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*

August 28, 2019

Abstract

We study a dynamic agency model in which firms have access to private evidence that predicts their short-term cash flows. Evidence is stochastic and cannot be fabricated, but it can be shrouded by the managers. When bad news are disclosed, the firm's investors find it optimal to remit current interest payments, which helps them insuring against bad luck. However, both when bad news are not preemptively disclosed and upon good news, the firm faces higher interest rates relative to the no-evidence case. On path, the value of firms that disclose more frequently is less sensitive to their performance. Their long-run leverage and default risk are lower, while their dividend-payout rates are higher. However, at the initial stage – where capital and information structures are jointly designed – default risk may increase with the availability of evidence, especially at low profitability firms. The benefit for firms from having evidence to disclose peaks at intermediate performance histories.

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), 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, storing and analyzing information. Adoption of the latest generation of such technologies, Big Data analytics, reached 59% of the firms surveyed by Dresner Advisory Services (2018) in 2018.[1] Firms have increased their access to evidence that leads performance, and are expected to disclose information to their investors more frequently. This paper studies the real effects of such enhanced disclosure opportunities in a dynamic agency setting, as well as the optimal patterns of technological adoption across firms.

A famous early adopter was the Germany national soccer team in 2014, during the World Cup that it eventually won. As Darcy Norman – who worked with that team – explains: "One of the key metrics we track is [...] how much power a player produces relative to their physiologic response to power. The more power a player generates during an exercise without burning too much energy, the more efficient and fit they are".[2]

Evidently, coaches can use the data to adjust drills and training goals, and to guide the choice of players for the next game. But the technology also allows them to better justify their choices with their principals. For example, a coach who faces a likely defeat may point to a quantitative metrics that helps to predict it, front-running the usual criticism of the choices of players and module that comes with a defeat. If there was always evidence to justify all choices, there would be unraveling à la Grossman (1981) and Milgrom (1981). Otherwise, in the more realistic case where evidence is stochastic, informed coaches have an incentive to be strategic in their disclosures, only revealing what favors them (Dye (1985), Shin (2003)). Thus, evidence brings about its own layer of conflict, which interacts with other agency problems. In a dynamic model, we study how this interaction changes the optimal contract and the allocation it implements.

As typical in the dynamic agency literature, we present our results in the context of a corporate finance problem. 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. Our key innovation is that we introduce the possibility for the firm to adopt an information technology.[3] Upon adoption, the technology produces stochastic evidence that predicts the

[1]The survey covers more than 5,000 firms globally. An additional 30% of the respondents signaled interest in adopting these technologies in the near future.

[2]As quoted in: How the Adidas miCoach System Has Helped Germany in the World Cup (2014).

[3]Similar dynamic agency models (without disclosure opportunities) have been studied by Bolton and Scharfstein (1990), Clementi and Hopenhayn (2006), DeMarzo and Fishman (2007), Biais et al. (2007). In the example of the coach of a sport team, it is perhaps more natural to think of the agency conflict as

one-period-ahead cash flows from then onwards.[4] While adoption is common knowledge, the realized evidence each period is privately observed by the manager, as in Dye (1985) or Shin (2003). The investors only observe evidence upon its voluntary disclosure.

The benefit of having access to evidence is that it may allow the investors to distinguish bad luck from bad behavior by the firm's manager, and so its disclosure is *always* optimally incentivized.This first observation, which is due to our optimal contracting approach, sets our analysis apart from the recent literature on dynamic evidence disclosure with fixed managerial compensation.[5] While in that literature bad news are not reported, in our model all the realized evidence is disclosed. Of course, this implies that the optimal contract must provide adequate incentives for such disclosures to happen.

Indeed, we find that the manager is optimally compensated for the disclosure of bad news. The result provides a novel rationale for 'pay without performance'. Existing explanations emphasize either the capture of boards by powerful executives (e.g., Bebchuk and Fried (2009)), or the need to motivate innovation by managers (Manso (2011)). Both stories do not make a distinction between bad performance with or without evidence to justify it. Our analysis suggests that the correlation between compensation and performance should be negative when bad evidence is disclosed, and positive otherwise. In our implementation, which follows DeMarzo and Fishman (2007), pay without performance consists of remitting the firm's short-term debt interest payment when the manager discloses that the bad performance is due to a transitory shock beyond his or her control.

However, and perhaps at first sight surprisingly, this effect is counterbalanced by a strictly higher interest rate faced by the firm both absent disclosure, and when cash flows are high.[6] Through the lens of our example, this means that while coaches are not punished for a defeat that they can prove to be due to circumstances beyond their control – such as an unusually high players' fatigue – the coaches also face a harsher punishment for a defeat that comes unwarranted, and they are rewarded less for their successes.

Interestingly, this 'show it and prove it' feature does not arise from the need to incentivize disclosure: at the optimal contract, the manager would have *strict* incentives to disclose bad news even if no-disclosure lead to less harsh consequences. Instead, it is driven by the investors' desire to insure the firm against inefficient liquidation, which is the negative consequence of bad luck. The firm's capital structure and performance

involving unobservable effort. However, as is well known, the two problems share the same qualitative predictions, and diversion is more tractable. This is the reason behind our modeling choice.

[4] All our results remain unchanged if evidence arises after the cash flows, and can certify them.

[5] See Acharya, DeMarzo and Kremer (2011), Guttman, Kremer and Skrzypacz (2014) and DeMarzo, Kremer and Skrzypacz (2019) where managers are assumed to maximize the stock price of their firms.

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

history determine its funding liquidity – that is, the amount of short-term funds it can borrow. The funding liquidity of the firm optimally adjusts over time so that its marginal value upon disclosure of bad news equals the expected marginal value across all other states. It follows that the marginal value of liquidity is strictly lower (i.e., its level is strictly higher) when bad news are disclosed, relative to when they are not, and that any pay granted upon the disclosure of bad news implies a lower pay across all other states.

Consistent with the view that evidence serves as self insurance against inefficient liquidation, on path the Pay-for-Performance Sensitivity (PPS) of the firm – which measures in percentage terms how much its expected value changes with the realized cash flow – decreases with the availability of evidence. This is intuitive, because pay-for-performance is the reason why inefficient liquidation occurs in such a model. Absent evidence, the PPS is a constant, which just makes the manager indifferent between diverting the cash flows or not. Evidence enables to lower the PPS without generating diversion incentives.

In addition, and unlike in previous models, the PPS increases in the firm's performance history. This is because as high cash flows accumulate liquidity in the firm, the probability of default decreases and evidence becomes less useful. The gap in compensation upon a low cash flow with and without disclosure reduces progressively, until it completely disappear at the dividend-payout boundary. Only at the boundary we observe the same degree of PPS as in Biais, Mariotti, Plantin and Rochet (2007) or DeMarzo and Sannikov (2006). This reconciles dynamic agency with the evidence in Bandiera, Guiso, Prat and Sadun (2015) that high-powered incentives are used by relatively more profitable firms.

In our sport example, it means that high-powered cash incentives for coaches should be used more, the better the past performance of their teams. Coaches who manage a team that has been very successful in the past face a larger variation in their expected compensation each game, relative to those who manage less successful teams. Such prediction critically depends on the presence of disclosure opportunities.

Because the firm is protected by limited liability, investors cannot simply rely on cash incentives in order to prevent diversion: they also need to liquidate their investment upon a sufficiently prolonged stream of bad performance. Similarly, to guarantee that a coach works in the interest of the team, a long streak of defeats should trigger the termination of his or her contract. So we ask: does evidence also reduce the need for liquidation, and therefore it lowers the firm's default probability? As it turns out, the answer depends on whether contracts are already in place, and there is some realized history of past performance, or whether we are considering the initial stage, at which contracts and information structures must be jointly designed.

Conditional on a given contract (or capital structure) and performance history, evidence reduces the probability that the firm will default. This follows immediately from our previous result that the pay-performance-sensitivity drops. Because the manager's lifetime pay (or, in our implementation, the firm's funding liquidity) is determined by its performance history, a lower pay-performance sensitivity implies that the funding liquidity level varies less, reducing default or the chance of exhausting the liquidity. Empirically, this implies that in secondary markets, when the firm's capital structure is unchanged, credit spreads should be negatively associated with the frequency of voluntary disclosures.

However, when the joint design of capital and information structures is considered, the probability of default may *increase* as the availability of evidence rises. Namely, the investors may optimally reduce the initial funding liquidity granted to the firm, and in so doing they set the firm on a path that may entail lower rates of firm survival. This happens because the marginal value of granting funding liquidity to the firm falls in the presence of evidence disclosure opportunities, as they help insuring against bad luck and default. In other words, evidence is a substitute for funding liquidity in dealing with the agency problems between investors and managers. Back to our sport example, this means that turnover among coaches may also *increase* with better information technologies.

