Voluntary disclosure, moral hazard and default risk *

Shiming Fu[1] and Giulio Trigilia[2]

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

October 15, 2019

Abstract

We study a dynamic moral hazard setting where the manager has private evidence that predicts the firms cash flows. When performance is low, bad news disclosure is rewarded by a lower borrowing cost relative to the no-evidence case. In contrast, no disclosure is associated with higher borrowing costs. On net, disclosure weakens the relation between cash incentives (or stock values) and short-term performance, which lowers the firm's default risk on path. However, the expectation of future disclosure might increase the optimal initial level of debt, to a point where the firm's default risk actually rises relative to the no-evidence case.

Key words: voluntary disclosure, credit spreads, default risk, dynamic moral hazard, funding liquidity, information technologies, real options

JEL classification: G32, D86, D61

*Previously circulated under the title: "Voluntary disclosure under dynamic moral hazard". We thank Dan Bernhardt, Michael Gofman, Ron Kaniel, Ilan Kremer, Meg Meyer, Alan Moreira, Dmitry Orlov, Heikki Rantakari, Cliff Smith, Vish Viswanathan, Neng Wang, 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), UIUC, Tsinghua (PBCSF), CRETA (Warwick), Conference on financial markets and corporate finance (SUFE), Rochester (Simon), FTG Summer (LBS), SED (Mexico City), NASMES (Davis), MWET (Drexel), AMES (Xiamen). All errors are our own.

1 Introduction

Over the last two decades, technological progress drastically reduced the costs of generating, analyzing and diffusing information. Corporate managers have access to better evidence that leads the performance of their firms, and investors increasingly expect them to disclose it.[1] Do such radical transformations have real effects? To answer, one faces three major challenges. First, disclosures are part of a broader set of actions taken by managers to govern their firms, such as raising financing, investing or deciding on payout policies. Second, managers with unfavorable news can stay put, conceal it and pretend to be uninformed. Thus, an adverse selection problem adds to the canonical moral hazard friction between managers and investors.[2] Third, the anticipation of future disclosures impacts today's firm valuation, which feeds back into the managers' current disclosure decisions.

To tackle these challenges, we embed a disclosure problem à la Dye (1985) or Jung and Kwon (1988) in a dynamic agency model close to DeMarzo and Fishman (2007) – from now onwards, DF07.[3] A principal – who collectively represents a group of outside investors – contracts with an agent who runs a firm. We refer to the agent as the firm's *manager* and to the principal as the firm's *investors*. The firm generates risky i.i.d. cash flows over time, that are non-verifiable and can be diverted by its management.[4] Our key innovation is that we introduce the possibility for the firm to invest in a (possibly costly) information technology. Upon adoption, the technology produces stochastic evidence that predicts the one-period-ahead cash flows from then onwards. While the adoption is common knowledge, the realized evidence each period is privately observed by the manager. Investors only observe the evidence upon its voluntary disclosure.

In order to describe our results, we adapt the implementation proposed by DF07 to our setting. The firm is optimally financed by a set of standard securities: equity, long-

[1]According to Dresner Advisory Services (2018), which surveys 5,000 firms globally, adopters of the latest such technologies – Big Data analytics and Artificial Intelligence – reached 59% in 2018, one the most frequent application being forecasting (See Mihet and Philippon (2018) and Farboodi, Mihet, Philippon and Veldkamp (2019)). Relatedly, Marcum (2018) suggests that 'Investors are demanding more than quarterly return reporting', and 'are no longer satisfied to collect a standardized Due Diligence Questionnaire'.

[2]As managers may be uninformed, unraveling à la Grossman (1981) and Milgrom (1981) breaks, as pointed out in Dye (1985) and Jung and Kwon (1988). See also Fishman and Hagerty (1990), Shavell (1994), Shin (2003), Acharya, DeMarzo and Kremer (2011) and Guttman, Kremer and Skrzypacz (2014). An additional source of strategic disclosure is the presence of disclosure costs, as in Verrecchia (1983).

[3]See also the seminal contribution by Bolton and Scharfstein (1990), as well as Clementi and Hopenhayn (2006), DeMarzo and Sannikov (2006) and Biais, Mariotti, Plantin and Rochet (2007). Recently, Fu and Krishna (2019) characterizes the optimal contracting when private information is persistent.

[4]As is well known, the results are qualitatively unchanged if diversion opportunities are replaced by private effort provision by managers, which affects the realized cash flows.

term debt, short-term debt. It starts off with a positive amount of leverage, and investors adjust it over time to discipline the manager and prevent cash diversion. The manager is optimally compensated with a fraction of the firm's stock, and vesting provisions do not allow him to sell the stock holdings as long as he is employed by the firm. The manager chooses whether to disclose evidence, pay dividends and/or repay the firm's debts at each point in time, in order to maximize the present discounted value of his payoff.

When cash flows are high, the manager uses retained earnings to repay the firm's debts moving forward. Dividends are disbursed only once the firm fully repays its short-term debt – that is, when its past performance has been sufficiently good. When cash flows are low, instead, the manager needs to borrow more to pay interest on the firm's existing debts. In this event, disclosure is especially useful because investors are unsure if the negative performance is due to bad luck (a negative cash-flow shock), or to bad management (diversion). Indeed, when the manager discloses bad news investors learn not only that the cash flow of the firm is going to be low, but also that this is not due to mismanagement. This leads to our **first result**: the manager is optimally rewarded for the disclosure of bad news, and the reward comes in the form of a reduction in borrowing costs. While such costs do not directly affect the manager's payoff, they do so indirectly. A lower cost of debt reduces the firm's default risk, which increases its stock value (that, at the optimal contract, the manager maximizes).

This effect can only be understood by modeling explicitly the firm's capital structure, as we do, and it is consistent with the empirical evidence that firms do disclose bad news (e.g., Skinner (1994), Enache, Li and Riedl (2018)) and that disclosure improves market liquidity and reduces borrowing costs (e.g., Balakrishnan, Billings, Kelly and Ljungqvist (2014) and Boone and White (2015)). The mechanism we propose is, to our knowledge, new in the literature. Alternatively, it has been argued that disclosure of bad news reduces litigation costs (see Skinner (1994) and Marinovic and Varas (2016)). We find the two hypothesis complementary, because the risks of litigation and default are two important downsides of debt financing. In addition, this compensation for the disclosure of bad news offers a possible alternative explanation behind 'pay without performance', other than the capture of boards by powerful executives (e.g., Bebchuk and Fried (2009)), or the need to motivate innovation by managers (Manso (2011)). The idea is that the value of the manager's stock holding – which accounts for a large fraction of his or her variable compensation (Murphy (2013)) – rises in bad states, conditional on there being disclosure that such states are transitory, and they do not depend on mismanagement.

Interestingly, pay without performance does not arise from the need to incentivize disclosure. In fact, the manager would have *strict* incentives to disclose bad news even if disclosure lead to a lower reward than the optimal one. In other words, the incentive constraint for bad news disclosure never binds in our model. What drives pay without performance, instead, is the investors' desire to insure the firm against inefficient liquidation. If the manager discloses bad news, the investors are willing to relieve the firms debt pressure, because doing so reduces future default risk and increases their expected payoff, without exacerbating the moral hazard problem. In this sense, our dynamic setting captures well an effect of disclosure that is absent in the existing literature, namely, the fact that disclosure increases the firm's funding liquidity – that is, the amount of short-term debt the firm can borrow before exhausting its debt capacity. While this occurs as a result of the investors' profit maximization problem, it is also socially efficient.

However, our **second result** is that evidence brings about a strictly higher interest rate whenever low cash flows are not preemptively disclosed.[5] Importantly, this feature does not arise from pooling as in strategic disclosure models (e.g., Dye (1985)).[6] Indeed, in our model there is no pooling: all news are disclosed on-the-equilibrium path. In contrast, it is an optimal response to deal with the moral hazard problem. There are two reasons for this. First, charging a higher interest rate is less costly to the firm moving forward. Whenever its manager will disclose bad news in the future, the firm's borrowing cost will fall and the investors gain from a lower probability of liquidation. Second, this higher interest rate is more effective to prevent diversion, because it is always charged when the performance is good, and otherwise it is only charged in the absence of disclosure.

Because of the two countervailing forces we just highlighted – the lower borrowing costs upon disclosure of bad news vs. the higher borrowing costs in other states – the natural question to ask is what the net effects are on the manager's compensation and on the firm's funding liquidity dynamics. To answer, we consider the effect disclosure has on the pay-for-performance sensitivity (PPS) of the optimal contract, which measures how stock valuations (and also the funding liquidity) vary with performance. Our **third result** is that pay-for-performance falls when the manager can disclose information, which means that disclosure reduces the severity of the agency conflict. This is intuitive: disclosure lowers the informational asymmetry between the manager and the investors, and so it dampens the variations in the firm's stock values and funding liquidity that are required by the optimal contract to deal with the agency problem.

A related, more subtle result is that, unlike what happens in the existing dynamic agency models, the PPS increases in the firm's cumulative performance history. This occurs because, as high cash flows accumulate funding liquidity in the firm, the probability

[5]Variations in rates can be implemented as covenants on the firm's debt (Smith and Warner (1979)).

[6]In the Dye model, the expected price conditional on no-disclosure decreases in the likelihood that managers are informed.This is because the equilibrium is in threshold strategies, and the threshold news (which managers are indifferent to disclose or not) falls with the likelihood that they are informed.

of default decreases and evidence becomes less and less useful. As a consequence, the gap in stock values upon a low cash flow with and without disclosure reduces progressively, until it completely disappear at the dividend-payout boundary, where we observe the same degree of PPS as in DF07 or Biais et al. (2007). This helps reconciling dynamic agency models with the evidence that high-powered incentives appear to be used by relatively better performed firms (e.g., Bandiera, Guiso, Prat and Sadun (2015)).

At this point, one may wonder if the presence of disclosure opportunities for managers is indeed *equivalent* to a reduction in agency costs, such as a better monitoring technology through which investors obtain evidence directly. Perhaps surprisingly, this is not the case. While reducing the agency friction would unambiguously lower the firm's default risk ex ante, our **fourth result** states that the probability of default may actually *increase* at the initial financing stage as the availability of evidence rises. In fact, we find that while disclosure reduces the probability of default *for a given level of the firm's debt*, the expectation that the firm will disclose in the future might *increase the debt level* to a point where the firm ends up on a path that entails *lower* rates of survival. This happens because the marginal value of granting funding liquidity to the firm ex ante falls with disclosure opportunities, as they help to hedge against bad luck. Importantly, this result underscores the importance of explicitly accounting for the feedback-loop between disclosure and financing policies, which naturally arises in a dynamic model.

Specifically, the relation between disclosure opportunities and default risk depends on a firm's profitability. There exists a profitability threshold such that default risk and disclosure are *negatively correlated* for all firms more profitable than that at the threshold. At such firms, both managers and investors benefit from a better disclosure environment.[7] In contrast, at low profitability firms disclosure opportunities and default risk may be *positively correlated*, in which case only the investors benefit from them. Managers are worse off, and would prefer not to be informed about the firm's performance.

Although liquidation helps reducing the information rents paid to the manager, low-profitability firms are more likely to generate low cash flows and default ex post. Therefore, absent evidence disclosure, it is costly for investors to commit to liquidate such firms by imposing high leverage. Evidence reduces this cost, because default is expected to occur less often ex post: only when poor performance is not guided by managerial disclosure. Thus, evidence makes it more desirable for investors to be aggressive in their liquidation strategies, leveraging up the firm. Thus, default risk may rise with the avail-

[7]This helps clarifying the important difference between our disclosure model and a monitoring setting, such as Fuchs (2007), Piskorski and Westerfield (2016), Smolin (2017), Zhu (2018) or Orlov (2019) While in most monitoring models the information is used to curb managerial rents, this need not be true when the realized evidence is private information of the managers and must be disclosed.

ability of evidence, implying that both the PPS and efficiency can decrease, contrary to the effect of a reduction in the agency friction itself.

Finally, we introduce disclosure costs through a real option that allows firms to invest in a better information technology. In this setting, we study which firms find it optimal to make the investment and adopt a strategy of frequent disclosure. This introduces dependency between the past firm's performance and the frequency of voluntary disclosures.[8] The analysis is complicated by the fact that the firm's value function now includes an American option with strike price equal to the disclosure cost, and benefits arising from the reduction in the firm's default risk. Our **fifth result** is that the set of firms that choose to incur the disclosure cost and increase their disclosure frequency 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 upper threshold, the benefits are too low, as the firm is far from its default boundary. As the disclosure cost falls, the thresholds diverge and the adoption region expands.[9]

Empirically, this implies that firms that experienced sufficiently negative or sufficiently positive performance from the time of initial financing onwards won't disclose much, and the bulk of voluntary disclosures occurs at firms that had intermediate performance histories. Theoretically, this pattern marks a substantial difference between the optimal adoption of physical investment options, such as those studied in DeMarzo, Fishman, He and Wang (2012), and that of information-related options, such as ours. While the value of physical options typically increases with the firm's past performance, the value of information-related options peaks at intermediate performance histories.

