

**Characterization of the first-best value.** We consider the case in which \( \theta_{ac} = 0 \) (the general case is essentially the same, but the notation is more cumbersome and is available upon request). From the equation \( rq = \max_{(a,e)} \{(\psi + e) - \frac{1}{2} \left( \theta_a a^2 + \theta_e e^2 \right) + (\phi + a)q \} \) and first-order conditions \( q = \theta_a e^{FB} \) and \( 1 = \theta_e e^{FB} \), we have a quadratic equation in \( q^2 \): \( \frac{1}{2\theta_a} q^2 - (r - \phi)q + (\psi + \frac{1}{2\theta_a}) = 0 \). The larger positive root, \( q = \theta_a(r - \phi) + \theta_a \sqrt{(r - \phi)^2 - \frac{2}{\theta_a}(\psi + \frac{1}{2\theta_a})} \), is the required solution. For \( q \) to be real, we need \( (r - \phi)^2 > \frac{2}{\theta_a}(\psi + \frac{1}{2\theta_a}) \).

**Appendix for Section 4.4**

**Lemma 1.** Let \( \tau_B \equiv \inf \{ t | \delta_t \leq \delta_B \} \) and \( a_t \) and \( e_t \) be the interior solution on \([\delta_B, \infty)\) given \( \theta \). For \( \theta \in \{ \psi, \phi, \rho, \sigma, \alpha, \gamma, \theta_a, \theta_e, C, \tau \} \), denote \( f_{E,\theta}(\delta) \) as the value function for that parameter value. Then

\[
\frac{\partial f_{E,\theta}(\delta)}{\partial \theta} = \mathbb{E}^\delta \left[ \int_0^{\tau_B} e^{-rt} \left( \frac{\delta (\psi + e_t - (1 - \tau)C) - g(a_t, e_t; \delta_t)}{\theta_a a_t + \theta_e e_t + (\phi + a_t)q} + \frac{\partial (\psi + e_t)}{\partial \theta} \delta_t f'_{E,\theta}(\delta_t) + \frac{1}{2} \frac{\partial^2 (\psi + e_t)}{\partial \theta^2} \delta_t^2 f''_{E,\theta}(\delta_t) \right) dt \right]
\]

with \( \beta_t^p = \theta_a a_t + \theta_e e_t + \theta_{ac} a_t \) and \( \beta_t^e = \theta_e e_t + \theta_{ac} a_t \).
Proof. Given any default threshold $\delta_B$, the investor’s payoff $f_E(\delta; \delta_B)$ solves the ODE

$$rf_E(\delta; \delta_B) = \delta(\psi + e) - (1 - \tau)C - g(a, e; \delta) + (\phi + a)\delta f_E(\delta; \delta_B) + \frac{1}{2}\sigma_A^2\delta^2 f_E''(\delta; \delta_B)$$

with boundary condition $f_E(\delta_B; \delta_B) = 0$. Here $a$ and $e$ are the optimal choices and $\beta^T = \theta_a a + \theta_{ae} e$ and $\beta^T = \theta_a + \theta_{ae} a$. Denote $\delta_B(\theta)$ as the optimal default threshold under $\theta$, then by definition $f_{E, \theta}(\delta) = f_E(\delta; \delta_B(\theta))$. Differentiating both sides of (33) with respect to $\theta$ and evaluate at $\delta_B = \delta_B(\theta)$, then together with the Envelope theorem, we get

$$r \frac{\partial f_{E, \theta}(\delta)}{\partial \theta} = \frac{\partial}{\partial \theta} \left( \delta(\psi + e) - (1 - \tau)C - g(a, e; \delta) + (\phi + a) \delta f_E(\delta; \delta_B) + \frac{1}{2}\sigma_A^2\delta^2 f_E''(\delta; \delta_B) \right)$$

with boundary condition $\frac{\partial f_{E, \theta}(\delta)}{\partial \theta} = 0$ evaluated at the optimal default threshold. Note that $FOC(a)$ and $FOC(e)$ are the derivative of the right-hand side of the HJB-equation with respect to the efforts. The objects are zero under the optimal choice of $a$ and $e$. The lemma then follows from the Feynman-Kac formula.