Specifically, three cases can arise. While the investors must always be better off with greater evidence availability, the benefit for the manager and for the firm (whose value is equal to the aggregate surplus generated) may vary. First, there is a *win-win-win* scenario, in which evidence reduces the deadweight losses associated to default, and the increased firm value is split between the manager and the investors. Second, there is a *win-lose-win* scenario in which the probability of default falls, but so does managerial pay. Third, there is a *win-lose-lose* scenario, in which the possibility to adopt the information technology reduces managerial pay while it increases the probability of a future default. In this scenario, evidence exacerbates the conflict between rent-extraction by the investors and efficiency in the utilization of the firm's information technology.

Because these three cases have widely different empirical implications, we dig further into the conditions required for them to occur. We find that a firm's profitability is key in this respect. In particular, there exists a profitability threshold such that the *win-win-win* case prevails for all firms more profitable than that at the threshold. High-profitability firms are able to use evidence disclosure in order to decrease the risk of default and termination, and both the firm's manager and the investors benefit from having access to it. This result is important to clarify the difference between our evidence disclosure model and a typical monitoring setting. While in most monitoring models the information would be used by investors to curb managerial pay, this need not be true when the information technology is decentralized and the realized evidence needs to be disclosed. To our knowledge, we are the first to theoretically point out such distinction. In the empirical literature, Mas (2016) shows that mandatory disclosure was associated with increased managerial pay at the introduction of the 1934 Securities and Exchange Act. However, we are not aware of work relating compensation to voluntary disclosure.

In contrast, the *win-lose-lose* case may occur at low profitability firms. Such firms are relatively more likely to generate low cash flows at any given period. Therefore, absent evidence disclosure, it is costly for the investors to aggressively bring them to liquidation after a streak of bad performances. Evidence makes it more desirable for investors to be aggressive, because liquidation only occurs when a bad performance is not accompanied by managerial disclosure. Thus, default risk may rise with the availability of evidence. It follows that, for the investors to be better off, managerial pay must drop. This could be related to the recent surge in bonds default rates (e.g., Becker and Ivashina (2019)).

Our model delivers clear-cut predictions also on the capital structure side, because disclosure opportunities affect both the firm's leverage and its dividend payout policy. As in all dynamic agency models, it is optimal to backload cash payments to the manager. The firm issues dividends only after a sequence of positive shocks, which implies that its dividend payout rates rise with its survival probability. Because firms that disclose more frequently have higher survival probabilities – conditional on any given performance history – it follows that they also display higher dividend payout rates. In contrast, because leverage is the highest when the firm is close to default and termination, higher survival probabilities imply lower long-run leverage ratios.

Empirically, these results suggest that studies of the real effects of disclosure should pay attention to default risk, which is tightly correlated with a firm's capital structure, in addition to discount rates. In fact, while the existing empirical work uncovers a positive causal effect of voluntary disclosure on the liquidity of a firm's securities and on its cost of capital,[7] such effects might be mitigated by the cash flow channel we identified, especially at low profitability firms, due to the higher default risk they face.

As for the patterns of technological adoption, the set of adopting firms consists of those that experienced intermediate performance histories. The region is characterized by two performance-related thresholds. Below the lower threshold, the value of the firm as a going concern is too low to justify spending resources on the technology. Above the upper threshold, the benefits are too low, as the firm is already far from its default boundary. As the cost of the technology falls, the thresholds diverge and the adoption

[7]See Francis, Nanda and Olsson (2008), Balakrishnan, Billings, Kelly and Ljungqvist (2014) and Boone and White (2015), and the surveys by Healy and Palepu (2001) and Leuz and Wysocki (2008). Bertomeu and Cheynel (2016) survey theory work that interprets such findings in CAPM-like models.

region expands. Importantly, this pattern of adoption is different from that of physical investment options, such as those studied in DeMarzo, Fishman, He and Wang (2012). While the value of physical options typically increases with past performance, the value of information-related options peaks at intermediate performance histories.

Finally, modeling information technologies as options is useful to clarify that they affect the valuation of both adopting and non-adopting firms. While the low-performance firms condition their adoption on receiving a stream of positive cash flow shocks, the high-performance ones keep the option as an insurance policy, intending to exercise it in the future only if they receive a sufficiently negative cash flow shock. As this increases valuations before adoption, to estimate the benefits of new technologies by comparing the performance of adopters and non-adopters is methodologically flawed. The non-adopters are fundamentally different from the firms that existed prior to the advent of the new information technologies.

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 discusses the impact of disclosure on the policy dynamics and on other variables of interest. Section 7 shows the patterns of information technology adoption. Section 9 implements our optimal contract by means of short and long-term debt, and equity. Section 8 discusses the initiation problem, when securities are issued. Section 10 concludes.

2 Literature Review

Our paper is related to several literatures. Theoretically, it builds on the dynamic agency model developed by Clementi and Hopenhayn (2006), Biais et al. (2007) and DeMarzo and Fishman (2007). A recent strand of papers on dynamic moral hazard introduced information production and dissemination possibilities, and studied their consequences on second best allocations (e.g., Fuchs (2007), Piskorski and Westerfield (2016), Smolin (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 also related to the theoretical work on voluntary disclosure (e.g., Beyer and Guttman (2012), Acharya, DeMarzo and Kremer (2011), Guttman, Kremer and Skrzypacz (2014), Marinovic and Varas (2016) and DeMarzo, Kremer and Skrzypacz (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.

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

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 $\Delta := h - l$, $p := P(x_t = h) \in (0, 1)$, and $\mu := \mathbb{E}(x_t)$. The firm is owned by a Principal (P) – who represents the firm's investors – and is operated by a Manager (M). Both P and M are risk-neutral and discount future consumption at the same rate $r \in (0, 1)$.[8]

Moral hazard. We introduce the possibility of moral hazard by assuming that M privately observes the realized cash flows $\{x_t\}$. By misreporting a good cash flow, claiming it to be bad, M can divert Δ output and obtains a private benefit of $\delta := \lambda\Delta$, where $\lambda \in (0, 1]$ represents the severity of the moral hazard problem.[9] 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 P can choose make a one-time 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 ($\pi = 0$).[10] To make this investment, P must spend a fixed cost of $c \geq 0$.[11] 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. Thus, IT adoption corresponds to the exercise of a one-time American option with infinite maturity that cannot be reversed, with strike price c.

Once the option is exercised, P expects evidence to be available with probability $\hat{\pi}$, but she never knows whether M acquired some evidence or not. So, at each date t, M chooses whether or not to *voluntarily disclose* the realized evidence to P. We denote the disclosure action by $a_t \in A := \{d, n\}$, where d stands for disclosure and n for non-disclosure. If M discloses the evidence, investors will predict the cash flow accurately. That is, $p(x_t = h|e_t = g) = p(x_t = l|e_t = b) = 1$.[12] Because (i) disclosure is always incentivized, and (ii) the availability of evidence is conditionally independent from the realized cash flow, no-disclosure has no impact on the investors' beliefs. That is, absent evidence disclosure, P predicts that cash flows are high with probability p.

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

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

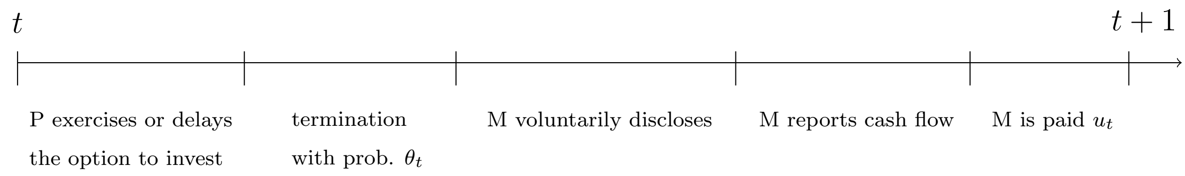
[10] The fact that in the absence of technological investment $\pi = 0$ is just a normalization. All our results go through unchanged if we assumed that, absent investment, the firm would have some positive $\pi' < \hat{\pi}$.

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

[12] Since the effects of evidence on our outcomes of interest are already non-monotone and complex with perfect evidence, it does not seem necessary to also consider imperfect correlation in this model.