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 clarifies that some of our results obtain in finite-horizon settings, but not all. It also conveys some intuition that helps to understand the full model. Section 5 introduces the infinite horizon model. Section 6 solves for the payout and liquidation boundaries, characterizes the firm dynamics, and shows the patterns of information technology adoption. Section 7 implements our optimal contract by means of short and long-term debt, and equity. Section 8 discusses the impact of disclosure on the policy dynamics and on other variables of interest. Section 9 discusses

[8]We are grateful to Adriano Rampini for suggesting this extension to us.

[9]Modeling this irreversible information option is already quite complex, and our attempts to make it reversible have proven very difficult to solve analytically. Instead of relying on numerical simulation, we decided to present the model that we could solve, to discuss the economic mechanisms at greater depth.

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 (2017), Zhu (2018) and Orlov (2019)). The distinguishing feature of our model is that, while other papers focus on monitoring technologies where the principal acquires information directly, in ours the information is observed by the agent and it must be disclosed.

Because we model information systems as technologies that produce disclosure opportunities for managers, à la Dye (1985) or Shin (2003), our work is clearly related to the extensive literature 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 (2019)). While these recent papers extended the Dye model dynamically, they differ from our setting in important ways. First, managerial compensation is exogenous, whereas we consider optimal compensation. This implies that the equilibria they characterize feature partial disclosure, whereas ours do not. Second, in some of these papers evidence is long-lived, and so they study not only what is being disclosed, but also when. In ours, evidence is short-lived.

Another 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 impacting the cash flows, improves the information available for the management to disclose. Contrary to the value of physical options, which typically increases in the firm's past performance, that on information-related options is non-monotonic. It peaks at intermediate performance histories.

We also contribute 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 private information and has the possibility 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. While this literature has flourished since Bull and Watson (2004), including the contributions of Koessler and Perez-Richet (2017), Hart, Kremer and Perry (2017) and Ben-Porath, Dekel and Lipman (2019), to our knowledge we are the first to incorporate evidence in a dynamic agency setting.

3 Environment

A firm produces i.i.d. cash flows $x_t \in \{h, l\}$ for $t = 1, 2, .., T$, where $h > l > 0$. Define $\Delta := h - l$, $p := \Pr.[x_t = h] \in (0, 1)$, and $\mu := \mathbb{E}(x_t)$. The firm is owned by a principal and is operated by a manager. Both the principal and the manager are risk-neutral and discount future consumption at the same rate $r \in (0, 1)$.[10]

Moral hazard. We introduce the possibility of moral hazard by assuming that the manager privately observes the realized cash flows $\{x_t\}$. By misreporting a good cash flow, claiming it to be bad, the manager can divert Δ output and obtains a private benefit of $\delta := \lambda\Delta$, where $\lambda \in (0, 1]$ represents the severity of the moral hazard problem.[11] By applying 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 the principal can choose to make a one-time and irreversible investment in an information technology that will produce evidence $e_t \in \{g, b\}$ each subsequent period with probability $\hat{\pi} \in (0, 1)$. To ease notation, π 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 ($\pi = 0$).[12] To make this investment, the principal must spend a fixed cost of $c \geq 0$.[13] Thus, adoption corresponds to the exercise of a one-time American option with infinite maturity and strike price c. 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. Formally, we assume that $\Pr.[x_t = h | e_t = g] = \Pr.[x_t = l | e_t = b] = 1$. Since the effects of evidence on our outcomes of interest will already be complex and non-monotone with perfect evidence, we do not consider imperfect correlation in this paper.

Once the option is exercised, investors expect evidence to be available with probability $\hat{\pi}$, but they never know if the manager is informed. So, at each date t, the manager chooses whether or not to *voluntarily disclose* the evidence. We denote the disclosure action by

[10] Common discounting is not needed to derive our qualitative results, but it simplifies the arguments.

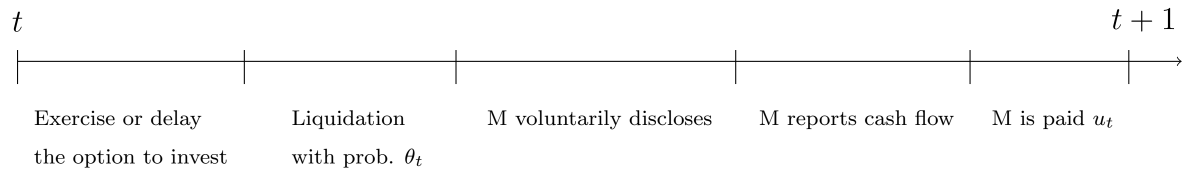
[11] The notation here is not redundant: the effects of Δ on allocations and contracts are slightly different from those of λ, in ways that we will emphasize while discussing the comparative statics.

[12] The fact that in the absence of technological investment $\pi = 0$ is just a normalization. All our results go through unchanged if we assumed that, absent investment, the firm would have some positive $\pi' < \hat{\pi}$.

[13] The cost can be thought of as the presented discounted value of the setup and maintenance expenses.

$a_t \in A := \{d, n\}$, where d stands for disclosure and n for non-disclosure.

Figure 1: Timing in period t (prior to adoption)

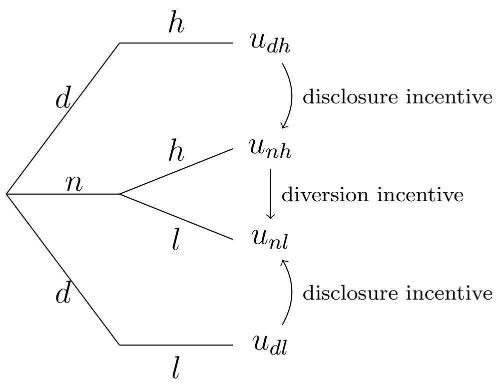


Contracting. To maximize investors' value, the principal offers the manager a contract that specifies, for every history of reports and disclosures, the probability of liquidating the firm $\theta_t \in [0, 1]$, and the cash compensation $u_t \geq 0$. If the firm is liquidated, both parties get their outside option payoff, which is normalized to zero. In the first best case, the firm is never liquidated and the firm value is $s^* := \frac{\mu(1+r)}{r}$. Figure 1 shows the timing of events in a generic period t, prior to adopting the information technology. Importantly, all our results go through unchanged under the alternative assumption that evidence arises after the cash flows, and can certify them.

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.

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 the manager are denoted by u_i, for each possible outcome $i \in \mathcal{H}_1$. The contract must provide two kinds of incentives: (i) to prevent the manager from diverting cash flows, which requires $u_{nh} \geq \delta + u_{nl}$; (ii) to disclose information if there is any, which requires both $u_{dh} \geq u_{nh}$ and $u_{dl} \geq u_{nl}$. Trivially, it is optimal to set $u_{nl} = u_{dl} = 0$, $u_{dh} = u_{nh} = \delta$, and not exercise the option.

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 the manager's disclosure action in the first period; the second and third elements denote the $t = 1$ and $t = 2$ realized cash flows, respectively.

As standard in a dynamic agency model, committing to liquidate the firm when $x_1 = l$ may be optimal, because it alleviates the diversion problem and reduces the information rent required for incentive compatibility to hold. In addition, it is trivial to show that it is always optimal to delay compensation to the terminal nodes, at $t = 2$. To highlight the role of evidence disclosure in the simplest case, we present here the results for a two-period model with costless evidence, leaving the costly-evidence case to the Appendix.[14] Note that when evidence is free, the information technology is always adopted at $t = 0$, without loss of generality – because it can be ignored at no cost, if necessary.

Proposition 1. *If $T = 2$ and $c = 0$ (i.e., evidence is costless), there exists an information threshold $\pi_T < 1$ and a profitability threshold $0 < p_T < 1$ such that, for $a \in A$:*

(a) If $\hat{\pi} < \pi_T$, the firm is never liquidated. Compensations are $u_{ahh} = \frac{(1+p)\delta}{p}$, $u_{alh} = \delta$;

(b) If $\hat{\pi} > \pi_T$, then the firm is liquidated at $t = 1$ when a low cash flow is reported and there is no disclosure. Compensations are $u_{ahh} = \frac{\delta}{p}, u_{dlh} = \delta$;

(c) If $\hat{\pi} = \pi_T$, the optimal contract is any mixture of the above ones.

Moreover, $\pi_T > 0$ if and only if $p < p_T$.

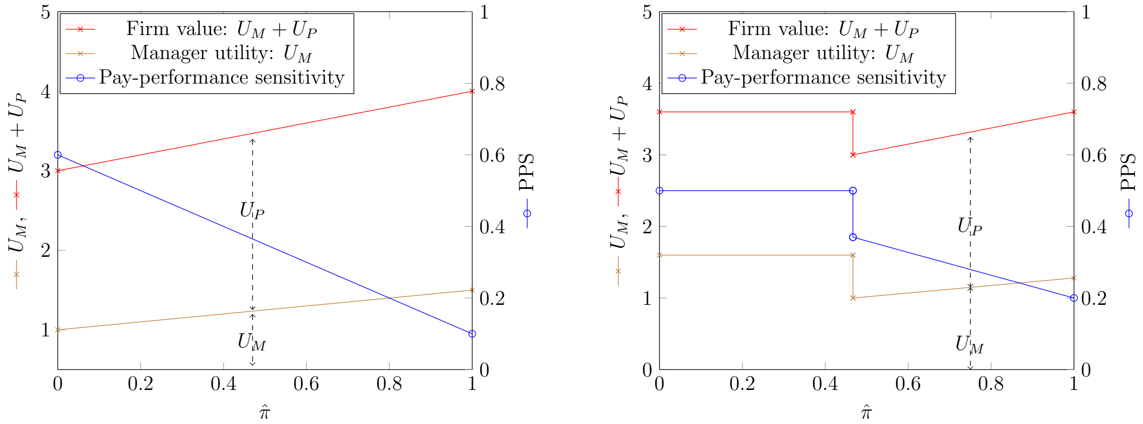
Proposition 1 clarifies a few implications of evidence on contracts and allocations. First, evidence is *only* useful if the firm is liquidated with some probability at $t = 1$. Therefore, when liquidation does not occur as in case (a) of the proposition, both the investors' and the manager's payoffs do not depend on $\hat{\pi}$: evidence is irrelevant.

Second, while investors always benefit from a better information technology – that is, a higher $\hat{\pi}$ – both the manager's compensation and the probability of default may rise or fall with it. Depending on the firm's profitability, as well as the availability of evidence, we have three possible scenarios. They are depicted in Figures 3a and 3b, where U_M

[14] In the other case of $c > 0$, the firm has a meaningful choice between enhancing its disclosure opportunities or not. But note that the choice to invest is never delayed to $t = 1$ in the two period model, because evidence only plays a role here when liquidation occurs later.

denotes the expected managerial compensation at $t = 0$, and U_P the expected investors' payoff.[15]

Figure 3: Comparative statics in the two-period model



(a) High profitability firms: $p \geq p_T$

(b) Low profitability firms: $p \leq p_T$

1. For highly profitable firms ($p \geq p_T$, Figure 3a), Proposition 1 implies that the threshold $\pi_T \leq 0$, and so only our case (b) is possible. The probability of default – equal to $(1-p)(1-\hat{\pi})$ – drops with $\hat{\pi}$. The expected compensation to the manager – equal to $p\delta(1+(1-p)\hat{\pi})$ – rises in $\hat{\pi}$. Investors gain because the revenues from a lower default probability more than offset the rise in managerial rent;

2. For low profitability firms ($p < p_T$, Figure 3b), Proposition 1 implies that $\pi_T > 0$, and therefore both (a) and (b) are possible. If the technology seldom generates evidence (i.e., $\hat{\pi} < \pi_T$), the probability of default is zero, and the managerial compensation is $2p\delta$, as if there was no evidence. Obviously, such technologies would never be adopted for any positive cost;

3. For low profitability firms ($p < p_T$, Figure 3b), if the technology frequently generates evidence (i.e., $\hat{\pi} \geq \pi_T$), then the probability of default rises from zero in the above case to $(1-p)(1-\hat{\pi})$, while the compensation drops from $2p\delta$ to $p\delta(1+(1-p)\hat{\pi})$. Investors gain because the revenues from lower managerial compensation more than offset the increase in the default probability.

The two-period model clarifies that evidence disclosure has heterogeneous effects on the firm value, through investors' liquidation strategy. On the one hand, default upon low

[15] Evidently, the sum of the two denoted the economic surplus generated by the firm, or the firm value, which is equal to the first-best surplus minus the deadweight losses due to liquidation.

cash flow at $t = 1$ is prevented by bad news disclosure, because investors are reassured that the firm's poor performance is not due to mismanagement (branch dl in Figure 2). On the other hand, evidence increases the probability that the firm will default when its poor performance at $t = 1$ is not preemptively disclosed by the manager (branch nl in Figure 2). Greater disclosure opportunities reduce the probability that the manager is uninformed (i.e., there is a lower chance of reaching nl). Thus, liquidation in the nl branch becomes less costly for investors. As the threat of liquidation is optimal to reduce managerial rents, investors rely more on it than they did in the no-evidence case.

