Uncertainty Shocks and Dynamic Compensation[*]

Felix Zhiyu Feng[†]

August 2014

Abstract

This paper studies optimal dynamic compensation when firms are subject to uncertainty shocks but have only limited ability to commit to long-term contracts. I analyze a continuous-time dynamic principal-agent model with private effort and regime switching in cash flow volatility and characterize the optimal managerial compensation and termination policy. In high volatility times, firms are forced to expedite payments to managers because sufficient deferred compensation is no longer credible. At the same time, contract length shortens, that is, termination becomes more likely. This relation between the timing of payments and expected contract length may explain the sizeable cash bonuses observed in crises times. In contrast, with full commitment firms defer compensation more when volatility is high.

JEL Classification: D86, M12, G34

Key Words: Uncertainty shock, Dynamic compensation, Limited commitment

"It is apparently going to take more than shrinking bank profits to put a big dent in Wall Street bonuses."

<div align="right">——The New York Times, February 29th, 2012</div>

1   Introduction

Uncertainty is at the heart of finance. Financial firms face more uncertainty relative to other types of firms and are more sensitive to uncertainty shocks, which have been identified by a growing body of research as the key contributing factor to the recent financial crisis[1]. It is unsurprising that financial firms rely heavily on incentive pay, which constitutes a majority of the overall cost of a firm as well as total employee compensation. While incentive pay is directly motivated by the presence of uncertainty, most of the extant research on managerial compensation has so far focused on its relationship with profitability instead of uncertainty and especially how compensation and uncertainty *dynamically* evolve with each other[2].

The current paper fills this gap by studying how uncertainty dynamically affects compensation through a continuous time principal-agent model. Due to a moral hazard problem, the optimal contract incentivizes the manager by promising delayed cash bonuses after the manager's performance exceeds a certain benchmark. When uncertainty increases, the manager's performance becomes a noisier signal and it is more difficult to gauge his true effort. As a result I show that the optimal contract raises the performance hurdle for awarding cash bonuses so managers are *less* likely to receive cash bonuses.

However, such result requires firms to fully commit to long-term incentive contracts, which is not always consistent with the nature of such contracts in practice. More specifically,

---

[1]See for instance Bloom (2009), Brunnermeier and Sannikov (2012), Atkeson et al. (2013), Di Tella (2013), He and Krishnamurthy (2013)

[2]There are several studies exploring the volatility and the sensitivity of pay-for-performance, such Lambert and Larcker (1987), Aggarwal and Samwick (1999), and Core and Guay (2002). See Prendergast (2002) for a survey on this topic. These studies usually do not focus on the correlation between volatility and the level of pay itself. Regarding the payment structure of financial firms, Axelson and Bond (2012), Bijlsma et al. (2012), Glode et al. (2012), Bond and Glode (2013), Bénabou and Tirole (2013), Glode and Lowery (2013), Thanassoulis (2012) examine the escalating use of incentive compensation for financial firms but their focus is not on the role of uncertainty. Kaplan and Rauh (2010), Philippon and Reshef (2012) show that high level incentive payments extend beyond a handful of top managerial elites and those payments remained substantial during the recent crisis

to maintain correct incentives, firms must commit in two ways: commit to making payments when due and commit to retaining managers until managers' performance is sufficiently poor. The latter type of commitment is generally infeasible in practice given the prevalence of at-will employment. Under US labor laws, firms can fire employees without having to establish just cause or give warning. Firms can also liquidate anytime, after which they are no longer liable for any future compensation promised to employees.

In light of these observations I re-examine the implication of uncertainty on compensation when firms have ability to unilaterally terminate contracts, which I refer to as principal's limited commitment. Limited commitment restrains firms' ability to use deferred compensation as incentives. Moreover, the concern over firms' commitment is the greatest when uncertainty is high and firm value is low. Therefore, I find that firms with limited commitment are forced to compensate their managers more immediately and managers may be *more* likely to receive cash bonuses during times of rising uncertainty.

