Optimal Short-Termism

Finance Working Paper N° 546/2018
June 2021

Dirk Hackbarth
Boston University, CEPR and ECGI

Alejandro Rivera
University of Texas at Dallas

Tak-Yuen Wong
National Tsing Hua University

ECGI Working Paper Series in Finance

Optimal Short-Termism

Working Paper N° 546/2018
June 2021

Dirk Hackbarth
Alejandro Rivera
Tak-Yuen Wong

We are grateful to Gustavo Manso (the editor), the associate editor, two anonymous referees, Hengjie Ai (discussant), Dominic Barton, Maria Cecilia Bustamante (discussant), Gilles Chemla, David De Angelis, Alex Edmans, Thomas Geelen (discussant), Zhiguo He, Ron Kaniel, Alon Kfir (discussant), Taejin Kim (discussant), Leonid Kogan, Robert Marquez, Erwan Morellec, Martin Szydlowski, Stephen Terry, Valery Polkovnichenko, Antoinette Schoar, Sarah Williamson, Jinqiang Yang, seminar participants at the Center for Economic Policy Research (CEPR) Third Annual Spring Symposium in Financial Economics (London), the second Conference on Corporate Policies and Asset Prices (London), the Asian Meeting of the Econometric Society (Hong Kong), Paris-Dauphine, Shanghai University of Finance and Economics, the Financial Management Association Asia/Pacific (Hong Kong), the Asian Finance Association (Tokyo), the China International Conference in Finance (Tianjin), the Econometric Society European Meetings (Cologne), the Texas Lone Star Conference, the Texas Economic Theory Conference (Dallas), the International Conference on Financial Risks and Their Management (Kyoto), the European Finance Association Meeting (Carcavelos), and members of the McKinsey Global Institute and Focusing Capital on the Long Term (FCLT) Global for helpful comments.

Abstract

This paper develops a dynamic contracting (multi-tasking) model of a levered firm. In particular, the manager selects long-term and short-term efforts, and shareholders choose optimal debt and default policies. Excessive short-termism ex-post is optimal for shareholders, because debt has an asymmetric effect: shareholders receive all gains from short-term effort but share gains from long-term effort. We find that grim growth prospects and shareholder impatience imply higher optimal levels of shorttermism. Also, an incentive cost effect and a real option effect create non-trivial patterns for the endogenous default threshold. Finally, we quantify agency costs of excessive short-termism, which underscore the economic significance of our results.

Keywords: Capital structure, Contracting, Multi-tasking.

JEL Classifications: D86, G13, G32, G33, J33.

Dirk Hackbarth
Professor of Finance
Boston University, Questrom School of Business
595 Commonwealth Avenue
Boston, MA 02215, United States
phone: +1 617 358 4206
e-mail: dirkhackbarth@gmail.com.

Alejandro Rivera*
Assistant Professor of Finance
University of Texas at Dallas, Naveen Jindal School of Management
800 W. Campbell Road
Richardson, Texas 75080-3021, United States
phone: +1 972 883 5048
e-mail: alejorivera1@gmail.com

Tak-Yuen Wong
Associate Professor of Financer
National Tsing Hua University, Department of Quantitative Finance
No. 101 Section 2 KuangFu Road
Hsinchu, 300044, Taiwan
phone: +886035162136
e-mail: etywong110@gmail.com

*Corresponding Author

Optimal Short-Termism*

Dirk Hackbarth† Alejandro Rivera‡ Tak-Yuen Wong§

June 10, 2021

Abstract

This paper develops a dynamic contracting (multi-tasking) model of a levered firm. In particular, the manager selects long-term and short-term efforts, and shareholders choose optimal debt and default policies. Excessive short-termism ex-post is optimal for shareholders, because debt has an asymmetric effect: shareholders receive all gains from short-term effort but share gains from long-term effort. We find that grim growth prospects and shareholder impatience imply higher optimal levels of short-termism. Also, an incentive cost effect and a real option effect create non-trivial patterns for the endogenous default threshold. Finally, we quantify agency costs of excessive short-termism, which underscore the economic significance of our results.

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

*We are grateful to Gustavo Manso (the editor), the associate editor, two anonymous referees, Hengjie Ai (discussant), Dominic Barton, Maria Cecilia Bustamante (discussant), Gilles Chemla, David De Angelis, Alex Edmans, Thomas Geelen (discussant), Zhiguo He, Ron Kaniel, Alon Kfir (discussant), Taejin Kim (discussant), Leonid Kogan, Robert Marquez, Erwan Morellec, Martin Szydlowski, Stephen Terry, Valery Polkovnichenko, Antoinette Schoar, Sarah Williamson, Jinqiang Yang, seminar participants at the Center for Economic Policy Research (CEPR) Third Annual Spring Symposium in Financial Economics (London), the second Conference on Corporate Policies and Asset Prices (London), the Asian Meeting of the Econometric Society (Hong Kong), Paris-Dauphine, Shanghai University of Finance and Economics, the Financial Management Association Asia/Pacific (Hong Kong), the Asian Finance Association (Tokyo), the China International Conference in Finance (Tianjin), the Econometric Society European Meetings (Cologne), the Texas Lone Star Conference, the Texas Economic Theory Conference (Dallas), the International Conference on Financial Risks and Their Management (Kyoto), the European Finance Association Meeting (Carcavelos), and members of the McKinsey Global Institute and Focusing Capital on the Long Term (FCLT) Global for helpful comments.

†Boston University Questrom School of Business, CEPR, and ECGI. Email: dirkhackbarth@gmail.com.
‡Naveen Jindal School of Management, University of Texas at Dallas. Email: alejorivera1@gmail.com
§Department of Quantitative Finance, National Tsing Hua University. Email: etywong110@gmail.com.

"The jury is still out on corporate short-termism." — Lawrence Summers
Financial Times, February 9, 2017

1 Introduction

Corporate short-termism has often been criticized for harming long-term performance.[1] 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 suboptimal for shareholders? Or rather, can the behavior of managers be a result of equity value-maximization that crucially depends on firm characteristics, such as current debt-equity ratio and future growth prospects?

This paper answers these and related questions by analyzing a contracting problem between a manager (agent) and shareholders (principal) of a company capitalized with debt and equity (He, 2011). Shareholders design incentives for the manager to exert long-term effort (growth) and short-term effort (profit). In our novel setup, a resource constraint that binds the agent's effort captures the essence of short-termism: an increase in short-term effort makes long-term effort more costly, thereby undermining long-term performance. Notably, ours is the first model in which capital structure (debt and default), incentive compensation, and two-dimensional (short-term and long-term) effort are jointly optimized in a dynamic setting. Hence, we make important, nontrivial progress toward understanding the interactions between debt and multi-tasking and, in particular, excessive short-termism (i.e., boosting short-term performance at the expense of long-term growth) within an optimal contracting framework.

Our model reveals a novel short-termism trade-off. 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 immediately receive all benefits from short-term effort, because it generates higher contemporaneous cash flows that are not transferred to debt in bankruptcy. That is, debt has an asymmetric effect. While it decreases long-term effort, it does not reduce the gains from short-term effort. Shareholders balance benefits of short-term effort (profitability) and long-term effort (growth) so that debt leads to excessive short-termism (i.e., higher short-term effort than for an unlevered firm).

[1] See McKinsey & Co.'s short-termism study or surveys by Poterba and Summers (1995) and Graham et al. (2005). Relatedly, an important, ongoing controversy is whether public firms face stock market pressures that induce short-termism in their investment policies (Asker et al., 2015; Feldman et al., 2018). Moreover, it is an open question whether no/zero short-termism is even the benchmark policy for equity holders.

We model two stochastic processes: the firm's size and its profitability. The firm's 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 have only a contemporaneous effect on cash flows (transitory shocks). In our multi-tasking framework, managerial effort can be allocated toward 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. Because 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 the costs of bankruptcy.

While our main result shows that boosting short-term performance at the expense of long-term growth 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 toward financial distress, the underinvestment problem increases; hence, it becomes increasingly desirable for shareholders to implement higher levels of short-term effort. Less long-term effort not only reduces the risk of investment benefits being largely reaped by bondholders but also makes it cheaper for shareholders to incentivize more short-term effort (short-termism). Higher levels of short-termism are optimal for shareholders, but detrimental to bondholders. Hence, short-termism is an indirect—in addition to underinvestment as a direct—agency cost of debt. Interestingly, a commitment to no underinvestment is also a commitment to no short-termism, but the converse is not true.

We emphasize that the main result on short-termism is driven by capital structure and debt overhang, and therefore applies generally to firms that are not subject to shareholder-manager conflicts of interest. By incorporating moral hazard, we acquire two benefits. First, we demonstrate that short-termism arises even if corporate managers are nonmyopic, and it is affected by shareholders' investment horizons. Second, the workhorse model of Leland (1994) attains an optimal leverage ratio of 60% to 70%. Managerial moral hazard resolves this issue by delivering a realistic optimal leverage. Hence, we can calibrate and quantify the economic impact of short-termism more precisely.

We now discuss four additional results. First, our model highlights endogeneity issues. 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 that 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.[2]

Second, the continuous-time, dynamic framework permits analytical comparative statics. Let us highlight the heterogeneous effect of permanent volatility and transitory volatility. To begin with, the volatility of permanent shocks (growth shocks) has a nonmonotonic effect on the default threshold and equity value. On the one hand, the real-option effect implies that shareholders tend to wait and delay default under a higher permanent volatility. This increases the shareholders' option value to default, thereby increasing equity value at the expense of bondholders. On the other hand, in the presence of moral hazard, the firm's size becomes a noisier signal about the agent's long-term effort when the firm's growth is more uncertain. Thus, incentive provision becomes costlier. The increased incentive cost implies lower cash-flow growth, a lower equity value, and a higher default threshold. Our calibration shows that the real-option effect dominates the incentive-cost effect for large volatility, resulting in an inverted-U-shaped default threshold. This nonmonotonicity provides a rationalization for the mixed evidence in regards to the risk-shifting hypothesis found in the empirical literature.[3]

However, the volatility of the transitory shocks (profit shocks) has only the incentive-cost effect: as the profitability becomes a noisier signal about the short-term effort, shareholders reduce short-term incentives. This reduction amplifies the loss absorbed by shareholders during financial distress. Therefore, equity value decreases and the default threshold increases in the volatility of the transitory shocks.[4]

Third, we adopt a baseline calibration and quantify the agency cost of debt, which we decompose 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 1%, where up to one

[2]Indeed, Kaplan (2017) finds little long-term evidence in favor of the so-called short-termism critique. For better identifying the critique's merits, it is crucial to have reliable proxies for differences in growth opportunities.

[3]For example, see Eisdorfer (2008) for evidence in favor of the risk-shifting hypothesis and Gilje (2018) for evidence against the risk-shifting hypothesis.

[4]We also show that a higher correlation of transitory and permanent shocks increases the risk borne by the manager, which increases the risk compensation. Therefore, correlation unambiguously increases the default threshold and reduces the equity value. In a dynamic liquidity management model, without agency frictions, 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.

half of this reduction is due to excessive short-termism.[5] However, contrary to standard intuition, managerial short-termism is not detrimental to equity value, but is in fact desirable. Short-termism is an indirect cost of debt overhang; hence, there are two related commitment problems, one for underinvestment and one for short-termism, which bondholders have to recognize at the outset.[6]

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 it optimal to implement higher short-term effort and lower long-term effort, which in turn will reduce equity value for both the patient, regular shareholders and the bondholders, both of whom 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 1% in equity value, 5% of debt value, and 2.5% 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]

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 a short horizon, and also when a firm is financed with short-term debt. Holmstrom and Tirole (1993) show that the liquidity of the stock market affects the efficiency of equity-based compensation in disciplining managerial myopia. Froot et al. (1992) discuss a potential link between the short-term horizon of shareholders and short-term managerial behavior. Compared to these works, our fully 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, thus inducing myopia. Edmans et al. (2012) and Marinovic and Varas (2019) show that equity vesting implements the manipulation-minimizing optimal dynamic contract. Zhu (2018) shows that a contract tracking the number of consecutive high outputs mitigates myopic agency. Thakor (2019) argues that short-termism may limit managers' information rent, because long-term projects could pro-

[5] However, if we use debt value instead of firm value as the distortion's point of reference—as suggested, e.g., by Mello and Parsons (1992)—the total agency cost exceeds 2%, of which more than 1% is due to short-termism.

[6] He (2011) implicitly assumes commitment to no short-termism and thus obtains higher ex-ante firm values. However, ex-post equity values are higher in our setting, because short-termism enhances equity value.

[7] To 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.

duce more noisy information regarding managers' ability. While these works focus on short-term actions that are value-destroying (e.g., earnings manipulation), our paper focuses on the case in 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 model setup is based on the literature on multi-tasking agency spawned by Holmstrom and Milgrom (1991). Static papers include Itoh (1993), Auriol et al. (2002), and Bernardo et al. (2004). These papers focus on incentive provisions but do not analyze the impact of capital structure. Our paper contributes to the growing literature on dynamic multi-tasking models.[8] Szydlowski (2019) studies project choices. In his model, all shocks are transitory, and thus, actions have short-lived symmetric effects on cash-flow duration. DeMarzo et al. (2014) and Wong (2019) show that high-powered incentives lead to increased risk-taking. However, the effects of risk choices in these models are short-term. Rivera (2020) studies the interaction between risk-shifting and moral hazard and finds that moral hazard amplifies risk-shifting incentives. Our paper is the first to explicitly model the firm's optimal capital structure when the two tasks have asymmetric effects on the term structure of cash flows.

Our paper also builds on recent research in dynamic corporate finance. Gorbenko and Strebulaev (2010) study the effect of temporary Poisson shocks in the Leland framework. Décamps et al. (2017) and Lee and Rivera (2021) study permanent and transitory shocks in an equity-financed firm with liquidity concerns. 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 over both the long term and the short term and relate them to the debt-overhang problem. In a contemporaneous paper, Gryglewicz et al. (2020) study short- and long-term efforts in a dynamic agency model of an all-equity-financed firm, which rationalizes asymmetric benchmarking and pay-for-luck observed in the data. Hence, our paper's contribution is complementary to theirs.

The paper proceeds as follows. Section 2 sets up the model. Section 3 characterizes the manager's problem. Section 4 analyzes the shareholder's problem and derives the qualitative results. Section 5 quantifies. Section 6 studies alternative setups. Section 7 concludes.

[8] Seminal papers in dynamic agency with a single action include DeMarzo and Sannikov (2006), Biais et al. (2007), and Biais et al. (2010).

2 Model Setup

Consider a continuous-time environment with infinite horizon. We model a levered firm with dynamic managerial moral hazard. At time 0, shareholders (the principal) raise long-term debt to finance the firm. Once the debt is in place, shareholders hire a manager (agent) to operate a project and implement the firm's investment (agent's effort) policy through a dynamic incentive contract. The agent multi-tasks: on the one hand, she exerts effort to grow the firm in the long run; on the other hand, she takes short-term actions to raise the firm's current profitability. The default time is chosen by the shareholders ex-post. Shareholders and bondholders are risk-neutral, and the manager is risk-averse. Everyone discounts cash flows at rate $r > 0$.

The project produces a stream of cash flows subject to both permanent and transitory shocks. Over $(t, t + dt)$, the after-tax cash flow is given by $dY_t = \delta_t dA_t$, where δ_t is the firm size and dA_t is the contemporaneous profitability. The firm size $\delta_t \in [\underline{\delta}, \infty)$ satisfies

$$d\delta_t = (\phi + a_t)\delta_t dt + \sigma_\delta \delta_t dZ_t^P, \qquad t < \underline{\tau} \tag{1}$$

$$\delta_t = \underline{\delta}, \qquad t \geq \underline{\tau}, \tag{2}$$

where $\underline{\tau} = \inf\{t : \delta_t \leq \underline{\delta}\}$.[9] Moreover, ϕ is the baseline growth rate of the firm size, $a_t \in [0, \bar{a}]$ is the unobservable long-term effort exerted by the agent, $\sigma_\delta > 0$ is the volatility, and Z_t^P is a standard Brownian motion. Because cash flows are proportional to δ_t; both the shock dZ_t^P and the agent's effort a_t increase the firm's size δ_t and have a permanent impact on the project's prospects.

The profitability A_t follows the controlled arithmetic Brownian motion process:

$$dA_t = (\psi + e_t)dt + \sigma_A dZ_t^T, \tag{3}$$

where ψ is the baseline drift rate of the profitability, $e_t \in [0, \bar{e}]$ is the unobservable short-term effort exerted by the agent, $\sigma_A > 0$ is the volatility, and Z_t^T is a standard Brownian motion. Given that the current cash flow depends on the contemporaneous profitability but not the future, the effort e_t and shock dZ_t^T have only a transitory effect on the firm. Permanent shock Z_t^P and transitory shock Z_t^T have a correlation coefficient $\rho \in [-1, 1]$, so that $\mathbb{E}_t \left[dZ_t^P dZ_t^T \right] = \rho dt$. Let Z_t^A be another standard Brownian motion that is independent of the permanent shock Z_t^p. We can express the

[9] That is, δ_t follows a controlled geometric Brownian motion process with absorbing boundary at $\underline{\delta}$. We impose this arbitrarily small absorbing boundary for technical reasons. It has no impact on our numerical results, where we set $\underline{\delta} = 10^{-1}$.

transitory shock as

$$dZ_t^T = \rho dZ_t^P + \sqrt{1-\rho^2} dZ_t^A. \tag{4}$$

That is, short-term profitability shocks dZ_t^T consist of profit-specific transitory shocks dZ_t^A and permanent shocks dZ_t^P.[10]

Shareholders observe the paths of cash flows $\{Y_t : t \geq 0\}$ and firm size $\{\delta_t : t \geq 0\}$. This implies that the incremental profitability dA_t is observable to the shareholders as well. Denote by $\hat{\mathbb{F}} = \{\mathcal{F}_t : t \geq 0\}$ the filtration generated by the public observables, that is, $\mathcal{F}_t = \sigma(\{\delta_s, A_s : s \leq t\})$ for all t. Let $\mathbb{F}$ be the smallest augmented filtration containing $\hat{\mathbb{F}}$.

The agent's 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 a quadratic form for the effort cost $g(a, e; \delta) = \frac{1}{2}\left(\theta_a a^2 + \theta_e e^2 + 2\theta_{ae} ae\right)\delta$. The tasks are asymmetric in cost, the cost is proportional to firm size δ, and the two efforts can be either complements ($\theta_{ae} \leq 0$) or substitutes ($\theta_{ae} \geq 0$). The latter case implies a resource constraint on the agent's total effort, capturing the essence of short-termism: an increase in short-term effort makes long-term effort more costly, thereby undermining long-term growth.[11] Lastly, γ is the coefficient of absolute risk aversion under CARA utility. Given a stream of consumption $\{c_t : t \geq 0\}$, 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, in which she can borrow and save at 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 $S_0 = 0$. As usual, this account is subject to a no-Ponzi-scheme condition.