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

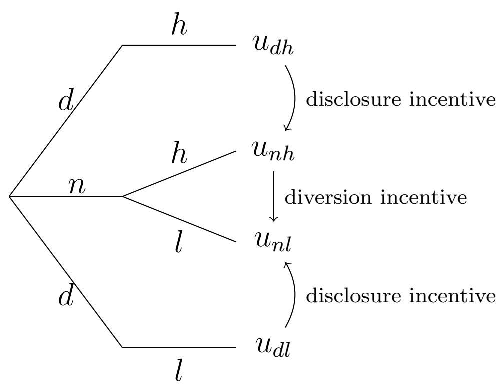
Figure 1: Timing in period t, prior to exercising the option



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 M are denoted by u_i, for $i \in \mathcal{H}_1$. The contract must provide two kinds of incentives: (i) to prevent the agent from diverting cash flows, which requires $u_{nh} \geq \delta + u_{nl}$; (ii) to disclose information, which requires both $u_{dh} \geq u_{nh}$ and $u_{dl} \geq u_{nl}$. Trivially, it is optimal for P to set $u_{nl} = u_{dl} = 0$ and $u_{dh} = u_{nh} = \delta$, and since $c > 0$ the option is never exercised.

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

As we proceed in the infinite horizon case, we break down the analysis in two parts: in the first, we characterize the optimal contract conditional on a given evidence-availability $\pi \in \{0, \hat{\pi}\}$; in the second, we compare the case of $\pi = 0$ (no adoption) with the case of $\pi = \hat{\pi}$ (adoption), and we derive the optimal adoption decision. Note that because $T = 2$, and evidence can only be useful at $t = 1$, adoption is never delayed in the two-period model. Either the option is exercised at $t = 0$, or it will never be exercised.

Proposition 1. *If $T = 2$, there exists a threshold $\bar{\pi} < 1$ such that, for $a \in A$:*

(a) If $\hat{\pi} < \bar{\pi}$, the firm is never terminated. M's compensation is $u_{ahh} = \frac{(1+p)\delta}{p}$, $u_{alh} = \delta$;

(b) If $\hat{\pi} > \bar{\pi}$, then the firm is terminated at $t = 1$ when a low cash flow is reported and there is no disclosure. M's compensation is $u_{ahh} = \frac{\delta}{p}$, $u_{dlh} = \delta$;

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

Moreover, $\bar{\pi} > 0$ if and only if $p < \bar{p}$.

Proposition 8 clarifies a few implications of evidence on contracts and allocations. First, evidence is *only* useful if the firm is terminated with some probability at $t = 1$, when the low cash flow arises. In case (a), when termination does not occur, both the firm and the investors value do not depend on $\hat{\pi}$: evidence is irrelevant again.

Second, moving from $\pi = 0$ to $\pi = \hat{\pi}$ we can have one of three scenarios:

1. For highly profitable firms (i.e., $p \geq \bar{p}$) the probability of default drops from $(1-p)$ to $(1-p)(1-\hat{\pi})$, while M's compensation rises from $p\delta$ to $p\delta(1+(1-p)\hat{\pi})$. P gains because the revenues from a lower default probability more that offset the rise in managerial rents. This is the scenario we labelled *win-win-win*;

2. For low profitability firms ($p < \bar{p}$) if the technology seldom generates evidence (i.e., $\hat{\pi} < \bar{\pi}$), the probability of default was and remains zero, and nothing changes. Jumping ahead for a moment, it is easy to see that such technologies would never be adopted for any positive cost $c > 0$;

3. For low profitability firms ($p < \bar{p}$) if the technology frequently generates evidence (i.e., $\hat{\pi} \geq \bar{\pi}$), the probability of default rises from zero to $(1-p)(1-\hat{\pi})$, while M's compensation drops from $2p\delta$ to $p\delta(1+(1-p)\hat{\pi})$. P gains because the revenues from lower managerial rents more than offset the increase in the default probability. This is the *win-lose-lose* scenario: both M's compensation and the surplus fall.

Summing up, while the investors always benefit from a better information technology – that is, a higher π – both M's compensation and the probability of default may rise or fall with it. These scenarios are depicted in Figures 3a and 3b, where U_M denotes the expected managerial compensation at $t = 0$, and U_P the expected principal's payoff.[13] To measure cash incentives, we plot the Pay-Performance Sensitivity (PPS), defined as:

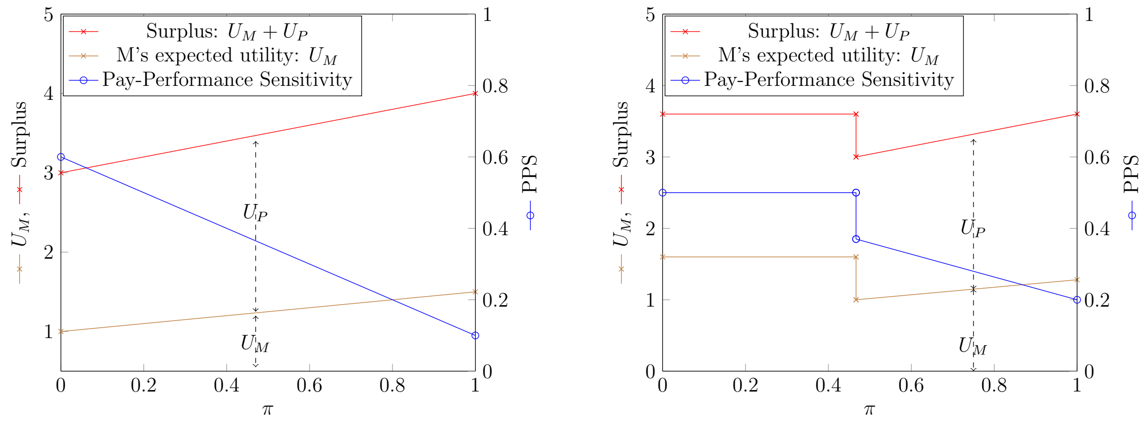
$$\text{PPS} := \frac{\mathbb{E}(U_M \,|\, x_1 = h) - \mathbb{E}(U_M \,|\, x_1 = l)}{\Delta} \tag{1}$$

The PPS measures in percentage terms how M's compensation changes with firm performance. In both figures, the PPS weakly drops with π, clarifying that the role of evidence is to dampen cash incentives by insuring both agents against bad states of the world.

The two-period model clarifies that evidence reduces the use of cash incentives, and it has heterogeneous effects on both the firm's value, and its probability of default. Nevertheless, some of our results cannot be understood by means of such a simple model. For instance, conditional on termination happening on-path, M's expected compensation when bad news are not disclosed at $t = 1$ is always equal to zero, irrespective of π. In contrast, as we shall see, in the full model such compensation decreases with π.

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

Figure 3: Comparative statics in the two-period model



(a) High Profitability Firms: $p \geq \hat{p}$ (b) Low Profitability Firms: $p \leq \hat{p}$

5 Infinite-horizon model

In this section, we first formulate the firm's problem in the infinite-horizon environment, and then characterize policies and their dynamic features.[14]

5.1 Contracting

As is well known, when shocks are i.i.d., the agent's continuation utility v is a state variable that summarizes all relevant information in any given history. For any state v, the contract specifies the probability of liquidating the firm at the beginning of the period θ, and then compensates M either with cash, or with promised utility contingent on M's actions. When evidence is disclosed, the contract pays $\mathbf{u}_d = (u_{dh}, u_{dl}) \in \mathbb{R}^2$ to M 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 M cash $\mathbf{u}_n = (u_{nh}, u_{nl}) \in \mathbb{R}^2$, and promises continuation utility $\mathbf{w}_n = (w_{nh}, w_{nl}) \in \mathbb{R}^2$.

Whether it is worthwhile to invest in the costly information technology or not, and if so, when to make the investment, all depend on the value that this option brings to the firm. To evaluate the moneyness of this evidence-generating option, we first consider the optimal contracting for the firm given the investment has already been made. We then step back and determine the optimal option exercise patterns.

[14]More rigorous arguments which guarantee that the recursive representation of our problem is appropriate are standard and so we leave them to the Appendix.

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

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

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

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 value. 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 M'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 P's perspective also maximizes firm value (i.e., surplus), given a utility v promised to M. [15] Thus, the optimal contract solves the following dynamic program:

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

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

The objective function of (S) reflects the fact that (i) with probability θ, termination 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 terminated, in which case the firm's value depends on both whether M 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 solves the case where the option is never exercised (or, equivalently, there is no option). This is because, in the latter case, the only relevant control variables are those conditional on no-disclosure. Thus, we use $s(v; \hat{\pi})$ and $s(v; 0)$ to denote the value functions in program (S) corresponding to these two possible cases.

5.2 Investment option

We next analyze the investment decision, i.e. the decision of whether and when to exercise the investment option. Suppose that, for a given history represented by v, the firm has not 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 terminated 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 M.