To analyze cash incentives, we define the firm's pay-performance sensitivity (PPS) as the measure of how compensation changes with the cash flows, in percentage terms:

$$\text{PPS} := \frac{\mathbb{E}(U_M \mid x_1 = h) - \mathbb{E}(U_M \mid x_1 = l)}{(1 + r)\Delta} \tag{1}$$

The first effect of evidence illustrated above implies that the expected managerial pay conditional on low cash flow becomes larger, which reduces the PPS. Indeed, figures 3a and 3b confirm that the PPS (weakly) drops with $\hat{\pi}$. In this respect, disclosure plays a role similar to that of alleviating the agency conflict λ in the existing literature.

However, the second effect we illustrated suggests the possibility that the firm defaults more frequently while the PPS of managerial pay goes down. This marks a substantial economic difference between introducing disclosure opportunities and reducing the agency conflict λ. Indeed, in Biais et al. (2007) and the other dynamic agency models without disclosure, a reduction in the agency conflict always reduces the firm's default probability.

While it is useful for developing some intuition, some of our results cannot be fully understood in the context of the two-period model. For instance, the managerial pay and the liquidation decision at the nl node of time 1 are the same on both sides of π_T, only jumping at π_T. In this sense, the two-period model cannot fully capture the effect of changing the firm's disclosure frequency on default probability as the performance history varies. To analyze these issues and to study the optimal patterns of adoption of information technologies, we must turn to the full model.

5 Infinite-horizon model

In this section, we use the recursive approach to formulate the firm's problems in the infinite-horizon environment. Whether it is worthwhile to invest in the costly information technology or not and, if so, when to make the investment, all depend on trading off the cost with the value that this option brings to the firm. To evaluate the *moneyness* of this evidence-generating option, we first consider the optimal contracting problem under the assumption that the investment has already been made. Then, we step back and determine the optimal exercise time for the option.

5.1 Contracting with evidence

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

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:

$$u_{dh} + \frac{w_{dh}}{1+r} \geq u_{nh} + \frac{w_{nh}}{1+r} \tag{IC_g}$$

Likewise, whenever the manager obtains bad evidence we have:

$$u_{dl} + \frac{w_{dl}}{1+r} \geq u_{nl} + \frac{w_{nl}}{1+r} \tag{IC_b}$$

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

$$u_{nh} + \frac{w_{nh}}{1+r} \geq \delta + u_{nl} + \frac{w_{nl}}{1+r} \tag{IC_n}$$

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

$$v = (1-\theta)\left[\pi E_d\Big(\mathbf{u}_d + \frac{\mathbf{w}_d}{1+r}\Big) + (1-\pi)E_n\Big(\mathbf{u}_n + \frac{\mathbf{w}_n}{1+r}\Big)\right], \tag{PK}$$

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

$$u_{dh}, u_{nh}, u_{dl}, u_{nl} \geq 0 \tag{LL}$$

Because the agents share the same discount factor, it follows that the optimal contract from the principal's perspective also maximizes firm value (i.e., surplus), given a utility v promised to the manager.[16] Thus, the optimal contract solves the following dynamic program:

$$\begin{aligned} s(v) = \max_{\theta, u_j, w_j} \; & (1-\theta)\Big\{\mu + \frac{1}{1+r}\big[\pi E_d(\mathbf{s}_d) + (1-\pi)E_n(\mathbf{s}_n)\big]\Big\} \\ \text{s.t.} \quad & (PK), (IC_g), (IC_b), (IC_n), (LL), \end{aligned} \tag{S}$$

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

The objective function of (S) reflects the fact that (i) with probability θ, liquidation takes place before the subsequent evidence and cash flow realize, in which case the firm's value drops to zero; and (ii) with probability $(1-\theta)$ the firm is not liquidated, in which case the firm's value depends on both whether the manager receives 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 function of program (S).

On the one hand, program (S) with $\pi = \hat{\pi}$ solves the firm's problem given the investment option has already been exercised. On the other, if $\pi = 0$, the program also solves the case where the option is never exercised (which corresponds to the case studied in DF07). This is because, in the latter case, evidence never arises and therefore the only relevant control variables are those conditional on no-disclosure. Given this observation, we use $s(v; \hat{\pi})$ and $s(v; 0)$ in the following analysis to denote the firm value with evidence and that with no evidence forever, respectively.

5.2 Investment decision

We now formulate the firm's problem with the investment decision, i.e., the decision of whether and when to exercise the investment option. Suppose that, for a given history

[16] As is well known, this does not imply that a contract that maximizes the principal's expected utility is socially optimal: in general, the principal starts the contract from a socially suboptimal initial condition.

represented by v, the firm has not exercised the option yet. The firm's value in this scenario is denoted by $f(v)$ and, evidently, no evidence will be disclosed today. If the firm is not liquidated right away – which occurs with probability $(1-\theta)$ – it obtains a cash flow today and proceeds to tomorrow's state, either w_{nh} or w_{nl}, depending on the cash flow reported by the manager.

Come tomorrow, the firm can either invest c and obtain the value of $s(w_{ni};\hat{\pi})-c$ (for $i=h,l$) from the subsequent date onwards, or delay investment again and obtain $f(w_{ni})$. It is easy to see that when $s(w_{ni};\hat{\pi})-c > f(w_{ni})$, the firm exercises the investment option tomorrow. Otherwise, it waits until at least one more period to invest. Thus, the firm's problem when the investment has not been undertaken yet can be formulated as follows:

$$f(v) = \max_{\theta,u_j,w_j} (1-\theta)\Big\{\mu + \frac{1}{1+r}E_n\big[\max(\mathbf{f}_n, \mathbf{s}_n(\hat{\pi})-c)\big]\Big\} \qquad (F)$$
$$\text{s.t.} \quad (PK), (IC_n), (LL)$$

where $\mathbf{f}_n = (f(w_{nh}), f(w_{nl}))$, $\mathbf{s}_n(\hat{\pi}) = (s(w_{nh};\hat{\pi}), s(w_{nl};\hat{\pi}))$, and $\pi = 0$ in (PK).

6 Policy characterization

In this section, we characterize the firm's problems. In the next we will examine which firms invest in the information-generating the option and under what conditions. 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 DF07 where evidence is never available (the extensive margin). This benchmark case corresponds to our model in which the option is never exercised.

6.1 Payout and liquidation

Because liquidation is inefficient, it is optimal to delay the manager's cash compensation until the continuation utility v is sufficiently large. Without loss of generality, we shall assume that the manager is paid by cash whenever the firm is indifferent between paying him right away, and delaying the payment. Formally, we define the *cash payout boundary* as the smallest continuation utility where the firm value reaches its first best. That is,

$$\bar{v} := \inf\{v : f(v) = s^* \; or \; s(v;\hat{\pi}) = s^*\} \qquad (2)$$

The definition implies that the firm value (with or without evidence) is strictly less than the first best s^* before the continuation utility reaches $\bar{v}$. While in general both the payout boundary and the manager's payoff dynamics may depend on the availability of evidence, the next result shows that the value $\bar{v}$ is actually a constant, irrespective of evidence availability and the manager's disclosure choice, while the cash compensation granted to the manager varies with π in the short-run.

Proposition 2. *The cash payout boundary is:*

$$\bar{v} = r^{-1}(1+r)p\delta. \tag{3}$$

Moreover, for $a \in \{d, n\}$, the optimal cash compensation is

$$u_{al}(v) = 0, \quad u_{ah}(v) = \max\left\{(1+r)\delta - (1+\hat{r})(\bar{v} - v),\ 0\right\} \tag{4}$$

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

Proposition 2 shows that the cash compensations are independent of the manger's disclosure actions. No cash payment is made to the manager whenever a low cash flow is reported. When the firm is one-step away from $\bar{v}$, the manager receives cash compensation upon reporting a high cash flow and such payment falls with the availability of evidence π. However, when the firm reaches $\bar{v}$, the cash compensation becomes independent of π. This is intuitive: the payoff boundary $\bar{v}$ is the smallest continuation utility at which all incentive constraints cease to bind. Once the boundary has been reached, liquidation never occurs moving forward, and therefore diversion incentives can be prevented by cash payments only (as if no evidence existed).

Before reaching the payout boundary, the manager is incentivized by variations in his promised continuation utilities. Liquidation may occur after a sequence of low cash flow shocks. When the continuation utility is low enough, the only way to align incentives and keep the compensation promises is to stochastically liquidate the firm at the beginning of the period. To capture these dynamics, we define the thresholds such that no liquidation occurs in the next n periods to be:

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

These thresholds are related to the previous definitions of liquidation probability θ and the payout boundary $\bar{v}$. If v is larger than v^1, the firm will not be liquidated in the current period, but may be liquidated in the next period. So 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 $\bar{v}$ is the limit of this sequence of thresholds v^∞. Indeed, liquidation never occurs if $v \geq \bar{v}$.

6.2 Firm dynamics

We first characterize the firm's problem (S). Recall that it conveniently solves two possible scenarios: the option has already been exercised ($\pi = \hat{\pi}$); and the benchmark case as in DF07 where the option is never exercised ($\pi = 0$). The firm value function $s(\cdot)$ in this problem is concave, as standard in the literature, because liquidation is inefficient. This further implies that any randomization of continuation utilities is costly for both the investors and the manger. Using concavity and the optimal conditions of the firm's problem (S), we can figure out which constraints bind and derive the optimal policies.

Lemma 1. *For any $v < \bar{v}$ 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 parties are not affected by evidence disclosure. Contingent on a good performance being reported, the manager does not divert and so there is no need to further condition the manager's payoffs on evidence disclosure. Thus, in the rest of the paper we do not distinguish the manager'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. It is important to stress that this property would *not* generally hold in a monitoring setting, in which the principal observes the evidence directly. In that case, the principal could further reduce the manager's payoff when evidence is available, without violating any constraint.

In contrast, the optimal contract provides strict incentives for the manager to disclose bad news: i.e., $w_{dl} > w_{nl}$. Punishing the manager for a bad performance is costly to the principal because it induces more inefficient liquidation. If the evidence shows that the bad performance is not caused by the manager's behavior but by bad luck, then there is no incentive benefit in punishing the manager for it. In other words, promising the manager higher utility in the state dl does not worsen the diversion problem, but improves efficiency by reducing the probability of liquidation. In this sense, pay for bad luck is not driven by the need for the principal to incentivize disclosure, but rather it is directly beneficial to the investors because it enables to lower both the pay-for-performance sensitivity, and the volatility of the agent's continuation utility.

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 $v \in (0, v^1]$: $\theta = \frac{v^1 - v}{v^1}$, $\;w_{nl} = 0$, $\;w_{dl} = v^1$, $\;w_h = \min\left\{\frac{r\bar{v}}{p}, \bar{v}\right\}$;*

- *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 $v^n(\pi) = [1 - (\frac{1}{1+\hat{r}})^n]\bar{v}$.*

If stochastic liquidation does not occur at the beginning of the period, the policies for $v < v^1$ are the same as those for $v = v^1$. In addition, firm value is linear in this region. Given this characterization, it is clear 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 the manager discloses evidence of transitory bad luck, the investors compensate for disclosure by raising the promised utility to v^1, independently from the continuation utility entering the period. The reward depends on $v^1 - v$.

If $v > v^1$, the manager is still rewarded for disclosing bad luck because the contract forgives the low performance today, and starts tomorrow as if the history was the same as before the current low cash flow. This policy does not affect moral hazard, because the manager cannot fabricate evidence. Thus, the desire to insure the firm against default implies that the expected marginal firm value tomorrow should be the same when bad news are disclosed, and when they are not (which means that either there is no evidence, or the evidence brings about good news). Moreover, optimality requires that the expected marginal firm value today and tomorrow should be the same. Based these two observations, we are able to conclude that the marginal firm value stays constant if bad news disclosed, which implies that $w_{dl} = v$.

Finally, the optimal contract rewards good luck. Proposition 3 shows that the ranking of continuation utility does not depend on the levels of v and π: the manager gets the largest continuation utility upon a high performance, the lowest one upon a low performance and no disclosure, and the middle one upon disclosure of bad news. This pattern implies that, on the fastest route to liquidation, the manager never discloses evidence and always reports low cash-flow. So, the n-period liquidation thresholds we previously defined (v^n) can be explicitly derived from the policy functions. Evidently, these thresholds and the policy dynamics all depend on the manager's disclosure behavior, as well as on the availability of evidence.

6.3 Investment policy

Now that we explored the role of evidence disclosure in the policy dynamics through contracting, we can bring back the disclosure cost c and analyze the optimal patterns of adopting the information technology across firms and histories.