As in Leland (1994), the firm issues a perpetual debt with a constant coupon rate C.[12] With a marginal corporate tax rate $\tau \in (0, 1)$, the tax shield per unit of time is τC. We denote by $\alpha \in (0, 1)$ the proportional bankruptcy cost parameter. The structural-credit-risk models have illustrated that

[10] As shown later, our main results rely on the substitutability of efforts. Correlation, when positive, is isomorphic to substitutability. Restrictions on the effort cost (to be specified below) capture this substitution effect directly.

[11] Our modeling approach avoids the complexity induced by time-varying effort persistence; for example, the decaying long-run effect of manipulation in Marinovic and Varas (2018) requires the optimal contract to track the agent's information rent as an additional state variable.

[12] See Section 6.1 for an extension of our model to the finite maturity case.

endogenous default by shareholders is an important mechanism in understanding credit risks. Let δ_B be the default threshold. If the firm's fundamentals (firm size) δ_t become sufficiently weak, especially after a sequence of negative permanent shocks, shareholders will default once $\delta_t \leq \delta_B$.

We assume that the shareholders can fulfill their promise to the agent at default. However, under our implementation of the optimal contract (see Section 4.1.2), shareholders don't have incentives to renege on their promises to the manager. As a result, the agent's continuation utility is irrelevant when choosing the optimal default strategy. Upon default, bondholders receive the liquidation value of the firm and continue to run the firm as an unlevered firm. For simplicity, we assume the agent continues to operate the firm after the change in ownership, with a continuation contract offered by the unlevered firm.

2.1 The Contracting Problem

At time 0, and right after debt is issued, the shareholders design a contract to maximize their expected discounted profits. Both shareholders and the agent can fully commit to the contract $\Gamma = \langle c, a, e, \tau_B \rangle$, which specifies the agent's wage process $\{c_t : t \geq 0\}$, the recommended effort policy $\{(a_t, e_t) : t \geq 0\}$, and the default (stopping) time τ_B; all processes are $\mathbb{F}$-adapted. Given a contract Γ, the agent solves

$$W_0(\delta_0, \Gamma) = \sup_{(\hat{c},\hat{a},\hat{e})} \mathbb{E}^{(\hat{a},\hat{e})} \left[\int_0^\infty -\frac{1}{\gamma} e^{-\gamma(\hat{c}_t - g(\hat{a}_t, \hat{e}_t; \delta_t))} e^{-rt} dt \right] \tag{5}$$
$$\text{s.t. } dS_t = rS_t dt + c_t dt - \hat{c}_t dt,\ S_0 = 0,\ S_t \geq 0,$$

and also subject to (1) and (3). In (5), $W_0(\delta_0, \Gamma)$ is the agent's time-0 utility under the contract Γ and $\mathbb{E}^{(a,e)}[\cdot]$ is the expectation taken under the probability measure $\mathcal{P}^{(a,e)}$ induced by the effort policy (a, e). The intertemporal budget constraint specifies the evolution of the agent's saving account: over $(t, t + dt)$, the change in saving dS_t is the accumulated interest $rS_t dt$ plus his wage $c_t dt$ minus his actual consumption $\hat{c}_t dt$.

A contract is said to be 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 policy obediently, $\hat{a}_t = a_t$ and $\hat{e}_t = e_t$; and does not save or withdraw from the bank account, $\hat{c}_t = c_t$. Formally, given debt

C, the shareholders' problem is to find an optimal contract:

$$\sup_{\Gamma} \mathbb{E}^{(a,e)} \left[\int_0^{\infty} e^{-rt} \left((\delta_t(\psi + e_t) - (1-\tau)C) \mathbf{1}_{t \le \tau_B \wedge \underline{\tau}} - c_t \right) dt \right] \tag{6}$$

$$\text{s.t. } \Gamma \text{ is ICNS and } W_0(\delta_0, \Gamma) \ge w_0, \tag{7}$$

where τ_B and $\underline{\tau}$ are stopping times that are $\mathbb{F}$-measurable, and $\mathbf{1}_{t \le \tau_B \wedge \underline{\tau}}$ is an indicator function. In the shareholders' problem, the first constraint requires the contract to be ICNS, and the second constraint is the agent's participation constraint for an initial outside option w_0.[13] Note that the objective of the contracting problem contains the postdefault consumption, which is driven by our assumption that the shareholders have the ability to fulfill the promise at default.

3 The Agent's Problem

In this section, we characterize the agent's behavior given an arbitrary contract and provide the necessary conditions for a contract to be ICNS. Following the dynamic contracting literature (e.g., Sannikov, 2008), we take the agent's continuation utility as the state variable. Given any contract $\Gamma = \langle c, a, e, \tau_B \rangle$, the agent's continuation utility on the equilibrium path is defined as

$$W_t(\delta_t, \Gamma) \equiv \mathbb{E}_t^{(a,e)} \left[\int_t^{\infty} -\frac{1}{\gamma} e^{-\gamma(c_s - g(a_s, e_s; \delta_s)) - r(s-t)} ds \right]. \tag{8}$$

Here, the agent's continuation utility depends on the firm size δ_t, which serves as a natural state variable, and the contract Γ induces continuation consumption $\{c_s : s \ge t\}$ and effort choices $\{(a_s, e_s) : s \ge t\}$. The expectation $\mathbb{E}_t^{(a,e)} [\cdot]$ is conditional on information $\mathcal{F}_t$.[14]

Proposition 1. *The necessary and sufficient conditions for a contract $\Gamma = \langle c, a, e, \tau_B \rangle$ to be ICNS are as follows. First, the dynamics of the agent's continuation utility satisfies*

$$dW_t = (rW_t - u(c_t, a_t, e_t))\, dt + (-\gamma r W_t) (\underbrace{\beta_t^P \sigma_\delta \delta_t dZ_t^P}_{\text{Permanent}} + \underbrace{\beta_t^T \sigma_A \delta_t dZ_t^T}_{\text{Transitory}}), \tag{9}$$

where the processes $\left\{(\beta_t^P, \beta_t^T) : t \ge 0\right\}$ are progressively measurable with respect to $\mathbb{F}$ and satisfy for any effort policies (a, e) the square integrability conditions:

$$\mathbb{E}^{(a,e)} \left[\int_0^{\infty} (\beta_t^P)^2 \delta_t^2 dt \right] < \infty \text{ and } \mathbb{E}^{(a,e)} \left[\int_0^{\infty} (\beta_t^T)^2 \delta_t^2 dt \right] < \infty, \tag{10}$$

[13] Similar to the revelation principle, it is without loss of generality that we can focus on ICNS contracts. This is because the principal has full commitment and can save on behalf of the agent.

[14] It is well known that with private savings, FOCs cannot guarantee the global optimality of the recommended policies (e.g., Kocherlakota, 2004). However, under CARA preferences without wealth effects, the first-order conditions (11) together with the no-saving conditions are sufficient for contracts to be ICNS.

Moreover, incentive-compatible effort policies satisfy the agent's first-order conditions

$$\beta_t^P = \theta_a a_t + \theta_{ae} e_t \text{ and } \beta_t^T = \theta_{ae} a_t + \theta_e e_t \tag{11}$$

at all times. If $\theta_a\theta_e \neq \theta_{ae}^2$, then the mapping $(\beta_t^P, \beta_t^T) \mapsto (a_t, e_t)$ defined by the first-order conditions is one-to-one and onto. Second, the no-savings condition requires that for any time t,

$$u_c(c_t, a_t, e_t) = -\gamma r W_t(\delta_t, \Gamma; S = 0), \tag{12}$$

where $W_t(\delta, \Gamma; S)$ is the continuation utility with a time-t saving balance S.

The novelty of our model is that the multi-tasking agent's effort choices affect both the firm size and contemporaneous profitability. Thus, the contract stipulates both short-term and long-term incentives to control the choices by exposing the agent to both permanent and transitory shocks. We measure incentives in monetary terms: the scaling factor $-\gamma r W_t$ in (9) translates a dollar into a unit of utility. Therefore, the sensitivity of W_t to fundamental shocks, $\beta_t^P \propto dW_t/d\delta_t$, captures the marginal monetary incentive for the long-term effort; and the sensitivity of W_t to profit shocks, $\beta_t^T \propto dW_t/dA_t$, represents the marginal monetary incentive for the short-term effort.

Given monetary incentives (β_t^P, β_t^T) that satisfy technical restrictions (10), the agent chooses long-term effort $\hat{a}_t$ and short-term effort $\hat{e}_t$ to maximize her continuation utility at each point in time. The first-order conditions (11) are local incentive compatibility constraints, which state that the agent's marginal benefit of both efforts must equal her marginal monetary cost of the respective effort. Under the assumption that $\theta_a\theta_e \neq \theta_{ae}^2$, the incentives (β_t^P, β_t^T) that satisfy (11) uniquely implement (a_t, e_t).

Lastly, the no-savings condition (12) is essentially the consumption Euler equation, which states that it is optimal for the agent to consume all the current wage, given that the saving balance is zero. This condition has two implications. First, given the CARA utility, it must be that $u(c_t, a_t, e_t) = rW_t$ and the drift of (9) vanishes. This implies that, under the integrability condition (10), both the continuation utility W_t and marginal utility of consumption $u_c = -\gamma r W_t$ evolve as martingales. Second, the agent's consumption must be

$$c_t = g(a_t, e_t; \delta_t) + \frac{-1}{\gamma} \ln(-\gamma r W_t). \tag{13}$$

Therefore, at each point in time, the contract compensates the agent for her effort cost in addition to the promised utility, which measures the agent's performance in the past. Substituting back these conditions yields that the continuation value of the agent is a martingale with dynamics given by

$$dW_t = (-\gamma r W_t)\left(\beta_t^P \sigma_\delta \delta_t dZ_t^P + \beta_t^T \sigma_A \delta_t dZ_t^T\right). \tag{14}$$

4 Dynamic Agency and Capital Structure

4.1 Optimal Contract in a Levered Firm

Given the dynamics of the agent's continuation utility and the necessary and sufficient conditions for incentive compatibility and no savings, we define the set of admissible contracts denoted by $\mathcal{C}$.

Definition 1. *A contract* $(\{\beta_t^P, \beta_t^T\}, \tau_B)$ *is admissible if*

- *it is adapted to the filtration* $\mathbb{F}$ *and satisfies the integrability conditions (10).*
- *the corresponding effort policies obtained according to (11) are such that* $a_t \in [0, \bar{a}\mathbf{1}_{\underline{\delta}<\delta_t<\overline{\delta}}]$ *and* $e_t \in [0, \bar{e}\mathbf{1}_{\underline{\delta}<\delta_t<\overline{\delta}}]$ *for all times* t*; where* $\bar{a}$, $\bar{e}$, $\underline{\delta}$, *and* $\overline{\delta}$ *are fixed (finite) parameters.*[15]
- *equations (1) and (14) have unique strong solutions.*

We can now write the Markov formulation of the original shareholder's problem (6)-(7):[16]

$$P(\delta_0, W_0) = \sup_{(\{\beta_t^P, \beta_t^T\}, \tau_B) \in \mathcal{C}} \mathbb{E}^{(a^*, e^*)}\left[\int_0^\infty e^{-rt}\left(\left(\delta_t(\psi + e_t) - (1-\tau)C\right)\mathbf{1}_{t \le \tau_B \wedge \mathcal{I}} - c_t\right)dt\right] \tag{15}$$

$$\text{with } (a^*, e^*) = (a_t^*, e_t^*) \text{ given by (11), and subject to (1), (13), and (14).} \tag{16}$$

$P(\delta_0, W_0)$ denotes the shareholders' value function when the initial continuation value and firm size are W_0 and δ_0, respectively. We first solve heuristically (15)-(16) and then formalize these results at the end of this section in Proposition 2, which we prove in Appendix A.2.

The shareholders' value function can be written as

$$P(\delta_t, W_t) = \underbrace{f(\delta)}_{\text{total equity}} - \underbrace{\frac{-1}{\gamma r}\ln(-\gamma r W_t)}_{\text{agent's certaintly equivalent}}. \tag{17}$$

[15] For technical reasons, the effort policies have to equal zero for very large values of δ. In our numerical simulations, we set $\overline{\delta} = 10^4$; this bound has no impact in our numerical results. Because $\underline{\delta}$ (resp. $\overline{\delta}$) is extremely small (resp. large), assumptions about such remote events have a negligible impact on value functions in the regime of interest due to discounting. For example, as seen in expression (6), we assume that shareholders get zero additional cash flows after hitting the absorbing boundary $\delta_t = \underline{\delta}$. However, we can assume that they instead get the unlevered firm value with observable effort described in Section A.2.2, and our numerical results stay effectively unchanged.

[16] See Décamps and Villeneuve (2019) for further details regarding the delicate treatment required to transform a dynamic contracting problem with an embedded real option into its Markovian representation.

The shareholders' problem is equivalent to the one that maximizes the total equity, which is independent of the agent's continuation utility W_t, given that (a, e) satisfies (11). The Hamilton-Jacobi-Bellman (HJB) equation for the principal's problem is

$$rf(\delta) = \max_{(a,e)} \left\{ \begin{array}{l} \delta(\psi + e) - (1-\tau)C - \frac{1}{2}\delta(\theta_a a^2 + 2\theta_{ae} ae + \theta_e e^2) + (\phi + a)\delta f'(\delta) + \frac{1}{2}\sigma_\delta^2 \delta^2 f''(\delta) \\ -\frac{\gamma r \delta^2}{2} \left((\theta_a a + \theta_{ae} e)^2 \sigma_\delta^2 + (\theta_e e + \theta_{ae} a)^2 \sigma_A^2 + 2\rho (\theta_a a + \theta_{ae} e)(\theta_e e + \theta_{ae} a) \sigma_\delta \sigma_A \right) \end{array} \right\}, \tag{18}$$

which is an ordinary differential equation (ODE). In the HJB, the first line is the expected cash flow net of the monetary cost of effort plus the expected capital gain due the changes in the firm size. The second line represents the incentive cost for the shareholders to induce both the long-term and the short-term effort. Incentive costs arise because the risk-averse agent is exposed to both the permanent and the transitory shocks, and thus, risk compensation is required. The ODE (18) is subject to the following value-matching, smooth-pasting, and transversality conditions:

$$f(\delta_B) = 0,\ f'(\delta_B) = 0,\ \lim_{\delta \to \infty} \left(f(\delta) - \delta \frac{\psi}{r - \phi} \right) = (1-\tau)\frac{C}{r}. \tag{19}$$

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 ψ per unit of capital, capitalized by the baseline growth rate of the firm ϕ. This is due to the fact that efforts converge to zero as δ goes to infinity.

Assuming an interior solution over the effort choices delivers solutions

$$a^*(\delta) = K_0 f'(\delta) + K_1 \text{ and } e^*(\delta) = K_1 f'(\delta) + K_2, \tag{20}$$

where K_0, K_1, and K_2 are given in Appendix A.2.1.

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 Appendix A.2.1, we compute $\hat{B}$ and $\hat{D}$ such that when the constraint binds at $a \geq 0$ (lower bound), the optimal policies are given by $e^*(\delta) = \frac{1}{\hat{D}}$ and $a^*(\delta) = 0$. If the constraint on long-term effort binds at $a \leq \bar{a}$ (upper bound), optimal policies are given by $e^*(\delta) = \frac{1-\bar{a}\hat{B}}{\hat{D}}$ and $a^*(\delta) = \bar{a}$. Because of the after-tax coupon payment $(1-\tau)C$, shareholders absorb more losses when the firm's fundamental (firm size) is weak: $\delta_t(\psi + e^*(\delta_t)) < (1-\tau)C$, net of the monetary compensation for the effort cost. Therefore, as δ_t falls to δ_B, shareholders default optimally and refuse to fulfill their debt obligations. Standard value-matching $f(\delta_B) = 0$ and smooth-pasting $f'(\delta_B) = 0$ conditions characterize the endogenous default threshold and its optimality.

Before stating our main result, we assume the following sufficient conditions on the model's parameters:

Condition 1. $\psi > \frac{1}{2}(\theta_a(\bar{a})^2 + 2\theta_{ae}\bar{a}\bar{e} + \theta_e(\bar{e})^2)$.

Condition 2. $\gamma < \frac{\psi(\psi - \frac{1}{2}(\theta_a(\bar{a})^2 + 2\theta_{ae}\bar{a}\bar{e} + \theta_e(\bar{e})^2))}{(1-\tau)Cr\Sigma(\bar{a},\bar{e})}$.

Proposition 2. *Assume Conditions 1 and 2 hold, then the value function for problem (15)-(16) can be written as (17), where $f(\delta)$ is an almost-everywhere twice-continuously differentiable function with linear growth that satisfies the ODE (18)-(19) whenever $\delta \geq \delta_B$, and $f(\delta) = 0$ otherwise. Thus, the value function is given by*

$$P(\delta, W) = \begin{cases} f(\delta) - \frac{-1}{\gamma r}\ln(-\gamma r W) & \text{if } \delta \geq \delta_B, \\ -\frac{-1}{\gamma r}\ln(-\gamma r W) & \text{otherwise.} \end{cases} \quad (21)$$

4.1.1 Debt Valuation and Optimal Leverage

Recall that once debt is in place, shareholders select the optimal long-term contract. Bondholders anticipate the effect of debt on shareholders' future behavior. Hence, in pricing the perpetual debt contract, creditors take the optimal effort policy $(a^*(\delta), e^*(\delta))$ in (20) and the default thresholds δ_B as given. For any coupon C, the debt value $D(\delta)$ satisfies the following ODE:

$$rD(\delta) = C + (\phi + a^*(\delta))\,\delta D'(\delta) + \frac{1}{2}\sigma_\delta^2\delta^2 D''(\delta), \quad (22)$$

with boundary conditions $D(\delta) \to \frac{C}{r}$ as $\delta \to \infty$, and

$$D(\delta_B) = (1-\alpha)f_u(\delta_B), \quad (23)$$

where $f_u(\delta)$ is the unlevered firm value characterized in Section 4.2.[17] Observe that in equation (22), long-term effort directly affects the expected capital gains of the debt contract. In contrast, the short-term action affects cash flows and shareholders' ability to absorb losses. Thus, both efforts affect the default boundary and debt and equity values.

