Voluntary disclosure, moral hazard and default risk *

Shiming Fu¹ and Giulio Trigilia²

¹Shanghai University of Finance and Economics
²University of Rochester, Simon Business School

June 19, 2019

Abstract

We introduce voluntary disclosure opportunities in a dynamic agency model with non-verifiable cash flows. Disclosure has two countervailing effects. When bad news arise, the firm’s management is given some slack by the investors, which lowers the optimal pay-for-performance sensitivity. However, paying for bad luck also reduces the value for investors of providing liquidity to the firm in the first place. Disclosure lowers the firm’s probability of default conditional on a given performance history: it is associated with lower leverage and higher dividend payout rates. However, its effects in primary markets are heterogeneous across firms. At some low profitability firms, more frequent disclosure is associated to a higher probability of default and lower managerial pay. At other firms, both relations are reversed. Firms with intermediate performance history benefit the most from increasing the arrival rate of information to disclose, and are expected to disclose more frequently.

Key words: voluntary disclosure, credit spreads, default risk, dynamic moral hazard, funding liquidity

JEL classification: G32, D86, D61

*Previously circulated under the title: “Voluntary disclosure under dynamic moral hazard”. We thank Dan Bernhardt, Michael Gofman, Ron Kaniel, Meg Meyer, Alan Moreira, Dmitry Orlov, Heikki Rantakari, Clifford Smith, Vish Viswanathan, Ming Yang, Adriano Rampini, Martin Szydlowski (discussant), Pavel Zruymov and the participants at: FTG (Kellogg), SFS Cavalcade (CMU), UNC Junior Finance Roundtable, Oxford (Nuffield College), CRETA (Warwick), Rochester (Simon), FTG Summer (LBS), SED (Mexico City), NASMES (Davis), MWET (Drexel). All errors are our own.
1 Introduction

Over the last two decades, technological progress drastically reduced the costs of generating, storing and analyzing information. According to Dresner Advisory Services (2018), which surveys 5,000 firms globally, adopters of the latest such technologies – broadly referred to as Big Data and Artificial Intelligence – surged from 17% in 2015 to 59% in 2018, while another 30% signaled the intention to adopt them in the future.\footnote{Despite the increase in adoption, there still is a variety of competing infrastructures, most of which are open-source. The most popular to this date are Spark, Kafka, MapReduce, Kubernetes and Yarn.} The most frequent applications are data optimization and forecasting, which means that adopting firms have increased their access to evidence that leads performance and are expected to disclose information to their investors more frequently. So, we ask: what are the disclosure patterns induced by these information technologies? And how does their adoption affect firms’ default risk, pay-for-performance sensitivity, leverage and dividend payout?

Our paper highlights the difference between information generated by costly third-party monitoring of a firm’s management, and information produced by a technology that the firm’s management itself can control. This distinction is key. In practice, arm’s length investors in a large firm have both limited incentives and ability to monitor, relying on the firm itself to provide basic information about the sources of its performance. Conversely, a firm’s management has access to information technologies that can sometimes shed light on its likely near-term future performance, indirectly conveying information to outside investors about how well the management itself is doing. The problem, of course, is that while this information may not be easily manipulable, it can be shrouded by managers, who have the discretion to not disclose information. The issue from the perspective of the outside investors is that no disclosure could be due either to the absence of evidence, or to its strategic concealment.

It is immediate that, if they can, a firm’s investors should design correct incentives for managers to reveal information.\footnote{This is an instance of Homlstrom’s informativeness principle (Holmstrom (1979)): investors can design a randomization procedure that implements the allocation that was optimal with lower evidence, committing to disregard managerial disclosures probabilistically.} We show how optimal incentive design differs when the goal is to provide management incentives to reveal information, both good and bad, when management receives it, versus the optimal design when monitoring is external to the firm and costly. While in a typical monitoring setting the information is used by investors to curb managerial rents,\footnote{See Fuchs (2007), Piskorski and Westerfield (2016), Smolin (2017), Zhu (2018) and Orlov (2019).} this need not be true in a disclosure setting. Especially at high profitability firms, disclosure opportunities can be accompanied by higher rents paid to management, which is optimal because rewarding disclosure lowers the deadweight losses
associated with default. More generally, we show that disclosure opportunities have heterogeneous real effects on financing and default, depending on whether they arise at times when firms are setting up their capital structure or not. These effects also vary predictably in the cross-section of firms and critically depend on market conditions.

Specifically, we build a model based on two ingredients. First, a dynamic agency conflict between firms and investors, due to the non-verifiable nature of cash flows. Specifically, managers can divert cash each period as in DeMarzo and Fishman (2007), which implies that investors use capital structure and the threat of default as means to prevent malfeasance by the firm’s insiders. Second, we introduce the option for firms to invest in a costly technology that randomly generates information about future performance. As in Dye (1985) or Shin (2003), once the technology is adopted, with some probability the firm’s manager privately observes a signal that predicts future cash flows. While the signal itself cannot be manipulated, it can be shrouded by the manager. The focus on optimal contracting distinguishes our paper from the recent disclosure literature.

The implementation proposed in DeMarzo and Fishman (2007) extends to our setting with minor adjustments. In equilibrium the firm starts off by borrowing – both short and long-term – and issuing equity. Dividends are paid after a sufficiently positive stream of earnings realizations, while persistent bad performance brings the firm to default and liquidation. Disclosure has the benefit of alleviating the moral hazard problem when cash flows are low. On the one hand, the manager can rely on evidence to show that a low performance is not due to diversion. On the other, the investors can condition their liquidity provision on the disclosed evidence without distorting moral hazard incentives. As a result, the funding liquidity of the firm at low cash flows depends on whether or not disclosure occurs. The firm’s funding liquidity can be thought of as the unused balance on its credit line, which evolves through time. In our model, the interest rate on funds withdrawn from the credit line depends on both the reported cash flows, and the evidence disclosed. Practically, such variations in rates can be thought of either as covenants attached to the short-term debt issuance (e.g., Smith and Warner (1979)), or as performance-sensitive debt (e.g., Manso, Strulovici and Tchistyi (2010)).

In particular, we find that – compared to the case of no evidence considered in DeMarzo and Fishman (2007) – interest rates are zero when a low cash flow is preemptively disclosed, while they are higher both when cash flows are low and there is no disclosure, and when cash flows are high. In other words, the optimal capital structure features pay

---

4See also Bolton and Scharfstein (1990), Biais et al. (2007), and DeMarzo and Sannikov (2006).

for verifiable bad luck. The result provides a novel rationale for ‘pay without performance’ (Bebchuk and Fried (2009)). Existing explanations tend to emphasize either capture of boards by powerful executives, or the desire of motivating innovation by managers – with the associated need to incentivize risk taking to compensate early failures (Manso (2011)). In contrast, we show that it may be in the shareholders’ interest to compensate the firm’s executives when bad performance is proven to be the consequence of bad luck. This goal is more likely to be achievable with the advent of technologies that facilitate the production and the analysis of performance-related quantitative evidence.

In light of the previous result, we can sign the effect of exercising the option on Pay-for-Performance Sensitivity (PPS): it is negative. To see why PPS must fall when the technology is adopted, consider first the PPS conditional on a signal being available. In this case, disclosure provides insurance to the firm in bad states of the world, which lowers the PPS. If, instead, the signal is not available, the PPS does not depend on the technology, exactly as in DeMarzo and Fishman (2007). Because the actual PPS is a convex combination of the two possibilities, it must decrease when the option is exercised. In addition, we find that – unlike in previous models such as Biais et al. (2007) or DeMarzo and Sannikov (2006) – in our model the PPS increases monotonically with the firm’s past performance history: PPS rises as performance improves.

Conditional on a given performance history, firms that adopted the technology and have a lower PPS are less likely to default, as one would expect. Indeed, credit spreads are smaller when insiders are more likely to have evidence to disclose. Empirically, this implies that – in secondary markets, when the firm’s capital structure is unchanged – credit spreads are negatively associated with the frequency of voluntary disclosures. This is consistent with the empirical findings of Balakrishnan et al. (2014), who consider firms facing an exogenous drop in their information environment and react by enhancing disclosure. Because major issuance events are unlikely to be frequent in their dataset, they are likely to capture primarily the effect of disclosure on secondary markets, which, consistent with our model, is positive.

Disclosure opportunities also affect the firm’s leverage and dividend payout. The firm issues dividends only after a sufficiently positive sequence of shocks, which implies that its dividend payout rates are positively correlated with its survival probability. Because firms that disclose information more frequently have higher survival probabilities – conditional on any given performance history – it follows that they also display higher dividend payout rates. In contrast, because leverage is the highest when the firm is close to liquidation, higher survival probabilities imply lower leverage. As a consequence, our model yields

---

6PPS is defined as the change in the insiders’ payoff over the change in cash flows across states.
predictions about the effect of voluntary disclosure on the firm’s capital structure in secondary markets.

Turning attention to primary markets – where capital structure choices are made – a trade-off arises. On the one hand, disclosure alleviates the moral hazard problem and increases a firm’s survival probability for any given level of liquidity. On the other, disclosure lowers the value of providing the firm with liquidity in the first place, and so it incentivizes investors to reduce the initial funding liquidity granted to the firm. As a result, the firm may well end up on paths that entail lower survival probability when the option is adopted. Which effect dominates depends critically on a firm’s profitability. We are not aware of work testing this hypothesis directly, but the result could partially account for the recent surge in average bond default rates (Becker and Ivashina (2019)).

In particular, we find that for high-profitability firms the beneficial effects dominate: disclosure yields higher survival probabilities, greater funding liquidity and lower credit spreads, as well as higher dividend payout rates and lower leverage. For low-profitability firms, instead, the negative effects prevail: credit spreads widen, leverage increases and dividend payouts fall. We solve analytically for a profitability threshold that separates these firms, but the threshold is only sufficient (not necessary) for non-monotonicity of credit spreads to arise. Firms more profitable than those at the threshold may be negatively impacted by the option for some parameter configurations.

Empirically, our results suggest that studies of the real effects of disclosure should pay attention to default risk, in addition to discount rates. While existing empirical work uncovered a positive causal effect of voluntary disclosure on the liquidity of a firm’s securities and on its cost of capital, our results suggest that such effects are mitigated by the negative impact disclosure has on expected cash flows at low profitability firms, which may be more likely to default when they are expected to disclose more frequently. In contrast, the two effects are amplified at high profitability firms, where disclosure not only lowers the cost of capital, but it also increases the expected cash flows generated by the firm due to its reduced probability of default.

As for the patterns of adoption, we predict that the set of adopting firms consists of those that experienced intermediate performance histories: the very profitable and the very unprofitable ones do not adopt the technology. The region is characterized by two performance-related thresholds. Below the lower threshold, the value of the firm as a going concern is too low to justify spending resources on the technology. Above the

---

7See Francis, Nanda and Olsson (2008), Balakrishnan, Billings, Kelly and Ljungqvist (2014) and Boone and White (2015), and the surveys by Healy and Palepu (2001) and Leuz and Wysocki (2008). Bertomeu and Cheynel (2016) survey theory work that interprets such findings in CAPM-like models.
upper threshold, the benefits of IT are too low, as the firm is already far from its default boundary. In the adoption region, the beneficial effect of disclosure outweighs the cost of exercising the option and investment is undertaken. As the cost of the technology falls, the thresholds diverge and the adoption region expands. Importantly, this pattern of adoption of our information-related real option is markedly different from that of alternative physical options, such as those studied in DeMarzo, Fishman, He and Wang (2012). Indeed, while in their model the value of the option is a monotonically decreasing function of past firm performance, in our model such value is non-monotonic: it is the highest for intermediate performance histories.