Come tomorrow, the firm can either invest c and obtain the value of $s(w_{ni}) - c$ (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 $s(w_{ni}) - c > f(w_{ni})$, the firm exercises the investment option tomorrow. Otherwise, it waits until at least one more period to invest. Thus, the firm's

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

problem when the investment has not been undertaken yet can be formulated as follows:

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

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

5.3 Initiation and payout

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

$$v_0 = \arg\max_v \{\max[f(v), s(v; \hat{\pi}) - c] - v\} \quad (2)$$

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

Because termination is inefficient, it is optimal to delay M's cash compensation until the continuation utility v is sufficiently large. Without loss of generality, we shall assume that M is paid by cash whenever the firm is indifferent between paying him or her right away, or 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,

$$\bar{v} := \inf\{v : f(v) = s^* \text{ or } s(v; \hat{\pi}) = s^*\} \quad (3)$$

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 M's payoff dynamics may depend on the availability of evidence, the next result shows that actually the value $\bar{v}$ is a constant, irrespective of evidence availability. In contrast, the cash compensation granted to the manager varies with both the option exercise strategy, and the level of $\hat{\pi}$ in the short-run.

Proposition 2. *The cash payout boundary is:*

$$\bar{v} = r^{-1}(1+r)p\delta. \quad (4)$$

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

$$u_{al}(v) = 0, \quad u_{ah}(v) = \max\big\{(1+r)\delta - (1+\hat{r})(\bar{v} - v),\ 0\big\} \quad (5)$$

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

Proposition 2 shows that the payout boundary $\bar{v}$ *does not* depend on the possibility of generating and disclosing evidence. Even when the cost of generating evidence is infinitely large (or $\hat{\pi}$ is infinitely small or zero), the boundary does not change. At. the cash payout boundary, no cash payment is made to M whenever a low cash flow is reported. When the firm is one-step away from $\bar{v}$, M 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 from 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, termination never occurs moving forward, and therefore evidence is no longer useful.

6 Impact of evidence disclosure

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

6.1 Policy characterization

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

Before reaching the payout boundary, M is incentivized by variations in her promised continuation values. If the continuation value grows high enough to reach $\bar{v}$, the firm is never terminated, all incentive constraints become slack, and the firm's value reaches the first best. When M's continuation value is at intermediate levels, termination may occur after a sequence of low cash flows that are not accompanied by voluntary disclosures by the managers. When the continuation value 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 characterize the dynamics, we define the thresholds such that no termination can possibly occur in the next n periods to be:

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

These values correspond to the lowest continuation values such that the firm can survive with certainty for at least n periods. For example, if $v > v^1$ the firm will not be terminated in the current (or one) period, but may be terminated in the next period. These thresholds are related to the previous definitions of termination probability θ and the payout boundary $\bar{v}$. Specifically, stochastic termination 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, termination never occurs if $v \geq \bar{v}$.

Because the firm can be terminated, any randomization of continuation utility is costly for both parties, implying that the firm's value $s(\cdot)$ is concave. Using concavity and the optimal conditions of the firm's problem (S), we can find 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 M and P are not affected by evidence disclosure. Contingent on a good performance being reported, M does not divert and so there is no need to further condition M's payoffs on evidence disclosure. Thus, in the rest of the paper we do not distinguish M's payoff across states dh and nh. Accordingly, we denote w_h and u_h as the continuation value and the cash payment, respectively, conditional on cash flows being high. Notice that this property would *not* generally hold in a monitoring setting, in which P observes the evidence directly. In that case, P could further reduce the payment to M when evidence is available, without violating any constraint.

In contrast, the optimal contract provides strict incentives for M to disclose bad news: i.e., $w_{dl} > w_{nl}$. Punishing M for a bad performance is costly to P because it induces more inefficient termination. If the evidence shows that the bad performance is not caused by M's behavior, but instead by bad luck, then M should not be punished. Promising M higher utility in the state dl does not worsen the diversion problem, but improves efficiency by reducing the probability of termination. 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 volatility of the agent's continuation utility, and the pay-for-performance sensitivity. As we mentioned in the introduction, this is akin the solution to a classical insurance problem for a fixed budget.

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

If stochastic termination does not occur at the beginning of the period, the policies $w_i(v)$ for $i \in \mathcal{H}_1$ are the same as $w_i(v^1)$. In addition, firm value in this region is linear. Given this characterization, we clearly see that – after a low performance – the contract possibly promises the manager a higher continuation utility when bad news is disclosed ($w_{dl} = v^1 > v$), deviating significantly from the case of $\pi = 0$ considered in previous work. In this region, whenever M discloses evidence of transitory bad luck P compensates for disclosure by raising the promised utility to v^1, independently from the continuation utility entering the period. The reward depends on $v^1 - v$.

In the region above v^1, M is still rewarded for disclosing bad luck: $w_{dl} = v$.[16] The contract forgives the low performance today, and starts tomorrow as if the history is the same as before the current low cash flow. This mechanism does not affect M's diversion incentives, because M cannot fabricate evidence. Moreover, because the volatility in M's continuation utility is costly for investors, it is optimal to set w_{dl} as close to v as possible.

Finally, as standard, the optimal contract rewards good luck. Proposition 3 shows that the ranking of continuation utility does not depend on the levels of v and π. M gets the largest continuation utility contingent on high performance, the lowest one contingent on low performance and no disclosure, and the middle one contingent on disclosure of bad news. This pattern implies that, on the fastest route to termination, M never discloses evidence and always reports low cash-flow. So, the n-period threshold can be explicitly derived from the policy functions. Evidently, both the termination thresholds and the policy dynamics depend on M's disclosure behavior and on the availability of evidence.

[16]Notice that to keep the continuation utility fixed across times effectively requires a payment from P to M, because of the time value of money.

6.2 Comparative statics

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

$$\text{PPS} := \frac{\mathbb{E}(v \,|\, x_1 = h) - \mathbb{E}(v \,|\, x_1 = l)}{\Delta} \tag{6}$$

This measure indicates – in percentage terms – how managerial compensation changes as a function of the firm's performance.

Turning to credit spreads, which measure the firm's default risk, the standard definition is: Credit spread $=$ (1 $-$ recovery rate) $\times$ Pr.[default], where the recovery rate denotes the fraction of the firm's value recovered by creditors upon default and/or liquidation. Because we normalized the recovery rate to zero, the expression simplifies to:

$$\text{Credit (or CDS) spread} = 1 - \frac{s}{s^*}, \tag{7}$$

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 P's and the M's expected payoffs). Importantly, this is distinct from the agency cost, which would consist in the sum: Credit spread + managerial rents.

Proposition 4. *When the availability of evidence $\hat{\pi}$ 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 increases with v for any given $\hat{\pi}$

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

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

To discuss Proposition 4 and the economic intuition behind it, start from claim (c), which analytically illustrates the properties of the policy function when the disclosure frequency π varies. The managerial payoff w_{dl} either remains the same as the beginning of period utility v, or it jumps up to v^1, regardless of the level of 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. This is akin the redistribution of income across states in an insurance problem. Figure 4 illustrates this pattern of how policies move as π increases for a given value of v. As π increases, the continuation utility is less likely to move downward, but its lowest value becomes worse.

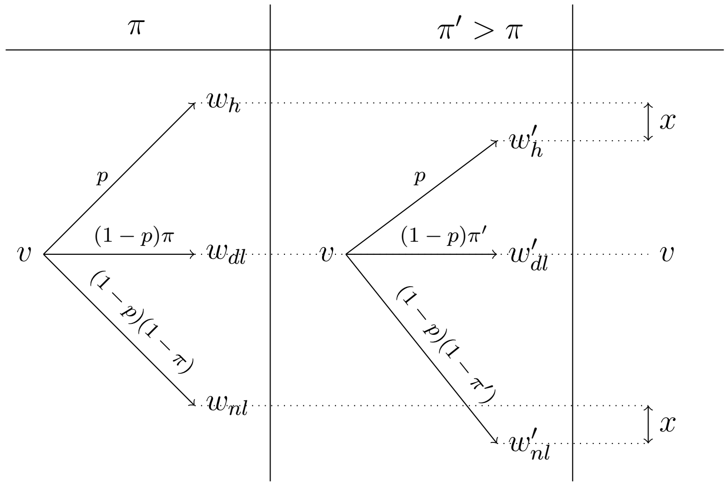


Figure 4: Impact of Evidence on Managerial Payoff

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

$$\text{PPS} = \lambda - \frac{\pi \hat{r} \left[\bar{v} - \max(v, v^1)\right]}{(1 + r)\Delta} \tag{8}$$

Obviously, this measure depends on both v and $\hat{\pi}$. Two effects drive the PPS 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 4 states, the dominant effect is *always* for evidence to lower the PPS, for every v. This is not surprising, as PPS is a necessary evil of optimal contracts: it prevents diversion, but it produces a positive default probability. An example is displayed in Figure 5. When $\pi = 0$, the PPS equals λ, regardless of v. This result holds in existing dynamic agency models without evidence – e.g., DeMarzo and Fishman (2007). For $\pi > 0$, the PPS is strictly lower than λ for every $v < \bar{v}$.