Given an initial firm size δ_0, initial shareholders choose coupon C to maximize levered firm value (ex-ante equity value) $TV(\delta_0; C) = f(\delta_0; C) + D(\delta_0, C)$ at time 0. That is, debt policy

[17] An alternative specification for the boundary condition at bankruptcy can be $D(\delta_B) = f_u((1-\alpha)\delta_B)$. Because $f_u(\delta)$ becomes asymptotically linear as $\delta \to \infty$, this alternative specification is equivalent to (23) for large values of δ. For small values of δ, we obtain that $(1-\alpha)f_u(\delta_B) \leq f_u((1-\alpha)\delta_B)$, implying effectively higher bankruptcy costs for smaller firms and encouraging lower leverage. We adopt the boundary condition in equation (23), because it helps our model capture the stylized fact that smaller firms have lower leverage, as documented in Frank and Goyal (2008).

maximizes equity value plus the proceeds from the debt issue. Then, they design the optimal long-term contract with the agent that implements the effort policy $(a^*(\delta), e^*(\delta))$, and they run the firm until they declare bankruptcy. We define the firm's optimal initial market leverage ratio as

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

4.1.2 Implementation

In our implementation, the agent maintains his certainty equivalent as the balance in his savings account $B_t = -\frac{1}{\gamma r}\ln(-\gamma r W_t)$. Under the optimal contract, the dynamics of the balance are given by

$$dB_t = \frac{1}{2}\gamma r\left((\beta_t^T)^2\delta_t^2\sigma_A^2 + (\beta_t^P)^2\sigma_\delta^2\delta_t^2 + 2\rho\beta_t^T\beta_t^P\delta_t^2\sigma_A\sigma_\delta\right)dt + \beta_t^T\delta_t\sigma_A dZ_t^T + \beta_t^P\delta_t\sigma_\delta dZ_t^P. \qquad (24)$$

In words, the agent's certainty equivalent is given by B_t at any point in time, and the shareholders adjust the balance continuously according to (24) in order to provide the appropriate incentives to the agent. The adjustment in (24) includes the drift term that reflects the incentive cost in (18), and the volatility terms that reflects the agent's exposure to both the transitory and permanent shocks.

In practice, the manager has a balance of B_t in his bank account, which allows for transfers between shareholders and the manager, according to (24). The manager's consumption (13) comes from two sources: $g(a_t, e_t; \delta_t)$ paid directly by shareholders and $rB_t = -\frac{1}{\gamma}\ln(-\gamma r W_t)$ paid from his savings account.[18] The market value of equity is $f(\delta) = P(\delta, W) + \frac{-1}{\gamma r}\ln(-\gamma r W_t)$. That is, market value of equity equals the principal's value $P(\delta, W)$ (if the agent had a continuation value of W and no private savings) plus the agent's private savings $\frac{-1}{\gamma r}\ln(-\gamma r W_t)$. If a new owner were to purchase the firm, they would pay $f(\delta)$, independently of the manager's current continuation value.

After default, bondholders take ownership of the firm and design a continuation contract for the manager that maximizes unlevered firm value. The continuation value of the manager is W_τ, which corresponds to a certainty equivalence of $B_\tau = \frac{-1}{\gamma r}\ln(-\gamma r W_\tau)$. Fortunately for the new owners (previous bondholders), the savings of the manager are exactly equal to B_τ. That is, from the perspective of the bondholders, it is "as if" the continuation contract had to deliver a certainty equivalence of zero dollars to the manager. As a result, debt value at default is independent of the manager's continuation value and matches our boundary condition $D(\delta_B) = (1-\alpha)f_u(\delta_B)$.[19]

[18] Theoretically, the CARA framework cannot rule out the possibility of $B_t \leq 0$. This case corresponds to the agent continuing to exert effort while drawing from a credit line from the bank. Such a situation is a consequence of our assumption that the agent has unlimited liability; it has been commonly assumed in the literature (e.g., He (2011) and Gryglewicz and Hartman-Glaser (2020)) in order to reduce the dimensionality of the problem.

[19] Under this implementation, we do not need to specify priorities at bankruptcy, since the deferred compensation

4.2 Optimal Contract in an Unlevered firm

We now characterize the optimal contract in an unlevered firm.[20] Because the unlevered firm is also run by a manager, the value of an unlevered firm $f_u(\delta)$ must satisfy (18) with $C = 0$. Thus, the unlevered firm loses the tax shield τC but does not default.

$$rf_u(\delta) = \max_{(a,e)} \left\{ \begin{array}{c} \delta(\psi + e) - \frac{1}{2}\delta(\theta_a a^2 + 2\theta_{ae} ae + \theta_e e^2) + (\phi + a)\delta f_u'(\delta) + \frac{1}{2}\sigma_\delta^2 \delta^2 f_u''(\delta) \\ -\frac{\gamma r \delta^2}{2} \left((\theta_a a + \theta_{ae} e)^2 \sigma_\delta^2 + (\theta_e e + \theta_{ae} a)^2 \sigma_A^2 + 2\rho (\theta_a a + \theta_{ae} e)(\theta_e e + \theta_{ae} a) \sigma_\delta \delta_A \right) \end{array} \right\}. \tag{25}$$

For an interior solution, we use the first-order conditions for the maximization of (25) to obtain[21]

$$a_u^*(\delta) = K_0 f_u'(\delta) + K_1 \text{ and } e_u^*(\delta) = K_1 f_u'(\delta) + K_2. \tag{26}$$

Using equation (26), one can show the marginal unlevered firm value $f_u'(\delta) \to \frac{\psi}{r-\phi}$ as $\delta \to \infty$. Together with the boundary condition $f_u(0) = 0$, the HJB-equation (25) can be solved numerically. For the special case where $\rho = 0$ and $\theta_{ae} = 0$, tasks are independent and the optimal efforts are

$$a_u^*(\delta) = \frac{f_u'(\delta)}{\theta_a \left(1 + \gamma r \theta_a \sigma_\delta^2 \delta\right)} \text{ and } e_u^*(\delta) = \frac{1}{\theta_e \left(1 + \gamma r \theta_e \sigma_A^2 \delta\right)}. \tag{27}$$

We can make two observations. First, since long-term effort $a_u^*(\delta)$ 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 $e_u^*(\delta)$ because it affects only the current profitability. Second, the optimal effort depends on the volatility: σ_δ affects $a_u^*(\delta)$ and σ_A affects $e_u^*(\delta)$. This is because the volatility of the processes A_t and δ_t measure how informative these processes are about e and a, respectively. The following Proposition formalizes the solution to the problem faced by the shareholders of an unlevered firm.

Proposition 3. *The value function for the shareholders of an unlevered firm (i.e., problem (15)-(16) with $C = 0$) is given by $P_u(\delta, W) = f_u(\delta) - \frac{-1}{\gamma r}\ln(-\gamma r W)$, where $f_u(\delta)$ is the unique twice-continuously differentiable solution with linear growth of the HJB (25).*

is held *outside* the firm. By contrast, in the direct interpretation of the optimal contract, in which the agent has no private savings and deferred compensation is held *inside* the firm, we would need to assume that the existing shareholders can fulfill their promise to the agent at bankruptcy: the deferred compensation balance B_t must be senior to the debt claim. In practice, the priority often depends on the details of compensation plans. "Secular Trust," as a funding vehicle for deferred compensations, is bankruptcy-proof. See footnote 10 of Sundaram and Yermack (2007), Gerakos (2010), and Section 4 of Edmans and Liu (2011) for details.

[20]In Appendix A.2.2, we characterize the first-best (no agency) value and effort policy of the unlevered firm.

[21]In Appendix A.2.1, we show how to deal with the cases in which the constraints on the effort policies bind.

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. It is well known that underinvestment results directly from debt. In contrast, short-termism is affected only by the presence of debt indirectly through the underinvestment problem. Moreover, the effect of debt on short-termism disappears when the costs for the shareholders of implementing a particular pair of efforts (a, e) are independent. Independence occurs when the cross term in the cost function is zero ($\theta_{ae} = 0$) and when the shocks are uncorrelated ($\rho = 0$). Proposition 4 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 avoiding underinvestment), this will suffice as a commitment device to avoid excessive short-termism (i.e., shareholders will indirectly be committing to the unlevered short-term-effort policy). Proposition 5 formalizes this result. Figure 1 summarizes the findings of this section.

Figure 1. Chain of Debt-Induced Distortions

Proposition 4. *Suppose that the cost of implementing long-term and short-term effort are independent from each other (i.e., $\theta_{ae} = 0$ and $\rho = 0$), 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 they do not fully internalize the benefit of this investment. The reason is that some of the cash flows generated from the 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.[22] 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

[22] 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 policy will have a permanent effect (long-term effort) or a transitory effect (short-term effort) on the firm's cash flows.

no distortion, and the optimal $e(\delta)$ is not affected by the presence of debt.

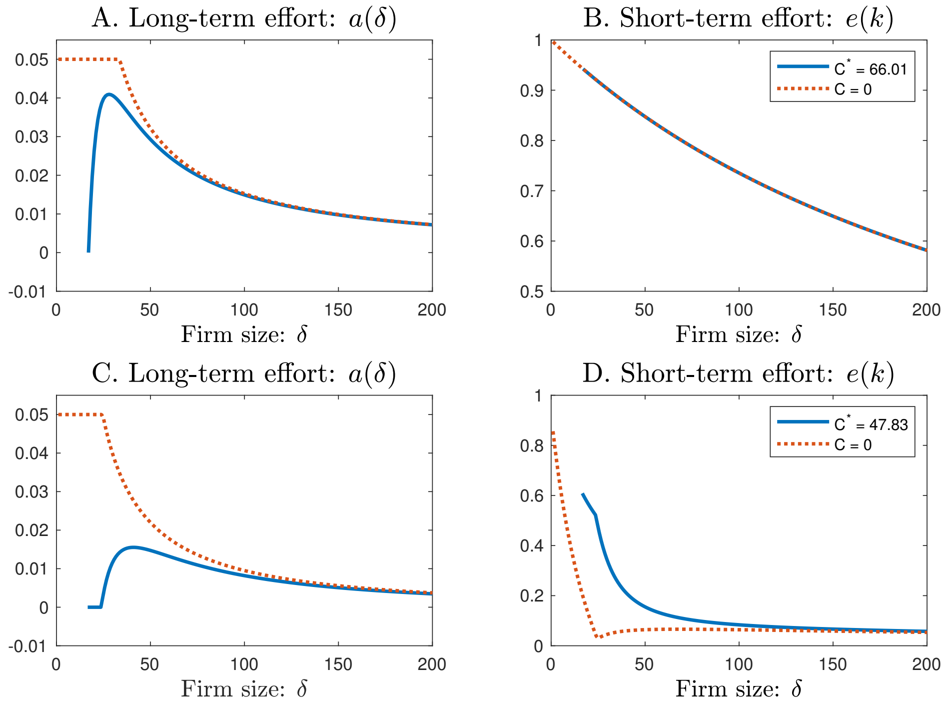

Figure 2. The Effect of Debt Overhang on the Optimal Effort

The dashed (solid) lines represent the optimal effort in the unlevered (optimally levered) firm. Actions are independent ($\theta_{ae} = 0$) in Panels A and B, and are substitutes ($\theta_{ae} = 1.5$) in Panels C and D. 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 $\psi = 1$.

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 a numerical illustration of Proposition 4 showing that short-term effort is identical in the levered and unlevered cases when 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 directly affected by debt overhang, since it increases contemporaneous cash flows but has no effect on future cash flows. Hence, shareholders find it optimal to incentivize the manager to focus on 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 on the short term when the probability is high that the company will not survive in the long term.[23] In the sequel, we refer to the *underinvestment problem* as the fact that debt induces shareholders to implement a lower long-term effort (i.e., $a(\delta) \leq a_u(\delta)$), and we refer to the *short-termism problem* as the fact that shareholders implement more short-term effort in the presence of debt (i.e., $e(\delta) \geq e_u(\delta)$).[24]

A fundamental reason for debt overhang is that shareholders lack commitment toward bondholders and choose an optimal policy that maximizes their value once debt is issued. We assume moral hazard in all cases, and we vary only the ability to commit to either short-term or long-term efforts, or to both efforts. To compare and contrast the base case without commitment to either effort described in Section 4.1, we consider the following three cases:

1. The no-debt-overhang case (denoted NO): In this case, shareholders commit to implementing the unlevered policies (a_u, e_u) for both short-term and long-term efforts. 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 $f_{NS}(\delta)$.[25]

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 $f_{NU}(\delta)$.

In terms of timing (instead of commitment), these cases can be understood as follows. The baseline case corresponds to the situation in which the incentives contract is signed *after* debt has been issued. By contrast, the NO case corresponds to the situation in which the incentives contract is signed *before* debt is issued. Finally, the NU (resp. NS) case captures a hybrid situation in which

[23]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 ρ is positive, the results are similar. 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., $\theta_{ae} > 0$).

[24]Gopalan et al. (2014) find that shorter-duration contracts are offered by firms with higher leverage. Similarly, Akins et al. (2018) document that firms tend to decrease pay horizon following a covenant violation. Both findings are consistent with our prediction that shareholders optimally implement higher short-termism during financial distress.

[25]In the NS case, shareholders commit to some amount of short-term effort, namely $e_u \geq 0$. This is different from assuming away short-term effort by setting $\theta_e = \infty$.

the portion of the contract pertinent to firm size (resp. cashflows) is signed *before* debt is issued, while the portion pertinent to cashflows (resp. firm size) is signed *after*. In other words, the levered firm holds one of the effort policies fixed and dynamically reoptimizes the other.

The rigorous statement of the problems that are solved by each of the cases above can be found in Appendix A.4.3. We can now state the second main result of this section.

Proposition 5. $f(\delta)$ *solves the no-underinvestment case (NU) if and only if it solves the no-overhang case (NO).*

Proposition 5 states that shareholder commitment to the unlevered long-term-effort policy automatically delivers commitment to the unlevered short-term-effort policy. 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 effort pair (a_u, e_u), committing to a_u would suffice to achieve this goal.[26] However, the converse is in general not true. That is, 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 underinvestment problem. When the costs of the actions are independent, the choice of short-term effort is unaffected by the presence of debt (Proposition 4). 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, preventing the distortion of debt on long-term effort would automatically eliminate the distortion on short-term effort (Proposition 5). Taken together, Propositions 4 and 5 uncover the asymmetric effect of debt on short-term vs. long-term effort, showing that excessive short-termism is an indirect agency cost of debt.

4.4 Comparative Statics

We now characterize the effect of parameters on the equity value $f(\delta)$ and the endogenous default threshold δ_B. Here, we focus on the ex-post value and default decision. That is, we treat the coupon as an exogenous parameter rather than accounting for the ex-ante capital-structure choice.

[26] This is useful, because committing to the unlevered firm policies is value-enhancing (see Section 5.1).

	Results independent of ρ					Only when $\rho \geq 0$					
	$\partial\psi$	$\partial\phi$	$\partial\rho$	∂C	$\partial\tau$	$\partial\gamma$	$\partial\theta_a$	$\partial\theta_e$	$\partial\theta_{ae}$	$\partial\sigma_A$	$\partial\sigma_\delta$
$\partial f(\delta)/$	+	+	−	−	+	−	−	−	−	−	?
$\partial\delta_B/$	−	−	+	+	−	+	+	+	+	+	?

Table 1. Analytical Comparative Statics

First, risk parameters σ_δ and σ_A differ in their effects on equity value and default threshold. An increase in the magnitude of the permanent shocks σ_δ makes firm size δ a noisier signal of the long-term effort a, thus increasing the incentive cost and accelerating default. However, there is an opposing effect: a higher volatility σ_δ also generates a real-option effect that delays the equity's default decision. Specifically, the terms involving σ_δ on the HJB-equation (18) (with $\rho = 0$) are

$$\frac{1}{2}\sigma_\delta^2\delta^2 f''(\delta) - \frac{1}{2}\gamma r\delta^2\left((\theta_a a + \theta_{ae} e)^2\sigma_\delta^2\right).$$

While the second term captures the incentive-cost effect, the first term captures the real-option effect, and it is increasing in σ_δ under the convexity of the shareholders' value.[27] The two countervailing forces generate an inverted U-shape for the default threshold as a function of σ_δ. Numerically, this result is illustrated on Panel A of Figure 3.

In contrast, an increase in the volatility of transitory shocks σ_A has no real-option effect. However, in general, the incentive cost of higher σ_A is ambiguous. When the correlation is negative, an increase in σ_A increases the cost of incentivizing short-term effort e but decreases the overall risk the agent is exposed to (via diversification). As a result, the overall effect on incentive costs is ambiguous. When $\rho \geq 0$, there is no diversification benefit and an increase in σ_A unambiguously reduces shareholder value. Panel B of Figure 3 illustrates these comparative statics with $\rho = 0$.

Second, when $\rho \geq 0$, an increase in θ_a (θ_e) increases the exposure of the agent to the permanent (temporary) shock $dZ^P(dZ^T)$. This leads to a higher incentive cost and lower equity value, which, in turn, increases the default threshold. Panels C and D of Figure 3 show that the default threshold is increasing in both short-term-effort and long-term-effort costs when $\rho = 0$. However, similar to σ_A, this result holds only when $\rho \geq 0$, because the tasks tend to be substitutes. If $\rho < 0$, the diversification benefit from exposure to negatively correlated shocks can lower the agent's *total* risk-exposure

[27] $f''(\delta) > 0$ in our numerical simulations. For the incentive-cost effect, see Holmstrom (1979) and a related analysis by Chaigneau et al. (2018) in a setting with limited liability.

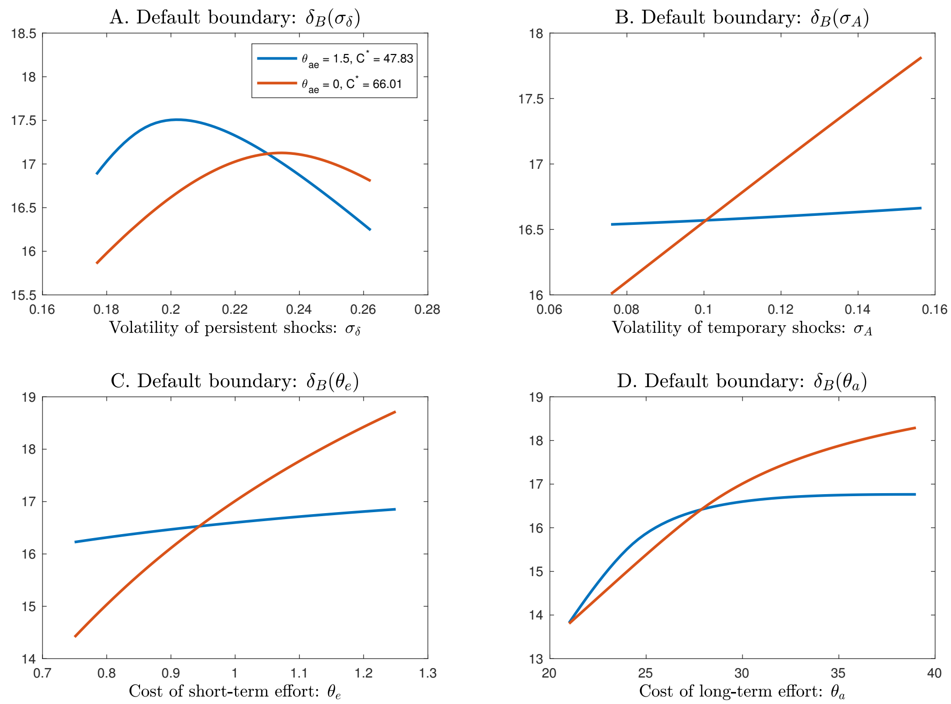

Figure 3. Comparative Statics for Default Boundary

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

and hence overall incentive cost. As a result, the effect on equity value is ambiguous when $\rho < 0$.