We also find that information technologies affect the valuation of some firms that did not adopt them yet. More precisely, there exist two other thresholds – one below the adoption region and one above it – such that firms located between these thresholds and the adoption region strictly benefit from having the option to delay their adoption. Concretely, low-performance firms wait to see their enterprise value increase, and condition their adoption on receiving a positive cash flow shock, while high-performance firms keep the option as an insurance policy, intending to adopt if they receive a sufficiently negative cash flow shock.

The paper unfolds as follows. Section 2 reviews the related literature. Section 3 presents the economic environment and the contract space. Section 4 considers two finite horizons versions of our model. A one-period example shows how evidence is irrelevant for static incentives, suggesting that, if evidence plays a role, it must be that it affects the dynamic incentive constraints. A two-period example shows how most of our results do not rely on the horizon being infinite, and conveys the intuition for some of our findings. Section 5 introduces the infinite horizon model. Section 6 discusses the impact of disclosure on the policy dynamics and on other variables of interest. Section 8 implements our optimal contract by means of short and long-term debt, and equity. Section 7 shows the patterns of information technology adoption. Section 9 discusses the initiation problem, when securities are issued. Section 10 concludes.

## 2 Literature Review

Our paper is related to several literatures. Theoretically, it builds on the dynamic agency model developed by Clementi and Hopenhayn (2006), Biais et al. (2007) and DeMarzo and Fishman (2007). A recent strand of papers on dynamic moral hazard introduced information production and dissemination possibilities, and studied their consequences on second best allocations (e.g., Fuchs (2007), Piskorski and Westerfield (2016), Smolin
The distinguishing feature of our model is that, while other papers focus on monitoring technologies where the principal acquires information directly, we assume that the realized information is observed by the agent and, to be payoff relevant, it must be voluntarily disclosed. This assumption captures the realistic feature that most firms’ investors are institutional or retail, and do not participate actively in the day-to-day operations of the firm.

Because we model information systems as technologies that produce disclosure opportunities for managers, à la Dye (1985) or Shin (2003), our work is also related to the theoretical work on voluntary disclosure (e.g., Beyer and Guttman (2012), Acharya, DeMarzo and Kremer (2011), Guttman, Kremer and Skrzypacz (2014), Marinovic and Varas (2016) and DeMarzo, Kremer and Skrzypacz (2017)). While these recent papers extended the Dye model to a dynamic setting, they differ from our setting in important ways. First, managerial compensation is exogenous, whereas we consider optimal compensation. This implies that their equilibria might feature strategic disclosure, whereas ours do not, due to Homlstrom’s informativeness principle (Holmstrom (1979)). Second, in some of these papers, evidence is potentially long-lived, and so the dimension of analysis is not only what is being disclosed, but also when managers disclose.

Our paper also relates to a recent literature that focuses on understanding the heterogeneous effects of information technologies on the cross-section of firms. In particular, Mihet and Philippon (2018) and Farboodi, Mihet, Philippon and Veldkamp (2019) focus on explaining the role of size in shaping adoption patterns and the consequences of adoption for the price informativeness of stock listed firms. Our work is complementary to this literature, in that we focus instead on how these technologies shape disclosure patterns by managers and, ultimately, on how they impact the informational landscape in which firms operate and interact with outside investors.

A related literature studies the consequences of real investment options for firms in dynamic agency models (e.g., DeMarzo, Fishman, He and Wang (2012), Bolton, Chen and Wang (2011)). Relative to this literature, we contribute by considering a different type of option which, instead of directly increasing the firm’s cash flows, improves the information available for the management to disclose the firm’s shocks to its investors. As we discussed in the introduction, such options have very different effects.

Our paper is also related to the literature emphasizing the possible negative real effects of a richer information environment. Most work on this topic assumes that the principal receives some information, but cannot commit to how the information is going to be used in determining some interim action (e.g., Crémer (1995), Meyer and Vickers (1997), Prat (2005) and Zhu (2018)). In contrast, in our model the agent receives the information and
needs to disclose it, while the principal has full commitment power.

Finally, at a more abstract level, our work is related to the literature discussing the role played by hard evidence in mechanism design problems. Since the seminal work of Bull and Watson (2004), the literature has recently flourished. Notable contributions are Koessler and Perez-Richet (2014), Hart, Kremer and Perry (2017) and Ben-Porath, Dekel and Lipman (2019). In this context, we are the first – to our knowledge – to consider the role of evidence in an otherwise standard dynamic agency setting.

3 Environment

A firm produces i.i.d. cash flows $x_t \in \{h, l\}$ for $t = 1, 2, \ldots, T$, where $h > l > 0$. Define $\Delta := h - l$, $p := P(x_t = h) \in (0, 1)$, and $\mu := \mathbb{E}(x_t)$. The firm is owned by a Principal (P) – who represents the firm’s investors – and is operated by a Manager (M). Both P and M are risk-neutral and discount future consumption at the same rate $r \in (0, 1)$.

Moral hazard. We introduce the possibility of moral hazard by assuming that M privately observes the realized cash flows $\{x_t\}$. By misreporting a good cash flow, claiming it to be bad, M can divert $\Delta$ output and obtains a private benefit of $\delta := \lambda \Delta$, where $\lambda \in (0, 1]$ represents the severity of the moral hazard problem. From the revelation principle, we can restrict communication protocols to direct messages that report $x_t$, and focus on the implementation of truthful reporting.

Evidence. We assume that P can choose to invest in an information technology that produces evidence $e_t \in \{g, b\}$ each period with probability $\hat{\pi} \in (0, 1)$. To ease notation, $\pi$ in the paper denotes a random variable that takes values of either 0 or $\hat{\pi}$, depending on whether the technology has been adopted ($\pi = \hat{\pi}$), or not ($\pi = 0$). To make this investment, P must spend a fixed cost of $c \geq 0$. Evidence consists of verifiable information that cannot be manipulated, and which perfectly predicts cash flow $x_t$: good evidence implies high cash flows, while bad evidence implies low cash flows. IT adoption effectively corresponds to the exercise of a one-time American option with infinite maturity that cannot be reversed, with strike price $c$.

Once the option is exercised, P expects evidence to be available with probability

\footnotetext{\(^8\)Common discounting is not needed to derive our qualitative results, but it simplifies the arguments.\(^9\)The notation here is not redundant: the effects of $\Delta$ on allocations and contracts are slightly different from those of $\lambda$, in ways that we will emphasize while discussing the comparative statics.\(^10\)The fact that in the absence of technological investment $\pi = 0$ is just a normalization. All our qualitative results go through unchanged if we assumed that, absent investment, the firm would have a positive $\pi < \hat{\pi}$.\(^11\)The cost can be thought of as the presented discounted value of the setup and maintenance expenses.}
\(\hat{\pi}\), but she never knows whether \(M\) possesses evidence or not. So, at each date \(t\), \(M\) chooses whether or not to *voluntarily disclose* the realized evidence to \(P\). We denote the disclosure action by \(a_t \in A := \{d, n\}\), where \(d\) stands for disclosure and \(n\) for non-disclosure. If \(M\) discloses the evidence, investors will predict the cash flow accurately. That is, \(p(x_t = h|e_t = g) = p(x_t = l|e_t = b) = 1\).\(^{12}\) Because (i) disclosure is always incentivized, and (ii) the availability of evidence is conditionally independent from the realized cash flow, non-disclosure has no impact on the investors’ beliefs. That is, absent evidence disclosure, \(P\) predicts that cash flows are high with probability \(p\).

**Contracting.** To maximize investors’ value, \(P\) offers \(M\) a contract that specifies, for every history of reports and disclosures, the probability of liquidating the firm \(\theta_t \in [0, 1]\), and the cash compensation \(u_t \geq 0\). In the first best case, the firm is never liquidated and has the value of \(s^* := \frac{\mu(1+r)}{r}\). If the firm is liquidated, both parties get their outside option payoff, which is normalized to zero. Figure 1 shows the timing of events in a generic period \(t\), prior to the exercise of the IT-investment option.

![Figure 1: Timing in period \(t\), prior to exercising the option](image)

4 **Finite-horizon model**

To highlight the key driving forces behind our results, we start with a static and a two-period versions of the model. For simplicity, we set \(r = 0\) in this section.

\(^{12}\)Since the effects of evidence on our outcomes of interest are already non-monotone and complex with perfect evidence, it does not seem necessary to also consider imperfect correlation in this model.
Figure 2: Event tree of static setting

One-period setting. Figure 2 draws the event tree when $T = 1$. The set of possible outcomes is $\mathcal{H}_1 := \{dh, dl, nh, nl\}$, and cash compensations to $M$ are denoted by $u_i$ for $i \in \mathcal{H}_1$. The contract must provide two kinds of incentives: (i) to prevent the agent from diverting cash flows, which requires $u_{nh} \geq \delta + u_{nl}$; (ii) to disclose information, which requires both $u_{dh} \geq u_{nh}$ and $u_{dl} \geq u_{nl}$. It is optimal for $P$ to set $u_{nl} = u_{dl} = 0$ and $u_{dh} = u_{nh} = \delta$, and since $c > 0$ the option is never exercised.

Two-period setting. It follows from the one-period case that at $t = 2$ evidence is irrelevant. So, the set of relevant final histories is $\mathcal{H}_2 := \{ahh, ahl, alh, all\}_{a \in A}$, where the first element $a \in \{d, n\}$ denotes $M$’s disclosure action in the first period; the second and third elements denote the $t = 1$ and $t = 2$ realized cash flows, respectively. As in most dynamic agency models (e.g., Biais et al. (2007)), committing to liquidate the firm when $x_1 = l$ may be optimal, because it alleviates the diversion problem and reduces the rents required for incentive compatibility to hold. Now, disclosures matter and the optimal contract is shaped by both the cost $c$ and the benefit $\pi$ of exercising the option.

Proposition 1. If $T = 2$, there exists a $\bar{c}$ such that if $c \geq \bar{c}$ the option is never exercised, while if $c < \bar{c}$ there exist two profitability thresholds $p$ and $\bar{p}$ such that $p < \bar{p}$ and:

(a) If $p \in [p, \bar{p}]$, the option is exercised and the probability of default is $(1 - p)(1 - \hat{\pi})$;

(b) If $p < p$, the option is not exercised and the firm’s probability of default is zero;

(c) If $p > \bar{p}$, the option is not exercised and the firm’s probability of default is $1 - p$.