The conclusion of this paper sheds light on some of the empirical findings regarding managerial compensation that are not driven simply by firms' profitability. In a cross-sectional study, Peters and Wagner (2014) show that higher industry-level equity volatility is strongly associated with higher managerial compensation, a result which the authors attribute to turnover risks. Similarly, Kaplan and Minton (2012) conduct time series analyses and find a downward trend in CEO tenure in the past decades and suggest that the increase in CEO compensation during such period may be explained the increasing dismissal risk. In addition to rationalizing the aforementioned empirical findings, this paper also contributes to the heated debated over managerial compensation during the recent crisis when, despite huge losses of company wealth, many financial firms still paid out sizable compensation their bankers and executives[3]. Different from the widely cited managerial entrenchment view, this paper argues that large bonuses *can* be the results of an *optimal* contract if firms do not have full commitment over managerial contracts. Finally, this paper reconciles optimal contract theory with observations from the "pay-for-luck" literature, such as Bertrand

---

[3]Merrill Lynch, for example, paid out a total of $3.6 billion in bonuses in the 2008 fiscal year despite having suffered losses of $27 billion; and Citigroup paid out $5.3 billion in bonuses after a $27.7 billion loss. See *Wall Street Journal*: Wall Street Compensation–'No Clear Rhyme or Reason'. July 30th, 2009. More detailed statistics can be found in the press release of the New York State Comptroller (2009) as well as in Frydman and Jenter (2010) and Kaplan (2012)

and Mullainathan (2001) and Axelson and Baliga (2009), who find no significant decline in managerial compensation during market downturns.

Economic literature has long recognized that firms do not possess full commitment to labor contracts[4]. Recent research shows that firms can also default on labor contracts because of limited or costly access to financial markets[5]. Besides intertemporal compensation most of these studies also assume firms cannot commit to terminal payments either such as the severance pay[6]. In practice, those payments at the time of contract termination, either due to managerial turnover or bankruptcy, are generally difficult to be made fully contingent and are subject to changes in the external enforcement environment[7]. Furthermore, I demonstrate that the key results of this paper are robust to introducing real costs associated with default or allowing contract renegotiation. In short, the assumption of limited commitment (that the principal cannot commit to retaining managers when firm value drops below a certain level) in this paper is made in a simple, extreme case to ensure tractability but is without loss of generality able to capture the important concern of imperfect contract in practice.

Different from comparative statics analysis, I model uncertainty shocks through stochastic regime switching between low and high volatility states. Under regime switching, firms optimally allocate managerial deferred compensation until the marginal value before the uncertainty shock is equal to the marginal value after the shock. These important dynamics are absent from simple comparative statics, which implicitly hold managerial deferred compensation constant when comparing different volatility levels. Moreover, the dynamics of regime switching point out that cash payments should not be confused with "reward." In fact, under limited commitment, the lifetime present value managers derive from a contract that results in more immediate payments during financial crises is lower. This is because

---

[4]There are many studies on risk sharing in labor contracts such as Thomas and Worrall (1988) Abreu et al. (1990), Ray (2002), Berk et al. (2010), Grochulski and Zhang (2011), Miao and Zhang (2013) which emphasize the lack of commitment from the firm side. More generally the relational contract literature such as Atkeson (1991), Levin (2003), Grochulski and Zhang (2013), and Opp and Zhu (2013) who study the lack of commitment from both contracting parties

[5]For instance, Ellul et al. (2014), Palacios and Stomper (2014).

[6]For instance, Berk et al. (2010) and Ai and Li (2013). Even if some severance payments can be included in managerial contracts, as long as they are not fully guaranteed there is still the concern of limited commitment and the argument of this paper still applies

[7]In recent cases such as the bankruptcy of Hostess Brands Inc. and Hawker Beechcraft Corp., the US Justice Department blocked proposals to grant extra bonuses to the executives of those companies.

managers are simultaneously subject to a higher probability of termination. In contrast, firms with full commitment defer compensation more when volatility is high, and managers may be consequently rewarded with a higher present value of total compensation.