The second part of claim (b) states that – as can be seen from Figure 5 again – the PPS increases with v, converging to λ as v approaches $\bar{v}$. 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 all the ws are linear in v, and the gap between w_h and w_{nl} is a constant independent from v. Because both increase with v at a slope strictly higher than one, and which depends on π, it follows that as the gap between w_{dl} and w_{nl} shrinks, the PPS must increase. Empirically, the two predictions that emerge are that: (i) better information technologies 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.

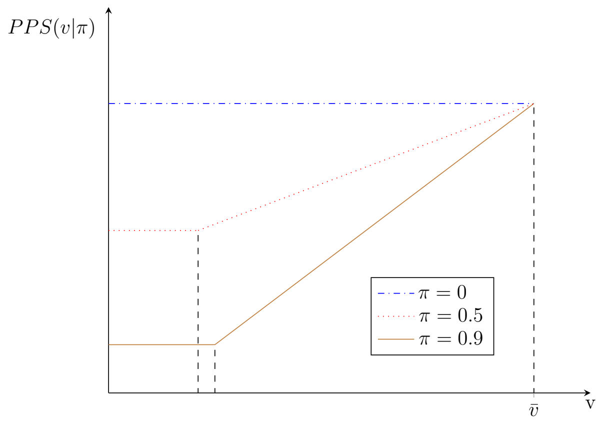


Figure 5: Pay-for-Performance Sensitivity

Once we established that evidence reduces the pay-performance sensitivity, it is not surprising that it increases firm value by reducing the probability of default *for every* $v < \bar{v}$, as claimed in part (a) of Proposition 4 and plotted in Figures 6 and 7 for some numerical simulations. Indeed, the only reason why the probability of default is positive in such a dynamic agency model is that the optimal compensation requires to be tighten to the firm's performance in order to prevent cash diversion. The figures stress this point simulating a set of possible histories for a cross section of firms that are identical in all dimensions, except for having different technologies for evidence disclosure.

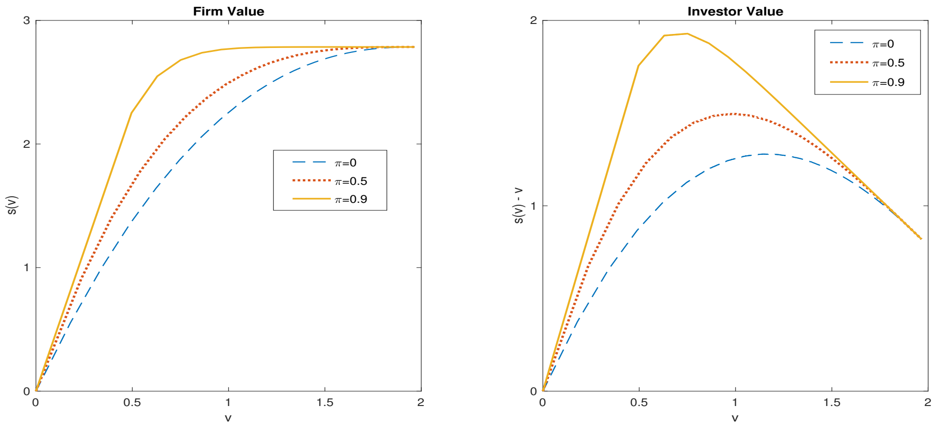


Figure 6: Firm and Investor Value

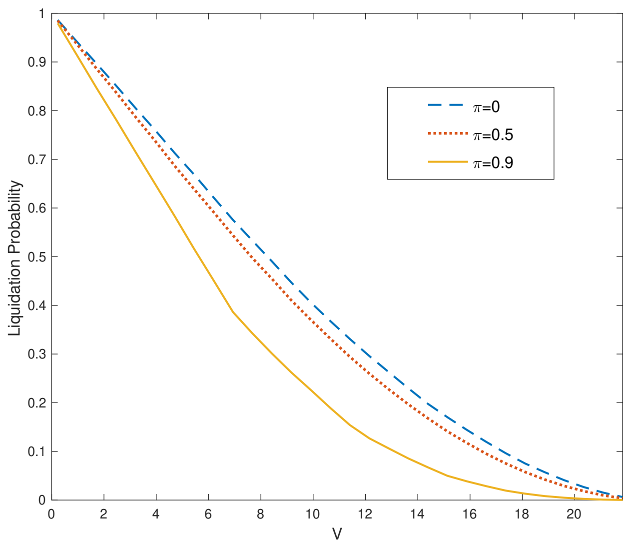


Figure 7: Simulated termination Probability

Finally, part (d) of Proposition 4 shows the downside of evidence: the n-period termination thresholds v^n all increase as evidence become more available. For a cross-section of firms starting at the same state v, the shortest time to be terminated drops as π rises. In other words, on the shortest way to termination, better evidence implies faster termination. In contrast, the quickest time to reach the cash payout boundary lengthens as π rises. These patterns are plotted in Figure 8. According to the characterization in Proposition 3, we know that the fastest way to termination on the equilibrium path occurs when a firm never discloses evidence and experiences a sequence of low cash flows.

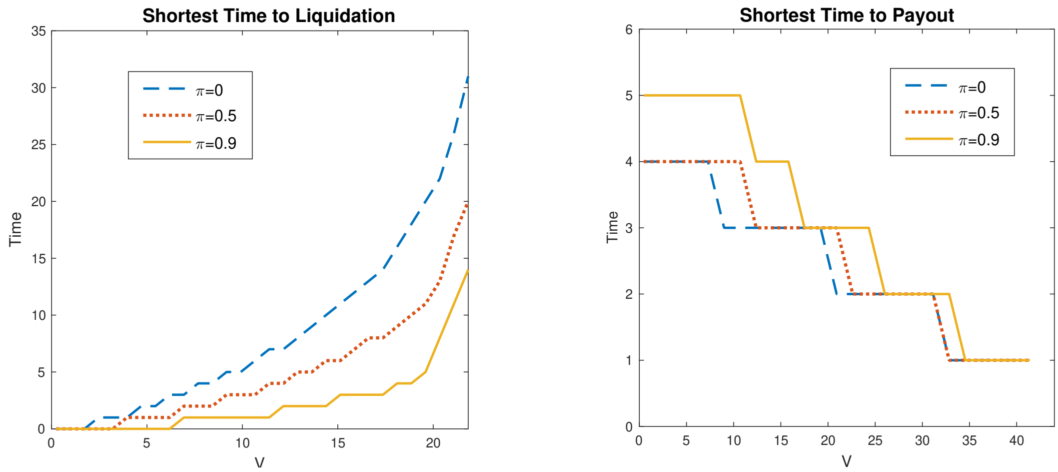


Figure 8: Simulated convergence time toward the two boundaries

7 Investment option

The value of evidence in our model comes from two channels. First, as described before, the availability of evidence increases firm value by avoiding the inefficient punishment of underperforming managers. Second, because the value of evidence is endogenous and it varies with M's continuation utility, delaying the payment of the adoption cost $c > 0$ can be valuable to the firm. Clearly, if the investment cost is too large, the option will never be exercised. Because $s(\cdot)$ is continuous in all arguments, and v belongs to a compact set, we can define the largest cost at which the option can be exercised as:

$$\bar{c} := \max_{v}\{s(v;\hat{\pi}) - s(v;0)\} \tag{9}$$

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

$$v^l = \inf_{v}\{f(v) \leq s(v;\hat{\pi}) - c\}, \qquad \text{and} \qquad v^h = \sup_{v}\{f(v) \leq s(v;\hat{\pi}) - c\} \tag{10}$$

Proposition 5. *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.*

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

When $c < \bar{c}$, the investment option can possibly be exercised in some state v. When v is close to $\bar{v}$, the probability of default is very small. Thus, the value of evidence is fairly small and $f(v)$ 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 value of exercising the option right away (i.e., $s(v;\hat{\pi}) - c$) is either tiny or negative. The following result summaries the patterns of adoptions as a function of past firm performance v at the optimal capital structure .