Because $\partial p / \partial c > 0$ and $\partial \bar{p} / \partial c < 0$, a reduction in the strike price of the option $c$ increases the probability of default of low profitability firms, while it increases the probability of default of high profitability firms.
In the two-period case, evidence might enable P to distinguish bad luck from bad behavior in the first period, and affect the optimal termination policy. Figure 3 shows that the option to invest at a strike price $c$ and produce evidence with probability $\pi$ attracts firms that are neither too profitable ($p \leq \bar{p}$), nor too unprofitable ($p \geq \underline{p}$). When the two value functions (conditional on whether the option is exercised or not) are tangent, the option is only exercised by firms with $p = \hat{p}$. If the cost drops to $c < \bar{c}$, the set of firms that exercise the option expands to $p \in [\underline{p}, \bar{p}]$. High profitability firms choose not to invest and terminate when $x_1 = l$. Low profitability firms, in contrast, never terminate and do not exercise the option either. If $c > \bar{c}$ the option is never exercised.

Figure 3: The set of firms that exercise the option as the strike price $c$ changes

Thus, a reduction in the strike price of the option $c$ leads to increased adoption and more disclosure by both profitable and unprofitable firms. However, Figure 4 shows that its effect on default probabilities and credit spreads is heterogeneous across firms. For high-profitability firms that switch to exercising the option (right-panel), default probabilities decrease as disclosure avoids inefficient termination. For low-profitability firms that switch (left-panel), the opposite occurs. While at a high cost $c$ they had a low (zero) default probability, as $c$ drops evidence provides a tool for P to reduce the rents paid to M, while not defaulting the firms when disclosure occurs. As a result, the firm’s default probability rises. Together with the cost $c$, this amounts to an increase in the deadweight losses and a reduction in the social surplus, even though it increases the investor’s payoff.
The two-period model is stylized, in that either the firm exercises the option immediately, or it never does. The timing does not allow one to study how the possibility of delaying the exercise time shapes the dynamics and the initial conditions. However, it highlights the complex interplay between the strike price of the option, the frequency of disclosure and the firm’s default risk. In the next section, we analyze these forces in the full infinite-horizon model, where the option to adopt the information technology can be exercised with delay.

5 Infinite-horizon model

In this section, we first formulate the firm’s problem in the infinite-horizon environment, and then characterize policies and their dynamic features.\textsuperscript{13}

5.1 Contracting

As is well known, when shocks are i.i.d., the agent’s continuation utility $v$ is a state variable that summarizes all relevant information in any given history. For any state $v$, the contract specifies the probability of liquidating the firm at the beginning of the period $\theta$, and then compensates M either with cash, or with promised utility contingent on M’s actions. When evidence is disclosed, the contract pays $u_d = (u_{dh}, u_{dl}) \in \mathbb{R}^2$ to M and promises continuation utility $w_d = (w_{dh}, w_{dl}) \in \mathbb{R}^2$, depending on whether the high or the low cash flow is reported. Similarly, when no evidence is disclosed, the contract pays M cash either $u_n = (u_{nh}, u_{nl}) \in \mathbb{R}^2$, and promises continuation utility $w_n = (w_{nh}, w_{nl}) \in \mathbb{R}^2$.

\textsuperscript{13}More rigorous arguments which guarantee that the recursive representation of our problem is appropriate are standard and so we leave them to the Appendix.
Whether it is worthwhile to invest in the costly information technology or not, and if so, when to make the investment, all depend on the value that this option brings to the firm. To evaluate the moneyness of this evidence-generating option, we first consider the optimal contracting for the firm given the investment has already been made. We then step back and determine the optimal option exercise patterns.

Given that the investment in the information technology has been made, evidence regarding future cash flows arrives with probability $\hat{\pi}$. Because our programming that solves the firm’s policies in this scenario also applies to the scenario where evidence never arises (or the investment option is never exercised), we use the variable $\pi$ to represent both, with the indication of $\pi = \hat{\pi}$ for the former and $\pi = 0$ for the latter.

Before we define the firm’s problem, we consider the diversion and disclosure incentive constraints. First, since the manager can always conceal evidence, any voluntary disclosure has to be contractually incentivized. Contracts may disregard evidence in some states of the world. However, because of Holmstrom’s informativeness principle, it only makes sense that evidence disclosure is either promoted, or overlooked; it should never be actively prevented. That is, whenever the manager obtains good evidence:

$$ u_{dh} + \frac{w_{dh}}{1+r} \geq u_{nh} + \frac{w_{nh}}{1+r} \quad (IC_g) $$

Likewise, whenever the manager obtains bad evidence we have:

$$ u_{dl} + \frac{w_{dl}}{1+r} \geq u_{nl} + \frac{w_{nl}}{1+r} \quad (IC_b) $$

Second, when the manager does not disclose good evidence, he can always report a low cash flow and divert $\Delta$. So, the diversion incentive compatibility demandszAA:

$$ u_{nh} + \frac{w_{nh}}{1+r} \geq \delta + u_{nl} + \frac{w_{nl}}{1+r} \quad (IC_n) $$

Any feasible contract must fulfill its promises and deliver the given continuation value. In other words, the optimal contract satisfies a promise-keeping constaint which requires:

$$ v = (1 - \theta) \left[ \pi E_d \left( u_d + \frac{w_d}{1+r} \right) + (1 - \pi) E_n \left( u_n + \frac{w_n}{1+r} \right) \right], \quad (PK) $$

where, to ease notation, we define M’s expected utility conditional on evidence disclosure as $E_a(u_n + \frac{w_n}{1+r}) = p(u_{ah} + \frac{w_{ah}}{1+r}) + (1-p)(u_{al} + \frac{w_{al}}{1+r})$ for $a = d, n$. In addition, contracts
must satisfy limited liability, i.e.:

\[ u_{dh}, u_{nh}, u_{dl}, u_{nl} \geq 0 \quad (LL) \]

Because the agents share the same discount factor, it follows that the optimal contract from P’s perspective also maximizes firm value (i.e., surplus), given a utility \( v \) promised to M.\(^ {14} \) Thus, the optimal contract solves the following dynamic program:

\[
s(v) = \max_{\theta, u_j, w_j} (1 - \theta) \left\{ \mu + \frac{1}{1 + r} \left[ \pi E_d(s_d) + (1 - \pi)E_n(s_n) \right] \right\}
\]

s.t. \((PK), (IC_g), (IC_b), (IC_n), (LL)\),

where \( s(v) \) denotes the expected firm value, \( s_a = (s(w_{ah}), s(w_{al})) \) for \( a = d, n \), and \( E_a(s_a) = ps(w_{ah}) + (1 - p)s(w_{al}) \) denotes the expected firm values conditional on possible disclosure actions.

The objective function of \((S)\) reflects the fact that (i) with probability \( \theta \), liquidation takes place before the subsequent evidence and cash flow realize, in which case the firm value drops to zero; and (ii) with probability \( (1 - \theta) \) the firm is not liquidated, in which case the firm value depends on whether M receives the leading evidence or not, and whether the cash flow is high or low. Because the two events are independent, we can express the expected firm value as that in the objective of \((S)\).

On the one hand, the program \((S)\) with \( \pi = \hat{\pi} \) solves the firm’s problem given the investment option has already been exercised. On the other hand, if \( \pi = 0 \), the program exactly solves the case where the option is never exercised (no evidence ever possible). This is because in the latter case, the only relevant control variables are those conditional on no disclosure. Hence, we use \( s(v; \hat{\pi}) \) and \( s(v; 0) \) to denote the value functions of \((S)\) for these respective cases.

### 5.2 Investment option

We next analyze the investment decision, i.e. the decision of whether and when to exercise the investment option. Suppose that, for a given history represented by \( v \), the firm has not yet exercised the option. The firm’s value in this scenario is denoted as \( f(v) \) and, obviously, no evidence will be disclosed today. If the firm is not liquidated – which occurs with probability \( (1 - \theta) \) – then it obtains the expected cash flow \( \mu \) today and proceeds

\(^{14}\)This does not imply that a contract that maximizes P’s expected utility is socially optimal: in general, P starts the contract from a socially suboptimal initial condition – we shall return to this point.
to tomorrow’s state of either $w_{nh}$ or $w_{nl}$, depending on the cash flow reported by M. Come tomorrow, the firm can either invest $c$ and obtain the value of $s(w_{ni}) - c$ (where $i = h, l$) from the subsequent date onwards, or delay investment again and obtain the value of $f(w_{ni})$. It is easy to see that when $s(w_{ni}) - c > f(w_{ni})$, the firm exercises the investment option tomorrow. Otherwise, it waits until at least one more period to invest. Hence, the firm’s problem when the investment has not been undertaken yet can be formulated as follows:

$$f(v) = \max_{θ, u_j, w_j} \left\{ \mu + \frac{1}{1 + \rho} E_n \left[ \max(f_n, s_n(\hat{\pi}) - c) \right] \right\}$$

s.t. $(PK), (IC_n), (LL)$

where $f_n = (f(w_{nh}), f(w_{nl}))$, and $\pi = 0$ in $(PK)$.

### 5.3 Initiation and payout

When the firm is initiated at time zero, $P$ promises a continuation utility $v_0$ to maximize its expected profits over the lifetime of the firm. That is,

$$v_0 = \arg \max_v \{ \max[f(v), s(v; \hat{\pi}) - c] - v \}$$

Clearly, at the outset, the firm may either exercise the option right away or wait to make the investment later.

Because liquidation is inefficient, it may be optimal to delay M’s cash compensation until the continuation utility $v$ is sufficiently large. Throughout the paper, we adopt the payout policy that M is paid by cash, if the firm is indifferent between paying him or delaying the payment, which is without of generality. Formally, we define the cash payout boundary as the smallest continuation utility where the firm value reaches its first best. That is,

$$\bar{v} := \inf \{ v : f(v) = s^* \text{ or } s(v; \hat{\pi}) = s^* \}$$

The definition implies that the firm value (with or without evidence) is strictly less than the first best $s^*$ before the continuation utility reaches $\bar{v}$. In general, both the payout boundary and M’s payoff dynamics may depend on the availability of evidence. The next result shows that actually the value $\bar{v}$ is a constant, irrespective of evidence availability, but the cash compensation varies with the option exercise strategy and level of $\hat{\pi}$ in the short-run.
Proposition 2. The cash payout boundary is:
\[ \bar{v} = r^{-1}(1 + r)\rho \delta. \] (3)

Moreover, for \( a \in \{d, n\} \), the optimal cash compensation is
\[ u_{ai}(v) = 0, \quad u_{ah}(v) = \max \{(1 + r)\delta - (1 + \hat{r})(\bar{v} - v), 0\} \] (4)

where \( \hat{r}(\pi) := \frac{r}{1-(1-p)\pi} \).

Proposition 2 shows that the payout boundary \( \bar{v} \) does not depend on the possibility of generating and disclosing evidence. Even when the cost of generating evidence is infinitely large (or \( \pi \) is infinitely small or zero), the boundary does not change. Then, no cash payment is made to M if low cash flow is reported. When the firm is one-step away from \( \bar{v} \), M receives cash compensation upon reporting high cash flow. In addition, the result says that the cash payment is reduced if evidence is more likely to be available. However, when the firm reaches \( \bar{v} \), the cash compensation is independent of evidence. The case of \( \pi = 0 \) in (4) indicates that the investment option has not been exercised. The investment option has no impact here, because after this cash payment, the firm is at \( \bar{v} \) where the investment will not be made.