The value of evidence in our model comes from two channels. First, the availability of evidence increases firm value by avoiding the inefficient punishment of underperforming managers. Second, because the value of evidence is endogenous and it varies with the manager's continuation utility, delaying the investment can be valuable to the firm. Thus, the technology adoption decision depends on two things: (i) the strike price c; and (ii) the endogenous variable v, which represents the accumulated firm performance history.

First, consider the cost dimension. From the firm's problem (F), we know that the value of exercising the option right away is $s(v;\hat{\pi})-c$, while that of delaying the investment is $f(v)$. As both value functions $s(\cdot;\hat{\pi})$ and $f(\cdot)$ are bounded above and below, if the cost c is too large it is never optimal to exercise the option. We define a threshold cost as:

$$\bar{c} := \max_{v}\{s(v;\hat{\pi}) - s(v;0)\} \tag{5}$$

The threshold $\bar{c}$ depends on the availability of evidence $\hat{\pi}$, as well as on other parameters such as the severity of the agency conflict λ, the profitability of the firm p, and so on.

When $c \geq \bar{c}$, the technology is never adopted. In this case, as the firm's value $f(v)$ is no less than its value when evidence never arises (i.e., $s(v;0)$), we know that the difference in firm value $s(v;\hat{\pi}) - f(v)$ is smaller than $s(v;\hat{\pi}) - s(v;0) \leq c$. Thus, the value of exercising the option $s(v;\hat{\pi}) - c$ is no larger than that of delaying $f(v)$, for any v. When $c < \bar{c}$, the investment option can be exercised in some state v. Otherwise, the firm value $f(v)$ becomes the same as its lower bound $s(v;0)$. The definition in (5) implies that there exists some v such that the value of exercising right away (i.e., $s(v;\hat{\pi}) - c$) is strictly larger than that of delaying (i.e., $f(v)$).

Second, consider the role played by the continuation value v. It is optimal to delay exercising the option either in the region close to $\bar{v}$, or in the region close to 0. The intuition is as follows. When v is close to $\bar{v}$, the probability of default is very small. Therefore, the firm value without evidence $f(v)$ gets 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, instead, the firm's value is so small that the option's NPV is either tiny or negative. To specify these two regions, we define two continuation-utility thresholds:

$$v^l := \inf_{v}\{f(v) \leq s(v;\hat{\pi}) - c\}, \qquad \text{and} \qquad v^h := \sup_{v}\{f(v) \leq s(v;\hat{\pi}) - c\} \tag{6}$$

The following result summarizes the patterns of adopting the information technology as a function of the investment cost c and the endogenous state variable v.

Proposition 4. *The firm never exercises the investment option if $c \geq \bar{c}$. Otherwise, there exists a nonempty interval $\mathbf{v} \subset (v^l, v^h)$ where investment option is exercised right* *away, while the option is delayed for* $v \in [0, v^l] \cup [v^h, \bar{v}]$.

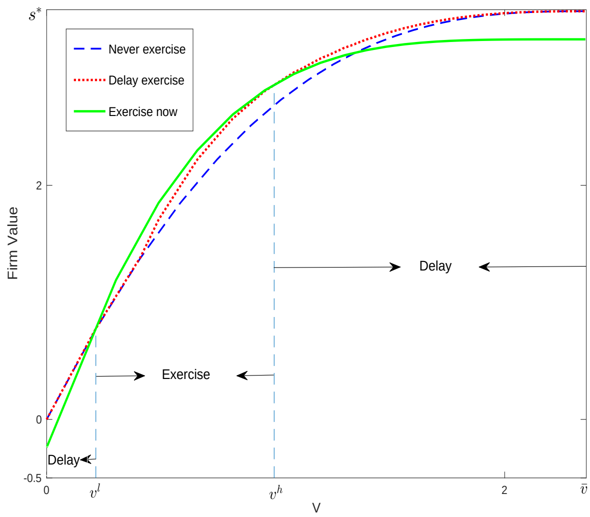


Figure 4: Option exercise region

Figure 4 illustrates the result of Proposition 4. 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 DF07. 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 investment. Investment is made if $v \in (v^l, v^h)$, and delayed if $v \geq v^h$ or $v \leq v^l$. The difference between the red and green line in Figure 4 reflects the option value. This difference is large at intermediate levels of v, because in these states of the world the firm is 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 option.

7 Capital structure implementation

Before presenting the main comparative statics of interest, which highlight the real effects of disclosure on the firm's dynamics, we show how to implement 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., DF07) the securities in our implementation only include equity, long-term debt, and a credit line (short-term debt).

Long-term debt claim is a perpetuity that pays a fixed coupon every period. Short-term debt (the credit line) defines the amount of credit that can be withdrawn by the firm anytime within some endogenous limit. The difference between the credit limit z and its balance m proxies the firm's debt capacity, or its *funding liquidity* level. As standard, equity 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, its interest rate. In our model, any balance on the credit line account is charged an interest rate $\hat{r}_i$, for $i \in \mathcal{H}_1 = \{dh, dl, nh, nl\}$, which is contingent on *both* performance and disclosure. In contrast to models such as DF07, investors sometimes charge a higher interest rate than r, or forgive the current period interest charge. The intuition is that, as anecdotal evidence suggests, lenders are willing to cut some slack to their borrowers when a temporary low performance is proven to be due to circumstances beyond the manager's control. For instance, banks routinely renegotiate their loans and prefer to delay payments than to force their borrowers into insolvency or liquidation proceedings. One way to implement variable interest rates depending on the disclosed evidence in practice may be through covenants on the firm's short-term debt (e.g., Smith and Warner (1979)). The following result summarizes a security design that implements the optimal contract.[17]

Proposition 5. *Under the following security and compensation design, the manager never diverts, and always discloses any available evidence.*

- *The manager holds λ 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 $\hat{r}_{i \neq dl} = \hat{r}(\pi)$.*

Firm cash flows are first used to repay the coupon and any credit line balance. The firm only issues dividends after it pays off the credit balance.

In this implementation, the credit balance m summarizes the history and serves as our state variable. It maps one-to-one to the state variable v in the contracting problem. Because the manager can borrow all the available credit and pay it out as dividend, the continuation value of the manager must be at least $\lambda(z - m)$. Moreover, the investors will not leave more informational rents than necessary to the manager. Thus, the following

[17]Evidently, as in all other security design problems, such design can never be unique.

condition must hold at any history:

$$v = \lambda(z - m) \tag{7}$$

The condition states that the scaled continuation utility in the contract (i.e., $\frac{v}{\lambda}$) can be represented by the firm's funding liquidity level $z - m$ in the implementation. Given this representation, as well as the policy dynamics in Proposition 3, it is straightforward to figure out how the firm's liabilities evolve. To illustrate, suppose that the firm starts a period with credit balance m. The interest rate on the credit line is $\hat{r}$, unless bad cash flow news is disclosed in which case the interest rate drops to zero. The credit balance in the following period – denoted as $m_{i\in\mathcal{H}_1}$ – will be

$$m_h = (1 + \hat{r})m - (1 + r)(\Delta - d) \tag{8}$$

$$m_{nl} = (1 + \hat{r})m \tag{9}$$

$$m_{dl} = m \tag{10}$$

where $d := \frac{u_h}{\lambda}$ is the dividend payout.

The firm's cash flow is first used to pay the long-term debt coupon l. If cash flows are good, the firm is charged an interest rate equal to $\hat{r}$ on its beginning balance m, but will have an additional cash of $(1 + r)(\Delta - d)$ to repay the short-term debt in the next period (independently from disclosure). Therefore, the new credit balance m_h follows (8). If a bad news is disclosed, then the firm can borrow at zero interest rate and the balance stays constant. Finally, if no evidence disclosed and low cash flow realizes, then $\hat{r}$ is charged on the balance m and the new balance m_{nl} follows (9).

As shown in Proposition 5, 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 the firm value reaches the first best. In that stage, evidence disclosure is payoff irrelevant. To prevent diversion, the manager needs to hold a fraction λ of the firm's shares.

What varies with the manager's disclosure decisions is the interest rates on short-term debt. This is because evidence disclosure affects the value of funding liquidity and its optimal level in the short-run. When bad news is disclosed, the firm has lower cash flow to repay its debt and is more likely to default. In this case, it is optimal for the investors to charge a lower interest rate. To compensate the investors, a higher interest rate is charged otherwise. It is easy to see from Proposition 5 that the average interest rate is exactly r, but the interest gap between disclosing bad news or not is $\hat{r}_{i \neq dl} - \hat{r}_{dl} = \hat{r}$, and it increases with $\hat{\pi}$. In other words, as the manager is more likely to be informed, investors on average still earn the risk free rate r, but the interest rate variation is larger.

8 Comparative statics

Having characterized the optimal policy functions and implemented them with standard securities, we can now examine how the firm value, default risk, managerial compensation and firm dynamics depend on the frequency of evidence disclosure. This analysis is natural in a dynamic model, because the effects of evidence endogenously depend on the manager's continuation utility v, which represents the firm's accumulated performance. However, we suspend for now the consideration as to how likely it is for a firm to find itself at a given v [18], which requires to explicitly account for the effect of future anticipated disclosures on current (i.e., date zero) optimal debt levels. Before presenting the results, it is useful to redefine the PPS measure in the context of the infinite-horizon setting, and to provide a formal definition of the firm's credit spread.

Recall that, following the existing literature, we defined the PPS in (1). In the full model, the manager's payoffs in the definition become the continuation utilities w_i for $i \in \mathcal{H}_1$. The PPS measure indicates – in percentage terms – how managerial compensation changes as a function of the firm's cash flows. In addition, given the implementation in Section 7, we know that the manager's compensation comes from his λ fraction of the firm's dividends. Thus, the firm's total equity is evaluated as $\frac{w_i}{\lambda}$, implying that $\frac{\text{PPS}}{\lambda}$ measures how the firm's equity value changes with its cash flows. To facilitate our analysis, we can also obtain the explicit form of the PPS from the policy characterization in Proposition 3 as follows:

$$\text{PPS} = \lambda - \frac{\pi \hat{r} \left[\bar{v} - \max(v, v^1) \right]}{(1+r)\Delta} \tag{11}$$

Turning to the credit spread, which measure the firm's default risk, the standard definition is: credit spread := (1 − recovery rate) × Pr.[default], where the recovery rate denotes the fraction of the firm's value recovered by creditors upon default and/or liqui-

[18] Note that any state $v \in [0, \bar{v}]$ is on-the-equilibrium path, regardless of the initial conditions of the problem. That is, there always exists a sequence of shocks that can take the firm from v_0 to any such v.

dation. Because we normalized the recovery rate to zero, the expression simplifies to:

$$\text{Credit (or CDS) spread} = 1 - \frac{s}{s^*}, \tag{12}$$

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 the manager's expected payoffs).

In this section, we first consider how the availability of evidence impacts the firm's problem (S). Then, we turn to its impact on program (F). The following comparative static results assume either that the firm has already exercised the investment option (i.e,, its manager receives evidence with probability $\hat{\pi}$ each period), or that it never adopts the information technology (as in DF07). Shocking π in the following proposition can be interpreted either as raising it from 0 to $\hat{\pi}$, or raising it from a lower $\hat{\pi}$ to a higher $\hat{\pi}$.

Proposition 6. *When the availability of evidence π rises, 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 (weakly) increases with v for any given π;

(c) w_h and w_{nl} both (weakly) fall, while w_{dl} stay constant;

(d) the n-period liquidation thresholds v^n fall, for $n = 1, 2, ...$

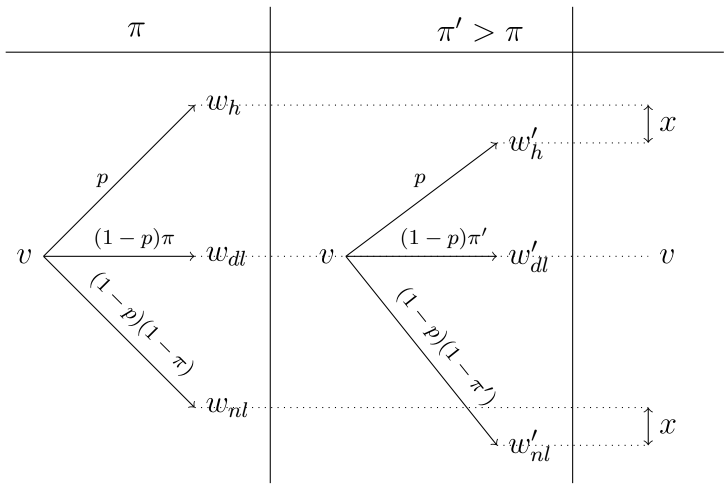


Figure 5: Impact of evidence on managerial payoff

To discuss Proposition 6 and the economic intuition behind it, we can start from claim (c), which analytically illustrates the properties of the policy function when the disclosure frequency π varies. The manager's 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 π. The diversion constraint binds, establishing that the gap $w_h - w_{nl}$ must be constant and equal to $(1 + r)\delta$. So, 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. Figure 5 illustrates this pattern for a given v. As π increases, the continuation utility is less likely to move downward, but its lowest value becomes worse.