Finally, we explore the role of the correlation between the two types of risks for default decisions. Panel A of Figure 4 shows that the default boundary is increasing in the correlation between the transitory shock and the permanent shock. Higher correlation increases the risk compensation for a given policy (a, e), inducing shareholders to default earlier. Panels B and C depict effort levels for three different values of ρ.

Interestingly, the comparative statics of θ_{ae} are similar to the ones of ρ. Panels D, E, and F chart comparative statics for the default boundary and the policies with respect to the substitutability/complementarity of efforts. As θ_{ae} increases, the two tasks become substitutes, and it becomes more costly to incentivize both efforts together: the marginal cost of increasing e for a given value of a is increasing in θ_{ae}. Hence, a higher substitutability is associated with a higher incentive cost, which accelerates default.

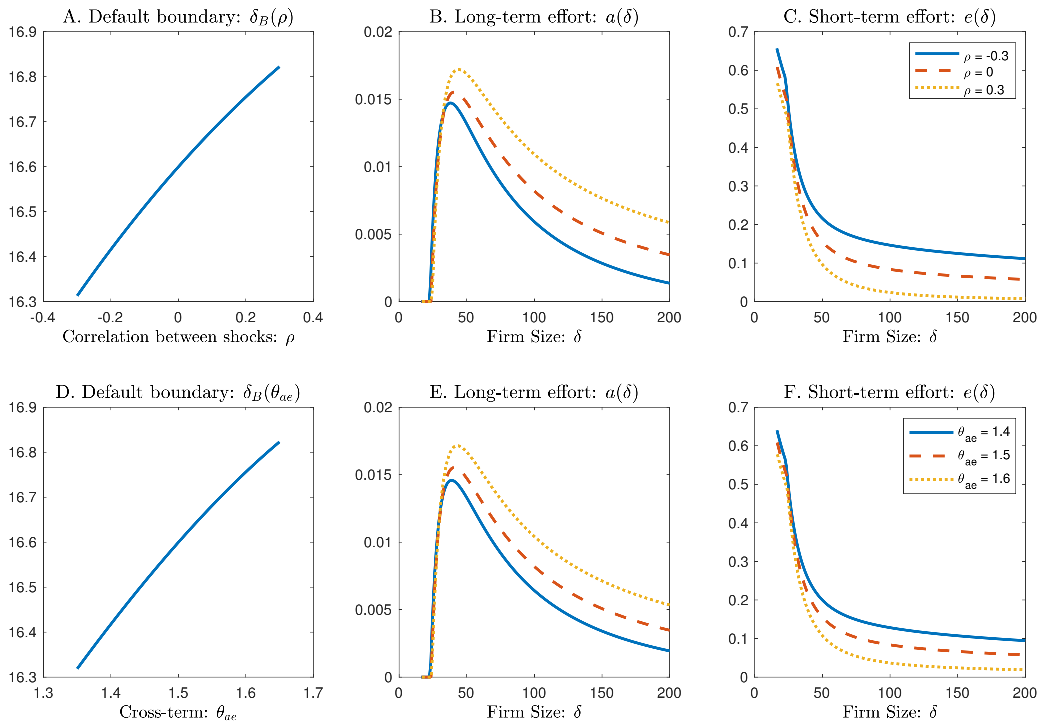

Figure 4. Effects of Correlation and Substitutability

We assume $C = 47.83$, $\delta_0 = 100$, $r = 0.05$, $\theta_a = 30$, $\theta_e = 1$, $\theta_{ae} = 1.5$, $\gamma = 5$, $\sigma_\delta = 0.25$, $\sigma_A = 0.12$, $\tau = 0.15$, $\alpha = 0.30$, $\phi = -0.005$, $\psi = 1$, ρ is either -0.3, 0, or 0.3, and θ_{ae} is either 1.4, 1.5, or 1.6.

5 Quantitative Analysis

This section develops model implications that are novel relative to, e.g., Leland (1994) or He (2011). First, we compute in our calibrated model the various sources of firm value, and we 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 a 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 our quantitative solutions, we adopt standard baseline parameter values: interest rate

$r = 5\%$, baseline growth rate of firm size $\phi = -0.5\%$, correlation $\rho = 0$, volatility of firm size $\sigma_\delta = 25\%$, volatility of 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 A.6, 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, for example, $\delta_0 = 100$, 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,

$$TB(\delta) = \mathbb{E}_t^{(a^*,e^*)}\left[\int_t^\chi e^{-r(s-t)}\tau C ds\right] \text{ and } BC(\delta) = \mathbb{E}_t^{(a^*,e^*)}\left[e^{-r(\chi-t)}\alpha f_u(\delta_B)\right], \tag{28}$$

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,

$$DC(\delta) = \mathbb{E}_t^{(a^*,e^*)}\left[\int_t^\chi e^{-r(s-t)} g(a_s, e_s; \delta_s) ds\right], \tag{29}$$

$$IC(\delta) = \mathbb{E}_t^{(a^*,e^*)}\left[\int_t^\chi e^{-r(s-t)}\frac{1}{2}\gamma r\left((\beta_s^T)^2\sigma_A^2\delta_s^2 + (\beta_s^P)^2\sigma(\delta_s)^2 + 2\rho\beta_s^P\beta_s^T\delta_s\sigma_A\sigma(\delta_s)\right)ds\right], \tag{30}$$

and

$$CC(\delta) = DC(\delta) + IC(\delta). \tag{31}$$

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

5.1.1 Quantifying the Cost of Debt Overhang

Next, we study the cost of debt overhang (i.e., the cost of underinvestment plus the cost of short-termism for firm value). When shareholders cannot commit to an effort policy, debt distorts the effort choice. To capture the cost, we adopt the no-overhang case (NO) in which shareholders commit to use the optimal efforts a_u and e_u without leverage, as described in Section 4.3. We denote the resulting equity, debt, total firm, and leverage values, respectively, by $f_{NO}(\delta)$, $D_{NO}(\delta)$, $TV_{NO}(\delta_0; C)$, and $ML_{NO}(\delta_0)$ (see Appendix A.4.3 for details).[28]

[28]The no-overhang case involves optimal unlevered long-term-effort and short-term-effort policies that do not maximize levered equity value. We emphasize that throughout this section we keep all the parameters governing the

Panel A: Optimal Leverage									
δ	C^*	$\delta_0/(r-\phi)$	$TV(\delta)$	$f(\delta)$	$D(\delta)$	$BC(\delta)$	$TB(\delta)$	$CC(\delta)$	δ_B
75	33.38	1363.63	1,654.88	1,043.06	611.82	14.93	86.53	189.44	10.89
100	47.83	1818.18	2,124.98	1,287.24	837.74	28.65	115.60	173.56	16.59
150	81.07	2727.27	3,070.32	1,732.44	1,337.88	58.29	180.22	160.98	29.81
Panel B: Unlevered Case									
δ	C	$\delta_0/(r-\phi)$	$TV(\delta)$	$f(\delta)$	$D(\delta)$	$BC(\delta)$	$TB(\delta)$	$CC(\delta)$	δ_B
75	0	1363.63	1,587.70	1,587.70	0	0	0	224.46	0
100	0	1818.18	2045.68	2045.68	0	0	0	228.06	0
150	0	2727.27	2,959.19	2,959.19	0	0	0	232.32	0

Table 2. Sources of Firm Value in the Base Case
This table shows the various sources of firm value for the base case (i.e., the case without commitment, where policies are optimally chosen by shareholders after debt has been issued) for three different firm sizes δ. 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 $\psi = 1$.

Table 3 displays firm values for three different initial firm sizes, $\delta_0 = 75, 100, 150$. Column 1 provides quantities when effort policies are optimally chosen by shareholders after debt has been issued (normalized relative to the unlevered unmanaged total firm value $\delta_0/(r-\phi) \equiv 100$ in parentheses). Column 2 gathers 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. Outputs in column 3 are for the base case but use the optimal coupon C^*_{NO} calculated under the no-overhang policies. Finally, column 4 reports the percentage change between columns 3 and 2.

A number of new results relative to, e.g., He (2011), emerge from Table 3. First, no-overhang effort policies increase 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, no-overhang effort policies reduce shareholder value by 0.53% (small firms) to 0.80% (large firms), because shareholders cannot implement their ex-post optimal policies. For any δ level, equity value is reduced (i.e., $f(\delta) > f_{NO}(\delta)$) and default hastens (i.e., $\delta_B < \delta_B^{NO}$). Intuitively,

managerial moral hazard problem constant, as to avoid conflating the effects of debt overhang and moral hazard on firm value.

	Base Case	No Overhang (NO)	Base Case with C^*_{NO}	Change
δ_0	$\delta_0 = 150$	$\delta_0 = 150$	$\delta_0 = 150$	
$\delta_0/(r-\phi)$	2727.27 (100)			
$f_u(\delta_0)$	2959.19 (108.50)			
$TV_{NO}(\delta_0)$	3070.32 (112.57)	3083.74 (113.07)	3070.25 (112.57)	0.43% (0.50)
$f_{NO}(\delta_0)$	1732.44	1693.48	1707.22	-0.80%
$D_{NO}(\delta_0)$	1337.88	1390.25	1363.02	1.99%
C^*_{NO}	81.07	83.08	83.08	
δ_B^{NO}	29.81	33.79	30.62	
$ML_{NO}(\delta_0)$	43.57	45.08	44.39	
$CS_{NO}(\delta_0)$	105.95	97.58	109.52	

	Base Case	No Overhang (NO)	Base Case with C^*_{NO}	Change
δ_0	100	100	100	
$\delta_0/(r-\phi)$	1818.18 (100)			
$f_u(\delta_0)$	2045.68 (112.51)			
$TV_{NO}(\delta_0)$	2124.98 (116.87)	2136.50 (117.51)	2124.70 (116.85)	0.55% (0.66)
$f_{NO}(\delta_0)$	1287.24	1236.50	1244.56	-0.64%
$D_{NO}(\delta_0)$	837.74	900.38	880.14	2.29%
C^*_{NO}	47.83	50.98	50.98	
δ_B^{NO}	16.59	20.20	17.84	
$ML_{NO}(\delta_0)$	39.42	42.13	41.42	
$CS_{NO}(\delta_0)$	70.93	66.22	79.22	

	Base Case	No Overhang (NO)	Base Case with C^*_{NO}	Change
δ_0	75	75	75	
$\delta_0/(r-\phi)$	1363.63 (100)			
$f_u(\delta_0)$	1587.70 (116.43)			
$TV_{NO}(\delta_0)$	1654.89 (121.35)	1665.98 (122.17)	1653.46 (121.25)	0.75% (0.92)
$f_{NO}(\delta_0)$	1043.06	943.30	948.36	-0.53%
$D_{NO}(\delta_0)$	611.82	722.67	705.10	2.49%
C^*_{NO}	33.38	40.04	40.04	
δ_B^{NO}	10.89	14.92	13.53	
$ML_{NO}(\delta_0)$	36.97	43.37	42.64	
$CS_{NO}(\delta_0)$	45.57	54.05	67.85	

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 $\delta_0/(r-\phi) \equiv 100$. 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 $\psi = 1$.

commitment prevents shareholders from cutting long-term and boosting short-term investments in financial distress. So, it is more costly to run a distressed firm, and earlier default is optimal.

Third, the gain in debt value from no-overhang policies ranges from 1.99% (large firms) to 2.49% (small firms). Shareholders default earlier, which increases recoveries, but also commit to more long-termism, which makes default less likely. Hence, debt value rises (i.e., $D_{NO}(\delta) > D(\delta)$). Note that the gain in debt value dominates the reduction in equity value and, as discussed above, total firm value goes up.

Finally, the effect of debt-overhang in terms of short-termism and lack of long-termism (i.e., underinvestment) is stronger for smaller firms. To mitigate the effect of debt overhang, optimal leverage is lower for small firms (column 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. However, shareholders cannot commit to the long-term-effort policy, which must be chosen optimally after the debt is in place. This case corresponds to the no-short-termism case (NS) with effort e_u described in Section 4.3.[29] 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).

[29] This case is, in spirit, closer to He (2011) in that only long-term effort is endogenous, but it features multiple efforts and associated compensation costs. Recall that the no-underinvestment (NU) case described in Section 4.3 assumes commitment to a_u but immediately implies e_u, so it is spanned by the no-overhang case (NO) in Section 5.1.1.

	Base Case	No Short-Termism (NS)	Base Case with C^*_{NS}	Change
δ_0	150	150	150	
$\delta_0/(r-\phi)$	2727.27 (100)			
$f_u(\delta_0)$	2959.19 (108.50)			
$TV_{NS}(\delta_0)$	3070.32 (112.57)	3078.63 (112.88)	3070.31 (112.57)	0.27% (0.31)
$f_{NS}(\delta_0)$	1732.44	1716.89	1725.51	–0.49%
$D_{NS}(\delta_0)$	1337.88	1361.74	1344.80	1.26%
C^*_{NS}	81.07	81.62	81.62	
δ_B^{NS}	29.81	31.60	30.03	
$ML_{NS}(\delta_0)$	43.57	44.23	43.80	
$CS_{NS}(\delta_0)$	105.95	99.41	106.92	

	Base Case	No Short-Termism (NS)	Base Case with C^*_{NS}	Change
δ_0	100	100	100	
$\delta_0/(r-\phi)$	1818.18 (100)			
$f_u(\delta_0)$	2045.68 (112.51)			
$TV_{NS}(\delta_0)$	2124.98 (116.87)	2132.24 (117.27)	2124.64 (116.85)	0.35% (0.42)
$f_{NS}(\delta_0)$	1287.24	1235.31	1240.01	–0.37%
$D_{NS}(\delta_0)$	837.74	896.92	884.63	1.38%
C^*_{NS}	47.83	51.32	51.32	
δ_B^{NS}	16.59	19.00	17.97	
$ML_{NS}(\delta_0)$	39.42	42.06	41.63	
$CS_{NS}(\delta_0)$	70.93	72.24	80.12	

	Base Case	No Short-Termism (NS)	Base Case with C^*_{NS}	Change
δ_0	75	75	75	
$\delta_0/(r-\phi)$	1363.63 (100)			
$f_u(\delta_0)$	1587.70 (116.43)			
$TV_{NS}(\delta_0)$	1654.89 (121.35)	1660.15 (121.74)	1654.12 (121.30)	0.36% (0.44)
$f_{NS}(\delta_0)$	1043.06	969.44	971.67	–0.23%
$D_{NS}(\delta_0)$	611.82	690.71	682.45	1.21%
C^*_{NS}	33.38	38.36	38.36	
δ_B^{NS}	10.89	13.30	12.86	
$ML_{NS}(\delta_0)$	36.97	41.60	41.25	
$CS_{NS}(\delta_0)$	45.57	55.45	62.08	

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 $\delta_0/(r-\phi) \equiv 100$. 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 $\psi = 1$.

In sum, commitment to the unlevered short-term policy increases firm value and debt value, but (and as expected) decreases equity value, which is why in reality we may not observe commitment. 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 ex-post harmful implications and, in particular, the expected net result for firm value.

5.1.3 Combined Cost of Debt Overhang and Short-Termism

In this section, we analyze the economic magnitude of the total agency cost. Our calibrated model implies a 1% loss of firm value, of which about half is due to excessive short-termism. As a percentage of debt value, however, the loss is nearly 2%, of which over 1% comes from excessive short-termism.[30]

Consistent with our results, the analysis of Hackbarth and Mauer (2012) indicates that agency costs between 0.5% and 1.0% of firm value are attributable to an *equilibrium feedback effect*. That is, once the firm's investment and financing decisions endogenously reflect bankruptcy costs, interest tax shields, and moral hazard, the residual (i.e., equilibrium) agency conflict will be small, because it has been *optimally minimized* by its equilibrium investment and financing decisions.

We illustrate this result graphically in Figure 5 by considering nonoptimal financial leverage levels of +10% and +20% deviations from the firm's equilibrium leverage of 40% in our base case. Panel A depicts the agency cost of debt (AC) as a function of firm size δ. The blue line corresponds to the total agency cost of debt (i.e., underinvestment plus excessive short-termism), while the red line correspond to the agency cost due to excessive short-termism. Panel B performs a similar exercise, but using debt value as the deflator. For each panel, the three dotted lines depict δ values of 75, 55, and 42, which correspond to leverage levels of 40%, 50%, and 60%, respectively. That is, we depict the agency costs implied at the optimal leverage for an initial firm size of 75, and we plot how such costs fluctuate throughout the lifetime of the firm. For moderately overlevered firms, the agency costs of short-termism and underinvestment range from 0.8%–1.6% and 1.2%–2.6%,

[30] Similarly, Mello and Parsons (1992) note that the magnitude of the agency cost of debt varies with its deflator (e.g., debt or equity value, first- or second-best firm value, levered or unlevered firm value); in their study, the agency cost is 0.8% of firm value but 4.3% of debt value: "From Table IV we read that the agency costs of this quantity of debt are $0.22 million, or eight-tenths of a percent of firm value. In terms of the amount of debt sold, however, these agency costs are close to 4.3%, a very large value. This should be compared to other costs such as underwriting fees and administrative expenses which are usually 1.3% of the value of a debt offering according to Mikkelson and Partch (1986)."