Proposition 6. *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 $v \in [0, v^l] \cup [v^h, \bar{v}]$.*

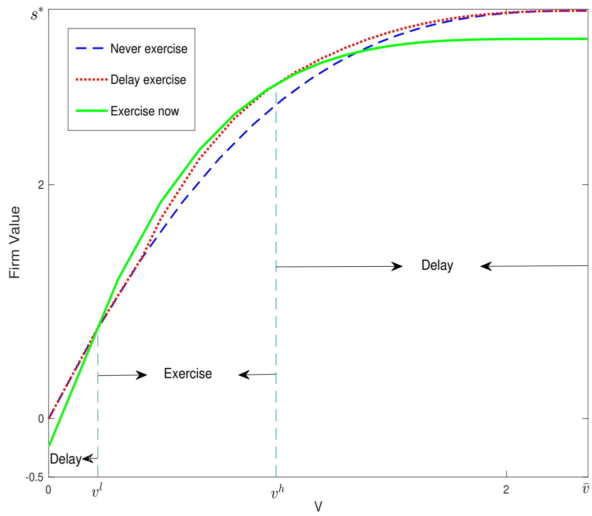


Figure 9: Option Exercise Region

Figure 9 illustrates the result of Proposition 6. It plots a numerical example of different firm values over continuation utility when $c < \bar{c}$. The green line plots the firm value if the option is never exercised which is $s(v;0)$ or the case of DeMarzo and Fishman (2007). The blue line plots the firm value if the investment option is exercised right away which is $s(v;\hat{\pi}) - c$. The red line plots the firm value $f(v)$ of delaying the 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 9 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 terminated or to reach the first best, not exercising the option.

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

8 Joint design of information and capital structures

Now that we have characterized the optimal contract, we can examine the agent's rent, the enterprise value, and the default probability at the issuance date. In particular, we analyze 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, so it was implicitly treating any history as equally likely, irrespective of the firm's characteristics. Now, instead, we jointly solve the contract and information design problems and so we explicitly account for the likelihood that any given history will arise as a function of the firm's covariates and the information technology it has access to.

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 – depends on the marginal value of increasing M's rents 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 firm 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 7. *There are two profitability theresholds $\underline{p} \leq \bar{p} \in (0, 1)$ such that:*

(i) if $p \leq \underline{p}$, then at the issuance date, the manager payoff v_0, and the firm value $\max\{f(v_0), s(v_0; \hat{\pi}) - c\}$ can both increase or decrease in the investment cost c;

(ii) if $p > \bar{p}$, then at the issurance date, the manager payoff and the firm value both 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.

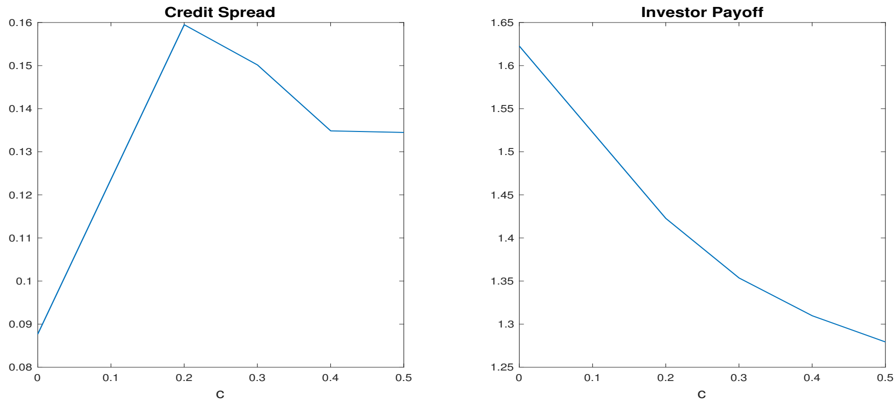


Figure 10: Credit Spreads and Investor Value

Intuitively, this occurs because to incentivize disclosure and prevent diversion P faces the trade-off of either loading on M's rents or raising the termination odds, both of which are costly. When the firm is likely to obtain low cash flows, the chance of terminating the firm is high and therefore termination is more costly. If the investment cost 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 termination concern. So the optimal policy loads on less rents (lower v_0).

To sharpen our understanding the consequences of the joint design of capital and information structure on investors' and managers' payoffs, as well as on the firm's credit spreads, we can revisit the two-period model. Recall that in such model an *exogenous* shift from $\pi = 0$ to some positive $\hat{\pi}$ could either increase or decrease the credit spreads, depending (among other things) on the firm's profitability p. Now, we ask what happens when such shift is *endogenous*, that is, when the adoption decision is optimal and the cost of adoption is explicitly considered. Of course, since termination may only occur at $t = 1$, the option would never be exercised with a delay, after $t = 0$. Thus, we only need to characterize whether or not the option is exercises right at the outset.

Proposition 8. *If $T = 2$, there exists a $\bar{c}$ such that if $c \geq \bar{c}$ the option is never exercised, while if $c < \bar{c}$ there exist two profitability thresholds $\underline{p}$ and $\bar{p}$ such that $\underline{p} < \bar{p}$ and:*

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

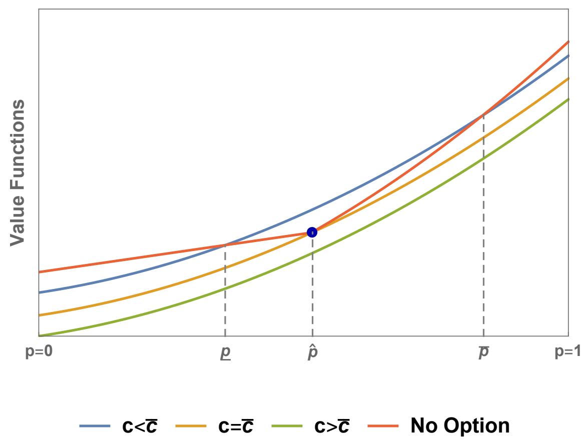
(b) If $p < \underline{p}$, the option is not exercised and the firm's probability of default is zero;

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

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

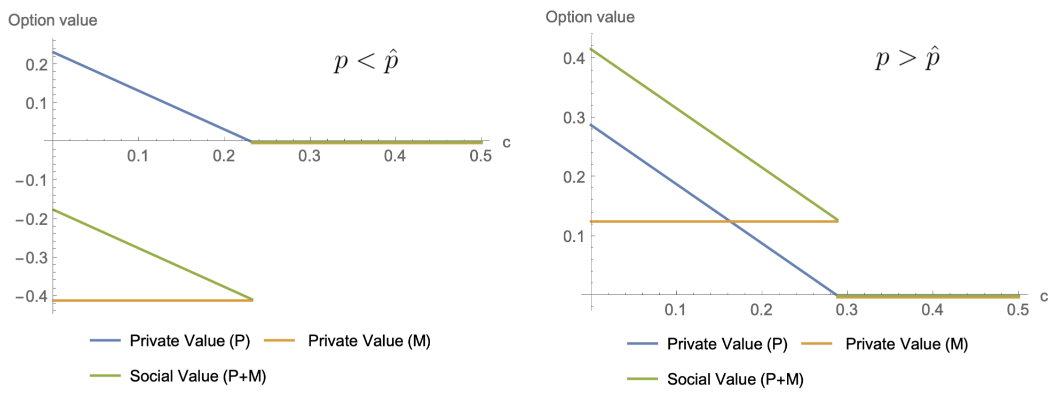
In the two-period case, evidence might enable P to distinguish bad luck from bad behavior in the first period, and it affects the optimal termination policy. Figure 11 shows that the option to invest at a strike price c and produce evidence with probability π attracts firms that are neither too profitable ($p \leq \bar{p}$), nor too unprofitable ($p \geq \underline{p}$). When the two value functions (conditional on whether the option is exercised or not) are tangent, the option is only exercised by firms with $p = \hat{p}$. If the cost drops to $c < \bar{c}$, the set of firms that exercise the option expands to $p \in [\underline{p}, \bar{p}]$. High profitability firms choose not to invest and terminate when $x_1 = l$. Low profitability firms, in contrast, never terminate and do not exercise the option either. If $c > \bar{c}$ the option is never exercised.

Figure 11: The set of firms that exercise the option in the two-period model



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 12 shows that its effect on default probabilities and credit spreads is heterogeneous across firms. For high-profitability firms that switch to exercising the option (right-panel), default probabilities decrease as disclosure avoids inefficient termination. For low-profitability firms that switch (left-panel), the opposite occurs. While at a high cost c they had a low (zero) default probability, as c drops evidence provides a tool for P to reduce the rents paid to M, while not defaulting the firms when disclosure occurs. As a result, the firm's default probability rises. Together with the cost c, this amounts to an increase in the deadweight losses and a reduction in the social surplus, even though it increases the investor's payoff.