This is intuitive: the payoff boundary \( \bar{v} \) is the smallest continuation utility at which incentive constraints cease to bind. Once that boundary has been reached, evidence is no longer useful. Because at \( \bar{v} \) the firm is never liquidated, cash compensation is not delayed further. As long as evidence is uncertain, the magnitude of cash compensation is \( \delta \), independent of evidence, whenever good performance reported, because M can always disguise good news and divert cash flows.

6 Impact of evidence disclosure

The decision to invest in the information technology depends on the value added by the availability of evidence IT brings about, net of the strike price \( c \). In this section we first characterize the firm’s problem given that the investment has already been made. In the next section we examine which firms exercise the option and under what conditions they exercise. To highlight the role of evidence disclosure, we consider what happens if the evidence is more or less available (the intensive margin), and then contrast the policies with the benchmark case as in DeMarzo and Fishman (2007) where evidence is never available (the extensive margin). This benchmark case corresponds to our model...
in which the option is never exercised.

6.1 Policy characterization

We first characterize the firm’s problem \((S)\). Recall that it solves two possible scenarios: the option already exercised \((\pi = \hat{\pi})\), and the option never exercised \((\pi = 0)\).

Before reaching the payout boundary, \(M\) is incentivized by variations in her promised continuation values. If the continuation value ever grows high enough, the firm is never liquidated, all constraints of the firm’s problem become slack, and firm value reaches the first best. When \(M\)’s continuation value is at intermediate levels, liquidation may occur after a sequence of low cash flows and no disclosures. When the continuation value is low enough, the only way to both align incentives and fulfill commitment is to stochastically liquidate the firm at the beginning of the period. To characterize the dynamics, we define the thresholds such that no liquidation can possibly occur in the next \(n\) periods to be:

\[
v^n := \inf\{v : \text{no liquidation in at least } n \text{ periods}\}, \quad \text{for } n = 0, 1, 2, ...
\]

These values correspond to the lowest continuation values such that the firm can survive with certainty for at least \(n\) periods. For example, if \(v > v^1\) the firm will not be liquidated in the current (or one) period, but may be liquidated in the next period. These thresholds are related to the previous definitions of liquidation probability \(\theta\) and the payout boundary \(\tilde{v}\). Specifically, stochastic liquidation at the beginning of any period is positive \((\theta(v) > 0)\) if and only if \(v < v^1\). In addition, the payout boundary \(\tilde{v}\) is the limit of this sequence of thresholds \(v^\infty\); indeed, liquidation ceases if \(v \geq \tilde{v}\).

Because the firm can be liquidated, any randomization of continuation utility is costly for both parties, implying that the firm value \(s(\cdot)\) is concave. Using concavity and the optimal conditions of the firm’s problem \((S)\), we can show which constraints bind and derive the optimal policies.

**Lemma 1.** For any \(v < \tilde{v}\) in the firm’s problem \((S)\), the constraints \((IC_g)\) and \((IC_n)\) bind while \((IC_b)\) holds as strict inequality.

One can immediately see that, contingent on the high cash flow being reported, evidence is payoff irrelevant: \(w_{dh} = w_{nh}\). In other words, as long as the investors receive a high cash flow, the payoffs to both \(M\) and \(P\) are not affected by evidence disclosure. Contingent on a good performance being reported, \(M\) does not divert and hence there is no need to further condition payoffs on evidence disclosure. Therefore in the rest of the paper we do not distinguish \(M\)’s payoff across states \(dh\) and \(nh\). Accordingly, we denote
$w_h$ and $u_h$ as the continuation value and the cash payment, respectively, conditional on cash flows being high. Notice that this property would not hold in a monitoring setting, in which $P$ observes the realized signal directly. In that case, $P$ could further reduce the payment to $M$ when evidence is available, without violating any incentive constraint.

In contrast, the optimal contract provides strict incentives for $M$ to disclose bad news: i.e., $w_{dl} > w_{nl}$. Punishing $M$ for a bad performance is costly to $P$ because it induces more inefficient liquidation. If the evidence shows that the bad performance is not caused by $M$’s behavior, but instead by bad luck, then $M$ should not be punished. Promising $M$ higher utility in the state $dl$ does not worsen the diversion problem, but improves efficiency by reducing the probability of liquidation.

Given the active constraints and the optimality conditions of the firm’s problem $(S)$, we obtain an explicit solution for the optimal policies:

**Proposition 3.** The optimal policies for the firm are as follows:

- For $v \in (0, v^1]$: $\theta = \frac{v^1 - v}{u} - v$, $w_{nl} = 0$, $w_{dl} = v^1$, $w_h = \min \left\{ \frac{r\bar{v}}{p}, \bar{v} \right\}$;
- For $v \in (v^1, \bar{v}]$: $\theta = 0$, $w_{nl} = v - \hat{r}(\bar{v} - v)$, $w_{dl} = v$, $w_h = \min \left\{ w_{nl} + \frac{r\bar{v}}{p}, \bar{v} \right\}$;
- The $n$-period liquidation thresholds are $v^n(\pi) = [1 - (\frac{1}{1+r})^n] \bar{v}$.

If stochastic liquidation does not occur at the beginning of the period, the policies $w_i(v)$ for $i \in H_1$ are the same as $w_i(v^1)$. In addition, firm value in this region is linear. Given this characterization, we clearly see that – after a low performance – the contract possibly promises the manager a higher continuation utility when bad news is disclosed ($w_{dl} = v^1 > v$), deviating significantly from the case of $\pi = 0$ considered in previous work. In this region, whenever $M$ discloses evidence of transitory bad luck $P$ compensates for disclosure by raising the promised utility to $v^1$, independently from the continuation utility entering the period. The reward depends on $v^1 - v$.

In the region above $v^1$, $M$ is still rewarded for disclosing bad luck: $w_{dl} = v$. The contract forgives the low performance today, and starts tomorrow as if the history is the same as before the current low cash flow. This mechanism does not affect the $M$’s diversion incentives, because $M$ can never mimic the type who discloses evidence that the cash flow will be low. Moreover, volatility in continuation utility is costly for investors because liquidation is inefficient. So, it is optimal to set $w_{dl}$ as close to $v$ as possible.

Finally, as standard, the optimal contract rewards good luck. Proposition 3 shows that the ranking of continuation utility does not depend on the levels of $v$ and $\pi$. $M$ gets

\[15\text{Notice that to keep the continuation utility fixed across times effectively requires a payment from } P \text{ to } M, \text{ because of the time value of money.}\]
the largest continuation utility contingent on high performance, the lowest one contingent on low performance and no disclosure, and the middle one contingent on disclosure of bad news. This pattern implies that, on the fastest route to liquidation, M never discloses evidence and always reports low cash-flow. So, the n-period threshold can be explicitly derived from the policy functions. Evidently, both the liquidation thresholds and the policy dynamics depend on M’s disclosure behavior and on the availability of evidence.

6.2 Comparative statics

Having characterized the optimal policy functions, we can examine how the optimal contract varies with the probability that the information technology produces evidence. These comparative statics highlight the impacts of the quality of the information technology. We focus on the impact of evidence disclosure on firm value, default risk, managerial compensation and firm dynamics. This analysis is specific to dynamic models because, regardless of the initial conditions of the problem, any \( v \in [0, \bar{v}] \) is on-the-equilibrium path. That is, there always exists a sequence of shocks that can take the firm from \( v_0 \) to any such \( v \). Before presenting the results, it is useful to provide a formal definition of both credit spreads and Pay-Performance Sensitivity (PPS). We follow the literature and define PPS as:

\[
PPS := \frac{\mathbb{E}(v \mid x_1 = h) - \mathbb{E}(v \mid x_1 = l)}{\Delta} \tag{5}
\]

This measure indicates in percentage terms how M’s compensation changes with firm performance. Normally, we have that Credit spread = \((1 - \text{recovery rate}) \times \Pr.[\text{default}]\), but because we normalized the recovery rate to zero,

\[
\text{Credit (or CDS) spread} = 1 - \frac{s}{s^*}, \tag{6}
\]

where \( s^* \) denotes the first best value of operating the firm, and \( s \) denotes the value at the constrained best, as implemented by our optimal contract (i.e., the sum of the principal’s and the manager’s expected payoff). Importantly, this is distinct from the agency cost, which would be defined as Credit spread + managerial rents.

**Proposition 4.** When the availability of evidence \( \hat{\pi} \) is higher, the optimal contract exhibits the following comparative statics, for any given \( v < \bar{v} \):

(a) firm value \( s \) increases or, equivalently, its credit spread falls;

(b) pay-performance sensitivity falls, while it increases with \( v \) for any given \( \hat{\pi} \)

(c) \( w_h \) and \( w_{nl} \) both weakly fall, while \( w_{dl} \) stay constant;
(d) the $n$—step liquidation threshold $v^n$ falls, for $n = 1, 2, ...$

![Figure 5: Firm and Investor Value](image)

Part (a) of Proposition 4 shows the overall effect of evidence on firm value and investor payoff. Given a promised utility to the manager, more evidence disclosure increases expected firm value $s(v)$. It follows that the investors’ expected payoff, $s(v) - v$, rises with $\hat{\pi}$, for any given value of $v$. Figure 5 plots numerical examples of these effects. The result also clarifies that the investors’ time-zero payoff (i.e., the highest value in each curve of the right plot) must increase with the availability of evidence. This is because investors choose where to start the firm so as to maximize their time-zero payoff.

![Figure 6: Simulated Liquidation Probability](image)

To clarify how evidence affects the probability of default in secondary markets, i.e., conditional on a given performance history $v$, we simulate the policy dynamics as shown
in Figure 6. The figure illustrates a cross section of firms that are identical except for having different chances of evidence disclosure. After certain histories, if these firms promise the same expected utility to their managers, then a firm with a higher chance of disclosing evidence has a lower probability of being eventually liquidated. This is the main reason why social efficiency and investors’ payoff both improve.

We can explicitly characterize the firm’s PPS from Proposition 3 as:

$$PPS = \lambda - \frac{\pi \hat{r} [\bar{v} - \max(v, v^1)]}{(1 + r)\Delta}$$  \hspace{1cm} (7)$$

Obviously, this measure depends on both $v$ and $\hat{\pi}$. An example is displayed in Figure 7. When $\pi = 0$, the PPS measure equals $\lambda$, regardless of $v$. This result holds in existing models of dynamic moral hazard without evidence disclosure – e.g., DeMarzo and Fishman (2007). Once we introduce evidence, a few effects emerge that are worth stressing. First, the PPS decreases with $\pi$, suggesting that cash flow incentives and disclosure are substitutes in the optimal contract. More evidence reduces PPS because the manager’s payoff is unassociated with low performance if the evidence predicts its bad luck. Second, the PPS also depends on $v$ itself. In particular it increases with $v$, converging to $\lambda$ as $v$ approaches $\bar{v}$. Larger PPS induces higher variations in continuation value, exacerbating the inefficient liquidation. When $v$ is small, liquidation is more severe, making a large PPS more costly for the firm. Allowing evidence disclosure significantly reduces PPS and the chance of liquidation. Evidence disclosure has no effect on PPS only at the payout boundary $\bar{v}$, where the firm has no chance of being liquidated.