Following our implementation, the firm's debt level varies in the opposite way to the continuation utility. Because the long-term debt is fixed, the firm's total debt amount is determined by its credit balance (m), which according to (7) is $z - \frac{w_i}{\lambda}$. So a lower continuation utility w_i means that the firm borrows more and has less funding liquidity moving forward. Through the implementation, claim (c) implies that with more evidence disclosure the debt level is less likely to go up, but the firm has to borrow more upon bad performance but no disclosure because of the higher interest rate charge, which dampens funding liquidity at the time when it's needed most.

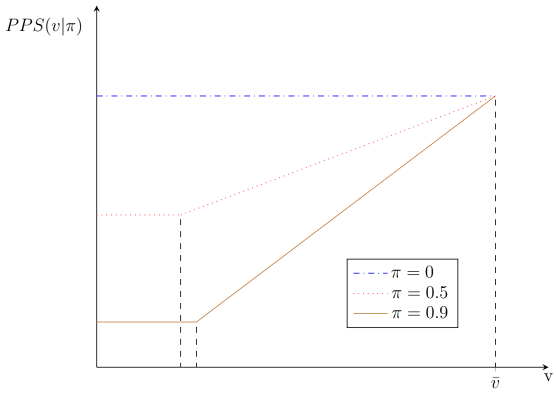


Figure 6: Pay-for-performance sensitivity

According to (11), the PPS measure depends on both v and π. Two effects drive the PPS to move in opposite directions. While pay for bad luck dampens it, the reduction in w_{nl} serves an offsetting role. However, as claim (b) of Proposition 6 states, the dominant effect is *always* for evidence to lower the PPS, for every given v. This is an intuitive result. While the PPS prevents diversion, it induces a positive probability of default. As the informational asymmetries between manager and investors decrease, because of evidence disclosure, the agency conflict becomes less severe and the PPS can be reduced, without introducing diversion incentives for the manager. An example is displayed in Figure 6. When $\pi = 0$, the PPS equals λ, regardless of v, which holds in existing dynamic agency models without evidence – e.g., DF07. When $\pi > 0$, the PPS is strictly lower than λ and decreases in evidence availability, for every $v < \bar{v}$.

The second part of claim (b) states that – as can be seen from Figure 6 again – the PPS increases with v, converging to λ as v approaches $\bar{v}$. This result verifies our intuition that evidence disclosure becomes less useful to reduce the agency conflict as the firm accumulates good performances (or, equivalently, its stock value rises). To understand this result, first note that evidence has no effect at the boundary $\bar{v}$, where the probability of default drops to zero. Thus, w_{dl} and w_{nl} converge to the same value at the boundary. Second, observe that w_i are all linear in v, and while the gap between w_h and w_{nl} is a constant, both increase with v at a slope strictly higher than one. It follows that the gap between w_{dl} and w_{nl} shrinks, and the PPS must increase. Empirically, the two predictions that emerge are that: (i) more evidence disclosure should lower the use of high-powered incentive compensation in general; and (ii) the effect should be stronger at firms that are experiencing low performance.

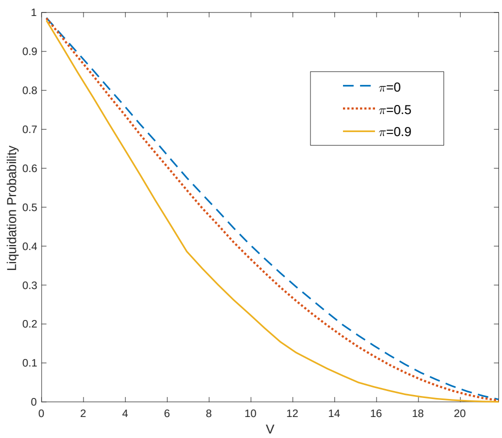


Figure 7: Simulated liquidation probability

Once we established that evidence reduces the PPS, it is intuitive that the continuation utilities (which correspond to the stock values in the implementation) vary less with cash flow shocks, leading to lower chances of default. Indeed, the reason why default occurs in such a dynamic agency model is that the optimal compensation is tighten to the firm's performance in order to prevent cash diversion. This mechanism unfortunately induces variations in stock values, which may shrink to zero after a sequence of bad cash flow shocks. For any given utility v promised to the manger, the numerical plot in Figure 7 shows that the probability of default decreases with the disclosure frequency. The reduced default risk also implies higher firm and investor values, as claimed in part (a) of Proposition 6. In addition, according to the definition in (12), higher firm value implies that the firm's credit spread drops with disclosure frequency, when we hold fix the amount of debt borrowed by the firm's manager.

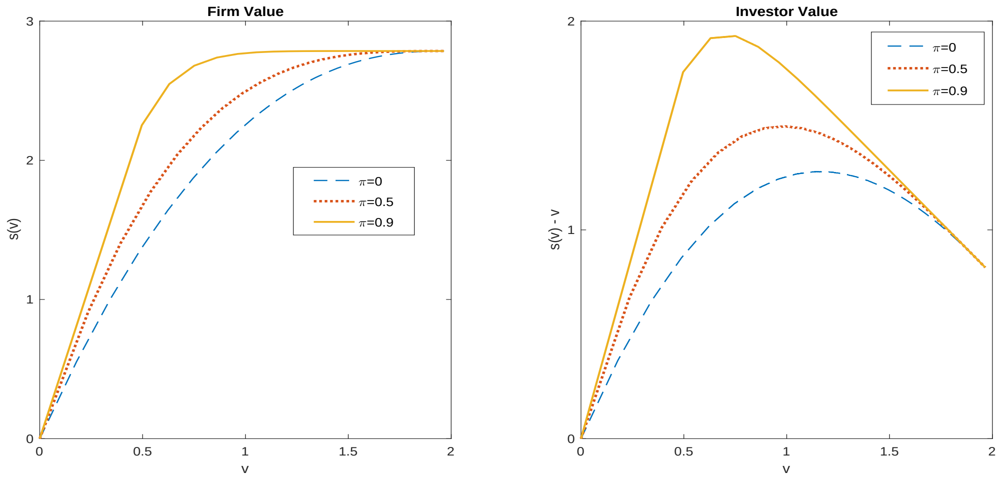


Figure 8: Firm and investor values

Finally, part (d) of Proposition 6 shows the downside of evidence: the n-period liquidation thresholds v^n all increase as disclosure become more frequent. For a cross-section of firms starting at the same state variable v (same amount of debt), the shortest time to default drops as π rises. In other words, on the shortest way to liquidation, better evidence implies faster default. According to Proposition 3, the fastest way to default on the equilibrium path occurs when a firm experiences a sequence of low-cash-flow shocks but does not have evidence to disclose. In contrast, the quickest time to reach the cash payout boundary lengthens as π rises. These patterns are plotted in Figure 9.

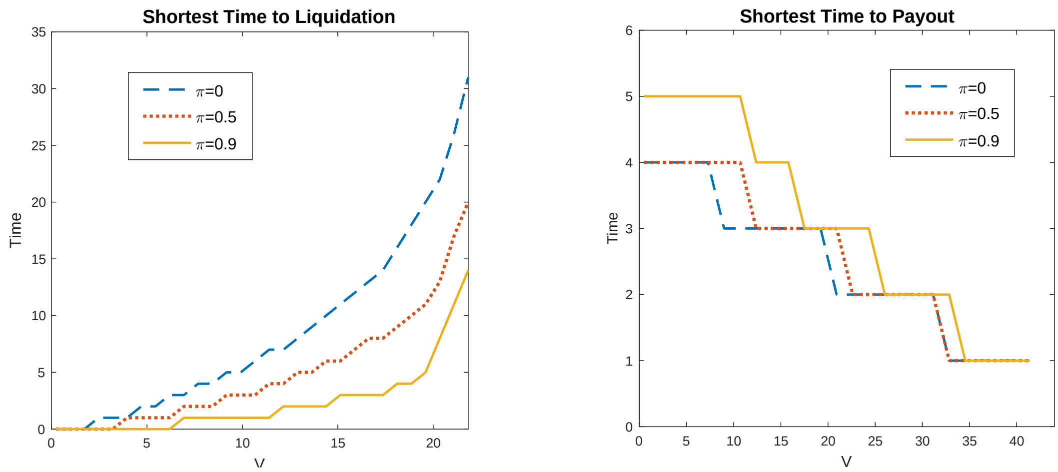


Figure 9: Simulated convergence time toward the two boundaries

The previous results illustrate how evidence disclosure opportunities impact the dynamic policies, given that the firm has already adopted the information technology (or will never adopt it, when $\pi = 0$). Now, we analyze how the likelihood of receiving evidence in the future and the cost of adopting the information technology affect the decision to exercise the investment option. Proposition 7 shows that, if evidence is more available, the firm is more likely to exercise the investment option overall, and the region in which adoption is delayed shrinks. If, instead, the investment cost is higher, the firm is more cautious and it is more likely to delay the investment.

Proposition 7. *The threshold cost is $\bar{c} > 0$. Both $\bar{c}$ and v^h increase in the evidence availability $\hat{\pi}$, while v^l decreases in $\hat{\pi}$. Furthermore, v^h decreases in the investment cost c, while v^l increases in c.*

9 Joint design of information and capital structures

Now that we have characterized the optimal contract, we can examine the managerial pay, the enterprise value, and the default probability at the issuance date. In particular, we study how these values change as the cost of adopting the technology c varies. The analysis in the previous sections was conducted for a given v, which implicitly treats any history as equally likely, irrespective of the firm's characteristics. At time zero, the investors optimally choose the initial leverage, taking into account the expected frequency of future disclosures by managers, and their value to the firm. At the same time, the investors optimally choose whether to exercise the investment option or not. Thus, we need to jointly solve the contracting and information design problems by explicitly accounting for the likelihood that any history arises, depending on the firm's covariates and the information technology it has access to.

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

$$v_0 = \arg\max_{v} \{\max[f(v), s(v; \hat{\pi}) - c] - v\} \tag{13}$$

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

In general, the time-zero properties are hard to characterize in any dynamic agency model, because they reflect the expectation of future firm performance and evidence disclosure. The firm's optimal starting point – which we labelled v_0 and represents its initial funding liquidity or, equivalently, its initial leverage – depends on the marginal value of increasing liquidity to the enterprise value. We find that, as evidence becomes available, this marginal value at v_0 can be either larger or smaller. On the one hand, for any v, better evidence implies lower default probability and so there is a greater surplus to be split between the manager and the investors. On the other, because evidence insures against bad luck, it reduces the value of providing a higher degree of funding liquidity v_0 to the firm in the first place. In other words, evidence substitutes cash as a means to solve the agency conflict. Which effects dominates depends on parameter values, as the next proposition states and Figure 10 depicts for a numerical simulation.

Proposition 8. *There exist two profitability thresholds $\underline{p} \in (0,1)$ and $\bar{p} \in (0,1)$ such that $\underline{p} \leq \bar{p}$ and, at the initial date $t = 0$:*

(i) If $p \leq \underline{p}$, both the equity value v_0, and the firm value $\max\{f(v_0), s(v_0;\hat{\pi}) - c\}$ can either jointly increase or jointly decrease in the investment cost c;

(ii) If $p > \bar{p}$, both the equity and and the firm's value decrease in the investment cost c.

It is not hard to see that, as generating evidence becomes cheaper, the surplus can rise. This is because the investors have an additional channel to govern the agency conflict vis à vis the firm's management. However, the opposite is also possible, and credit spreads may actually *increase* as disclosure opportunities improve. That is, managers who are expected to have access to better evidence may actually be worse off than the less informed ones, and the firms they run may be less likely to survive in the long run.

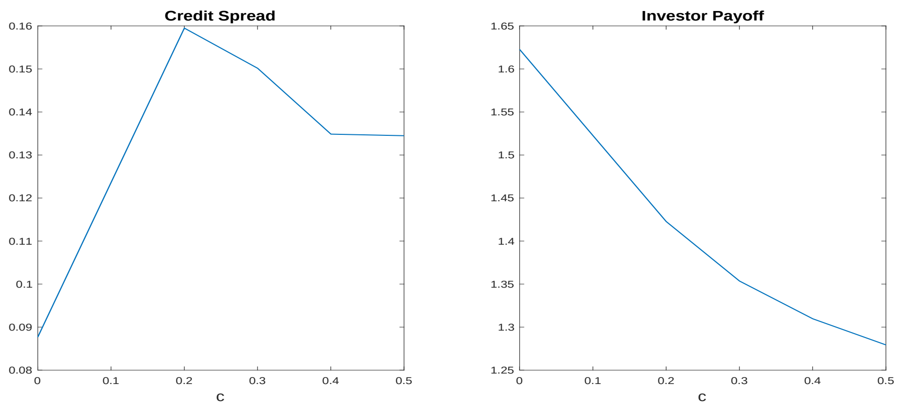


Figure 10: Credit spreads and investor value

Intuitively, this occurs because to incentivize disclosure and prevent diversion the principal faces the trade-off of either loading on the manager's rents or raising the liquidation 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 liquidation is more costly. If the investment cost is $c > \bar{c}$, the optimal policy loads more on managerial rents (larger v_0). As the cost drops, evidence is more likely to be produced and disclosed, which alleviates the liquidation concern. So the optimal policy loads on less rents (lower v_0).