Figure 12: Values of adoption in the two-period model



9 Capital Structure Implementation

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

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

Within this set of securities, disclosure affects the evolution of the credit line and. In particular, it determines its interest rate. In our model, any balance on the credit line account is charged an interest rate $\hat{r}_i$, for $i \in \mathcal{H}_1$, that is contingent on *both* performance and disclosure. In contrast to models such as DeMarzo and Fishman (2007), investors may sometimes charge a higher interest rate than r, or they can forgive part or all of 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 their 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 credit account balance reflects any repayment at the beginning of each period before the firm cash flow realizes. The following result summarizes a security design that implement the optimal contract.[17]

Proposition 9. *Under the following security and compensation design, the manager always discloses any evidence that might be available, and cash flows are used to either repay coupon and credit balance or to issue dividends.*

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

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

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

$$v = \lambda(z - m) \tag{11}$$

which must hold at any history. Given this relation, as well as the policy dynamics in Proposition 3, we can figure out how the firm's short-term debt evolves, which further implies the interest rates specified in Proposition 9.

To further illustrate the mechanisms, let us consider how the credit balance evolve over time. Suppose that the firm starts certain period with credit balance m. It pays

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

the long-term debt coupon l from the cash flow. The interest rate on the credit line is a constant value $\hat{r}$ unless bad cash flow news is disclosed when the interest rate becomes zero. The credit balance in the following period denoted as $m_{i\in\mathcal{H}_1}$ will be

$$m_h = (1+\hat{r})m + (1+r)(d_h - \Delta) \tag{12}$$

$$m_{nl} = (1+\hat{r})m \tag{13}$$

$$m_{dl} = m \tag{14}$$

where $d_h = \frac{u_h}{\lambda}$ is the dividend payout. If a bad news is disclosed, then interest is forgiven in the current period, and the new balance will stay the same. If the high cash flow realizes, the firm is charged a interest rate of $\hat{r}$ but will have $(1+r)\Delta$ more cash to repay the short-term debt in the next period (independently from disclosure). Therefore, the new credit balance m_h follows (12). If low cash flow realizes and no evidence disclosed, then $\hat{r}$ is charged toward the beginning balance m and the new balance m_{nl} follows (13).

As shown in Proposition 9, 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 terminated and surplus reaches the first best. In that stage, evidence is disclosure is payoff irrelevant. The necessary way to incentivize the manager is for him to hold λ fraction of dividend payments.

However, the interest rates of the credit line affect the evolution of the firm's short-term debt holding. Since evidence disclosure does affect firm liquidity in the short-run transition, the interest rates must vary with the manager's disclosure decisions. The variation in interest rates is essentially to incentivize the manager to disclose bad news. It is easy to see from Proposition 9 that the average interest rate is exactly r, but the interest gap between disclosing bad news or not is $\hat{r}_{dl} - \hat{r}_{nl} = \hat{r}$, and it increases with π. In other words, as the probability of the manager possessing evidence increases, investors on average still earn the risk free rate r, but they will design a larger interest rate variation to induce disclosure of bad news.

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 termination 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 termination odds and the minimal time required to reach the termination boundary, as well as inducing volatility spikes in continuation utilities for managers and in termination 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.

Becker, B. and V. Ivashina, "Disruption and Credit Markets," *Available at SSRN 3327798*, 2019.

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

Bertomeu, J. and E. Cheynel, "Disclosure and the cost of capital: a survey of the theoretical literature," *Abacus*, 2016, *52* (2), 221–258.

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.

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

Francis, J., D. Nanda, and P. Olsson, "Voluntary disclosure, earnings quality, and cost of capital," *Journal of accounting research*, 2008, *46* (1), 53–99.

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.

Healy, P. M. and K. G. Palepu, "Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature," *Journal of accounting and economics*, 2001, *31* (1-3), 405–440.

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

Leuz, C. and P. D. Wysocki, "Economic consequences of financial reporting and disclosure regulation: A review and suggestions for future research," *Available at SSRN 1105398*, 2008.

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.

Mas, A., "Does Disclosure affect CEO Pay Setting? Evidence from the Passage of the 1934 Securities and Exchange Act," *NBER Working Paper*, 2016.

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.

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.

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

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.

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

A Appendix

(1) The finite horizon model

It is useful to first prove Proposition 8. Proposition 1 will follow as a corollary.

Proof of Proposition 8. The $T = 1$ case taught us two facts: (i) whether P has exercised or not the option, this has no effect on the last period implementable payoffs; (ii) P could find it optimal to exercise the option at the beginning only if she terminates in the first period, after a low state is reported and no evidence is disclosed. So, there are only three policies to consider: TT (P does not exercise the option and terminates in the first period when $x = l$), NT (P does not exercise the option and never terminates) and OT (P exercises the option in the first period and terminates only when when $x = l$ and there is no disclosure). Wlog we can set all payments when the last cash flow is l to zero.

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

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

Under the policy OT, we need to determine three payments: u_{dlh} and u_{ahh} (for $a \in \{n, d\}$). However, from the disclosure IC we can see that $u_{nhh} = u_{dhh} := u_{hh}$, and so the problem reduces to solving for u_{hh} and u_{dlh}. For similar reasons as before, $u_{dlh} = \delta$. 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, $u_{hh} = \delta/p$; M's utility is $U_M(OT) = p\delta(1+(1-p)\hat{\pi})$; P's utility is $U_P(OT) = (2-(1-p)(1-\hat{\pi}))(l+p\Delta) - p\delta(1+(1-p)\hat{\pi}) - c$.

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

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

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

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

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

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

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

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

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

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

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 $\bar{\pi}$ is given by the value of π such that $U_P(OT|c=0) = U_P(NT)$. □

(2): The 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's easy to see that $\mathcal{F}$ so defined is a complete metric space. Define the Bellman operator $T : \mathcal{F} \to \mathcal{F}$ as:

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

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

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

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's easy to check that this policy satisfies all the constraints of (T). In addition, under this feasible policy we know

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

Hence, we must have $Tq(v) = s^*$. □

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

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

First, consider the program (S), and let $\{\theta, u_i, w_i\}_{i \in \mathcal{H}_1}$ be the optimal policy at $\bar{v}$. Note that, to achieve first best, all continuation values $w_{i \in \mathcal{H}_1}$ must be no less than v_1, and the liquidation probability θ is zero. In addition, from the constraints (IC_g),(IC_b), (IC_n), and (LL), we must have

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

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

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

Now condiser the program (F). Let $\{\theta, u_i, w_i\}_{i \in \mathcal{H}_1}$ be the optimal policy at $\bar{v}$. To achieve first best, the investment option must never be exercised, and moreover, the policy has to satisfy that $\theta = 0$ and $w_{nh}, w_{nl} \geq v_2$. So similarly, we can show that $v_2 \geq \frac{(1+r)p\delta}{r}$. In addition, since $f(v) \geq s(v; 0)$, Proposition A.1 implies that $f(v) = s^*$ 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:

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

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

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

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

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. The complementary slackness then implies $\alpha_b = 0$, a contradiction.

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

Third, we show that the Lagrangian multiplier $\alpha_n > 0$. Suppose not. Then from the above result, it has to be that $\alpha_g > 0 = \alpha_n$. Then the first oder conditions (FOC_{nh}) and (FOC_{nl}) together imply that $s'(w_{nh}) > s'(w_{nl})$. Hence, $w_{nh} < w_{nl}$ 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 $s'(w_{dh}) > s'(w_{nh})$ which further implies that $w_{dh} < w_{nh}$. This forms a contradiction with (IC_g).

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

Proof of Proposition 3. The proof is divided into two parts. Part (a) shows that the 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 auhxiliar 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 can 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. So the policy must satisfy the first order conditions of (S). Summing (FOC_{dh}) (FOC_{dh}) (FOC_{nl}), and using the envelope condition (EN), we can obtain

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

Then 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 $g'(\tilde{v}) > s'(\tilde{v})$, a contradiction. Hence, $g(v) = s(v)$ 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 $i \in \mathcal{H}_1$. Appenrently, this policy satisfies the constraints of (S). Since $s(v) = s^*$ when $v \geq \bar{v}$, we know 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 constriants of (S) implies that $w_{dl}(\underline{v} + \varepsilon) < \underline{v}$ for sufficiently small $\varepsilon > 0$. Then we have $s'[w_{dl}(\underline{v} + \varepsilon)] > s'(\underline{v} + \varepsilon)$, which is contradicted with (FOC_{dl}) at $\underline{v} + \varepsilon$. Hence, we must have $\underline{v} = v^1$. Then (PK) 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

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

Hence, $v^j = \frac{1}{1+\hat{r}}[v^{j-1} + \hat{r}\bar{v}]$. 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 $v^1 = \frac{\hat{r}}{1+\hat{r}}\bar{v}$. Finally, we can obtain the n-period liquidation threshold expressions simply by induction. □

Proof of Proposition 4. Consider $\pi = 0, \hat{\pi}$ as a parameter of the firm's problem (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 (S), the envelope condition with respect to π is obtained as

$$\frac{(1+r)s_\pi(v)}{1-p} = s(w_{dl}) - s(w_{nl}) - s'(w_{dl})(w_{dl} - w_{nl}) \tag{A.20}$$

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

Part(b). Take any $v \in [v^1, \bar{v}]$. According to the definition in (6), the pay-performance sensitivity 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, PPS is decreasing in π and increasing in v.