Previous results give the explicit forms of the policies $\{w_i\}_{i \in H_1, \theta}$, and the n-period liquidation boundary $v^n$. Part (c) of Proposition 4 analytically illustrates the properties of the firm evolution when the disclosure frequency $\pi$ varies.

![Figure 7: Pay-for-Performance Sensitivity](image-url)
The result shows the impact of disclosure on how continuation utilities evolve. The managerial payoff $w_{dl}$ either remains the same as the beginning of period utility $v$, or it jumps up to $v^1$, regardless of the level of $\pi$. From the promise-keeping constraint, we know that both $w_h$ and $w_{nl}$ must fall (weakly if $v < v^1$), because the continuation utility is more likely to stay at $v$. The diversion incentive constraint binds, establishing that the gap $w_h - w_{nl}$ must be constant and equal to $(1 + r)\delta$. Figure 8 illustrates this pattern of how policies move as $\pi$ increases for a given value of $v$ (or a given history). As $\pi$ increases, the continuation utility is less likely to move downward, but the lowest value becomes worse.

Finally, part (d) of Proposition 4 shows that the availability of evidence has a downside: the n-period liquidation thresholds $v^n$ all increase as evidence become more available. For the cross-section of firms starting at the same continuation utility, the shortest time to be liquidated becomes shorter as $\pi$ goes up. In other words, on the fastest way to liquidation, a firm with a better evidence technology can be liquidated faster. In contrast, the shortest time to reach the payout boundary becomes longer as $\pi$ goes up. These patterns are plotted in Figure 9, which verifies the previous illustration that the lowest continuation utility becomes lower as $\pi$ increases. According to the characterization in Proposition 3, we know that the fastest way to liquidation on the equilibrium path occurs when a firm never discloses evidence and experiences a sequence of low cash flows.
7 Investment option

The value of evidence in our model comes from two channels. First, as described before, the availability of evidence increases firm value by avoiding inefficient punishment on M. Second, because the value of evidence is endogenous and varies with continuation utility, the optional delay of paying the cost can be valuable to the firm. Clearly, if the investment cost is too large, then the option is never exercised. The largest cost at which the investment option is ever exercised is given by:

$$\bar{c} = \max_v \{s(v; \hat{\pi}) - s(v; 0)\}$$  \hspace{1cm} (8)

Moreover, the option is not exercised right away either in the region close to $\bar{v}$ or in the region close to 0. Accordingly, there are two thresholds that reflect these two regions.

$$v^l = \inf_v \{f(v) \leq s(v; \hat{\pi}) - c\}, \quad \text{and} \quad v^h = \sup_v \{f(v) \leq s(v; \hat{\pi}) - c\}$$  \hspace{1cm} (9)

**Proposition 5.** The threshold cost is $\bar{c} > 0$. Both $\bar{c}$ and $v^h$ increase in the evidence availability $\hat{\pi}$, but $v^l$ decreases in $\hat{\pi}$. Further, $v^h$ decreases in the investment cost $c$, while $v^l$ increases in $c$.

If the firm never exercises the investment option, its value is $s(v; 0)$ given a history represented by $v$. Obviously, the firm value with the investment option $f(v)$, as defined in (F), is no less than its baseline value of $s(v; 0)$. Therefore, the value of exercising the option $s(v; \hat{\pi}) - f(v)$ is smaller than $s(v; \hat{\pi}) - s(v; 0)$, which is further smaller than the
cost if \( c > \bar{c} \). Hence, the option is never exercised. The value \( \bar{c} \) reflects \( \hat{\pi} \) and other parameters including the severity of agency \( \lambda \), the profitability of the firm \( p \), and so on.

When \( c < \bar{c} \), the investment option is possibly exercised at certain \( v \). When the continuation utility \( v \) is close to \( \bar{v} \), the probability of default becomes very small. Evidence that prevents the inefficient termination becomes not that valuable. Hence, the value of delaying investment \( f(v) \) becomes very close to the first best level \( s^* \), which is strictly larger than the value of exercising the option right away \( s(v; \hat{\pi}) - c \). When \( v \) is close to 0, the firm value is small with or without evidence disclosure. The value of exercising the option right away \( s(v; \hat{\pi}) - c \) becomes tiny or even negative, which is smaller than the value of waiting to invest. The following result summaries the possibility of exercising the investment option, and when to exercise it.

**Proposition 6.** The firm never exercises the investment option if \( c \geq \bar{c} \). Otherwise, there exists a nonempty interval \( v \subset (v^l, v^h) \) where investment option is exercised right away, while the option is delayed for \( v \in [0, v^l] \cup [v^h, \bar{v}] \).

Figure 10 illustrates the result of Proposition 6. It plots a numerical example of different firm values over continuation utility when \( c < \bar{c} \). The green line plots the firm value if the option is never exercised which is \( s(v; 0) \) or the case of DeMarzo and Fishman (2007). The blue line plots the firm value if the investment option is exercised right away which is \( s(v; \hat{\pi}) - c \). The red line plots the firm value \( f(v) \) of delaying the IT investment. The investment is made if \( v \in (v^l, v^h) \), but delayed if \( v \geq v^h \) or \( v \leq v^l \). The
difference between the red and green line in Figure 10 reflects the option value of the IT investment. This difference is large at intermediate levels of \( v \), because for these levels of continuation utility, the firm is likely to spend more time in the constrained stage and therefore more likely to exercise the investment option eventually. On the left and right tails, this difference shrinks, because the firm is likely to be liquidated or to reach the first best, not exercising the investment option. When this option value becomes large and surpasses the value of investing right away, it is better to delay and wait.

Proposition 5 shows that if evidence is more available, then the firm is more likely to exercise the investment option overall, and the region of delay to invest shrinks. If instead the investment cost is higher, then the firm is more cautious, or more likely to delay the investment until its accumulative performance moves to a smaller middle range.

8 Capital Structure Implementation

This section implements the optimal contract using standard financial securities. To facilitate comparison with dynamic moral hazard models that do not have the possibility of disclosure (e.g., DeMarzo and Fishman (2007)) the securities in our implementation only include equity, long-term debt, and a credit line (short-term debt).

The long-term debt claim is essentially a perpetuity that pays a fixed coupon every period forever. The credit line defines the amount of credit that can be withdrawn by the firm anytime within the (endogenous) limit \( z \). The firm’s debt capacity, which is the difference between the credit limit and its balance, proxies the firm’s liquidity level. Finally, the equity component is a claim against the firm’s dividend payments. Given any capital structure, the manager controls the firm’s liquidity and payout policies. More precisely, the manager determines how and when to withdraw from (or repay to) the credit line, and how and when to pay dividends.

Within this set of securities, disclosure affects the evolution of the credit line and. In particular, it determines its interest rate. In our model, any balance on the credit line account is charged an interest rate \( \hat{r}_i \), for \( i \in \mathcal{H}_1 \), that is contingent on both performance and disclosure. In contrast to models such as DeMarzo and Fishman (2007), here investors may sometimes charge a higher interest rate than \( r \), or they can forgive part or all of the current period interest charge.

The credit account balance reflects any repayment at the beginning of each period before the firm cash flow realizes. The following result summarizes a security design that implement the optimal contract.\(^\text{16}\)

\(^{16}\)Evidently, as in all other security design problems, such design can never be unique.
Proposition 7. Under the following security and compensation design, the manager always discloses any evidence that might be available, and cash flows are used to either repay coupon and credit balance or to issue dividends.

- The manager holds $\lambda$ fraction of the firm equity;
- the long-term debt coupon is $l$;
- the credit line has limit $z = \frac{\bar{v}}{\lambda}$, and contingent interest rate $\hat{r}_{dl} = 0$ and $\hat{r}_{i\neq dl} = \hat{r}(\pi)$.

The firm only issues dividends after it pays off the credit balance and the coupon.

In the implementation, the credit balance or borrowed short-term debt, denoted as $m$, summarizes the history and functions as the state variable. It maps one-to-one to the state variable $v$ of the firm’s problem $(S)$. On the one hand, the manager can borrow all the available credit and pay it out as dividend. Thus, the continuation value of the manager in the firm must be at least $\lambda(z - m)$. On the other hand, the investors will not leave more information rent (in the form of liquidity) than necessary to the manager. Hence the continuation utility of the manager must be

$$v = \lambda(z - m)$$

which must hold at any history. Given this relation, as well as the policy dynamics in Proposition 3, we can figure out how the firm’s short-term debt evolves, which further implies the interest rates specified in Proposition 7.

To further illustrate the mechanisms, let us consider how the credit balance evolve over time. Suppose that the firm starts certain period with credit balance $m$. It pays the long-term debt coupon $l$ from the cash flow. The interest rate on the credit line is a constant value $\hat{r}$ unless bad cash flow news is disclosed when the interest rate becomes zero. The credit balance in the following period denoted as $m_{i} \in H_{t}$ will be

$$m_{h} = (1 + \hat{r})m + (1 + r)(d_{h} - \Delta)$$
$$m_{nl} = (1 + \hat{r})m$$
$$m_{dl} = m$$

where $d_{h} = \frac{\mu}{\lambda}$ is the dividend payout. If a bad news is disclosed, then interest is forgiven in the current period, and the new balance will stay the same. If the high cash flow realizes, the firm is charged a interest rate of $\hat{r}$ but will have $(1 + r)\Delta$ more cash to repay the short-term debt in the next period (independently from disclosure). Therefore, the
new credit balance $m_h$ follows (11). If low cash flow realizes and no evidence disclosed, then $\hat{r}$ is charged toward the beginning balance $m$ and the new balance $m_{nl}$ follows (12).

As shown in Proposition 7, one important feature of our model is that the equity holdings, the long-term debt coupon, and the credit limit do not depend on the availability of evidence: only the short-term interest rate does. The equity holdings determine how the residual cash flow (or dividends) are split between the manager and investors. In our model, when the firm starts paying out dividends, it has no possibility of being liquidated and surplus reaches the first best. In that stage, evidence is disclosure is payoff irrelevant. The necessary way to incentivize the manager is for him to hold $\lambda$ fraction of dividend payments.

However, the interest rates of the credit line affect the evolution of the firm’s short-term debt holding. Since evidence disclosure does affect firm liquidity in the short-run transition, the interest rates must vary with the manager’s disclosure decisions. The variation in interest rates is essentially to incentivize the manager to disclose bad news. It is easy to see from Proposition 7 that the average interest rate is exactly $r$, but the interest gap between disclosing bad news or not is $\hat{r}_{dl} - \hat{r}_{nl} = \hat{r}$, and it increases with $\pi$. In other words, as the probability of the manager possessing evidence increases, investors on average still earn the risk free rate $r$, but they will design a larger interest rate variation to induce disclosure of bad news.

9 At issuance date

Now that we have characterized the optimal contract, we can examine the agent’s rent, the firm value, and the default probability at the issuance date (or time zero). In particular, we will show how these values change as the cost of producing evidence varies.

**Proposition 8.** If $p$ is smaller than some $\hat{p} \in (0, 1)$, then at the issuance date, the agent’s rent $v_0$, the firm value $\max\{f(v_0), s(v_0; \hat{\pi}) - c\}$, and the credit spread can be all increasing or all decreasing, in the investment cost $c$.