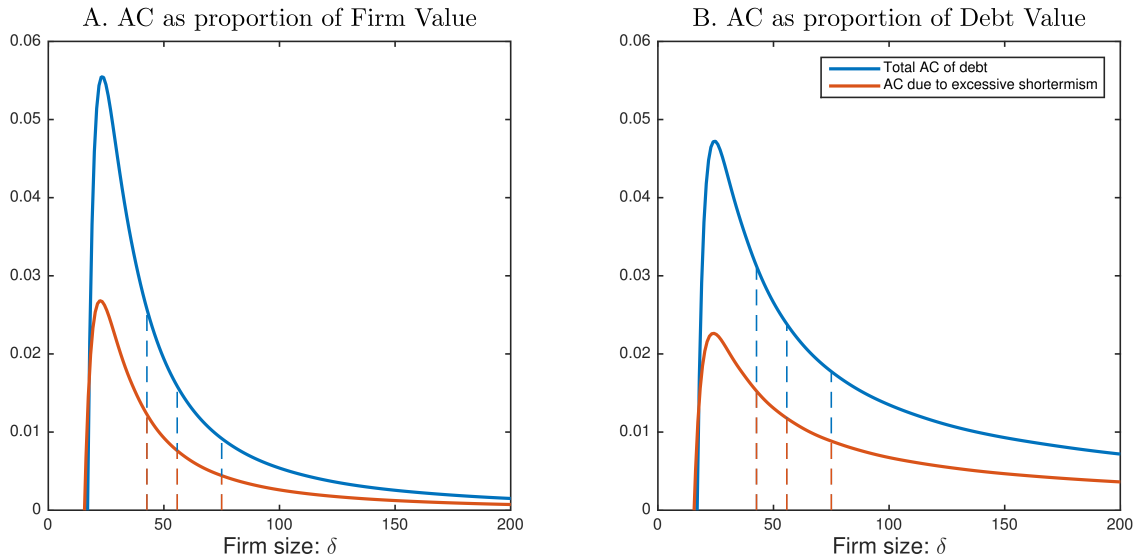

Figure 5. Agency Cost of Debt

We assume $r = 0.05$, $\theta_a = 30$, $\theta_e = 1$, $\theta_{ae} = 1.5$, $\gamma = 5$, $\sigma_\delta = 0.25$, $\sigma_A = 0.12$, $\rho = 0$, $\tau = 0.15$, $\alpha = 0.30$, and $\psi = 1$. The dashed lines at 75, 55, and 42 represent 40%, 50%, and 60% leverage.

respectively. Observe, however, these estimates are nearly twice as large for highly levered firms.

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 ϕ and interpret a large (low) value of ϕ as having bright (grim) prospects, because this firm is expected to experience fast (slow) growth in the future.

Figure 6 shows that firms with bright prospects focus a lot more on the long term and put little effort into the short term. Conversely, firms with grim prospects optimally focus on the short term and put little effort into the long term.[31] 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 ϕ) optimally focus on the long term. Consistent with Summers (2017), success is thus not only the result but also the cause of long-term investment (less short-termism).

[31] Consider the example of IBM. IBM's sales in 2017 were at the same level as they were in 1997. Over this period, it reduced costs and cut investment by half. Our model suggests that such decisions are not suboptimal and myopic. Rather, they are optimal responses to the new generation of technology firms that have taken over the industry.

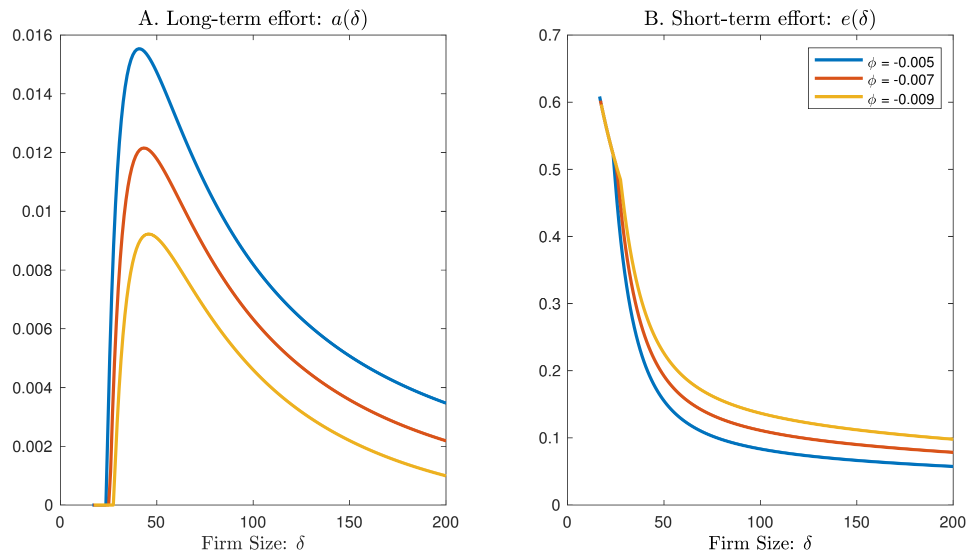

Figure 6. Effect of Growth Rate

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 $\psi = 1$.

5.3 Impact of Investor Horizon on Optimal Policies

In this section, we study the impact of investors' horizons 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 maintain the assumption that the manager's discount rate is r, but we allow the shareholders to have discount rate $r^S = r + \lambda$. As λ increases, shareholders become more impatient, which we interpret as having a shorter investment horizon.[32]

Figure 7 depicts the optimal investment policies for three different values of r^S. Panel A shows that, as the 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 δ. 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),

[32]While investor discount rates are not directly observable, it is conceivable that heterogeneity among investors' fee structure reflect heterogeneity in discount rates.

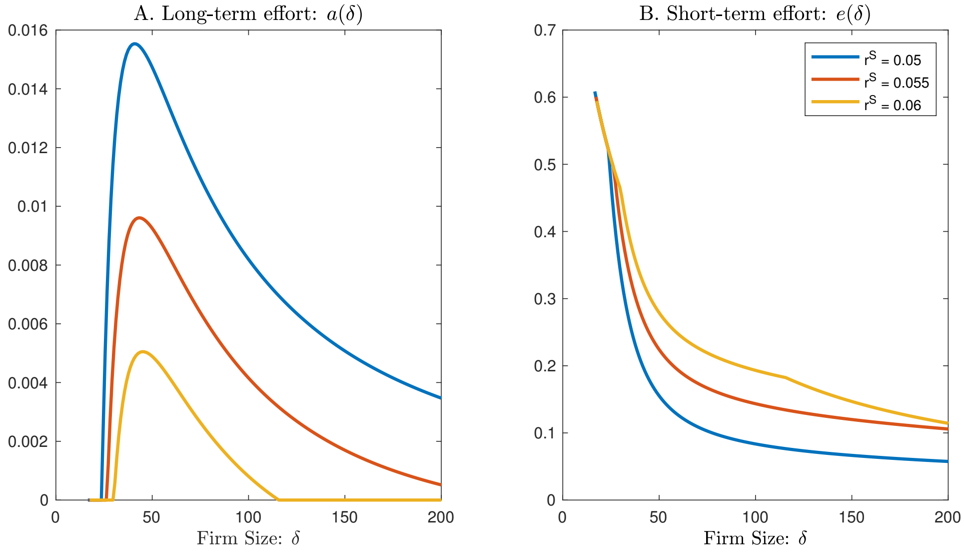

Figure 7. Effect of Investor Horizon

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.

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.

Panel A: $\delta_0 = 75, C^* = 33.38$						
r^S	$TV(\delta)$	% change $TV(\delta)$	$f(\delta)$	% change $f(\delta)$	$D(\delta)$	% 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	-1.25	1,283.03	-0.32	815.23	-2.68
6%	2,072.73	-2.45	1,275.90	-0.88	796.82	-4.88

Table 5. Firm Value for Different Investor Horizons

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

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 7. That is, we compute the effort policies that shareholders with discount rates

$r^S = 5\%$, 5.5%, and 6% would deem optimal, then we 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.[33]

Table 5 shows that implementing policies dictated by investors with shorter investment horizons reduces both equity and debt prices. It is intuitive that more underinvestment 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.[34]

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 8 depict comparative statics of equity value and credit spreads for different values of σ_δ. Equity value can be increasing in 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 our parameterizations, the real-option effect dominates and shareholders benefit from higher permanent risk σ_δ. Intuitively, shareholders face an asymmetric payoff from higher volatility. On the upside, they can benefit if higher σ_δ leads to a large value of δ, but they can exercise their option to default if it leads to a low value of δ. The higher probability of default negatively impacts debt value and leads to higher credit spreads on debt. These conflicting views between shareholders and bondholders with regards to the desirability of increasing risk are known in the literature as asset substitution or the risk-shifting problem.

[33] 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. We interpret the discount rate wedge $r^S - r > 0$ as a reduced-form way of capturing managerial moral hazard in the delegated wealth management industry, and manifesting itself in the form of suboptimally high discount rates. Moreover, our result that giving control of corporate policies to short-termistic investors leads to additional and significant reductions in value is consistent with the agenda of FCLT Global.

[34] 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 (2018)). However, investors with shorter time horizons may, in practice, be precisely those who are better at identifying firms with higher baseline growth rates ϕ that we analyzed in the previous section (see, e.g., Edmans (2017)).

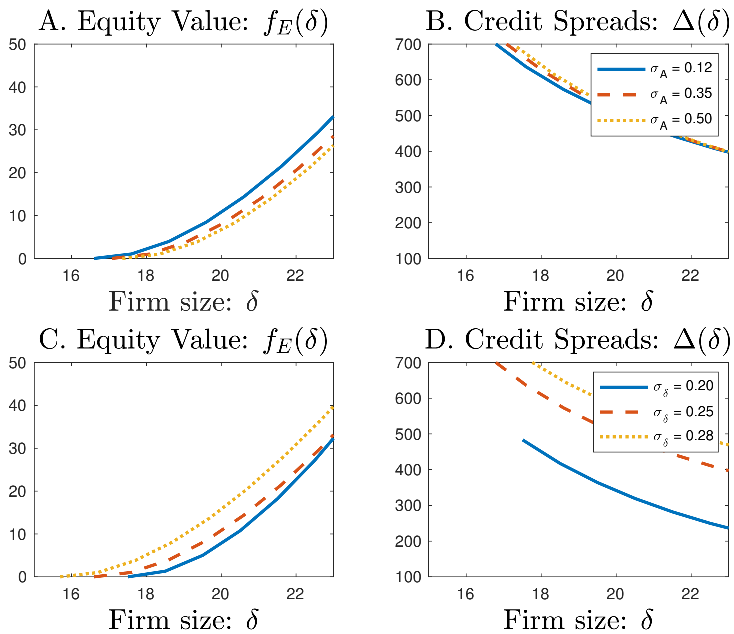

Figure 8. Comparative Statics of Equity Value and Credit Spreads

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

In contrast, transitory risk σ_A does not yield such an effect. Panels C and D of Figure 8 depict comparative statics of shareholder value and credit spreads for different values of σ_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 σ_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).[35]

Our results rationalize the finding that managers do not rank concerns about risk-shifting among 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

[35] See Ho-Eom et al. (2004) for evidence on credit spreads observed in reality and performance of various structural models in predicting actual spreads.

thus bondholders would know that it is also in shareholders' interest to minimize it.

6 Extensions and Alternative Setups

6.1 Finite Maturity Debt

Following Leland (1998), we extend our model to consider finite maturity debt in a stationary environment. The firm has debt with constant principal P and pays a constant coupon C at each moment in time. It rolls over a fraction m of its total debt. Therefore, any finite-maturity debt policy is fully characterized by (C, m, P). In the absence of bankruptcy, the average maturity equals $1/m$.

Figure 9 illustrates the optimal amount of long-term effort $a(\delta_t)$ and short-term effort $e(\delta_t)$ for a firm financed with infinite maturity debt (blue line), a firm financed with finite maturity debt (yellow line), and an all-equity-financed firm (dashed red line). Shorter debt maturity features two opposing effects. First, it induces more rollover payments that increase the endogenous default boundary, which directly exacerbate underinvestment and short-termism. Second, the repricing of newer debt encourages shareholders to increase firm size, hence mitigating underinvestment.[36] In our model, the first effect dominates for low values of δ, while the second effect dominates for higher values. Importantly, our key finding that excessive short-termism is an indirect agency cost of debt is qualitatively robust to the introduction of finite maturity debt, and is not specific to our baseline specification of infinite maturity debt without rollover payments (i.e., $m = 0$).[37]

6.2 Shareholder Limited Commitment

If shareholders can commit to a long-term contract with the manager, it is without loss of generality that one can search for an optimal contract within the space of incentive-compatible contracts that induce no private savings. However, if shareholders have limited commitment (i.e., if they renege on their promises to the manager when his future wages are large), a limited commitment constraint would have to be imposed on the optimal contract:

$$B_t \leq \bar{B}(\delta) \tag{32}$$

[36]See Diamond and He (2014) for a detailed discussion on the way the repricing mechanism alleviates the underinvestment problem.

[37]Our results are also robust to the introduction of performance-sensitive debt (PSD) a la Manso et al. (2010); they are available upon request.

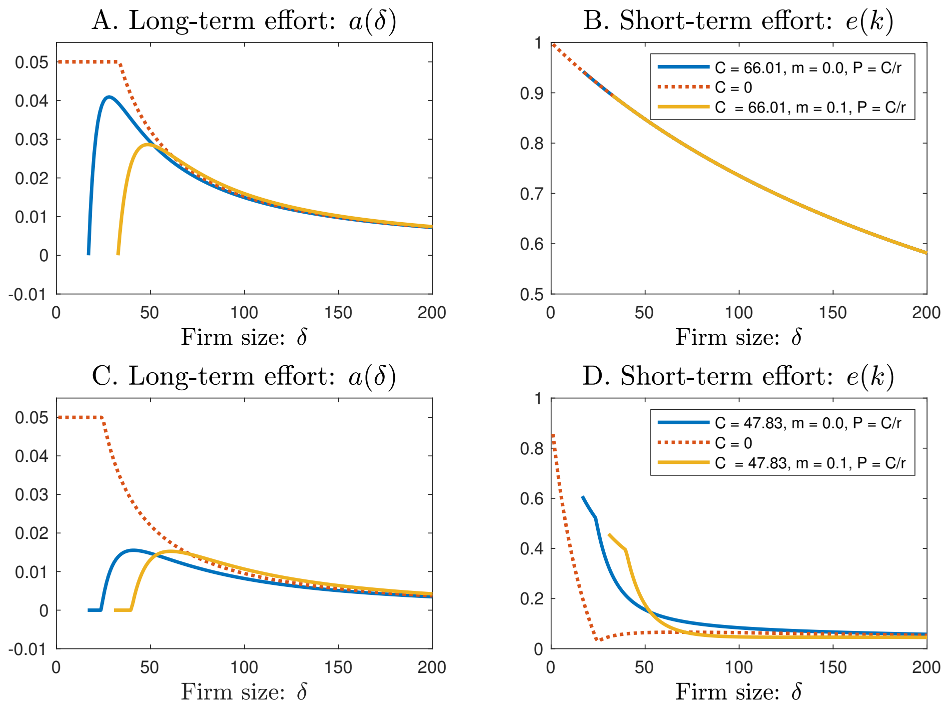

Figure 9. The Effect of Finite-Maturity Debt on the Optimal Effort

The red dashed, solid blue, and solid yellow lines, respectively, represent the unlevered, optimally levered with infinite maturity, and finite maturity cases. Actions are independent $\theta_{ae} = 0$ in Panels A and B, and are substitutes $\theta_{ae} = 1.5$ in Panels C and D. 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 $\psi = 1$.

Our implementation of the optimal contract does not require commitment from shareholders, because our optimal contract allows the manager to have private savings. His private savings are in fact equal to his certainty equivalent B_t. This contract is equivalent to the optimal contract, since it induces the same consumption on the manager and the same incentives. But from the perspective of the shareholders, they do not require commitment to deliver a high continuation value to the manager after good performance. The compensation after good performance has already been transferred to the manager's bank account, and shareholders cannot access it, since two-way transfers between the manager's savings account are governed by the rule specified in (24) that depends on the history of output and firm size.

An important question is what would happen if shareholders could not commit to the manager and still had to implement a no-savings contract? A full treatment of this question is beyond the scope of this paper. Intuitively, shareholders would reduce incentives after good performance. After good performance, B_t grows and the limited commitment constraint (32) tightens. Because the drift of B_t is increasing in the incentives to exert effort (equation (24)), shareholders would optimally decrease the drift of B_t by implementing lower effort levels.[38]

7 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 is 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. Put differently, there is a certain lack of long-termism that is nevertheless optimal for shareholders.

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 suboptimally short horizons focus excessively on short-term profitability at the expense of long-term growth. Our results indicate that such distortions significantly reduce firm value. Thus, designing capital-gains-tax policies aimed at lengthening investors' horizons can lead to an improvement in social welfare.

Finally, our analysis suggests that debt covenants that restrict the long-term-effort policy of the firms are more effective at maximizing firm value than covenants that restrict short-termism. Also, compensation contracts with inside debt can serve as a commitment device for shareholders to incentivize long-term effort during financial distress. We leave these questions for future research.

[38] See Ai and Li (2017) and Bolton et al. (2018) for models of two-sided limited commitment. They explore the implications for investment and risk-management when the agent *and* the principal have limited commitment.

References

Asker, John, Joan Farre-Mensa, and Alexander Ljungqvist, 2015, Corporate Investment and Stock Market Listing: A Puzzle?, *Review of Financial Studies* 28(2), 342-390.

Akins, Brian, Jonathan Bitting, David De Angelis, and Maclean Gaulin, 2018, The Salience of Creditors' Interests and CEO Compensation, Working Paper.

Auriol, Emmanuelle, Guido Friebel, and Lambros Pechlivanos, 2002, Career Concerns in Teams, *Journal of Labor Economics* 20(2), 289-307.

Bernardo, Antonio, Hongbin Cai, and Jiang Luo, 2004, Capital Budgeting in Multidivision Firms: Information, Agency, and Incentives, *Review of Financial Studies* 17(3), 739-767.

Biais, Bruno, Thomas Mariotti, G. Plantin,and J.C. Rochet, 2007, Dynamic security design: Convergence to continuous time and asset pricing implications, *Review of Economic Studies* 74(2), 345-390.

Biais, Bruno, Thomas Mariotti, Jean-Charles Rochet, and Stéphane Villeneuve, 2010, Large risks, limited liability, and dynamic moral hazard, *Econometrica* 78(1), 73-118.

Bolton, Patrick, Hui Chen, and Neng Wang, 2011, A Unified Theory of Tobin's Q, Corporate Investment, Financing, and Risk Management, *Journal of Finance* 66(5), 1545-1578.

Bolton, Partick, José Scheinkman, and Wei Xiong, 2006, Executive Compensation and Short-Termist Behaviour in Speculative Markets, *Review of Economic Studies* 73(3), 577-610.

Byun, Seong K., Valery Polkovnichenko, and Michael J. Rebello, 2017, Dynamics of Firm Savings and Investment with Temporary and Persistent Shocks, Working paper.

Chaigneau, Pierre, Alex Edmans, and Daniel Gottlieb, Does Improved Information Improve Incentives, 2018, *Journal of Financial Economics* 130(2), 291-307.

Décamps, Jean Paul, Sebastian Gryglewicz, Erwan Morellec, and Stephane Villeneuve, 2017, Corporate Policies with Permanent and Transitory Shocks, *Review of Financial Studies* 30(1), 162-210.

Décamps, Jean-Paul, and Stéphane Villeneuve, 2019, A two-dimensional control problem arising from dynamic contracting theory, *Finance and Stochastics* 23(1), 1-28.