We wish to highlight two implications of this result. First, it clarifies that a disclosure model is quite different from a monitoring model, in which the role of evidence is to reduce the agency conflict. While in a monitoring story a reduction of agency costs reflects always in a higher firm value and lower default risk, in a disclosure model the opposite might happen, especially at low profitability firms. Thus, the interaction between disclosure and capital structure is complex and varies predictably across firms. Second, it clarifies that disclosure plays two potentially conflicting roles, which can only be seen through the lengths of a contracting framework such as ours. For a given leverage and performance history of the firm, disclosure reduces default risk and positively affects the firm's value. However, by reducing the costs of debt, disclosure may provide incentives for investors to leverage up the firm, to a point where actually the overall default risk rises relative to the no-evidence case.

10 Conclusions

We study the implications of embedding voluntary disclosure of evidence in an otherwise standard dynamic agency model with non-verifiable cash flows. The model captures two key empirical regularities: (i) technological progress 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.

Evidence reduces the 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 investors trade-off higher liquidation odds with a lower managerial pay level.

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.

References

Acharya, V., P. DeMarzo, and I. Kremer, "Endogenous information flows and the clustering of announcements," *American Economic Review*, 2011, *101* (7), 2955–2979.

Balakrishnan, K., M. B. Billings, B. Kelly, and A. Ljungqvist, "Shaping liquidity: On the causal effects of voluntary disclosure," *Journal of Finance*, 2014, *69* (5), 2237–2278.

Bandiera, O., L. Guiso, A. Prat, and R. Sadun, "Matching firms, managers, and incentives," *Journal of Labor Economics*, 2015, *33* (3), 623–681.

Bebchuk, L. A. and J. M. Fried, *Pay without performance: The unfulfilled promise of executive compensation*, Harvard University Press, 2009.

Ben-Porath, E, E. Dekel, and B. L. Lipman, "Mechanisms with evidence: Commitment and robustness," *Econometrica*, 2019, *87* (2), 529–566.

Beyer, A. and I. Guttman, "Voluntary disclosure, manipulation, and real effects," *Journal of Accounting Research*, 2012, *50* (5), 1141–1177.

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

Bolton, P. and D. S. Scharfstein, "A Theory of Predation Based on Agency Problems in Financial Contracting," *American Economic Review*, 1990, *80* (1), 93–106.

_ **, H. Chen, and N. Wang**, "A unified theory of Tobin's q, corporate investment, financing, and risk management," *Journal of Finance*, 2011, *66* (5), 1545–1578.

Boone, A. L. and J. T. White, "The effect of institutional ownership on firm transparency and information production," *Journal of Financial Economics*, 2015, *117* (3), 508–533.

Bull, J. and J. Watson, "Evidence disclosure and verifiability," *Journal of Economic Theory*, 2004, *118* (1), 1–31.

Clementi, G. L. and H. A. Hopenhayn, "A theory of financing constraints and firm dynamics," *Quarterly Journal of Economics*, 2006, *121* (1), 229–265.

Crémer, J., "Arm's length relationships," *Quarterly Journal of Economics*, 1995, *110* (2), 275–295.

DeMarzo, P. and M. J. Fishman, "Optimal long-term financial contracting," *Review of Financial Studies*, 2007, *20* (6), 2079–2128.

_ **and Y. Sannikov**, "Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model," *Journal of Finance*, 2006, *61* (6), 2681–2724.

DeMarzo, P. M., I. Kremer, and A. Skrzypacz, "Test Design and Minimum Standards," *American Economic Review*, 2019, *109* (6), 2173–2207.

_ **, M. J. Fishman, Z. He, and N. Wang**, "Dynamic agency and the q theory of investment," *Journal of Finance*, 2012, *67* (6), 2295–2340.

Dresner Advisory Services, "Big Data Analytics Market Study," *Technical Report*, 2018.

Dye, R. A., "Disclosure of nonproprietary information," *Journal of accounting research*, 1985, pp. 123–145.

Enache, L., L. Li, and E. Riedl, "The disclosure of good versus bad news: evidence from the biotech industry," *Available at SSRN 3290822*, 2018.

Farboodi, M., R. Mihet, T. Philippon, and L. Veldkamp, "Big Data and Firm Dynamics," *NBER Working Paper*, 2019.

Fishman, M. J. and K. M. Hagerty, "The Optimal Amount of Discretion to Allow in Disclosure," *The Quarterly Journal of Economics*, 1990, *105* (2), 427–44.

Fu, S. and R. V. Krishna, "Dynamic financial contracting with persistent private information," *RNAD Journal of Economics*, 2019, *50* (2), 418–452.

Fuchs, W., "Contracting with repeated moral hazard and private evaluations," *American Economic Review*, 2007, *97* (4), 1432–1448.

Grossman, S. J., "The informational role of warranties and private disclosure about product quality," *Journal of Law and Economics*, 1981, *24* (3), 461–483.

Guttman, I., I. Kremer, and A. Skrzypacz, "Not only what but also when: A theory of dynamic voluntary disclosure," *American Economic Review*, 2014, *104* (8), 2400–2420.

Hart, S., I. Kremer, and M. Perry, "Evidence games: Truth and commitment," *American Economic Review*, 2017, *107* (3), 690–713.

Jung, W-O and Y. K. Kwon, "Disclosure when the market is unsure of information endowment of managers," *Journal of Accounting research*, 1988, pp. 146–153.

Koessler, F. and E. Perez-Richet, "Evidence reading mechanisms," *Social Choice and Welfare*, 2017, pp. 1–23.

Manso, G., "Motivating innovation," *Journal of Finance*, 2011, *66* (5), 1823–1860.

Marinovic, I. and F. Varas, "No news is good news: Voluntary disclosure in the face of litigation," *RAND Journal of Economics*, 2016, *47* (4), 822–856.

Meyer, M. A. and J. Vickers, "Performance comparisons and dynamic incentives," *Journal of Political Economy*, 1997, *105* (3), 547–581.

Mihet, R. and T. Philippon, "The Economics of Big Data," 2018.

Milgrom, P. R., "Good news and bad news: Representation theorems and applications," *Bell Journal of Economics*, 1981, pp. 380–391.

Murphy, K. J., "Executive Compensation: Where We Are, and How We Got There," *Handbook of the Economics of Finance SET: Volumes 2A & 2B: Corporate Finance and Asset Pricing*, 2013, p. 211.

Orlov, D., "Optimal design of internal disclosure," *Mimeo*, 2019.

Piskorski, T. and M. M. Westerfield, "Optimal dynamic contracts with moral hazard and costly monitoring," *Journal of Economic Theory*, 2016, *166*, 242–281.

Prat, A., "The wrong kind of transparency," *American Economic Review*, 2005, *95* (3), 862–877.

Shavell, S., “Acquisition and disclosure of information prior to sale,” *The RAND Journal of Economics*, 1994, pp. 20–36.

Shin, S. H., “Disclosures and asset returns,” *Econometrica*, 2003, *71* (1), 105–133.

Skinner, D. J., “Why firms voluntarily disclose bad news,” *Journal of accounting research*, 1994, *32* (1), 38–60.

Smith, C. W. and J. B. Warner, “On financial contracting: An analysis of bond contracts,” *Journal of Financial Economics*, 1979, *7*, 117–161.

Smolin, A., “Dynamic evaluation design,” *Available at SSRN 3051703*, 2017.

Verrecchia, R. E., “Discretionary disclosure,” *Journal of accounting and economics*, 1983, *5*, 179–194.

Zhu, J. Y., “Better Monitoring... Worse Productivity?,” *Mimeo*, 2018.

A Appendix

A-1. Two-period model

In this part, we consider the general case of the two-period model where $c \geq 0$. In other words, we explicitly consider the technology adoption decision. Proposition 1 in the paper will simply follow as a special case when one sets the cost parameter $c = 0$. We use the subscript T in the notation to highlight that the variables refer to the two-period model.

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

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

(b) If $p < \underline{p}_T$, the option is not exercised and the firm's probability of default is zero;

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

Because $\partial \underline{p}_T/\partial c > 0$ and $\partial \bar{p}_T/\partial c < 0$, a reduction in the strike price of the option c increases the probability of default of low profitability firms, while it decreases the probability of default of high profitability firms.

Proof. The $T = 1$ case taught us two facts: (i) whether the principal has exercised the option or not has no effect on the last period payoffs; (ii) the principal could find it optimal to exercise the option at the beginning only if the firm is sometimes liquidated in the first period. So, there are only three policies to consider: TT (the option is not exercised and the firm is liquidated in the first period when $x_1 = l$), NT (the option is not exercised and the firm is never liquidated) and OT (the option is exercised and the firm is liquidated in the first period only when there is no disclosure and $x_1 = l$). Without loss, 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 the manager 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 $u_{hh} = \delta/p$; the manager's utility at this policy is $U_M(TT) = p\delta$; the principal'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 $u_{hh} \geq \delta(1 + p)/p$; the manager's utility at this policy is $U_M(NT) = 2p\delta$; the principal'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, $u_{dlh} = \delta$. 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, $u_{hh} = \delta/p$; the manager's utility is $U_M(OT) = p\delta(1 + (1-p)\hat{\pi})$; the principal'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 p_T such that:

$$p_T := \frac{\sqrt{4\Delta l + (l + \delta - \Delta)^2} - (l + \delta - \Delta)}{2\Delta} \tag{A.14}$$

It is straightforward to show that $p_T \in (0,1)$. If $p > p_T$, the principal strictly prefers TT; if $p < p_T$, the principal strictly prefers NT; if $p = p_T$, the principal is indifferent between the two policies (or any randomization between them).

Second, either the principal exercises the option at p_T, or she never does. So, fixing $p = p_T$ and comparing $U_P(NT)$ and $U_P(OT)$ yields the threshold c_T:

$$c_T := \frac{\lambda\hat{\pi}}{2l^2\Delta}\left[l^2(1+\Delta^2) + \delta^2 + l(1-\Delta)2\delta - (\delta + l(1-\Delta))\sqrt{l^2(1+\Delta)^2 + 2l(1-\Delta)\delta + \delta^2}\right]$$

Focusing on $c < \bar{c}_T$, we need to consider two cases. If $p > p_T$, 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 $p \leq \bar{p}_T$, where:

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

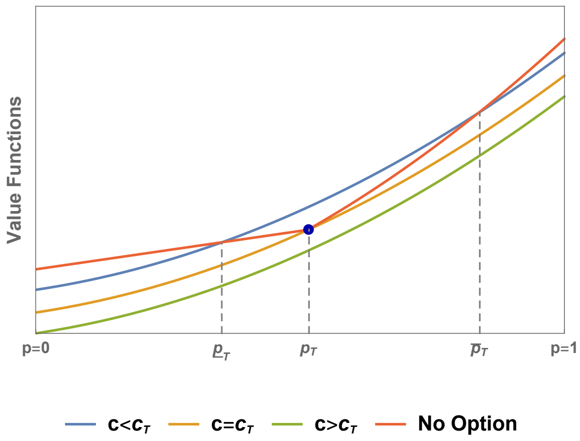
Finally, if $p < p_T$, we need to compare $U_P(NT)$ and $U_P(OT)$. We find that $U_P(OT) \geq U_P(NT)$ if and only if $p \geq \underline{p}_T$, where:

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

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

Figure 11 plots the firm value as a function of p, at different c and different adoption strategies. The three parallel curves obtain under the assumption the technology is adopted at $t = 0$. At $c = c_T$, the curve conditional on investing is tangent with that conditional on not investing only at p_T, meaning the technology is adopted only at p_T. If the cost drops to $c < c_T$, the set of firms that invest expands to $p \in [\underline{p}_T, \bar{p}_T]$. High profitability firms ($p > \bar{p}_T$) do not to invest, and liquidate when $x_1 = l$. Low profitability firms ($p < \underline{p}_T$), in contrast, never liquidate and do not invest either. If $c > c_T$, the technology is never adopted. Thus, a reduction in the strike price of the option c leads to increased adoption and more disclosure by both profitable and unprofitable firms.