In general, the time-zero properties are hard to characterize, because they reflect the expectation of all future firm performances and evidence disclosures. Note that the firm’s optimal starting point $v_0$ depends on the marginal value of raising continuation utility to the firm value. The key feature of our model is that as evidence becomes more available this marginal value (at $v_0$) can be either larger or smaller. On the one hand, the continuation utility is less likely to move downward as $\pi$ increases, making the marginal value of continuation utility smaller. In this case, the firm initiates by granting the agent
lower expected rent, which may drive the initial firm value downward. On the other hand, if there is no disclosure, the continuation utility can drop to a lower level, leading to a smaller firm value. The marginal value of continuation utility can become larger to hedge such situation. In this case, the firm initiates at larger $v_0$, which implies a larger firm value at issuance.

It’s not hard to understand that as generating evidence becomes cheaper the firm value can be higher because P has an additional channel to govern the agency conflict. But the opposite can also be possible. In fact, Proposition 8 confirms the intuition obtained in the two-period model that the credit spread of low profitability firms may actually increase as disclosure becomes possible. Managers who are expected to have access to evidence may be worse off than the less informed ones.

![Credit Spread and Investor Payoff](image)

Figure 11: Credit Spreads and Investor Value

Intuitively, this occurs because to incentivize disclosure and prevent diversion P faces the trade-off of either loading on M’s rents or raising the termination odds, both of which are costly. When the firm is likely to obtain low cash flows, the chance of terminating the firm is high and therefore termination is more costly. If the investment cost $c > \bar{c}$, the optimal policy loads more on managerial rents (larger $v_0$). As the cost drops, evidence is possibly produced and disclosed, which alleviates the termination concern. So the optimal policy loads on less rents (lower $v_0$).

## 10 Conclusions

We study the implications of embedding voluntary disclosure of evidence à la Dye (1985) in an otherwise standard dynamic agency model with non-verifiable cash flows, similar
to Biais et al. (2007). The model captures three key empirical regularities: (i) technological progress (as well as regulation) increasingly promotes the use of evidence about performance; (ii) evidence is decentralized, namely, it is typically better observed and understood by a firm’s management, than by its arm’s length stakeholders. So, its disclosure needs to be costly incentivized; (iii) evidence has a forward looking dimension: it is useful to predict future cash flows.

The optimal contract can still be implemented by simple securities, as proposed in DeMarzo and Fishman (2007): the firm borrows both short and long term, and retains a fraction of its equity. The main difference here is that the interest rate has to depend on both performance and evidence disclosure; in particular, when bad news are disclosed the one-period interest charges are forgiven by the principal. In all other circumstances, the interest rate is higher the more widespread evidence availability is.

We find that the presence of evidence reduces the optimal pay-for-performance sensitivity, because it enables the investors to condition their short-term liquidity prevision on both the reported cash flows and the evidence produced by the management. If the managers can convince the investors that a temporary negative performance is due to bad luck, as opposed to bad behavior, the investors can cut the firm some slack and accept a temporary relief on interest payments.

While this beneficial effect of evidence reduces the firm’s credit spread in secondary markets, when no capital structure decisions are made, the result may reverse in primary markets. Here, both the firm’s initial liquidity and its credit spread might be non-monotonic functions of disclosure. Namely, better evidence might lower firm value at the constrained optimal allocation, exacerbating the conflict between rent extraction by the principal and efficiency. This occurs especially at low profitability firms, because better evidence reduces the marginal value of providing initially financial slack to the firm, so that P trades-off higher liquidation odds with a lower managerial pay level.

We characterize the policy dynamics and show that evidence brings about two countervailing dynamic forces. On the one hand, it increases the persistence of a firm’s liquidity conditional on the firm experiencing poor performance that can be surely attributed to bad luck (i.e., conditional on bad news being disclosed). On the other hand, by requiring higher interest rates on the short-term debt, it deteriorates faster the firm’s liquidity when bad news are not disclosed and performance is poor.

Our numerical simulations suggest that while generating a relatively small increase in stakeholder’s value, evidence can dramatically reduce efficiency, increasing the liquidation odds and the minimal time required to reach the liquidation boundary, as well as inducing volatility spikes in continuation utilities for managers and in liquidation odds.
Importantly, the inefficiency induced by more frequent evidence disclosure that we derive arises in a model where the principal has full commitment power; it does not depend on the presence of time inconsistencies such as limited commitment.

We believe that our results might be a useful benchmark both to construct more complicated models that take seriously the possibility of limited commitment, and empirical exercises that seek to establish the relation between a firm’s liquidity and either the frequency of its disclosure, or its price impact.

References


A Appendix

Proof of Lemma 1. The $T = 1$ case taught us two facts: (i) whether $P$ has exercised or not the option, this has no effect on the last period implementable payoffs; (ii) $P$ could find it optimal to exercise the option at the beginning only if she terminates in the first period, after a low state is reported and no evidence is disclosed. So, there are only three policies to consider: $TT$ ($P$ does not exercise the option and terminates in the first period when $x = l$), $NT$ ($P$ does not exercise the option and never terminates) and $OT$ ($P$ exercises the option in the first period and terminates only when when $x = l$ and there is no disclosure). Wlog we can set all payments when the last cash flow is $l$ to zero.

Under the policy $TT$, there is one payment to determine: $u_{hh}$, that is, the payment to $M$ after two successes. The payment satisfies two ICs: (i) at date 2, $u_{hh} \geq \delta + u_{hl} = \delta$; (ii) at date 1: $pu_{hh} \geq \delta$. It follows that $u_{hh} = \delta/p$. $M$’s utility at this policy is $U_M(TT) = p\delta$; $P$’s utility is $U_P(TT) = (1 + p)(l + p\Delta) − p\delta$.

Under the policy $NT$, we need to determine two payments: $u_{lh}$ and $u_{hh}$. While $u_{lh}$ only satisfies $u_{lh} \geq \delta$, $u_{hh}$ satisfies both $u_{hh} \geq \delta$ (at date 2) and $pu_{hh} \geq \delta + p\delta$ (at date 1, where we plugged the optimal $u_{lh} = \delta$). It follows that $u_{hh} \geq \delta(1 + p)/p$. $M$’s utility at this policy is $U_M(NT) = 2p\delta$; $P$’s utility is $U_P(NT) = 2(l + p\Delta) − 2p\delta$.

Under the policy $OT$, we need to determine three payments: $u_{dlh}$ and $u_{ahh}$ (for $a \in \{n, d\}$). However, from the disclosure IC we can see that $u_{nhh} = u_{dhh} := u_{hh}$, and so the problem reduces to solving for $u_{hh}$ and $u_{dlh}$. For similar reasons as before, $u_{dh} = \delta$. As for $u_{hh}$, it must be the same as in policy $T$, because the only feasible deviation from $x = h$ is to claim that $x = l$ without disclosing evidence. So, $u_{hh} = \delta/p$. $M$’s utility is $U_M(OT) = p\delta(1 + (1 − p)\hat{\pi})$; $P$’s utility is $U_P(OT) = (2 − (1 − p)(1 − \hat{\pi}))(l + p\Delta) − p\delta(1 + (1 − p)\hat{\pi}) − c$.

First, when comparing $TT$ and $NT$ we obtain a threshold $\hat{p}$ such that:

$$\hat{p} := \frac{l(1 − \Delta) + \delta − \sqrt{4l^2\Delta + (l(1 − \Delta) − \delta)^2}}{−2l\Delta}$$

If $p > \hat{p}$, $P$ strictly prefers $TT$; if $p < \hat{p}$, $P$ strictly prefers $NT$; if $p = \hat{p}$, $P$ is indifferent between the two policies (or any randomization of the two policies).

Second, either $P$ exercises the option at $\hat{p}$, or she never does. So, fixing $p = \hat{p}$ and comparing $U_P(NT)$ and $U_P(OT)$ yields the threshold $\bar{c}$:

$$\bar{c} := \frac{\lambda\pi}{2l^2\Delta} \left[ l^2(1 + \Delta^2) + \delta^2 + l(1 − \Delta)2\delta − (\delta + l(1 − \Delta))\sqrt{l^2(1 + \Delta)^2 + 2l(1 − \Delta)\delta + \delta^2} \right]$$

Focusing on $c < \bar{c}$, we need to consider two cases. If $p > \hat{p}$, we need to compare $U_P(TT)$
and $U_P(OT)$. We find that $U_P(OT) \geq U_P(TT)$ if and only if $p \leq \bar{p}$, where:

$$\bar{p} := \frac{l\pi(\Delta - 1) - \pi \delta + \sqrt{\pi(4\Delta(1 - \lambda))(l\pi - c) + \pi(l(1 - \Delta) + \delta)^2})}{2\pi \Delta(1 - \lambda)}$$

Finally, if $p < \hat{p}$, we need to compare $U_P(NT)$ and $U_P(OT)$. We find that $U_P(OT) \geq U_P(NT)$ if and only if $p \geq p$, where:

$$p := \frac{(1 - \pi)l(\Delta - 1) - \delta + \sqrt{(l(1 - \Delta - \delta)(1 - \pi))^2 + 4(c + l(1 - \pi))(l\Delta(1 - \pi) + \pi \delta)}{2(l\Delta(1 - \pi) + \pi \delta))}$$

Note that $c$ enters $\bar{p}$ under the square-root and has a negative sign, while it enters $p$ only under the square-root with a positive sign. It follows that $\partial \bar{p}/\partial c < 0$ and $\partial p/\partial c > 0$.

To proceed with the proofs for the infinite-horizon model, let us first show some basic properties of the surplus function. Let $C$ be the space of continuous and bounded functions on the domain $R_+$. Let $F := \{q \in C : 0 \leq q \leq s^*\}$ be endowed with the 'sup' metric where $s^* = \frac{(1 + r)p\delta}{r}$ is the first best surplus. It’s easy to see that $F$ so defined is a complete metric space. Define the Bellman operator $T : F \to F$ as:

$$Tq(v) = \max_{\theta, \pi, u_d, w_i} (1 - \theta)\mu + \frac{1 - \theta}{1 + r} \left\{ \pi [pq(w_{dh}) + (1 - p)q(w_{dl})] + (1 - \pi)[pq(w_{nh}) + (1 - p)q(w_{nl})] \right\}$$

s.t. $(PK), (IC_g), (IC_b), (IC_n), (LL)$

It’s standard to show that the Bellman operator $T : F \to F$ is well defined and the constraint set is convex. Moreover, we can show the Bellman operator has the following property.

**Lemma A.1.** Let $F_1 = \{q \in F : q(v) = s^* \text{ for all } v \geq \frac{(1 + r)p\delta}{r}\}$. If $q \in F_1$, then $T1 \in F_1$.

**Proof.** Take any $q \in F_1$ and $v \geq \frac{(1 + r)p\delta}{r}$. Consider the following policy:

$$\theta = \pi = u_{dl} = u_{nl} = 0, \quad u_{dh} = u_{nh} = \frac{v}{p} - \frac{\delta}{r}, \quad w_i = \frac{(1 + r)p\delta}{r} \forall i \in H_1$$

It’s easy to check that this policy satisfies all the constraints of the firm’s problem $(S)$. In addition, under this feasible policy we know

$$(Tq)(v) \geq \mu + \frac{q(v)}{1 + r} = \frac{(1 + r)\mu}{r} = s^*$$

35
Hence, we must have \((Tq)(v) = s^*\).

**Proposition A.1.** The unique fixed point of \(T\), which we call \(s(v)\), is increasing, concave, and satisfies \(s(v) = s^*\) for any \(v \geq \frac{(1+r)p\delta}{r}\).