Décamps, Jean Paul, Thomas Mariotti, Jean-Charles Rochet, and Stephane Villeneuve, 2011, Free Cash Flow, Issuance Costs, and Stock Prices, *Journal of Finance* 66(5), 1501-1544.

DeMarzo, Peter, and Yuliy Sannikov, 2006, Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Model, *Journal of Finance* 61(6), 2681-2724.

DeMarzo, Peter M., Dmitry Livdan, and Alexei Tchistyi, 2014, Risking other people's money: Gambling, limited liability, and optimal incentives, working paper.

DeMarzo, Peter, Michael Fishman, Zhiguo He, and Neng Wang, 2012, Dynamic Agency and the Q Theory of Investment, *Journal of Finance* 67(6), 2295-2340.

Diamond, Douglas, and Zhiguo He, 2014, A Theory of Debt Maturity: the Long and Short of Debt Overhang, *The Journal of Finance* 69(2), 719-762.

Edmans, Alex, and Qi Liu, 2011, Inside Debt, *Review of Finance* 15(1), 75-102.

Edmans, Alex, 2017, The Answer to Short-Termism Isn't Asking Investors to Be Patient, *Harvard Business Review.*

Edmans, Alex, Xavier Gabaix, Tomasz Sadzik, and Yuliy Sannikov, 2012, Dynamic CEO Compensation, *Journal of Finance* 67(5), 1603-1647.

Eisdorfer, Assaf, 2008, Empirical Evidence of Risk-Shifting in Financially Distressed Firms, *Journal of Finance* 63(2), 609-637.

Froot, Kenneth, Andre Perold, and Jeremy Stein, 1992, Shareholder Trading Practices and Corporate Investment Horizons, *Journal of Applied Corporate Finance* 5(2), 42-58.

Feldman, Naomi E., Laura Kawano, Elena Patel, Nirupama Rao, Michael Stevens, and Jesse Edgerton, 2018, The Long and Short of It: Do Public and Private Firms Invest Differently?, Working Paper.

Frank, Murray Z., and Vidhan K. Goyal, Trade-off and Pecking Order Theories of Debt, 2008, Handbook of Empirical Corporate Finance. Elsevier, 135-202.

Gerakos, Joseph, 2010, CEO Pensions: Disclosure, Managerial Power, and Optimal Contracting, Pension Research Council Working Paper No. WP2007-5.

Giannetti, Mariassunta, and Xiaoyun Yu, 2018, Adapting to Radical Change: The Benefits of Short-Horizon Investors, Working paper.

Gilje, Erik, 2016, Do Firms Engage in Risk-Shifting? Empirical Evidence, *Review of Financial Studies* 29(11), 2925-2954.

Goldstein, Robert, Nengjiu Ju, and Hayne Leland, 2001, An EBIT-Based Model of Dynamic Capital Structure, *Journal of Business* 74(4), 483-512.

Gopalan, Radhakrishnan, Todd Milbourn, Fenghua Song, and Anjan Thakor, 2014, Duration of executive compensation, *The Journal of Finance* 69(6), 2777-2817.

Gorbenko, Alexander, and Ilya Strebulaev, 2010, Temporary versus Permanent Shocks: Explaining Corporate Financial Policies, *Review of Financial Studies* 23(7), 2591-2647.

Graham, John, and Campbell Harvey, 2001, The Theory and Practice of Corporate Finance: Evidence from the Field, *Journal of Financial Economics* 60(2-3), 187-243.

Graham, John, Campbell Harvey, and Shiva Rajgopal, 2005, The Economic Implications of Corporate Financial Reporting, *Journal of Accounting and Economics* 40(1-3), 3-73.

Gryglewicz, Sebastian, and Barney Hartman-Glaser, 2020, Investment timing and incentive costs, *The Review of Financial Studies* 33(1), 309-357.

Gryglewicz, Sebastian, Simon Mayer, and Erwan Morellec, 2020, Agency Conflicts over Short- versus Long-Termism in Corporate Policies, *Journal of Financial Economics* 136(3), 718-742.

Gryglewicz, Sebastian, Loriano Mancini, Erwan Morellec, Enrique Schroth, and Philip Valta, 2017, Transitory vs. Permanent shocks: Explaining Corporate Savings and Investment, Working paper.

Eom, Young Ho, Jean Helwege, and Jing-zhi Huang, 2004, Structural models of corporate bond pricing: An empirical analysis, *Review of Financial Studies* 17(2), 499-544.

Hackbarth, Dirk, and David C. Mauer, 2012, Optimal Priority Structure, Capital Structure, and Investment, *The Review of Financial Studies* 25(3), 747-796.

Hackbarth, Dirk, Jianjun Miao, and Erwan Morellec, 2006, Capital Structure, Credit Risk, and Macroeconomic Conditions, *Journal of Financial Economics* 82(3), 519-550.

He, Zhiguo, 2009, A Model of Dynamic Compensation and Capital Structure, Working Paper.

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

Holmstrom, Bengt, 1979, Moral Hazard and Observability, *The Bell journal of economics* 10(1), 74-91.

Holmstrom, Bengt, and Paul Milgrom, 1991, Multitasking Principal-Agent Analysis: Incentive Contracts, Asset Ownership, and Job Design, *Journal of Law, Economics, and Organization* 7, 286-295.

Holmstrom, Bengt, and Jean Tirole, 1993, Market Liquidity and Performance Measurement, *Journal of Political Economy* 101(4), 678-709.

Itoh, Hideshi, 1993, Incentives to Help in Multi-agent Situations, *Econometrica* 59(3), 611-636.

Karatzas, Ioannis, and Steven Shreve, 1991, *Brownian Motion and Stochastic Calculus*, Springer-Verlag: New York, Second Edition.

Kocherlakota, Narayana R., 2004, Figuring Out the Impact of Hidden Savings on Optimal Unemployment Insurance, *Review of Economic Dynamics* 7(3), 541-554.

Kaplan, Steven, 2017, Are U.S. Companies Too Short-Term Oriented? Some Thoughts, Working Paper.

Leland, Hayne, 1994, Corporate Debt Value, Bond Covenants, and Optimal Capital Structure, *Journal of Finance* 49(4), 1213-1252.

Lee, Seokwoo, and Alejandro Rivera, 2021, Extrapolation Bias and Robust Dynamic Liquidity Management, Forthcoming at *Management Science.*

Manso, Gustavo, 2008, Investment Reversibility and Agency Cost of Debt, *Econometrica* 76(2), 437-442.

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

Mello, Antonio, and John Parsons, 1992, Measuring the Agency Cost of Debt, *Journal of Finance* 47(5), 1887-1904.

Miao, Jianjun, and Alejandro Rivera, 2016, Robust Contracts in Continuous Time, *Econometrica* 84(4), 1405-1440.

Mikkelson, Wayne, and Megan Partch, 1986, Valuation Effects of Security Offerings and the Issuance Process, *Journal of Financial Economics* 15(1-2), 31-60.

Myers, Stewart, 1977, Determinants of Corporate Borrowing, *Journal of Financial Economics* 5(2), 147-175.

Poterba, James, and Lawrence Summers, 1995, A CEO Survey of U.S. Companies' Time Horizons and Hurdle Rates, *Sloan Management Review* 37, 43-53.

Rivera, Alejandro, 2020, Dynamic Moral Hazard and Risk-Shifting Incentives in a Leveraged Firm, *Journal of Financial and Quantitative Analysis* 55(4), 1333-1367.

Sannikov, Yuliy, 2008, A Continuous-Time Version of the Principal-Agent Problem, *Review of Economic Studies* 75(3), 957-984.

Shleifer, Andrei, and Robert Vishny, 1990, Equilibrium Short Horizons of Investors and Firms, *American Economic Review Papers and Proceedings* 80(2), 148–153.

Summers, Larry, 2017, Is Corporate Short-Termism Really a Problem? The Jury's Still Out, *Harvard Business Review.*

Sundaram, Rangarajan, and David Yermack, 2007, Pay Me Later: Inside Debt and Its Role in Managerial Compensation, *Journal of Finance* 62(4), 1551-1558.

Stein, Jeremy, 1988, Takeover Threats and Managerial Myopia, *Journal of Political Economy* 96(1), 61-80.

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

Strulovici, Bruno, and Martin Szydlowski, 2015, On the Smoothness of Value Functions and the Existence of Optimal Strategies in Diffusion Models, *Journal of Economic Theory* 159, 1016-1055.

Szydlowski, Martin, 2019, Incentives, Project Choice, and Dynamic Multitasking, *Theoretical Economics* 14(3), 813-847.

Von Thadden, Ernst-Ludwig, 1996, Long-Term Contracts, Short-Term Investment and Monitoring, *Review of Economic Studies* 62(4), 557-575.

Thakor, Richard T., 2019, Short-termism, Managerial Talent, and Firm Value, Working Paper.

Wong, Tak-Yuen, 2019, Dynamic Agency and Endogenous Risk-Taking, *Management Science* 65(9), 4032-4048.

Zhu, John, 2018, Myopic Agency, *Review of Economic Studies* 85(2), 1352-1388.

A Appendix

A.1 Appendix for Section 3

A.1.1 Change of Measure

We want to show that each policy (a, e) induces a probability measure. Fix a probability measure $\mathcal{P}^0$ under a zero-effort policy at all times such that

$$d\delta_t = \phi\delta_t dt + \sigma_\delta \delta_t dW_t^P, \text{ and } dA_t = \psi dt + \sigma_A dW_t^T,$$

where W_t^P and W_t^T are correlated standard Brownian motions under this measure $\mathcal{P}^0$ with $\mathbb{E}^0 \left[dW_t^P dW_t^T\right] = \rho dt$. We can decompose the correlated Brownian motions so that

$$\begin{pmatrix} W_t^P \\ W_t^A \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ -\frac{\rho}{\sqrt{1-\rho^2}} & \frac{1}{\sqrt{1-\rho^2}} \end{pmatrix}}_{\equiv M} \begin{pmatrix} W_t^P \\ W_t^T \end{pmatrix},$$

is a bi-dimensional standard Brownian motion with mutually independent components.

Now, we apply Girsanov's Theorem (theorem 5.1, Karatzas and Shreve (1991)) to $\left(W_t^P, W_t^A\right)'$. Let

$$\Theta_t(a,e) = \left(\frac{a_t}{\sigma_\delta}, \frac{e_t}{\sigma_A}\right), \text{ and } \tilde{\Theta}(a,e) = M\Theta_t(a,e) = \left(\frac{a_t}{\sigma_\delta}, -\frac{\rho}{\sqrt{1-\rho^2}}\frac{a_t}{\sigma_\delta} + \frac{1}{\sqrt{1-\rho^2}}\frac{e_t}{\sigma_A}\right),$$

and define

$$\xi_t(a,e) = \exp\left(\int_0^t \tilde{\Theta}_s(a,e) \begin{pmatrix} dW_s^P \\ dW_s^A \end{pmatrix} - \frac{1}{2}\int_0^t ||\tilde{\Theta}_s(a,e)||^2 ds\right),$$

where $||\cdot||$ is the Euclidean norm in $\mathbb{R}^2$. Under our regularity conditions, the Novikov's condition

$$\mathbb{E}^0 \left[\exp\left(\frac{1}{2}\int_0^t ||\tilde{\Theta}_s(a,e)||^2 ds\right)\right] < \infty$$

holds, then the process ξ_t is a martingale under measure $\mathcal{P}^0$. In particular, $\mathbb{E}^0\xi_t = \mathbb{E}^0\xi_0 = 1$, and the process Γ_t is a density process that defines the probability measure $\mathcal{P}^{(a,e)}$ via the Radon-Nikodym derivative $\frac{d\mathcal{P}^{(a,e)}}{d\mathcal{P}^0}|_{\mathcal{F}_t} = \xi_t(a,e)$. By Girsanov's Theorem, the process

$$\begin{pmatrix} Z_t^P \\ Z_t^A \end{pmatrix} = \begin{pmatrix} W_t^P \\ W_t^A \end{pmatrix} - \begin{pmatrix} \int_0^t \frac{a_t}{\sigma_\delta} ds \\ \int_0^t \left(-\frac{\rho}{\sqrt{1-\rho^2}}\frac{a_s}{\sigma_\delta} + \frac{1}{\sqrt{1-\rho^2}}\frac{e_s}{\sigma_A}\right) ds \end{pmatrix},$$

is a bi-dimensional standard Brownian motions under measure $\mathcal{P}^{(a,e)}$. Note that Z_t^P and Z_t^A are independent because W_t^P and W_t^A are. We can now calculate

$$\begin{pmatrix} Z_t^P \\ Z_t^T \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}}_{=M^{-1}} \begin{pmatrix} Z_t^P \\ Z_t^A \end{pmatrix},$$

where the second row is (4). So, the followings are increments of a standard Brownian motion under $\mathcal{P}^{(a,e)}$:

$$dZ_t^P = \frac{d\delta_t - (\phi + a_t)dt}{\sigma_\delta \delta_t}, \text{ and } dZ_t^T = \frac{dA_t - (\psi + e_t)dt}{\sigma_A},$$

with $\mathbb{E}^{(a,e)}\left[dZ_t^P dZ_t^T\right] = \rho dt$.

A.1.2 Proof of Proposition 1

Necessary Conditions: First, we provide the recursive representation of the agent's continuation utility (9). Fix an arbitrary contract $\Gamma = \langle c, a, e \rangle$. Recall that $\mathcal{P}^{(a,e)}$ denote the probability measure induced by the effort policy (a, e), and $\mathbb{F} = \{\mathcal{F}_t : t \geq 0\}$ is the filtration generated by A_t and δ_t, that is $\mathcal{F}_t = \sigma(\{A_s, \delta_s : s \leq t\})$. Thus, $\mathbb{F}$ represents public information. We assume incentive compatibility $\hat{a}_t = a_t$ and $\hat{e}_t = e_t$ and no-savings $\hat{c}_t = c_t$ whenever no confusion arises. Define the agent's lifetime expected utility as

$$V_t(\delta_t, \Gamma) \equiv \mathbb{E}_t^{(a,e)}\left[\int_0^\infty e^{-rs} u(c_s, a_s, e_s) ds\right] = \int_0^t e^{-rs} u(c_s, a_s, e_s) ds + e^{-rt} W_t(\delta_t, \Gamma), \tag{A.1}$$

where $\mathbb{E}_t^{(a,e)}[\cdot]$ is the expectation taken under the measure $\mathcal{P}^{(a,e)}$ conditional on the information $\mathcal{F}_t$; and $W_t(\delta_t, \Gamma)$, defined in (8), is the agent's continuation utility under the principal's information. Moreover, the process $\{V_t(\delta_t, \Gamma) : t \geq 0\}$ is a square-integrable $\mathcal{F}_t$-martingale under measure $\mathcal{P}^{(a,e)}$.

Then, the Martingale Representation Theorem (Theorem 4.2, Karatzas and Shreve (1991)) implies that there exists a process $\left\{b_t^P, b_t^A : t \geq 0\right\}$ that is progressively measurable with respect to $\mathbb{F}$ such that for any time $t > 0$,

$$\begin{aligned} V_t =& V_0 + \int_0^t (-\gamma r W_s) e^{-rs} b_s^P \sigma_\delta \delta_s dZ_s^P + \int_0^t (-\gamma r W_s) e^{-rs} b_s^A \sigma_A \delta_s dZ_s^A \\ =& V_0 + \int_0^t (-\gamma r W_s) e^{-rs} \left(b_s^P - \frac{\rho}{\sqrt{1-\rho^2}} \frac{\sigma_A}{\sigma_\delta} b_s^A\right) \sigma_\delta \delta_s dZ_s^P + \int_0^t (-\gamma r W_s) e^{-rs} \frac{b_s^A}{\sqrt{1-\rho^2}} \sigma_A \delta_s dZ_s^T, \end{aligned} \tag{A.2}$$

where in the second line, we substitute out dZ_t^A using equation (4). Setting

$$\beta_t^P = b_t^P - \frac{\rho}{\sqrt{1-\rho^2}} \frac{\sigma_A}{\sigma_\delta} b_t^A, \text{ and } \beta_t^T = \frac{b_t^A}{\sqrt{1-\rho^2}},$$

for all t, and differentiating the two expressions (A.1) and (A.2) for V_t with respect to t, we obtain (9).

Next, for the necessary condition for no-savings. Given an arbitrary contract Γ, consider the optimal continuation policy $\{c_s, a_s, e_s : s \geq t\}$; and suppose the agent is given extra savings at time t. It is well known that the wealth effect is absent with CARA preferences, then by the argument in He (2011), the new optimal continuation policy is to follow the original continuation effort choices and consume an extra amount rS now and at every time in the future $s > t$, 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 utility

$$W_t(\delta_t, \Gamma; S) = e^{-\gamma r S} W_t(\delta_t, \Gamma; 0), \tag{A.3}$$

where $W_t(\delta_t, \Gamma; 0)$ is the agent's continuation utility without savings as defined in (8).

Given an effort policy, the agent's problem (5) implies a necessary condition for consumption on the no-savings path: the agent's marginal utility of consumption must equal her marginal utility 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 (A.3), 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

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

which is the necessary condition (12) for the contract Γ to induce no savings.

Lastly, we provide a heuristic argument for the necessary condition for incentive compatibility. For any given contract Γ, the agent's problem (5) implies that at any point in time, the agent chooses the long-term and short-term effort to maximize the sum of her instantaneous utility $u(c_t, a_t, e_t)dt$ and the expected change in her continuation utility W_t. For any deviation $(\hat{a}_t, \hat{e}_t)$ from the recommended effort (a_t, e_t), then (9) implies that under the probability measure $\mathcal{P}^{(\hat{a},\hat{e})}$ induced by the deviation policy,

$$\mathbb{E}^{(\hat{a},\hat{e})}\left[dW_t\right] = (-\gamma r W_t)\left(\beta_t^P(\hat{a}_t - a_t) + \beta_t^T(\hat{e}_t - e_t)\right)\delta_t dt.$$

Thus, incentive compatibility requires

$$(a_t, e_t) = \underset{(\hat{a}_t, \hat{e}_t)}{\arg\max}\left\{u(c_t, \hat{a}_t, \hat{e}_t) + (-\gamma r W_t)\left(\beta_t^P(\hat{a}_t - a_t) + \beta_t^T(\hat{e}_t - e_t)\right)\delta_t\right\}. \tag{A.4}$$

Then, because $u_c(c_t, a_t, e_t) = -\gamma r W_t$ by (12), the first-order necessary conditions of the problem (A.4) are $\beta_t^P = g_a(a_t, e_t; \delta_t)/\delta_t$ and $\beta_t^T = g_e(a_t, e_t; \delta_t)/\delta_t$. Using the effort cost function, we obtain (11). If $\theta_a\theta_e \neq \theta_{ae}^2$, then the linear transformation

$$\begin{pmatrix} \theta_a & \theta_{ae} \\ \theta_{ae} & \theta_e \end{pmatrix} : \begin{pmatrix} a_t \\ e_t \end{pmatrix} \to \begin{pmatrix} \beta_t^P \\ \beta_t^T \end{pmatrix}$$

is invertible; hence, the first-order conditions (11) define a one-to-one and onto mapping between (a_t, e_t) and (β_t^P, β_t^T) under a suitably defined domain.