This paper makes several contributions to the literatures on contract theory, managerial compensation, and corporate governance. On the modeling side, to my knowledge this is the first paper that jointly considers agency, limited commitment, and regime switching. It fits in the growing literature of continuous-time dynamic agency models such as DeMarzo and Sannikov (2006), Sannikov (2008), He (2009), Biais et al. (2010) and is the closest to De-Marzo and Sannikov (2006) in modeling the cash flow process and the moral hazard problem. The regime switching technique for continuous time models is adopted from Hoffmann and Pfeil (2010), Piskorski and Tchistyi (2010), Bolton et al. (2013), and DeMarzo et al. (2012). Different from these models which focus mainly on stochastic regimes of profitability, this paper considers regimes of cash flow volatility[8]. Although the key results are generally replicable with stochastic profitability, volatility is more suitable for analyzing the compensation of financial firms which predominantly consists of contingent incentive payments rather than fixed salary payments.

The theoretical results of this paper generate testable empirical hypotheses: conditional on negative uncertainty shocks, commitment-constrained firms make larger immediate payments and have higher managerial turnover relative to unconstrained firms. While the empirical literature on corporate governance generally takes low total compensation and high pay-for-performance sensitivity as indicative of good governance, this paper shows the importance of considering the level and structure of compensation under the context of market uncertainty. Total compensation and pay-for-performance sensitivity are sensible proxies for firm governance only when firms have no commitment issue, which may not always be true, especially during spells of high volatility.

In terms of policy implication, this paper puts forth a caveat to the popular perception that the high compensation observed in the recent financial crisis is a sign that managers are entrenched and that the current compensation structure is largely suboptimal (or even

---

[8]There is another strand of research on endogenous volatility choice in the theory of asset management, such as Ou-yang (2003), Lioui and Poncet (2014), Cvitanić et al. (2014), but the focus of this paper is exogenous volatility shocks

corrupt). This public notion has spurred political activism to regulate and reform compensation practices[9]. However, without taking into account firms' commitment ability, policies intended to align managers' incentives with those of investors could actually backfire. For instance, TARP recommends deferring compensation to executives and the Dodd-Frank Act gives shareholders the right to disapprove any golden-parachute compensation to executives. However, executives' valuation of deferred compensation depends crucially on their assessment of firms' ability to commit to making future payments. With escalating uncertainty, it is difficult for firms to maintain managers' confidence, so managers must be paid with more immediacy. Thus, during times of financial crisis, following the recommendation to defer compensation and restrict retention payments would actually do more harm than good by undermining managerial incentives.

Finally, this model also reveals the possibility that different levels of agent effort could be implemented by the optimal contract under different levels of cash flow uncertainty. As uncertainty increases, the value of an incentive compatible contract decreases. If the uncertainty shock is sufficiently severe, the optimal contract could shift from being incentive compatible to one that allows shirking. The agent stops receiving cash payments and is instead compensated through his private benefit from shirking, and due to lower managerial effort, expected cash flows fall. Thus, if a crisis becomes sufficiently severe, managers will stop receiving bonuses. This, however, should be of little comfort because it implies managerial indifference and corresponding low productivity. This finding also suggests the potential edogeneity between profitability and volatility, which challenges empirical studies on executive compensation that treat the two as independent.

The remainder of the paper is organized as follows: Section 2 describes the main model with fixed uncertainty level and solves for the optimal contract under limited commitment. Section 3 introduces uncertainty shocks and derives the model's implications for dynamic compensation. Section 4 discusses two major extensions: security implementation and shirking in equilibrium. Section 5 concludes.

---

[9]See, for instance, the Troubled Asset Relief Program (TARP): Executive Compensation Rules & Guidance and the Dodd-Frank Wall Street Reform and Consumer Protection Act

2 Model

In this section I describe the model in a standard principal-agent environment with only one regime of uncertainty. I first solve the optimal contract assuming that the principal has full commitment power. After discussing the limitation of this strong assumption, I then solve the optimal contract but suppose limited commitment by the principal.

2.1 Basic Environment

Time is continuous. A principal, representing a firm, must hire an agent, representing a manager, to run a project. Both the principal and the agent are risk neutral. The cash flow Y_t of the project follows

$$dY_t = \mu(e_t)dt + \sigma dZ_t \ ,$$

where Z_t is a standard Brownian motion; $\mu(e_t)$ is the expected growth rate of cash flow depending on the agent's effort. σ is the cash flow volatility, which measures the level of uncertainty a firm faces when taking on such a project.