**Proof.** It is easy to see that \(T\) is monotone (whereby \(q_1 \leq q_2\) implies \(Tq_1 \leq Tq_2\)) and satisfies discounting (wherein \(T(q + a) = Tq + \delta a\)). Then the Blackwell’s theorem implies \(T\) is a contraction mapping on \(\mathcal{F}\) and hence has a unique fixed point in \(\mathcal{F}\). Let \(\mathcal{F}_2 = \{q \in \mathcal{F} : q(v)\) is increasing and concave for all \(v \in R_+\}\). It’s standard to show that \(T\) maps from \(\mathcal{F}_2\) to \(\mathcal{F}_2\). Combining Lemma A.1, we must have that the unique fixed point of \(T\) lies in \(\mathcal{F}_1 \cap \mathcal{F}_2\).

**Proof of Proposition 2.** Given the fact that \(s(v)\) reaches first best for a large enough \(v\), we can define \(v_1 = \inf\{v : s(v) = s^*\}\) and \(v_2 = \inf\{v : f(v) = s^*\}\).

First, consider the program \((S)\), and let \(\{\theta, u_i, w_i\}_{i \in \mathcal{H}_i}\) be the optimal policy at \(\bar{v}\). Note that, to achieve first best, all continuation values \(w_i \in \mathcal{H}_i\) must be no less than \(v_1\), and the liquidation probability \(\theta\) is zero. In addition, from the constraints \((IC_g),(IC_b), (IC_n)\), and \((LL)\), we must have

\[
(1 + r)w_{dl} + w_{dl} \geq (1 + r)u_{nl} + w_{nl} \geq v_1 \quad (\text{A.14})
\]
\[
(1 + r)w_{dh} + w_{dh} \geq (1 + r)u_{nh} + w_{nh} \geq (1 + r)\delta + v_1 \quad (\text{A.15})
\]

Then \((PK)\) implies \(v_1 \geq p\delta + \frac{\nu}{1+r}\), or \(v_1 \geq \frac{(1+r)p\delta}{r}\). Moreover, the conclusion of Proposition A.1 implies that \(v_1 \leq \frac{(1+r)p\delta}{r}\). Hence, we must have \(v_1 = \frac{(1+r)p\delta}{r}\).

Now consider the program \((F)\). To achieve first best, the investment option must be never exercised, and moreover, the policies \(w_i \geq v_2\). So similarly, we can show that \(v_2 \geq \frac{(1+r)p\delta}{r}\). In addition, since \(f(v) \geq s(v;0)\), Proposition A.1 implies that \(v_2 \leq \frac{(1+r)p\delta}{r}\). Therefore, we must have \(v_2 = \frac{(1+r)p\delta}{r}\).

To characterize the policies, we specify the first order conditions and the envelope condition of program \((S)\) as follows. Denote \(\eta\) as the Lagrangian multiplier of \((PK)\). Moreover, let \(\alpha_g, \alpha_b, \alpha_n\) be the multipliers of \((IC_g)\), \((IC_b)\), and \((IC_n)\), respectively. Then the first order conditions are:

\[
(1 - \theta)\pi p s'(w_{dh}) = (1 - \theta)\pi p\eta - \alpha_g \quad (\text{FOC}_{dh})
\]
\[
(1 - \theta)\pi (1 - p) s'(w_{dl}) = (1 - \theta)\pi (1 - p)\eta - \alpha_b \quad (\text{FOC}_{dl})
\]
\[
(1 - \theta)(1 - \pi) s'(w_{nh}) = (1 - \theta)(1 - \pi)\eta + \alpha_g - \alpha_n \quad (\text{FOC}_{nh})
\]
\[
(1 - \theta)(1 - \pi)(1 - p) s'(w_{nl}) = (1 - \theta)(1 - \pi)(1 - p)\eta + \alpha_b + \alpha_n \quad (\text{FOC}_{nl})
\]
The envelope condition is:

\[ s'(v) = \eta \quad \text{(EN)} \]

**Proof of Lemma 1.** Take any \( v < \bar{v} \), and let \( \{\theta, \pi, w_{i \in H_1}\} \) be the optimal policies of the program \((S)\) with \( \pi = \hat{\pi} \).

First, we show that \( \alpha_b = 0 \). Suppose not. Then by the first order conditions \((FOC_{dl})\) and \((FOC_{nl})\), we must have \( s'(w_{dl}) < s'(w_{nl}) \), which further implies that \( w_{dl} > w_{nl} \) by the concavity of \( s(\cdot) \). In other words, the constraint \((IC_b)\) holds as strict inequality. The complementary slackness then implies \( \alpha_b = 0 \), a contradiction.

Second, we show that it cannot be the case that \( \alpha_g = \alpha_n = 0 \). Suppose this is true. Then all the incentive constraints are not binding. Therefore, the surplus \( s(v) \) should be the same as if we solve the firm’s problem \((S)\) with only the promise keeping constraint \((PK)\). Since \( w_i = (1 + r)v \) is feasible, we know from the objective of \((S)\) that \( s(v) \geq s^* \), a contradiction with \( v < \bar{v} \).

Third, we show that \( \alpha_n > 0 \). Suppose not. Then from the above result, it has to be that \( \alpha_g > 0 = \alpha_n \). Then the first order conditions \((FOC_{nh})\) and \((FOC_{nl})\) together imply that \( s'(w_{nh}) > s'(w_{nl}) \). Hence, \( w_{nh} < w_{nl} \), contradicting with \((IC_n)\).

Fourth, we show that \( \alpha_g > 0 \). Suppose not. Then by \((FOC_{dh})\) and \((FOC_{nh})\), we know \( s'(w_{dh}) > s'(w_{nh}) \) which further implies that \( w_{dh} < w_{nh} \). This forms a contradiction with \((IC_g)\).

Last, using the facts of \( \alpha_n > 0 = \alpha_b \), we can conclude from \((FOC_{dl})\), \((FOC_{nl})\), and \((EN)\) that \( s'(w_{dl}) = s'(v) < s'(w_{nl}) \). Hence, concavity implies \( w_{dl} > w_{nl} \).

**Proof of Proposition 3.** The proof is divided into two parts. Part (a) shows that the optimal policies described in the Proposition satisfy all the constraints and optimality conditions of the firm’s problem \((S)\). Part (b) derives the \( n \)-period liquidation thresholds.

**Part (a).** Take any \( v \leq \bar{v} \), and let \( \{\theta, w_h, w_{dl}, w_{nl}\} \) be the optimal policies specified in the proposition. It’s easy to verify that they satisfy all the constraints of \((S)\). Using the result of Lemma 1, the first order conditions \((FOC_{dh})\) to \((FOC_{nl})\), and the envelop condition \((EN)\), we can eliminate all the Lagrange multipliers and obtain

\[
[1 - (1 - p)\pi]s'(v) = ps'(w_h) + (1 - p)(1 - \pi)s'(w_{nl})
\]  
(A.16)

This optimality condition along with \((PK)\) and \((IC_n)\) jointly determine the optimal policies. Hence, we are left to show that \((A.16)\) always holds under the proposed policies.

First, consider the case of \( v \geq v^1 \). In this region, \( \theta = 0 \). Plug in \( w_h, w_{dl}, w_{nl} \) to the
objective of \((S)\) and then differentiate it with respect to \(v\) to obtain

\[
(1 + r)s'(v) = p(1 + \hat{r})s'(w_h) + (1 - p)\pi s'(v) + (1 - p)(1 - \pi)(1 + \hat{r})s'(w_{nl}) \tag{A.17}
\]

Then plug in the expression \(\hat{r} = \frac{r}{1 - (1 - p)\pi}\) to (A.17) and rearrange will give exactly (A.16).

Second, in the case of \(v < v^1\), we can plug in \(\theta, w_{dl}, w_{nl}\) to the objective of \((S)\) to obtain

\[
s'(v) = s(v^1) = s'(v^1).
\]

Since \(w_{dl}, w_{nl}\) do not vary with continuation utility when it moves from \(v^1\) to \(v\) and since (A.16) holds at \(v^1\), it must also hold at \(v\).

**Part (b).** Notice that at the n-period thresholds the following must hold:

\[w_{nl}(v^1) = 0,\]

and \(w_{nl}(v^j) = v^{j-1}\) for \(j \geq 2\). According to the optimal policy of \(w_{nl}\), the latter implies

\[v^{j-1} = w_{nl}(v^j) = v^j - \hat{r}(\bar{v} - v^j)\]

Hence, \(v^j = \frac{1}{1 + \hat{r}}[v^{j-1} + \hat{r}\bar{v}]\). Moreover, by the optimal policies and \((PK)\), we have

\[(1 + r)v^1 = p(1 + r)\delta + (1 - p)\pi v^1 = r\bar{v} + (1 - p)\pi v^1\]

which implies \(v^1 = \frac{p}{1 + \hat{r}}\bar{v}\). Finally, we can obtain the threshold expression simply by induction. \(\Box\)

**Proof of Proposition 4.** Part (a). Consider any \(\pi < 1\) as a parameter of the firm’s problem \((S)\). Take any continuation value \(v\) at the no-liquidation region \([v^1, \bar{v}]\). Let \(w_{dl}, w_{nl}\) be the optimal policies at \(v\). Denote \(s_\pi(v)\) to be the derivative of \(s\) with respect to \(\pi\) at the fixed value \(v\). Then the envelop condition with respect to \(\pi\) is

\[
\frac{(1 + r)s_\pi(v)}{1 - p} = s(w_{dl}) - s(w_{nl}) - s'(w_{dl})(w_{dl} - w_{nl}) \tag{A.18}
\]

Since \(s(\cdot)\) is concave, \(w_{nl} < w_{dl}\), and \(s'(w_{dl}) < s'(w_{nl})\) (see the last part of the Lemma 1 proof), we must have \(s_\pi(v) > 0\). In addition, since \(s(\cdot)\) is linear in \(v\) for any \(v < v^1\), continuity of \(s(\cdot)\) implies that \(s_\pi(v) > 0\) for \(v < v^{(1)}\).

**Part (b).** See the proof in Proposition 2.

**Part (c).** Take any \(v \in [v^1, \bar{v}]\) (no-liquidation region). According to the definition in
(5), the pay-performance sensitivity can be calculated as

\[
PPS = \frac{w_h + (1 + r)w_h - [\pi w_{dl} + (1 - \pi)w_{nl}]}{(1 + r)\Delta}
\]

\[
= \frac{\pi w_{nl} + (1 + r)\delta - \pi w_{dl}}{(1 + r)\Delta}
\]

\[
= \frac{\pi[v - \hat{r}(\bar{v} - v)] + (1 + r)\delta - \pi v}{(1 + r)\Delta}
\]

\[
= \lambda - \frac{\pi \hat{r}(\bar{v} - v)}{(1 + r)\Delta}
\]

The second line is from the fact of \((IC_n)\) being equality, while the third line is from plugging in the policy expressions of \(w_{dl}, w_{nl}\). Obviously, PPS is decreasing in \(\pi\) and increasing in \(v\).