Sufficient Conditions: We show that the necessary conditions stated in Proposition 1 are indeed sufficient. That is, the contract induces incentive-compatible efforts and no private savings. Fix an arbitrary policy $(\hat{c}, \hat{a}, \hat{e})$, and consider the gain process for the agent

$$G_t^{(\hat{c},\hat{a},\hat{e})} = \int_0^t -\frac{1}{\gamma} e^{-\gamma(\hat{c}_s - g(\hat{a}_s, \hat{e}_s; \delta_s))} e^{-rs} ds + e^{-rt} e^{-\gamma r S_t} W_t(\delta_t, \Gamma; 0),$$

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 utility under the recommended effort policy (a, e) follows

$$dW_t = (-\gamma r W_t)\left(\beta_t^T \delta_t\left(dA_t - (\psi + e_t)dt\right) + \beta_t^P\left(d\delta_t - (\phi + a_t)\delta_t dt\right)\right),$$

with $W_t = W_t(\delta_t, \Gamma; 0)$. Now, consider the dynamics of $G_t^{(\hat{c},\hat{a},\hat{e})}$:

$$\begin{aligned} e^{rt} e^{\gamma r S_t} dG_t^{(\hat{c},\hat{a},\hat{e})} = &-\frac{1}{\gamma} e^{-\gamma(\hat{c}_t - g(\hat{a}_t, \hat{e}_t; \delta_t))} e^{\gamma r S_t} dt - rW_t - \gamma r W_t(rS_t + c_t - \hat{c}_t)dt \\ &+ (-\gamma r W_t)\left(\beta_t^T \delta_t\left(\hat{e}_t - e_t\right)dt + \beta_t^P \delta_t\left(\hat{a}_t - a_t\right)dt\right) \\ &+ (-\gamma r W_t)\left(\beta_t^T \delta_t \sigma_A dZ_t^T + \beta_t^T \delta_t \sigma_\delta dZ_t^P\right), \end{aligned}$$

where Z_t^P and Z_t^T are standard Brownian motions under the probability measure $\mathcal{P}^{(\hat{a},\hat{e})}$ induced by the agent's policy $(\hat{a}, \hat{e})$. Maximizing the drift with respect to $(\hat{c}_t, \hat{a}_t, \hat{e}_t)$ show that it is a concave problem in which the first-order conditions guarantee optimality. Moreover, the necessary conditions for ICNS contracts (11) and (12) ensures that $\mathbb{E}_t\left[e^{rt} e^{\gamma r S_t} dG_t^{(\hat{c},\hat{a},\hat{e})}\right] \leq 0$, and with equality for the optimal policy $(c + rS, a, e)$. Hence, we obtain that

$$e^{rt} e^{\gamma r S_t} dG_t^{(\hat{c},\hat{a},\hat{e})} = \text{non-positive drift} + (-\gamma r W_t)\left(\beta_t^T \delta_t \sigma_A dZ_t^T + \beta_t^T \delta_t \sigma_\delta dZ_t^P\right).$$

Integrating this equation yields

$$G_T^{(\hat{c},\hat{a},\hat{e})} \leq G_0^{(\hat{c},\hat{a},\hat{e})} + \int_0^T e^{-rt} e^{-\gamma r S_t} (-\gamma r W_t)\left(\beta_t^T \delta_t \sigma_A dZ_t^T + \beta_t^T \delta_t \sigma_\delta dZ_t^P\right),$$

for all T. We impose a transversality condition $\lim_{T\to\infty} \mathbb{E}\left[e^{-rT} e^{-\gamma r S_T} W_T\right] = 0$ on the feasible paths of S_t to ensure the second term on the right-hand side of the above inequality is indeed a martingale. Under the candidate optimal policy the transversality condition holds trivially since $S_T = 0$ and W_T is a martingale.

In fact, under the recommended policy the second term becomes $\int_0^T e^{-rt}dW_t$, which is a martingale (since W_t is a martingale). Letting $T \to \infty$, taking expectations, and recalling that $G_0^{(\hat{c},\hat{a},\hat{e})} = W_0$ imply that for any feasible policy $(\hat{c},\hat{a},\hat{e})$

$$\mathbb{E}^{(\hat{a},\hat{e})}\left[\int_0^\infty -\frac{1}{\gamma}e^{-\gamma(\hat{c}_t-g(\hat{a}_t,\hat{e}_t;\delta_t))}e^{-rt}dt\right] \leq W_0.$$

For a general policy the transversality conditions 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.

A.2 Appendix for Section 4.1

In this section, we first show that problem (15)-(16) can be rewritten as a one-dimensional control problem. Then, we show that the one dimensional problem satisfies the assumptions of Theorems 4 in Strulovici and Szydlowski (henceforth, SS(15)). Next, we show that the continuation region for this problem is of the form $[\delta_B, \infty)$. Finally, we collect results to prove Proposition 2.

Step 1: Reduce a Two-Dimensional Problem to a One-Dimensional Problem

Proposition 6. *The value function $P(\delta, W)$ of problem (15)-(16) satisfies*

$$P(\delta, W) = f(\delta) - \frac{-1}{\gamma r}\ln(-\gamma r W), \tag{A.5}$$

where the function $f(\delta)$ is defined on $[\underline{\delta}, \infty)$ by

$$f(\delta) = \sup_{(\beta_t^P,\beta_t^T,\tau_B)\in\mathcal{C}} H(\delta;\beta_t^P,\beta_t^T,\tau_B) \tag{A.6}$$

with

$$H(\delta;\beta_t^P,\beta_t^T,\tau_B) = \mathbb{E}^{(a^*,e^*)}\left[\int_0^{\tau_B\wedge\underline{\tau}} e^{-rt}\left((\delta_t(\psi+e_t)-(1-\tau)C)\right)dt - \int_0^\infty e^{-rt}\left(g(a_t,e_t,\delta_t)+\frac{1}{2}\gamma r\delta_t^2\Sigma_t\right)dt\right], \tag{A.7}$$

where $\Sigma_t = (\beta_t^P)^2\sigma_\delta^2 + 2\rho\beta_t^P\beta_t^T\sigma_\delta\sigma_A + (\beta_t^T)^2\sigma_A^2$ and where we recall that

$$d\delta_t = (\phi + a_t)\delta_t dt + \sigma_\delta\delta_t dZ_t^P, \tag{A.8}$$

and (a^*,e^*) are given by equation (11) as functions of (β^P,β^T).

Proof. First, let $B_t = -\frac{1}{\gamma r}\ln(-\gamma r W_t)$ denote the certainty equivalent of the agent's continuation utility. Fix an admissible control $(\beta_t^P,\beta_t^T,\tau_B)$, then

$$\begin{aligned}
&\mathbb{E}^{(a^*,e^*)}\left[\int_0^\infty e^{-rt}\left((dY_t-(1-\tau)Cdt)\,\mathbf{1}_{t\leq\tau_B\wedge\underline{\tau}} - c_t dt\right)\right]\\
&= \mathbb{E}^{(a^*,e^*)}\left[\int_0^{\tau_B\wedge\underline{\tau}} e^{-rt}\left(dY_t-(1-\tau)Cdt\right) - \int_0^\infty e^{-rt}\left(g(a_t,e_t;\delta_t)+rB_t\right)dt\right] && \text{(A.9)}\\
&= H(\delta;\beta_t^P,\beta_t^T,\tau_B) + \mathbb{E}^{(a^*,e^*)}\left[\int_0^\infty e^{-rt}\left(\frac{1}{2}\gamma r\delta^2\Sigma_t - rB_t\right)dt\right] && \text{(A.10)}
\end{aligned}$$

where the first equality follows from $c_t = g(a_t,e_t;\delta_t)+rB_t$ by (13) and the second from the definition of H. By Ito's formula, we can show that

$$B_t = B_0 + \int_0^t \frac{1}{2}\gamma r\Sigma_s\delta_s^2 ds + \int_0^t \beta_s^P\sigma_\delta\delta_s dZ_s^P + \int_0^t \beta_s^T\sigma_A\delta_s dZ_s^T. \tag{A.11}$$

Applying Ito's formula to $e^{rt}B_t$, integrating, and taking expectations under the integrability conditions (10) yields:

$$\mathbb{E}^{(a^*,e^*)}\left[\int_0^\infty e^{-rt}\left(\frac{1}{2}\gamma r\delta^2\Sigma_t - rB_t\right)dt\right] = -B_0 \tag{A.12}$$

Taking suprumum in (A.10) over the admissible controls $(\beta_t^P, \beta_t^T, \tau_B)$ we obtain:

$$\sup_{(\beta_t^P,\beta_t^T,\tau_B)} \mathbb{E}^{(a^*,e^*)}\left[\int_0^\infty e^{-rt}\left(\left(dY_t - (1-\tau)Cdt\right)\mathbf{1}_{t\le\tau_B\wedge\underline{\tau}} - c_t dt\right)\right] = \sup_{(\beta_t^P,\beta_t^T,\tau_B)} H(\delta;\beta_t^P,\beta_t^T,\tau_B) - B_0 \tag{A.13}$$

which is equation (A.5), as desired. □

We first notice that any optimal policy must set the implied effort $a_t = e_t = 0$ for all $t > \tau_B \wedge \underline{\tau}$. Hence, problem (A.6) can be rewritten as:

$$f(\delta) = \sup_{(\beta_t^P,\beta_t^T,\tau_B)\in\mathcal{C}} \mathbb{E}^{(a^*,e^*)}\left[\int_0^{\tau_B\wedge\underline{\tau}} e^{-rt}\left(\left(\delta_t(\psi+e_t) - (1-\tau)C\right) - g(a_t,e_t,\delta_t) - \frac{1}{2}\gamma r\delta_t^2\Sigma_t\right)dt\right] \tag{A.14}$$

Step 2: Check Assumptions of Theorem 4 in SS(15).

We now proceed to show that problem (A.14) satisfies assumptions 1-4 required for Theorem 4 of SS(15). Assumption 1 holds because $(a,e) \in [0,\overline{a}]\text{x}[0,\overline{e}]$ which implies (β^P, β^T) belong to a compact subset of $\mathbb{R}^2$. Assumption 2 is essentially Lipschitz continuity for the drift, the volatility, and the payoff function. It holds setting $K = (\phi+\overline{a}) + \sigma_\delta + (\psi + \overline{e} + (\theta_a\overline{a}^2 + 2\theta_{ae}\overline{ae} + \theta_e\overline{e}^2) + \gamma r\overline{\delta}((\theta_a\overline{a}+\theta_{ae}\overline{e})^2\sigma_\delta^2 + (\theta_e\overline{e}+\theta_{ae}\overline{a})^2\sigma_A^2 + 2\rho\left(\theta_a\overline{a}+\theta_{ae}\overline{e}\right)\left(\theta_e\overline{e}+\theta_{ae}\overline{a}\right)\sigma_\delta\sigma_A))$. Assumption 3 corresponds to linear growth conditions and a uniform lower bound on the volatility. For this assumption we needed to restrict $\delta_t \ge \underline{\delta} > 0$. Thus, this assumption holds with $\underline{\sigma} = \underline{\delta}\sigma_\delta$, $K_1^\mu = 0$, $K_2^\mu = (\phi+\overline{a})$, $K^\sigma = \sigma_\delta$, and $K^f = (1-\tau)C + (\phi+\overline{a}) + \sigma_\delta + (\psi+\overline{e}+(\theta_a\overline{a}^2 + 2\theta_{ae}\overline{ae} + \theta_e\overline{e}^2) + \gamma r\overline{\delta}((\theta_a\overline{a}+\theta_{ae}\overline{e})^2\sigma_\delta^2 + (\theta_e\overline{e}+\theta_{ae}\overline{a})^2\sigma_A^2 + 2\rho\left(\theta_a\overline{a}+\theta_{ae}\overline{e}\right)\left(\theta_e\overline{e}+\theta_{ae}\overline{a}\right)\sigma_\delta\sigma_A))$. Assumption 4 holds since the obstacle in our optimal stopping problem $g(\delta) = 0$.

Step 3: Characterize Continuation and Stopping Regions.

We can now apply Theorem 4 in SS(15). From i) we know that the value function $f(\delta)$ is finite and has linear growth. From ii) we know that $f(\delta)$ is continuously differentiable on $[\underline{\delta},\infty)$ and satisfies $f(\delta) \ge 0$. From iii) denoting by $\mathbb{Y}$ the subset where $f(\delta) = 0$ (i.e., the stopping region), then $\mathbb{Y}$ is closed and $f(\delta)$ solves (18) in $[\underline{\delta},\infty)]/\mathbb{Y}$ (i.e., the continuation region).

Under the parametric restrictions imposed in Conditions 1 and 2, the following proposition shows that the continuation region has the conjectured structure.

Proposition 7. *The stopping region $\mathbb{Y} = [\underline{\delta}, \delta_B]$ for some $\underline{\delta} \le \delta_B < \infty$.*

Proof. Suppose for a contradiction that the stopping region is disconnected. Let (δ_X, δ_Y) be an interval in the continuation region separating two portions of the stopping region. We immediately know that $f(\delta_X) = f(\delta_Y) = 0$ and $f'(\delta_X) = f'(\delta_Y) = 0$, since $f(\delta)$ is continuously differentiable. Now, rewrite (18) as

$$rf(\delta) = G(\delta, f'(\delta)) + \frac{1}{2}\sigma_\delta^2\delta^2 f''(\delta) \tag{A.15}$$

where

$$G(\delta,q) = \max_{(a,e)}\left\{\delta(\psi+e) - (1-\tau)C - \frac{1}{2}\delta(\theta_a a^2 + 2\theta_{ae}ae + \theta_e e^2) + (\phi+a)\delta q - \frac{1}{2}\gamma r\delta^2\Sigma(a,e)\right\} \tag{A.16}$$

Evaluating (A.15) at δ_X and δ_Y yields

$$0 = G(\delta_X, 0) + \frac{1}{2}\sigma_\delta^2\delta_X^2 f''(\delta_X), \tag{A.17}$$

$$0 = G(\delta_Y, 0) + \frac{1}{2}\sigma_\delta^2\delta_Y^2 f''(\delta_Y). \tag{A.18}$$

Since $f''(\delta_X) \geq 0$ (else $f(\delta_X + \epsilon) < 0$) and $f''(\delta_Y) \geq 0$ (else $f(\delta_Y - \epsilon) < 0$), it follows that $G(\delta_X, 0) \leq 0$ and $G(\delta_Y, 0) \leq 0$. We now make the following claim.

Claim: $G(\delta, 0) \leq 0$ for all $\delta \in [\delta_X, \delta_Y]$.

Next, let $\tilde{\delta}$ be a maximizer of $f(\delta)$ in $[\delta_X, \delta_Y]$. This implies that $f(\tilde{\delta}) > 0$, $f'(\tilde{\delta}) = 0$ and $f''(\tilde{\delta}) < 0$. Evaluating (A.15) at $\tilde{\delta}$ yields:

$$rf(\tilde{\delta}) = G(\tilde{\delta}, 0) + \frac{1}{2}\sigma_\delta^2\delta^2 f''(\tilde{\delta}) \leq 0 \tag{A.19}$$

which contradicts the fact that $f(\tilde{\delta}) > 0$. □

It remains to prove our claim.

Proof. Suppose for a contradiction that $G(\delta, 0) > 0$ for some $\delta \in [\delta_X, \delta_Y]$. Let $\hat{\delta}$ be a maximizer of $G(\delta, 0)$ in $[\delta_X, \delta_Y]$. This implies that $G_\delta(\hat{\delta}, 0) = 0$. Applying the envelope theorem in (A.16) and isolating $\hat{\delta}$ yields

$$\hat{\delta} = \frac{\psi + e^* - \frac{1}{2}\delta(\theta_a(a^*)^2 + 2\theta_{ae}a^*e^* + \theta_e(e^*)^2)}{\gamma r\Sigma(a^*, e^*)} \tag{A.20}$$

Moreover, it follows that

$$\hat{\delta} > \frac{\psi - \frac{1}{2}\delta(\theta_a(\bar{a})^2 + 2\theta_{ae}\bar{a}\bar{e} + \theta_e(\bar{e})^2)}{\gamma r\Sigma(\bar{a}, \bar{e})} > \frac{(1-\tau)C}{\psi}, \tag{A.21}$$

where Condition 1 guarantees positivity of the numerator and the second inequality follows from Condition 2. Finally, notice that setting $(a, e) = (0, 0)$ implies that $G(\delta, 0) \geq \delta\psi - (1-\tau)C$ for all δ. Since $G(\delta_Y, 0) \leq 0$, it follows that

$$\hat{\delta} \leq \delta_Y \leq \frac{(1-\tau)C}{\psi}, \tag{A.22}$$

which contradicts (A.21). □

Step 4: Proof of Proposition 2:

Proof. From Proposition 6, it follows that the shareholder value function is given by (A.5). Applying Theorem 4 in SS (2015) to (A.14) showed that $f(\delta)$ is nonnegative, finite, has linear growth, and is continuously differentiable. Moreover, by Proposition 7 $f(\delta)$ corresponds to the solution of the ODE (18)–(19), where the value matching follows from continuity of $f(\delta)$ and the smooth pasting from being continuously differentiable. Hence, we conclude the value function of problem (15)–(16) is given by (21).

□

A.2.1 Solution for Optimal Effort Policies in the HJB

The optimization on the RHS of the HJB (18) can be rewritten as

$$\max_{a\in[a,\bar{a}], e\in[0,\bar{e}]} \left\{ af'(\delta) + e - (1/2)a^2 \underbrace{(\theta_a + \gamma r\delta(\theta_a^2\sigma_\delta^2 + \theta_{ae}^2\sigma_A^2 + 2\theta_{ae}\theta_a\sigma_\delta\sigma_A\rho))}_{\hat{A}} \right.$$

$$- (1/2)e^2 \underbrace{(\theta_e + \gamma r \delta(\theta_{ae}^2 \sigma_\delta^2 + \theta_e^2 \sigma_A^2 + 2\theta_{ae}\theta_e \sigma_A \sigma_\delta \rho))}_{\hat{D}}$$

$$\left. -ae \underbrace{(\theta_{ae} + \gamma r \delta(\theta_a \theta_{ae} \sigma_\delta^2 + \theta_e \theta_{ae} \sigma_A^2 + (\theta_a \theta_e + \theta_{ae}^2)\sigma_\delta \sigma_A \rho))}_{\hat{B}} \right\}.$$

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 λ_1 and λ_2 the respective multipliers. We first calculate the unconstrained optimum by taking the FOC with respect to a and e to obtain

$$a^* = \frac{f'(\delta)\hat{D} - \hat{B}}{\hat{A}\hat{D} - \hat{B}^2}, \quad e^* = \frac{\hat{A} - f'(\delta)\hat{B}}{\hat{A}\hat{D} - \hat{B}^2}. \tag{A.23}$$

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 $(0, \hat{e}, \lambda_1, 0)$ satisfies the Kuhn-Tucker conditions (where $\hat{e} = 1/D$ and $\lambda_1 = \frac{\hat{D}f'(\delta) - \hat{B}}{\hat{D}}$): 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 λ_1 is satisfied, and λ_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 $\hat{e} = \frac{1 - \bar{a}\hat{B}}{\hat{D}}$ and $\lambda_2 = \frac{\bar{a}(\hat{A}\hat{D} - \hat{B}^2) - (f'(\delta)\hat{D} - \hat{B})}{\hat{D}}$): 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 λ_2 is satisfied, and λ_1 is zero.