Figure 11: The set of firms that invest in the technology



Proof of Proposition 1. Now, let's suppose that in the two-period model the firm is exogenously endowed with the information technology. So, set $c = 0$. In this case, the threshold p_2 is as defined in (A.14), and π_2 is given by the value of π such that $U_P(OT|c = 0) = U_P(NT)$. □

A-2. Infinite-horizon model

To proceed with the proofs for the infinite-horizon model, let us first show some basic properties of the firm value function. Let $\mathcal{C}$ be the space of continuous and bounded functions on the domain R_+. Let $\mathcal{F} := \{q \in \mathcal{C} : 0 \leq q \leq s^*\}$ be endowed with the 'sup' metric where $s^* = \frac{(1+r)\mu}{r}$ is the first best surplus. It is straightforward to see that $\mathcal{F}$ so defined is a complete metric space. Define the Bellman operator $\Gamma : \mathcal{F} \to \mathcal{F}$ as:

$$\begin{aligned} \Gamma q(v) = & \max_{\theta,\pi,u_i,w_i} \ (1-\theta)\mu \\ & + \frac{1-\theta}{1+r}\Big\{\pi[pq(w_{dh}) + (1-p)q(w_{dl})] + (1-\pi)[pq(w_{nh}) + (1-p)q(w_{nl})]\Big\} \\ & \text{s.t.} \quad (PK), (IC_g), (IC_b), (IC_n), (LL) \end{aligned} \tag{T}$$

As standard, the Bellman operator $\Gamma : \mathcal{F} \to \mathcal{F}$ is well defined and the constraint set is convex. Moreover, we can show the Bellman operator has the following property.

Lemma A.1. *Let $\mathcal{F}_1 = \{q \in \mathcal{F} : q(v) = s^* \text{ for all } v \geq \frac{(1+r)p\delta}{r}\}$. If $q \in \mathcal{F}_1$, then $\Gamma q \in \mathcal{F}_1$.*

Proof. Take any $q \in \mathcal{F}_1$ and $v \geq \frac{(1+r)p\delta}{r}$. Consider the following policy:

$$\theta = u_{dl} = u_{nl} = 0, \quad u_{dh} = u_{nh} = \frac{v}{p} - \frac{\delta}{r}, \quad w_i = \frac{(1+r)p\delta}{r} \; \forall i \in \mathcal{H}_1$$

It is easy to check that this policy satisfies all the constraints of (T). In addition, under this feasible policy we have that:

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

Thus, it must be the case that $\Gamma q(v) = s^*$. □

Proposition A.2. *The fixed point of* Γ*, which is unique and called* $s(\cdot)$*, 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 Γ is monotone (i.e., $q_1 \leq q_2 \;\Rightarrow\; \Gamma q_1 \leq \Gamma q_2$) and satisfies discounting (i.e., $\Gamma(q+a) = \Gamma q + \delta a$). As a consequence, the Blackwell's theorem implies that Γ is a contraction mapping on $\mathcal{F}$, and so it has a unique fixed point in $\mathcal{F}$. Let $\mathcal{F}_2 = \{q \in \mathcal{F} : q(v)$ is increasing and concave for all $v \in R_+\}$. As standard, Γ maps from $\mathcal{F}_2$ to $\mathcal{F}_2$. By Lemma A.1, it follows that the unique fixed point of Γ lies in $\mathcal{F}_1 \cap \mathcal{F}_2$. □

Proof of Proposition 2. Given 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 $\{\theta, u_i, w_i\}_{i\in\mathcal{H}_1}$ be the optimal policy at $\bar{v}$. To achieve first best, all the continuation values $w_{i\in\mathcal{H}_1}$ must be no less than v_1, and the liquidation probability θ must be zero. In addition, from the constraints (IC_g),(IC_b), (IC_n), and (LL), we have:

$$(1+r)u_{dl} + w_{dl} \geq (1+r)u_{nl} + w_{nl} \geq v_1 \tag{A.15}$$

$$(1+r)u_{dh} + w_{dh} \geq (1+r)u_{nh} + w_{nh} \geq (1+r)\delta + v_1 \tag{A.16}$$

The (PK) constraint implies that $v_1 \geq p\delta + \frac{v_1}{1+r}$, or $v_1 \geq \frac{(1+r)p\delta}{r}$. Moreover, Proposition A.2 implies that $v_1 \leq \frac{(1+r)p\delta}{r}$. Thus, we must have $v_1 = \frac{(1+r)p\delta}{r}$.

Now consider program (F). Let $\{\theta, u_i, w_i\}_{i\in\mathcal{H}_1}$ be the optimal policy at $\bar{v}$. To achieve first best, we need that (i) the investment option is never exercised; (ii) the policy satisfies $\theta = 0$ and $w_{nh}, w_{nl} \geq v_2$. As before, we can show that $v_2 \geq \frac{(1+r)p\delta}{r}$. In addition, since $f(v) \geq s(v; 0)$, Proposition A.2 implies that $f(v) = s^*$ for $v \geq \frac{(1+r)p\delta}{r}$, which means $v_2 \leq \frac{(1+r)p\delta}{r}$. 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 η as the Lagrangian multiplier of (PK). Moreover, let α_g, α_b, α_n be the multipliers of (IC_g), (IC_b), and (IC_n), respectively. Then the first order conditions are:

$$(1-\theta)\pi p s'(w_{dh}) = (1-\theta)\pi p\eta - \alpha_g \tag{FOC_{dh}}$$

$$(1-\theta)\pi(1-p)s'(w_{dl}) = (1-\theta)\pi(1-p)\eta - \alpha_b \tag{FOC_{dl}}$$

$$(1-\theta)(1-\pi)ps'(w_{nh}) = (1-\theta)(1-\pi)p\eta + \alpha_g - \alpha_n \tag{FOC_{nh}}$$

$$(1-\theta)(1-\pi)(1-p)s'(w_{nl}) = (1-\theta)(1-\pi)(1-p)\eta + \alpha_b + \alpha_n \tag{FOC_{nl}}$$

The envelope condition is:

$$s'(v) = \eta \tag{EN}$$

Proof of Lemma 1. Take any $v < \bar{v}$, and let $\{\theta, w_{i\in\mathcal{H}_1}\}$ be the optimal policy of the program (S) with $\pi = \hat{\pi}$.

First, we show that the Lagrangian multiplier $\alpha_b = 0$. Suppose not. Then, by the first oder 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. Complementary slackness then implies $\alpha_b = 0$, a contradiction.

Second, we show that it cannot be that $\alpha_g = \alpha_n = 0$. Suppose this is true. Then all the incentive constraints are not binding. Therefore, the firm's value $s(v)$ is the same as the one which solves problem (S) with only the promise keeping constraint (PK), and $w_i = (1+r)v$ for $i \in \mathcal{H}_1$ is feasible. Accordingly, the objective of (S) implies that $s(v) \geq s^*$. This is in contradiction with $v < \bar{v}$.

Third, we show that the Lagrangian multiplier $\alpha_n > 0$. Suppose not. Then, from the above result, it must be that $\alpha_g > 0 = \alpha_n$. The first oder conditions (FOC_{nh}) and (FOC_{nl}) together imply that $s'(w_{nh}) > s'(w_{nl})$. Thus, $w_{nh} < w_{nl}$ by concavity, contradicting with (IC_n).

Fourth, we show that the Lagrangian multiplier $\alpha_g > 0$. Suppose not. Then, by (FOC_{dh}) and (FOC_{nh}), we know $s'(w_{dh}) > s'(w_{nh})$. This further implies that $w_{dh} < w_{nh}$, and is in contradiction with (IC_g).

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

Proof of Proposition 3. The proof is divided into two parts. Part (a) shows that the proposed policy in the Proposition is optimal for problem (S). Part (b) derives the n-period liquidation thresholds.

Part (a). First, consider the case of $v \geq v^1$. Let $\theta = 0$, $w_{nl} = v - \hat{r}(\bar{v} - v)$, $w_{dl} = v$, $w_h = w_{nl} + (1+r)\delta$, and $u_i = 0$ for $i \in \mathcal{H}_1$. We can verify that they satisfy all the constraints of (S). Define an auxiliary function as follows:

$$g(v) = \mu + \frac{1}{1+r}\Big\{ps(w_h) + (1-p)[\pi s(v) + (1-\pi)s(w_{nl})]\Big\} \tag{A.17}$$

We now show that if $g(v) = s(v)$ at some $v \in [v^1, \bar{v}]$, then we must have $g'(v) = s'(v)$. Using the definition above, we derive:

$$(1+r)g'(v) = (1+\hat{r})[ps'(w_h) + (1-p)(1-\pi)s'(w_{nl})] + (1-p)\pi s'(v) \tag{A.18}$$

Given that $g(v) = s(v)$, we know the specified policy is optimal at such v. Thus, the policy must satisfy the first order conditions of (S). Summing (FOC_{dh}) (FOC_{dh}) (FOC_{nl}), and using the envelope condition (EN), we obtain:

$$ps'(w_h) + (1-p)(1-\pi)s'(w_{nl}) = [1-(1-p)\pi]s'(v) \tag{A.19}$$

Now we can plug (A.19) into (A.18) and rearrange to arrive at $g'(v) = s'(v)$.

Obviously, $g(v) \leq s(v)$. We can also verify that and $g(\bar{v}) = s(\bar{v})$. Suppose there is some $\hat{v} \in (v^1, \bar{v})$ such that $g(\hat{v}) < s(\hat{v})$. Then, there must exist some $\tilde{v}$ such that $g(\tilde{v}) = s(\tilde{v})$ and $g'(\tilde{v}) > s'(\tilde{v})$, which is a contradiction. As a consequence, $g(v) = s(v)$ and the specified policy is optimal.

Let $\hat{w}_i = \min\{w_i, \bar{v}\}$, and $\hat{u}_i = (1+r)(w_i - \hat{w}_i)$, for $i \in \mathcal{H}_1$. Evidently, this policy satisfies all the constraints of (S). Because $s(v) = s^*$ when $v \geq \bar{v}$, we also know that the modified policy is optimal.

Second, consider the case of $v < v^1$. Let $\underline{v} := \inf\{v : \theta(v) = 0\}$. From ($PK$) and ($IC_n$), we know $\underline{v} > 0$. Now suppose $\underline{v} < v^1$. The constraints of (S) imply that $w_{dl}(\underline{v} + \varepsilon) < \underline{v}$ for a sufficiently small $\varepsilon > 0$. Then, we have $s'[w_{dl}(\underline{v} + \varepsilon)] > s'(\underline{v} + \varepsilon)$, which is contradicted by (FOC_{dl}) at $\underline{v} + \varepsilon$. Thus, we must have $\underline{v} = v^1$. As a result, the (PK) constraint implies $\theta(v) = \frac{\min\{v, v^1\}}{v^1}$.

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:

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

Thus, $v^j = \frac{1}{1+\hat{r}}[v^{j-1} + \hat{r}\bar{v}]$. Moreover, by the optimal $w_h(v^1)$, $w_{dl}(v^1)$, 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 $v^1 = \frac{\hat{r}}{1+\hat{r}}\bar{v}$. We get the n-period liquidation thresholds by induction. □

Proof of Proposition 4. If $c = 0$, 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, observe that in the region close to the payout boundary $\bar{v}$, not exercising the option is optimal. This is because $f(\bar{v}) = s^* > s(\bar{v}; \hat{\pi}) - c$. So, v^h as in (6) 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 default boundary 0. This is because $f(0) = 0 > s(0; \hat{\pi}) - c$. Thus, v^l as in (6) is well defined, and the option is not exercised for $v \in [0, v^l]$.

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

In the case of $c > \bar{c}$, $s(v; \hat{\pi}) - c < s(v; 0)$ for any v, by the definition of $\bar{c}$ in (5). Thus, $s(v; \hat{\pi}) - c < f(v)$ for any v, and the option is never exercised. □

Proof of Proposition 5. We show that these securities implement both the payout, and the evolution of the optimal contract.

First, consider the case of $m = 0$, or the cash payout region. The manager's total payoff is λ fraction of the firm payout which is $\lambda d_h = u_h$. According to (8), (9) and (10), the subsequent credit line balance will all be zero, or $m_h = m_{nl} = m_{dl} = 0$. It follows that the continuation utilities in the contract are $w_h = w_{nl} = w_{dl} = \bar{v}$ by (7). So, in the implementation, the manager gets the same cash as at the optimal contract, and the firm always stays at first best, having zero probability of default.

Second, consider the case of $m > 0$. Given this credit balance, we can use (7) to (10) to derive the credit line balance of the subsequent period as:

$$m_{nl} = \frac{[1 + \hat{r}(\pi)](\bar{v} - v)}{\lambda}, \quad m_{dl} = \frac{\bar{v} - v}{\lambda}, \quad m_h = m_{nl} - (1 + r)\Delta \tag{A.20}$$

So, by (7), the manager's total payoff (by withdrawing all available credit) becomes

$$w_{nl} = v - \hat{r}(\pi)(\bar{v} - v), \quad w_{dl} = v, \quad w_h + (1 + r)u_h = w_{nl} + (1 + r)\delta, \tag{A.21}$$

which is the same as in Proposition 3. Note this derivation includes two scenarios, depending on whether the investment option is exercised or not. In the latter case, the relevant credit balance is simply m_{nl} or m_h, and the manager's payoff is w_{nl} or $w_h + (1 + r)u_h$. These values are obtained by setting $\pi = 0$ in (A.20) and (A.21). Thus, by Lemma A.2, we know that these payoffs are the same as in the optimal contract with the investment option. □

Proof of Proposition 6. Consider $\pi \in \{0, \hat{\pi}\}$ as a parameter of program (S).