Now consider any $v < v^1$. Since our PPS measure is only defined when the firm is not liquidated at the beginning of the period, we can 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, PPS decreases in π.

Parts (c) and (d). These are easy to obtain from the result of Proposition 3. □

Proof of Proposition 5. For $v \in (0, \bar{v})$, since $\hat{\pi} > 0$ and Proposition 4 shows that $s(v)$ strictly increases in π, we must have $s(v; \hat{\pi}) > s(v; 0)$. Then the continuity of $s(\cdot)$ implies $\bar{c}$ in (9) is well defined, and $\bar{c} > 0$. Moreover, the threshold cost $\bar{c}$ increases in $\hat{\pi}$ because $s(v; \hat{\pi})$ does.

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

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

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

Now consider any $0 < c < \bar{c}$. First, not that in the region close to the boundary $\bar{v}$, not exercising the option is optimal. This is because $f(\bar{v}) = s^* > s(\bar{v}; \hat{\pi}) - c$. Then v^h as in (10) is well defined, and the option is not exercised for $v \in [v^h, \bar{v}]$.

Second, not exercising the option is optimal in the region close to the boundary 0. This is becasue $f(0) = 0 > s(0; \hat{\pi}) - c$. Hence, v^l as in (10) is well defined, and the option is not exercised for $v \in [0, v^l]$.

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

In the case of $c > \bar{c}$, we have $s(v; \hat{\pi}) - c < s(v; 0)$ for any v, by the definition of $\bar{c}$ in (9). Then $s(v; \hat{\pi}) - c < f(v)$ for any v, implying the option is never exercised. □

Lemma A.2. *When the investment option is not exercised as in (F), the relevant optimal continuation utility w_{nh} and w_{nl} are the ones 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 be the same as if we solve (F) without (IC_n). Then since $w_{nh} = w_{nl} = (1+r)v$ and $\theta = 0$ is feasible, we know from the objective of (F) that $f(v) \geq s^*$, a contradiction with $v < \bar{v}$.

Then the optimal $w_{nh}(v)$ 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 $\pi = 0$. □

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*

$$a_0 = 0,\ b_i = \frac{\mu(r+p)(1+\hat{r})}{rp\delta}[\hat{r}(1-p)(1-\pi)/r]^i \tag{A.21}$$

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 the firm is not liquidated in the beginning of the period. From this observation, we can derive $s(\cdot)$ by induction.

When $v \in (0, v^1]$, the objective of (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 $s(v^1) = \frac{(1+r)(r+p)\mu}{r[1+r-(1-p)\pi]}$. Hence, $b_0 = \frac{s(v^1)}{v^1} = \frac{\mu(r+p)(1+\hat{r})}{rp\delta}$.

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

Note that the slope b_i in (A.21) is a function of p and π. In the following, we denote this slope as $b_i(\pi; p)$, and denote the derivative of this slope with respect to π as $b_i'(\pi; p)$.

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

Proof. From the expression of $b_i(\pi; p)$ in (A.21), 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 $p > \frac{r}{r+i(1+r)}$. □

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

$$g_1(p) =: \frac{2(1-p)^2(1-p)}{\lambda}\left(1 + \frac{l}{p\Delta}\right) \tag{A.22}$$

$$g_2(p) =: (6\sqrt{p}p + 4p + 3)p \tag{A.23}$$

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.22), we know that $g_1(\cdot)$ is strictly decreasing and that $g_1(\underline{p}) = \frac{\mu(1-\underline{p})^2}{\underline{p}\delta} h(\underline{p})$. Hence, statement (a) holds. To see 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, since $p \leq 1/9$, we have $\frac{3}{3\sqrt{p}+2} \geq 1$, which implies the second inequality in statement (ii) holds. □

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

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

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

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

$$\frac{\mu(1-p)^j}{p\delta} h(p) \geq 1 > \frac{\mu(1-p)^{j+1}}{p\delta} h(p) \tag{A.24}$$

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

$$\frac{\mu(1-p)^j}{p\delta} h(\sqrt{p}) \leq \frac{\mu(1-p)^{j+1}}{p\delta} h(p) < 1 \tag{A.25}$$

Then (A.24) (A.25) together imply there exists $\tilde{r} \in [p, \sqrt{p})$ such that $\frac{\mu(1-p)^j}{p\delta} h(\tilde{r}) = 1$.

Second, from the above derivation, we know that $\frac{\mu(1-p)^j}{p\delta} h(\tilde{r}) > \frac{\mu(1-p)^{j+1}}{p\delta} h(p)$. Hence, we have $h(\tilde{r}) > (1-p)h(p)$, which further implies that $\frac{p}{\tilde{r}} > \frac{2(1-p^2)}{1+\tilde{r}} - 1$. Moerover, because $\tilde{r} < \sqrt{p}$, Lamma A.5 (b) implies that $2(1-p^2) > \frac{3(1+\tilde{r})^2}{3\tilde{r}+2}$, which is equivalent to $\frac{2(1-p^2)}{1+\tilde{r}} - 1 > \frac{1}{3\tilde{r}+2}$. Hence, we know that $\frac{p}{\tilde{r}} > \frac{1}{3\tilde{r}+2}$, or $p > \frac{\tilde{r}}{\tilde{r}+2(1+\tilde{r})}$. Then we can obtain part (b) by $j \geq 2$. □

Proof of Proposition 7. Note that if $c = 0$, the firm exercises the option at initiationm, and optimally chooses

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

On the contrary, if $c = \bar{c}$ the firm never exercises its investment option, and at initiation, it optimally chooses

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

Take any $p \leq \underline{p}$. Consider the values j and $\tilde{r}$ in Lemma A.6, and let $r = \tilde{r}$. Then we have $s'(\tilde{v}_0; 0) = b_j(0; p) = 1$, and $\tilde{v}_0 = v^{j+1}(0)$. Moreover, given the result of 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. This further implies that $\hat{v}_0 \leq v^j(\hat{\pi}) < v^{j+1}(0) = \tilde{v}_0$. And because $s(v; \pi)$ is continuous in both v and π. A small change in π and a downward jump of initial continuatin utility from $\tilde{v}_0$ to $\hat{v}_0$ implies $s(\hat{v}_0; \hat{\pi}) < s(\tilde{v}_0; 0)$. So the firm value and the manager payoff at initiation must both have a increasing part as the investment cost increases from 0 to $\bar{c}$.

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

Proof of Proposition 9. We will show that the arrangement implements the payout and the evolution of the optimal contract.

First, consider the case of $m = 0$, or the cash payout region. The agent's total payoff is λ fraction of the firm payout which is $\lambda d_h = u_h$. And according to (12) (13) (14), the subsequent credit line balance will all be zero, or $m_h = m_{nl} = m_{dl} = 0$, which implis that the continuation utilities in th the contract are $w_h = w_{nl} = w_{dl} = \bar{v}$ by (11). So in the implementation, the agent gets the same cash as in the contract, and the firm always stays at first best, having no probability of defaulting.

Second, consider the case of $m > 0$. Given this credit balance, we can use (11) to (14) to derive the credit line balance of the subsequent period as:

$$m_{nl} = \frac{[1 + \hat{r}(\pi)](\bar{v} - v)}{\lambda}, \quad m_{dl} = \frac{\bar{v} - v}{\lambda}, \quad m_h = m_{nl} - (1 + r)\Delta \tag{A.26}$$

So by (11), the agent's ttotal payoff (by withdrawing all available credit) becomes

$$w_{nl} = v - \hat{r}(\pi)(\bar{v} - v), \quad w_{dl} = v, \quad w_h + (1+r)u_h = w_{nl} + (1+r)\delta \tag{A.27}$$

which is the same as in Proposition 3. Note this derivation includes the two senarios of whether the investment option is exercised or not. In the latter case, we only have the relevant credit balance to be m_{nl} or m_h, and correspondingly, the agent's payoff to be w_{nl} or $w_h + (1+r)u_h$. These values are obtained by setting $\pi = 0$ in (A.26) and (A.27). Hence, by Lemma A.2 we know these payoffs are the same as in the optimal contract with the investment option. □