Letting $K_0 = \frac{\hat{D}}{\hat{A}\hat{D} - \hat{B}^2}$, $K_1 = \frac{-\hat{B}}{\hat{A}\hat{D} - \hat{B}^2}$, and $K_2 = \frac{\hat{A}}{\hat{A}\hat{D} - \hat{B}^2}$ in A.23 yields the condensed expressions for the optimal effort policies in (20).

A.2.2 Unlevered Firm Contract With Observable Effort

Suppose there is no agency problem (i.e., the agent's actions are perfectly observable) and the firm has no debt. In this case, the firm value $f_{FB}(\delta)$ satisfies

$$rf_{FB}(\delta) = \max_{(a,e)} \left\{ \delta(\psi + e) - \frac{1}{2}\left(\theta_a a^2 + \theta_e e^2 + 2\theta_{ae} ae\right)\delta + (\phi + a)\delta f'_{FB}(\delta) + \frac{1}{2}\sigma_\delta^2 \delta^2 f''_{FB}(\delta) \right\}. \tag{A.24}$$

Since the cash flows and expected capital gains are proportional to δ, 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 (A.24), we have a equation that determines the unknown coefficient q:

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

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 efforts in this case are independent of firm size. Second, permanent shocks, transitory shocks, and their correlation do not affect q; 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 δ_t, but a positive transitory shock $dZ_t^T > 0$ has no effect on firm value.

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

A.3 Appendix for Section 4.2

A.3.1 Proof of Proposition 3

We apply Proposition 6 to write the problem as an equivalent one-dimensional control problem. The result follows from applying Theorem 1 of Strulovici and Szydlowski (2015). Existence and uniqueness can be guaranteed in this instance because there is no optimal stopping for the unlevered case.

A.4 Appendix for Section 4.3

A.4.1 Proof of Proposition 4

We will show that the optimal level of short-term effort in the case when $\rho = \theta_{ae} = 0$ is given by $e^* = \frac{1}{\hat{D}}$, where in this case $\hat{D} = \theta_e(1 + \gamma r\delta\theta_e\sigma_A^2)$. 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 $\rho = \theta_{ae} = 0$ implies $\hat{B} = 0$. Second, substituting $\hat{B} = 0$ in (A.23) yields $e^* = \frac{1}{\hat{D}}$, as expected. A similar argument shows that $e^* = \frac{1}{\hat{D}}$ for the case in which the constraints on a bind.

A.4.2 Proof of Proposition 5

Suppose that $f(\delta)$ solves (A.27), which using our condensed notation for $\hat{A}$, $\hat{B}$ and $\hat{D}$ can be rewritten as

$$rf(\delta) = \left\{ \begin{array}{c} a_u(\delta)f'(\delta) + e_u(\delta) - \frac{a_u^2(\delta)\hat{A}}{2} - \frac{e_u^2(\delta)\hat{D}}{2} - a_u(\delta)e_u(\delta)\hat{B} \\ +\delta\psi - (1-\tau)C - +\phi\delta f'(\delta) + \frac{1}{2}\sigma_\delta^2\delta^2 f''(\delta) \end{array} \right\}, \tag{A.25}$$

subject to

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

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

$$a_u(\delta), e_u(\delta) \in arg\max_{a,e} \left\{ a(\delta)f_u'(\delta) + e(\delta) - \frac{a^2(\delta)\hat{A}}{2} - \frac{e^2(\delta)\hat{D}}{2} - a(\delta)e(\delta)\hat{B} \right\},$$

which implies that

$$e_u(\delta) \in arg\max_e \left\{ a_u(\delta)f'(\delta) + e(\delta) - \frac{a_u^2(\delta)\hat{A}}{2} - \frac{e^2(\delta)\hat{D}}{2} - a_u(\delta)e(\delta)\hat{B} \right\}.$$

Therefore, $f(\delta)$ satisfies

$$rf(\delta) = \max_{e\in[0,\bar{e}]} \left\{ \begin{array}{c} a_u(\delta)f'(\delta) + e - \frac{a_u^2(\delta)\hat{A}}{2} - \frac{e^2\hat{D}}{2} - a_u(\delta)\hat{B} \\ +\delta\psi - (1-\tau)C - +\phi\delta f'(\delta) + \frac{1}{2}\sigma_\delta^2\delta^2 f''(\delta) \end{array} \right\}, \tag{A.26}$$

and $\delta_B^{NO} = \delta_B^{NU}$ satisfies the boundary conditions by assumption. Therefore, $f(\delta)$ satisfies (A.29). The proof of the converse is essentially identical.

A.4.3 Three Adjoint Problems

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-effort 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:

$$rf_{NO}(\delta) = \max_{\delta_B^{NO}} \left\{ \begin{array}{c} \delta(\psi + e_u) - (1-\tau)C - \frac{\left(\theta_a a_u^2 + 2\theta_{ae}a_u e_u + \theta_e e_u^2\right)\delta}{2} + (\phi + a_u)\delta f'_{NO}(\delta) + \frac{1}{2}\sigma_\delta^2\delta^2 f''_{NO}(\delta) \\ -\frac{\gamma r\delta^2}{2}\left((\theta_a a_u + \theta_{ae}e_u)^2\sigma_\delta^2 + (\theta_e e_u + \theta_{ae}a_u)^2\sigma_A^2 + 2\rho(\theta_a a_u + \theta_{ae}e_u)(\theta_e e_u + \theta_{ae}a_u)\sigma_\delta\sigma_A\right) \end{array} \right\} \tag{A.27}$$

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

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

where a_u and e_u are given by (26) 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 e_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

$$ML_{NO}(\delta_0) \equiv \frac{D_{NO}(\delta_0; C^*_{NO}(\delta_0))}{f_{NO}(\delta_0; C^*_{NO}(\delta_0)) + D_{NO}(\delta_0; C^*_{NO}(\delta_0))},$$

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:

$$rf_{NS}(\delta) = \max_{\substack{\delta_B^{NS} \\ a\in[0,\bar{a}]}} \left\{ \begin{array}{c} \delta(\psi + e_u) - (1-\tau)C - \frac{\left(\theta_a a^2 + 2\theta_{ae}ae_u + \theta_e e_u^2\right)\delta}{2} + (\phi + a)\delta f'_{NS}(\delta) + \frac{1}{2}\sigma_\delta^2\delta^2 f''_{NS}(\delta) \\ -\frac{\gamma r\delta^2}{2}\left((\theta_a a + \theta_{ae}e_u)^2\sigma_\delta^2 + (\theta_e e_u + \theta_{ae}a)^2\sigma_A^2 + 2\rho(\theta_a a + \theta_{ae}e_u)(\theta_e e_u + \theta_{ae}a)\sigma_\delta\sigma_A\right) \end{array} \right\} \tag{A.28}$$

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

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

where e_u is given by the solution of (26) 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 $e_{u.}$. 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

$$ML_{NS}(\delta_0) \equiv \frac{D_{NS}(\delta_0; C^*_{NS}(\delta_0))}{f_{NS}(\delta_0; C^*_{NS}(\delta_0)) + D_{NS}(\delta_0; C^*_{NS}(\delta_0))},$$

where C_{NS} denotes the optimal coupon.

Case 3: 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 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:

$$rf_{NU}(\delta) = \max_{\substack{\delta_B^{NU} \\ e\in[0,\bar{e}]}} \left\{ \begin{array}{c} \delta(\psi+e) - (1-\tau)C - \frac{(\theta_a a_u^2 + 2\theta_{ae} a_u e + \theta_e e^2)\delta}{2} + (\phi + a_u)\delta f'_{NU}(\delta) + \frac{1}{2}\sigma_\delta^2 \delta^2 f''_{NU}(\delta) \\ -\frac{\gamma r \delta^2}{2}\left((\theta_a a_u + \theta_{ae} e)^2 \sigma_\delta^2 + (\theta_e e + \theta_{ae} a_u)^2 \sigma_A^2 + 2\rho(\theta_a a_u + \theta_{ae} e)(\theta_e e + \theta_{ae} a_u)\sigma_\delta \sigma_A\right) \end{array} \right\} \tag{A.29}$$

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

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

where $a_u(\delta)$ is given by the solution of (26), and the subscript NU stands for *no underinvestment.*

A.5 Appendix for Section 4.4

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

$$\frac{\partial f_\theta(\delta)}{\partial\theta} = \mathbb{E}\left[\int_0^{\tau_B} e^{-rt} \left(\begin{array}{c} \frac{\partial(\delta_t(\psi+e_t) - (1-\tau)C - g(a_t, e_t; \delta_t))}{\partial\theta} + \frac{\partial(\phi+a_t)}{\partial\theta}\delta_t f'_\theta(\delta_t) + \frac{1}{2}\frac{\partial\sigma_\delta^2}{\partial\theta}\delta_t^2 f''_\theta(\delta_t) \\ -\frac{1}{2}\frac{\partial}{\partial\theta}\left(\gamma r \delta_t^2\left((\beta_t^P)^2\sigma_\delta^2 + 2\rho\sigma_\delta\sigma_A\beta_t^P\beta_t^T + (\beta_t^T)^2\sigma_A^2\right)\right) \end{array} \right) dt \,\middle|\, \delta_0 = \delta\right]$$

with $\beta_t^P = \theta_a a_t + \theta_{ae} e_t$ and $\beta_t^T = \theta_e e_t + \theta_{ae} a_t$.

Proof. Given any default threshold δ_B, the investor's payoff $f(\delta; \delta_B)$ solves the ODE

$$\begin{aligned} rf_(\delta; \delta_B) =& \delta(\psi+e) - (1-\tau)C - g(a, e; \delta) + (\phi + a)\delta f'(\delta; \delta_B) + \frac{1}{2}\sigma_\delta^2\delta^2 f''(\delta; \delta_B) \\ &- \frac{\gamma r \delta^2}{2}\left((\beta^P)^2\sigma_\delta^2 + 2\rho\sigma_\delta\sigma_A\beta^P\beta^T + (\beta^T)^2\sigma_A^2\right) \end{aligned} \tag{A.30}$$

with boundary condition $f(\delta_B; \delta_B) = 0$. Here a and e are the optimal choices and $\beta^P = \theta_a a + \theta_{ae} e$ and $\beta^T = \theta_e e + \theta_{ae} a$. Denote $\delta_B(\theta)$ as the optimal default threshold under θ, then by definition $f_\theta(\delta) = f_\theta(\delta; \delta_B(\theta))$.

Differentiating both sides of (A.30) with respect to θ and evaluate at $\delta_B = \delta_B(\theta)$, then together with the Envelope theorem, $\frac{\partial f_\theta(\delta)}{\partial\theta} = \frac{\partial f_\theta(\delta;\delta_B)}{\partial\theta}|_{\delta_B=\delta_B(\theta)}$, we get

$$\begin{aligned} r\frac{\partial f_\theta(\delta)}{\partial\theta} =& \frac{\partial}{\partial\theta}\left(\delta(\psi+e)-(1-\tau)C-g(a,e;\delta)\right)+\frac{\partial(\phi+a)}{\partial\theta}\delta f_\theta'(\delta)+\frac{1}{2}\frac{\partial\sigma_\delta^2}{\partial\theta}\delta^2 f_\theta''(\delta) \\ &-\frac{1}{2}\frac{\partial}{\partial\theta}\left(\gamma r\delta^2\cdot\left((\beta^P)^2\sigma_\delta^2+2\rho\sigma_\delta\sigma_A\beta^P\beta^T+(\beta^T)^2\sigma_A^2\right)\right)+(\phi+a)\delta\frac{\partial}{\partial\delta}\left[\frac{\partial f_\theta(\delta)}{\partial\theta}\right] \\ &+\frac{1}{2}\sigma_\delta^2\delta^2\frac{\partial^2}{\partial\delta^2}\left[\frac{\partial f_\theta(\delta)}{\partial\theta}\right]+\frac{\partial a}{\partial\theta}\cdot FOC(a)+\frac{\partial e}{\partial\theta}\cdot FOC(e) \end{aligned}$$

with boundary condition $\frac{\partial f_\theta(\delta)}{\partial\theta} = 0$ 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: $\frac{\partial f(\delta)}{\partial\psi} = \mathbb{E}\left[\int_0^{\tau_B} e^{-rt}\delta_t dt|\delta_0=\delta\right] > 0$, $\frac{\partial f(\delta)}{\partial\phi} = \mathbb{E}\left[\int_0^{\tau_B} e^{-rt}\delta_t f'(\delta_t)dt|\delta_0=\delta\right] > 0$
- Agency: $\frac{\partial f(\delta)}{\partial\gamma} = \mathbb{E}\left[\int_0^{\tau_B} e^{-rt}\left(-\frac{r}{2}\delta_t^2\cdot((\beta_t^P)^2\sigma_\delta^2+2\rho\sigma_\delta\sigma_A\beta_t^P\beta_t^T+(\beta_t^T)^2\sigma_A^2)\right)dt|\delta_0=\delta\right] < 0$ if $\rho \geq 0$
- Coupon and tax: $\frac{\partial f(\delta)}{\partial C} = \mathbb{E}\left[\int_0^{\tau_B} e^{-rt}(-(1-\tau))dt|\delta_0=\delta\right] < 0$, $\frac{\partial f(\delta)}{\partial\tau} = \mathbb{E}\left[\int_0^{\tau_B} e^{-rt}Cdt|\delta_0=\delta\right] > 0$
- Volatilities and correlation: $\frac{\partial f(\delta)}{\partial\sigma_A} = \mathbb{E}\left[\int_0^{\tau_B} e^{-rt}\left(-\gamma r\delta_t^2\right)\left(\rho\sigma_\delta\beta_t^P\beta_t^T+(\beta_t^T)^2\sigma_A\right)dt|\delta_0=\delta\right] < 0$ if $\rho \geq 0$ and ambiguous otherwise, $\frac{\partial f(\delta)}{\partial\sigma_\delta} = \mathbb{E}\left[\int_0^{\tau_B} e^{-rt}\left(\sigma_\delta\delta_t^2 f''-\gamma r\left((\beta_t^P)^2\sigma_\delta+\rho\sigma_A\beta_t^P\beta_t^T\right)\right)dt|\delta_0=\delta\right]$, and $\frac{\partial f(\delta)}{\partial\rho} = \mathbb{E}\left[\int_0^{\tau_B} e^{-rt}\left(-\gamma r\sigma_\delta\sigma_A\beta_t^P\beta_t^T\right)dt|\delta_0=\delta\right] < 0$
- Effort costs: $\frac{\partial f(\delta)}{\partial\theta_a} = \mathbb{E}\left[\int_0^{\tau_B} e^{-rt}(-1)\left(\frac{1}{2}a_t^2\delta_t+\gamma r\delta_t^2(\beta_t^P\sigma_\delta^2+\rho\sigma_\delta\sigma_A\beta_t^T)a_t\right)dt|\delta_0=\delta\right] < 0$ if $\rho \geq 0$ and ambiguous otherwise, $\frac{\partial f(\delta)}{\partial\theta_e} = \mathbb{E}\left[\int_0^{\tau_B} e^{-rt}(-1)\left(\frac{1}{2}e_t^2\delta_t+\gamma r\delta_t^2(\rho\sigma_\delta\sigma_A\beta_t^P+\beta_t^T\sigma_A^2)e_t\right)dt|\delta_0=\delta\right] < 0$ if $\rho \geq 0$ and ambiguous otherwise, and

  $$\frac{\partial f(\delta)}{\partial\theta_{ae}} = \mathbb{E}\left[\int_0^{\tau_B} e^{-rt}(-1)\left(a_te_t\delta_t+\gamma r\delta_t^2\left(\beta_t^P\sigma_\delta^2e_t+\rho\sigma_\delta\sigma_A(\beta_t^Te_t+\beta_t^Pa_t)+\beta_t^T\sigma_A^2a_t\right)\right)dt|\delta_0=\delta\right] < 0$$

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

Lemma 2. *Let δ_{Bi} be the optimal default boundary under parameter θ_i. If $\theta_2 > \theta_1$ implies $f_{\theta_2}(\delta) > f_{\theta_1}(\delta)$ for all δ, then $\delta_{B2} < \delta_{B1}$.*

Proof. Suppose not, $\delta_{B2} \geq \delta_{B1}$. First, $f_{\theta_2}(\delta_{B2}) = 0 > f_{\theta_2}(\delta_{B1})$ by the optimality of and value-matching at δ_{B2}. Second, $f_{\theta_2}(\delta_{B1}) > f_{\theta_1}(\delta_{B1})$ by the hypothesis. Therefore, $f_{\theta_1}(\delta_{B1}) < 0$ and the choice of δ_{B1} violates limited liability for the market value of equity $f(\delta) \geq 0$.

A.6 Appendix for Section 5: Calibration

We calibrate the parameter values θ_{ae} and θ_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

$$\frac{\theta_a a_u^2(\delta)\delta}{2} = \frac{\theta_e e_u^2(\delta)\delta}{2} = 2\theta_a e_u(\delta)a_u(\delta)\delta$$

for $\delta = 80$, which is right in the middle of the various values of δ_0 that we use in our numerical examples.

about ECGI

The European Corporate Governance Institute has been established to improve *corporate governance through fostering independent scientific research and related activities.*

The ECGI will produce and disseminate high quality research while remaining close to the concerns and interests of corporate, financial and public policy makers. It will draw on the expertise of scholars from numerous countries and bring together a critical mass of expertise and interest to bear on this important subject.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.

Electronic Access to the Working Paper Series

The full set of ECGI working papers can be accessed through the Institute's Web-site (www.ecgi.global/content/working-papers) or SSRN:

Finance Paper Series	http://www.ssrn.com/link/ECGI-Fin.html
Law Paper Series	http://www.ssrn.com/link/ECGI-Law.html