From this lemma, we obtain the following partial derivatives:

- Drift terms: $\frac{\partial f_E(\delta)}{\partial \tau} = \mathbb{E}^\delta \left[ \int_0^\tau e^{-\tau t} \delta dt \right] > 0$, $\frac{\partial f_E(\delta)}{\partial \psi} = \mathbb{E}^\delta \left[ \int_0^\tau e^{-\tau t} \delta f_E(\delta) dt \right] > 0$
- Agency: $\frac{\partial f_E(\delta)}{\partial \theta} = \mathbb{E}^\delta \left[ \int_0^\tau e^{-\tau t} (-1 - 1 - \tau) dt \right] < 0$, $\frac{\partial f_E(\delta)}{\partial \psi} = \mathbb{E}^\delta \left[ \int_0^\tau e^{-\tau t} C dt \right] > 0$
- Coupon and tax rate: $\frac{\partial f_E(\delta)}{\partial a} = \mathbb{E}^\delta \left[ \int_0^\tau e^{-\tau t} ((\beta_t^T)^2 \sigma_A^2 + 2\sigma_A \sigma_A \beta_t^T + (\beta_t^T)^2 \sigma_A^2) dt \right] < 0$
- Coupon and tax rate: $\frac{\partial f_E(\delta)}{\partial a} = \mathbb{E}^\delta \left[ \int_0^\tau e^{-\tau t} ((\beta_t^T)^2 \sigma_A^2 + 2\sigma_A \sigma_A \beta_t^T + (\beta_t^T)^2 \sigma_A^2) dt \right] < 0$
- Effort costs: $\frac{\partial f_E(\delta)}{\partial a} = \mathbb{E}^\delta \left[ \int_0^\tau e^{-\tau t} \left( \frac{1}{2} \sigma_A^2 \delta \right) + \gamma r \delta^2 \left( \beta_t^T \sigma_A^2 + 2\sigma_A \sigma_A \beta_t^T \right) dt \right] < 0$ if $\rho \geq 0$ and ambiguous otherwise, $\frac{\partial f_E(\delta)}{\partial a} = \mathbb{E}^\delta \left[ \int_0^\tau e^{-\tau t} \left( \frac{1}{2} \sigma_A^2 \delta \right) + \gamma r \delta^2 \left( \beta_t^T \sigma_A^2 + 2\sigma_A \sigma_A \beta_t^T \right) dt \right] < 0$ if $\rho \geq 0$ and ambiguous otherwise, and

$$\frac{\partial f_E(\delta)}{\partial a} = \mathbb{E}^\delta \left[ \int_0^\tau e^{-\tau t} \left( \frac{1}{2} \sigma_A^2 \delta \right) + \gamma r \delta^2 \left( \beta_t^T \sigma_A^2 + 2\sigma_A \sigma_A \beta_t^T \right) dt \right] < 0$$

if $\rho \geq 0$ and ambiguous otherwise.

Lemma 2. Let $\delta_{B_1}$ be the optimal default boundary under parameter $\theta_i$. If $\theta_2 > \theta_1$ implies $f_{E, \theta_2}(\delta) > f_{E, \theta_1}(\delta)$ for all $\delta$, then $\delta_{B_2} < \delta_{B_1}$.

Proof. Suppose not, $\delta_{B_2} \geq \delta_{B_1}$. First, $f_{E, \theta_2}(\delta_{B_2}) = 0 > f_{E, \theta_1}(\delta_{B_1})$ by the optimality of and value-matching at $\delta_{B_2}$. Second, $f_{E, \theta_2}(\delta_{B_1}) > f_{E, \theta_1}(\delta_{B_1})$ by the hypothesis. Therefore, $f_{E, \theta_1}(\delta_{B_1}) < 0$ and the choice of $\delta_{B_1}$ violates limited liability.