Part (a). Take any continuation value $v \in [v^1, \bar{v})$. Let w_{dl}, w_{nl} be the optimal policies at v. Denote $s_\pi(v)$ as the derivative of the firm value s with respect to π at the given v. Then, in program (S), the envelope condition with respect to π reads:

$$\frac{(1+r)s_\pi(v)}{1-p} = s(w_{dl}) - s(w_{nl}) - s'(w_{dl})(w_{dl} - w_{nl}) \tag{A.22}$$

Because $s(\cdot)$ 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, as $s(\cdot)$ is linear in v for $v < v^1$, continuity of $s(\cdot)$ implies $s_\pi(v) > 0$ for $v < v^1$.

Part(b). Take any $v \in [v^1, \bar{v}]$. According to the definition in (1), the PPS can be calculated as

$$\begin{aligned}
\text{PPS} &= \frac{w_h + (1+r)u_h - [\pi w_{dl} + (1-\pi)w_{nl}]}{(1+r)\Delta} \\
&= \frac{\pi w_{nl} + (1+r)\delta - \pi w_{dl}}{(1+r)\Delta} \\
&= \frac{\pi[v - \hat{r}(\bar{v} - v)] + (1+r)\delta - \pi v}{(1+r)\Delta} \\
&= \lambda - \frac{\pi \hat{r}(\bar{v} - v)}{(1+r)\Delta}
\end{aligned}$$

The second line is from the equality of (IC_n), while the third line is from plugging in the policy expressions of w_{dl}, w_{nl}. Obviously, the PPS decreases in π and increases in v.

Now consider any $v < v^1$. As our PPS measure is only defined when the firm is not liquidated at the beginning of the period, we simply get:

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

Obviously, in this case, the PPS decreases in π.

Parts (c) and (d). These follow immediately from Proposition 3. □

Proof of Proposition 7. For $v \in (0, \bar{v})$, $\hat{\pi} > 0$ and Proposition 6 imply that $s(v)$ strictly increases in π, and so $s(v; \hat{\pi}) > s(v; 0)$. Continuity of $s(\cdot)$ implies that $\bar{c}$ in (5) is well defined, and $\bar{c} > 0$. Moreover, $\bar{c}$ increases in $\hat{\pi}$, because $s(v; \hat{\pi})$ does.

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

When c increases, similar argument implies that the drop in $s(v; \hat{\pi})$ is larger than that in $f(v)$. Thus, v^h becomes smaller and v^l shrinks. $\square$

Lemma A.2. *When the investment option is not exercised as in (F), the relevant optimal continuation utilities w_{nh} and w_{nl} are the ones defined in Proposition 3 by setting $\pi = 0$.*

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 also solve (F) when (IC_n) is not imposed. It follows that the policy $w_{nh} = w_{nl} = (1 + r)v$ and $\theta = 0$ is feasible. However, we know from the objective of (F) that $f(v) \geq s^*$, which is in contradiction with $v < \bar{v}$.

So, the optimal $w_{nh}(v)$ and $w_{nl}(v)$ are jointly determined by (PK) and the equality of (IC_n), resulting in the same expressions as in Proposition 3 with $\pi = 0$. $\square$

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 $i = 0, 1, 2...$, and the coefficients satisfy*

$$a_0 = 0,\ b_i = \frac{\mu(r + p)(1 + \hat{r})}{rp\delta}[\hat{r}(1 - p)(1 - \pi)/r]^i \tag{A.23}$$

Proof. According to the policy characterization in Proposition 3, when $p \leq r$ we have $w_h(v) \geq \bar{v}$ 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 it not being 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 (S) implies

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

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

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 (S) implies:

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

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 yields $b_i = \frac{\hat{r}}{r}(1 - p)(1 - \pi)b_{i-1}$. The result follows by induction. $\square$

Note that the slope b_i in (A.23) is a function of p and π. In what follows, we denote this slope as $b_i(\pi; p)$, and the derivative of this slope with respect to π as $b_i'(\pi; p)$.

Lemma A.4. *Given the discount rate r and any $i \geq 1$, $b_i'(0;p) < 0 \iff p > \frac{r}{(i+1)r+i}$.*

Proof. From the expression of $b_i(\pi;p)$ in (A.23), we get

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

It's easy to see that $b_i'(0;p) < 0$ if and only if $p > \frac{r}{(i+1)r+i}$. □

To facilitate the following proof, we define $h(r) = (1+\frac{p}{r})(1+r)$ and functions $g_i(p)$ where $i = 1, 2$ to be

$$g_1(p) =: \frac{2(1-p)^2(1-p)}{\lambda}\left(1 + \frac{l}{p\Delta}\right) \tag{A.24}$$

$$g_2(p) =: (6\sqrt{p}p + 4p + 3)p \tag{A.25}$$

and the unique p_i to be the value that satisfies $g_i(p_i) = 1$. We then define a threshold profitability value $\underline{p} =: \min\{p_1, p_2, \frac{1}{9}\}$.

Lemma A.5. *For any $p \leq \underline{p}$, we have:*

(a) $\frac{\mu(1-p)^2}{p\delta}h(p) \geq 1$;

(b) $2(1-p)^2 \geq \frac{3(1+\sqrt{p})^2}{3\sqrt{p}+2} \geq (1+\sqrt{p})^2$.

Proof. By the definition in (A.24), we know that $g_1(\cdot)$ is strictly decreasing and that $g_1(p) = \frac{\mu(1-p)^2}{p\delta}h(p)$. Thus, statement (a) holds. To get result (b), note that for any $p \leq \underline{p}$ we have:

$$g_2(p) - 1 = 3(1+\sqrt{p})^2 - 2(1-p^2)(3\sqrt{p}+2) \leq 0$$

Rearrange to get $2(1-p)^2 \geq \frac{3(1+\sqrt{p})^2}{3\sqrt{p}+2}$. Moreover, because $p \leq 1/9$, we have $\frac{3}{3\sqrt{p}+2} \geq 1$, which implies that the second inequality in statement (ii) holds. □

Lemma A.6. *For any $p \leq \underline{p}$, there exist $j \geq 2$, $p \leq \tilde{r} < \sqrt{p}$ such that*

(a) $\frac{\mu(1-p)^j}{p\delta}h(\tilde{r}) = 1$;

(b) $p > \frac{\tilde{r}}{(j+1)\tilde{r}+j}$.

Proof. Take any $p \leq \underline{p}$. First, by part (a) of Lemma A.5, there must exist some $j \geq 2$ such that:

$$\frac{\mu(1-p)^j}{p\delta}h(p) \geq 1 > \frac{\mu(1-p)^{j+1}}{p\delta}h(p) \tag{A.26}$$

Because $(1-p)h(p) \geq h(\sqrt{p})$ by (b) of Lemma A.5, we also have:

$$\frac{\mu(1-p)^j}{p\delta}h(\sqrt{p}) \leq \frac{\mu(1-p)^{j+1}}{p\delta}h(p) < 1 \tag{A.27}$$

Then, (A.26) (A.27) together imply that there exists a $\tilde{r} \in [p, \sqrt{p})$ such that $\frac{\mu(1-p)^j}{p\delta}h(\tilde{r}) = 1$.

Second, from the previous derivation we conclude that $\frac{\mu(1-p)^j}{p\delta}h(\tilde{r}) > \frac{\mu(1-p)^{j+1}}{p\delta}h(p)$. It follows that $h(\tilde{r}) > (1-p)h(p)$, which further implies that $\frac{p}{\tilde{r}} > \frac{2(1-p^2)}{1+\tilde{r}} - 1$. Moeover, because $\tilde{r} < \sqrt{p}$, Lemma A.5 (b) implies that $2(1-p^2) > \frac{3(1+\tilde{r})^2}{3\tilde{r}+2}$, which is equivalent to $\frac{2(1-p^2)}{1+\tilde{r}} - 1 > \frac{1}{3\tilde{r}+2}$. Thus, $\frac{p}{\tilde{r}} > \frac{1}{3\tilde{r}+2}$, or $p > \frac{\tilde{r}}{3\tilde{r}+2}$. Part (b) follows from the fact that $j \geq 2$. □

Lemma A.7. *Let π be either 0 or $\hat{\pi}$. As $p \uparrow 1$:*

(a) the liquidation threshold $v^1(\pi) \to \delta$, and $s(v;\pi) \uparrow s^*$ for any $v \geq v^1(\pi)$;

(b) $s'_+[v^1(\pi);\pi] = \lim_{v \downarrow v^1(\pi)} s'(v;\pi) \downarrow 0$.

Proof. First, from Proposition 3 we know that $v^1(\pi) = \frac{\hat{r}\bar{v}}{1+\hat{r}}$, which converges to δ as $p \to 1$. Second, for any $v \geq v^1$, from the objective of problem (S) we have:

$$s(v;\pi) \geq \mu + \frac{p}{1+r}s(w_h;\pi) \geq \mu + \frac{p}{1+r}s(v;\pi),$$

because $\theta = 0$. It follows that $\frac{(1+r)\mu}{1+r-p} \leq s(v;\pi) \leq s^* = \frac{(1+r)\mu}{r}$. Thus, as $p \uparrow 1$, we have $s(v;\pi) \uparrow s^*$. Moreover, because this result holds for any $v \geq v^1(\pi)$, it must be that $\lim_{v \downarrow v^1(\pi)} s'(v;\pi) = 0$. □

Lemma A.8. *When the investment option is not exercised, the liquidation threshold is $v^1(0) = p\delta$. Moreover, as $p \uparrow 1$, we have $f'_+[v^1(0)] \downarrow 0$.*

Proof. Note that the firm solves problem (F) when the option is not exercised. From Lemma A.2, we know that, at the liquidation boundary $v^1(0)$, continuation utilities are $w_{nl} = 0$ and $w_h = \delta$. Thus, the (PK) constraint implies that $v^1(0) = p\delta$. Then, for any $v \geq v^1(0)$, the objective of (F) implies:

$$f(v) \geq \mu + \frac{p}{1+r}f(w_h) \geq \mu + \frac{p}{1+r}f(v)$$

or $f(v) \geq \frac{(1+r)\mu}{1+r-p}$. Because $f(v) \leq \frac{(1+r)\mu}{r}$, we must have $f(v) \uparrow s^*$ as $p \uparrow 1$. Because this result holds for any $v \geq p\delta$, the result follows. □

Given the Lemmas A.7 and A.8, we can define a profitability threshold as follows:

$$\bar{p} = \inf\{\hat{p} : f'_+[v^1(0)] < 1, s'_+[v^1(\pi);\pi] < 1, \forall p > \hat{p}\}$$

and conclude that $\bar{p} < 1$.

Proof of Proposition 8. Note that, for any given p, if c is sufficiently close to zero the firm exercises the option at initiation and optimally chooses:

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

Thus, there exists a $\varepsilon > 0$ such that, when $c \in [0,\varepsilon]$, the manager payoff is kept at $\bar{v}_0$, and the firm's value $s(\hat{v}_0;\hat{\pi}) - c$ strictly decreases in c.

On the contrary, if $c \geq \bar{c}$, the firm never exercises its investment option. At initiation, it optimally chooses:

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

Take any $p \leq \underline{p}$. What remains to be shown for part (i) is that the manager's payoff and the firm's value can both increase in c. Consider the values j and $\tilde{r}$ in Lemma A.6, and let $r = \tilde{r}$. It follows that $s'(\tilde{v}_0;0) = b_j(0;p) = 1$, and $\tilde{v}_0 = v^{j+1}(0)$. Moreover, given Lemma A.4, we know that if $\hat{\pi}$ is sufficiently small, then $b_j(\hat{\pi};p) < 1 = b_j(0;p)$, because $b'_j(0;p) < 0$ (by part (b) of Lemma A.6). Therefore, $\hat{v}_0 \leq v^j(\hat{\pi}) < v^{j+1}(0) = \tilde{v}_0$. Because $s(v;\pi)$ is continuous in both v and π, a small change in π and a downward jump of initial continuation utility from $\tilde{v}_0$ to $\hat{v}_0$ implies $s(\hat{v}_0;\hat{\pi}) < s(\tilde{v}_0;0)$. Thus, the manager's payoff and the firm's value at initiation must both have an increasing part as the investment cost increases from 0 to $\bar{c}$.

Now consider part (ii), by taking any $p > \bar{p}$. In this case, the initial continuation utility chosen is either $\hat{v}_0 = v^1(\hat{\pi})$, if the option is exercised, or $\tilde{v}_0 = v^1(0)$, if it is not. Moreover, the value function, which can be either $s(\cdot;\hat{\pi}) - c$ or $f(\cdot)$, (weakly) decreases in the cost c. So, as c rises above 0, both the value function and the chosen continuation utility drop. As a result, firm value decreases at the initial continuation utility. □