Now consider the liquidation region. Since our PPS measure is only defined when the firm is not liquidated in the beginning of the period, we can simply replace the \(v\) in the above derivation by \(v^{(1)}\) and get

\[
PPS = \lambda - \frac{\pi \hat{r}(\bar{v} - v^1)}{(1 + r)\Delta} = \frac{\lambda(1 + r - \hat{\pi})}{1 - (1 - p)\bar{\pi} + r}
\]

Obviously, in this case, PPS does not depend on \(v\) and decreases in \(\hat{\pi}\).

Proof of Proposition 5. For \(v \in (0, \bar{v})\), since \(\hat{\pi} > 0\) and Proposition 4 shows that \(s(v)\) strictly increases in \(\pi\), we must have \(s(v; \hat{\pi}) > s(v; 0)\). Then the continuity of \(s(\cdot)\) implies \(\bar{c}\) in (8) is well defined, and \(\bar{c} > 0\). Moreover, the threshold cost \(\bar{c}\) increases in \(\hat{\pi}\) because \(s(v; \hat{\pi})\) does.

When \(\hat{\pi}\) increases, note that the increase of \(s(v; \hat{\pi})\) is lager than that of \(f(v)\). This is because by the problem (\(F\)) the increase of \(f(v)\) is due to future increase in firm value when the option is exercised. Hence, \(v^h\) becomes larger and \(v^l\) becomes smaller.

When \(c\) increases, similar argument implies that the decrease of \(s(v; \hat{\pi})\) is lager than that of \(f(v)\). Hence, \(v^h\) becomes smaller and \(v^l\) becomes larger.

Proof of Proposition 6. If \(c = 0\), then the option is exercised right away because \(s(v; \hat{\pi}) \geq s(v; 0)\) for any \(v\). In this case, \(v^l = 0\) and \(v^h = \bar{v}\).

Now consider any \(0 < c < \bar{c}\). First, not that in the region close to the boundary \(\bar{v}\), not exercising the option is optimal. This is because \(f(\bar{v}) = s^* > s(\bar{v}; \hat{\pi}) - c\). Then \(v^h\) as in (9) is well defined, and the option is not exercised for \(v \in [v^h, \bar{v}]\).

Second, not exercising the option is optimal in the region close to the boundary \(0\).
This is because \( f(0) = 0 > s(0; \hat{\pi}) - c \). Hence, \( v^l \) as in (9) is well defined, and the option is not exercised for \( v \in [0, v^l] \).

Third, suppose the option is also delayed for \( v \in (v^l, v^h) \). Then the option is never exercised, and \( f(v) = s(v; 0) \) for all \( v \in [0, \bar{v}] \). However, since \( c < \bar{c} \), there are some \( v \) such that \( s(v; \hat{\pi}) - c > s(v; 0) = f(v) \), implying exercise the option is optimal at such \( v \). This is contradiction.

In the case of \( c > \bar{c} \), we have \( s(v; \hat{\pi}) - c < s(v; 0) \) for any \( v \), by the definition of \( \bar{c} \) in (8). Then \( s(v; \hat{\pi}) - c < f(v) \) for any \( v \), implying the option is never exercised.

**Lemma A.2.** When the investment option is not exercised, the relevant optimal continuation utility \( w_h \) and \( w_{nl} \) are the ones in Proposition 3 by setting \( \hat{r} = r \).

**Proof.** Take any \( v < \bar{v} \). We first show that the \((IC_n) \) constraint in problem \((F) \) must hold as equality at \( v \). Suppose not. Then \( f(v) \) should be the same as if we solve \((F) \) without \((IC_n) \). Then since \( w_{nh} = w_{nl} = (1 + r)v \) and \( \theta = 0 \) is feasible, we know from the objective of \((F) \) that \( f(v) \geq s^* \), a contradiction with \( v < \bar{v} \).

Then the optimal policy at \( v \) is jointly determined by \((PK) \) and the equality of \((IC_n) \), resultin in the expressions in Proposition 3 with \( \pi = 0 \).

**Proof of Proposition 7.** We will show that under the arrangement implements the payout and the evolution of the optimal contract.

First, in the cash payout region the balance of credit line becomes zero or \( m = 0 \). The agent’s total payoff is \( \lambda \) fraction of the firm payout with is \( \lambda d_h = u_h \). And according to (11) (12) (13), the subsequent credit line balance will all be zero, or \( m_h = m_{nl} = m_{dl} = 0 \), whichimplis \( w_h = w_{nl} = w_{dl} = \bar{v} \) by (10). So in the implementation, the agent gets the same cash as in the contract, and the firm always stays at \( \bar{v} \), having no probability of defaulting.

Second, in the none-cash-payout region, given the current credit balance is \( m \), we can use (10) to (13) to derive the subsequent credit line balance as:

\[
m = m_{dl} = \frac{\bar{v} - v}{\lambda} \\
m_{nl} = \frac{[1 + \hat{r}(\pi)](\bar{v} - v)}{\lambda}, \quad m_h = m_{nl} - (1 + r)\Delta
\]

So by (10), M’s continuation utility (by withdrawing all available credit) becomes

\[
w_{dl} = v, \quad w_{nl} = v - \hat{r}(\pi)(\bar{v} - v), \quad w_h = w_{nl} + (1 + r)\delta
\]

which is the same as in Proposition 3. Note this derivation includes the both senarios of
whether the investment option is exercised or not. In the latter case, we only have the subsequent credit balance to be \( m_n \) or \( m_h \), and the continuation utility to be \( w_{nl} \) or \( w_{lh} \). These values are obtained by setting \( \pi = 0 \) in the above expressions. Hence, by Lemma A.2 we know these continuation utilities are the same as in the option contract.

**Lemma A.3.** In the case of \( p \leq r \), we have \( s(v) = a_i + b_i v \) for any \( v \in [v^i, v^{i+1}] \), where \( i = 0, 1, 2 \ldots \), and the coefficients satisfy

\[
a_0 = 0, \quad b_i = \frac{\mu(r + p)(1 + \hat{r})}{r p \delta}[\hat{r}(1 - p)(1 - \pi)/r]^i
\]

(A.19)

**Proof.** According to Proposition 3, when \( p \leq r \) we will have \( w_{hl}(v) \geq \bar{v} \) for any \( v > 0 \). In other words, the firm will immediately reach the payout boundary \( \bar{v} \) after any high cash-flow shock, conditional on the firm is not liquidated in the beginning of the period. From this observation, we can derive \( s(\cdot) \) by induction.

When \( v \in (0, v^1) \), the objective of (A.1) implies

\[
s(v) = \frac{v}{v^1} \left( \frac{1}{1 + r} \mu + \frac{1}{1 + r} [ps^* + (1 - p)\pi s(v)] \right)
\]

from which we get \( s(v^1) = \frac{(1 + r)(r + p)\mu}{r(1 + r - (1 - p)\pi)} \). Hence, \( b_0 = \frac{s(v^1)}{v^1} = \frac{\mu(r + p)(1 + \hat{r})}{r p \delta} \).

When \( v \in [v^i, v^{i+1}] \) for \( i \geq 1 \), we have \( w_{nl}(v) = (1 + \hat{r})v - \hat{r}\bar{v} \). Then the objective of (A.1) implies

\[
[1 + r - (1 - p)\pi]s(v) = (1 + r)\mu + ps^* + (1 - p)(1 - \pi)s(w_{nl})
\]

and therefore

\[
a_i + b_i v = \frac{(1 + r)\mu + ps^*}{1 + r - (1 - p)\pi} + \frac{(1 - p)(1 - \pi)}{1 + r - (1 - p)\pi} \{a_{i-1} + b_{i-1}[(1 + \hat{r})v - \hat{r}\bar{v}]\}
\]

Equating the coefficients, we get \( b_i = \frac{\hat{r}}{\hat{r}}(1 - p)(1 - \pi)b_{i-1} \). The result is then obtained by induction.

According to (A.19), the slope \( b_i \) is a function of \( p \) and \( \pi \). We denote it as \( b_i(\pi; p) \), and denote the derivative of \( b_i \) with respect to \( \pi \) as \( b'_i(\pi; p) \).

**Lemma A.4.** Take any \( i \geq 1 \). There exists \( p \leq r \) such that \( b'_i(0; p) > 0 \). Moreover, \( b'_i(0; p = r) < 0 \).
Proof. From the expression of $b_i$ in (A.19), we get

$$b'_i(0; p) = \frac{\mu(r + p)}{rp\delta}(1 - p)^i[(1 - p)r - i(1 + r)p]$$

Clearly, $b'_i(0; p) > 0$ if $p$ is close to zero, while $b'_i(0; p = r) < 0$.

Lemma A.5. There exists some $j \geq 1$ and $p = r = \bar{r}$, such that $b_j(0) = 1$.

Proof. From the expression of $b_i$ in (A.19), we get

$$b_i(0; p = r) = \frac{\mu(r + p)}{rp\delta}(1 - p)^i$$

Clearly, $b_i(0; p = r)$ decreases in $r$ and decreases in $i$. And $b_0(0; p = r) = 2$ and $b_\infty(0; p = r) = 0$. Hence, there must exist $j$ and $\bar{r}$ such that $b_i(0; p = \bar{r}) = 1$.

Proof of Proposition 8. Note that if $c = \bar{c}$ the firm never exercises its investment option, and at initiation, the investors choose

$$\hat{v}_0 = \arg\max_v \{s(v; 0) - v\}$$

Consider the values $j$ and $\bar{r}$ in Lemma A.5, and let $p = r = \bar{r}$. Then we must have $s'(\hat{v}_0; 0) = 1$, and $\bar{v}_0 = v^{j+1}$. Given the result of Lemma A.4, we know there exists $\bar{\pi} > 0$ such that $b_j(\pi; \bar{r}) < 1$ for any $\pi < \bar{\pi}$.

On the contrary, if $c = 0$, the firm exercises the option at initiation. Hence,

$$\check{v}_0 = \arg\max_v \{s(v; \bar{\pi}) - v\}$$

If $\bar{\pi}$ is sufficiently small and $\bar{\pi} < \bar{\pi}$, the above argument implies that $\check{v}_0 \leq v^j < \hat{v}_0$. And because $s(v; \pi)$ is continuous in both $v$ and $\pi$. A small change in $\pi$ and a downward jump of initial continuin utility from $\check{v}_0$ to $\bar{v}_0$ implies $s(\check{v}_0; \bar{\pi}) < s(\check{v}_0; 0)$. So the firm value and M’s rent at initiation must both have a increasing part in $c$.

Now consider the case where $p$ is close to zero. Lemma A.4 implies $s'(\check{v}_0; 0) < s'(\check{v}_0; \bar{\pi})$ for a sufficiently small $\bar{\pi}$. Hence, $\check{v}_0 > \hat{v}_0$, which further implies $s(\check{v}_0; \bar{\pi}) > s(\check{v}_0; 0)$, since $s(\cdot; \pi)$ is increasing in $\pi$. Therefore, the firm value and M’s rent at initiation must both have a decreasing part in $c$. 

42