The agent controls the cash flow growth rate by choosing a binary effort level $e_t \in \{\bar{e}, \underline{e}\}$, representing "working" and "shirking", respectively. I assume $\mu(\bar{e}) = \mu$ and $\mu(\underline{e}) = \mu - C$ where $C > 0$, that is, shirking results in lower expected cash flow. However, the agent enjoys a private benefit λC whenever he shirks, where $\lambda \in (0, 1]$ measures the degree of agency problem in this model. Effort is private to the agent: the principal can observe Y_t but not e_t. The principal discounts future cash flows by r and the agent by $\gamma > r$, so the agent is more impatient.[10]

For now, I assume that the principal can commit to any contract once it is signed, but the agent is protected by limited liability with an outside option whose value R is normalized to 0. Limited liability implies that payments to the agent must be non-negative and that

---

[10]The asymmetry of discount rates is essential for a non-trivial incentive compatible contract to exist in this type of model. Because of the agent's private effort, the principal must delay cash payments to the agent until his cumulative performance exceeds a certain threshold. Having the agent discount future cash income heavier ensures that providing incentive through such delaying is costly for the principal so the principal will not want to delay payments forever. However, once I impose the principal's commitment constraint, this additional constraint leads to the existence of an optimal contract even for the case where $r=\gamma$, which I describe in Section 4

contract termination is necessary for incentive provision. The principal has an outside option L which she receives whenever the contract is terminated. Both the principal and the agent take no further action after the contract termination, which eliminates reputation concerns. One can interpret contract termination here as a firm's liquidation and exit from the market permanently or, equivalently, as a firm's replacement of managers, where L is the normalized net profit from contracting with a new manager.

I also assume for most part of this paper that the principal prefers to induce the agent to work. This is true as long as C, the cost of shirking, is high enough. In Section 5 I discuss how sufficiently large uncertainty shocks can allow shirking in equilibrium because incentive provision is too costly when uncertainty is high.

Let $\mathscr{F}_t$ be the filtration generated by the cash flow history. A contract which specifies a compensation process $\{I_t\}_{t\geq 0}$ from the principal to the agent, a termination time τ, and a recommended effort process e_t defines the agent's continuation utility W_t:

$$W_t = E\left[\int_t^\tau e^{-\gamma(s-t)}\left(dI_s + \lambda C \mathbb{1}_{\{e_t=\underline{e}\}}dt\right)|\mathscr{F}_t\right] .$$

where $\mathbb{1}_{\{e_t=\underline{e}\}}$ is an indicator function that takes value 1 if $e_t = \underline{e}$ and zero other wise. W_t simply measures the present value of all expected future payments and can be intuitively interpreted as the agent's "wealth". Similarly, the contract defines the principal's valuation of the project V_t which is the expectation of total future cash flow minus the payment to the agent plus the liquidation value when the contract is terminated.

$$V_t = E\left[\int_t^\tau e^{-r(s-t)}\left(dY_t - dI_s\right) + e^{-r(\tau-t)}L|\mathscr{F}_t\right] ,$$

So far the basic environment is identical to the one in DeMarzo and Sannikov (2006). I therefore briefly review their solution for the optimal contract to offer a benchmark for later comparison. Using the martingale method developed by Sannikov (2008), there exists $\mathscr{F}_t$ measurable processes β_t such that W_t evolves according to

$$dW_t + dI_t = \gamma W_t - \lambda C \mathbb{1}_{\{e_t=\underline{e}\}}dt + \beta_t(dY_t - (\mu - C\mathbb{1}_{\{e_t=\underline{e}\}})dt) . \tag{1}$$

Equation (1) means the principal can compensate the agent with either immediate payments dI or future payments dW. Payments are sensitive to the agent's performance, i.e. the realized cash flows dY_t, and the sensitivity is measured by β_t. DeMarzo and Sannikov (2006) shows that the contract is incentive compatible if and only if $\beta_t \geq \lambda$ for all $t < \tau$. Intuitively, β_t represents the agent's "skin in the game" in the incentive compatible contract or the sensitivity of the agent's continuation utility to the realized cash flow. The sensitivity must be no less than λ which is the amount of private benefit the agent can derive per unit of cash flow he looses from shirking.

Under the optimal incentive compatible contract the principal's valuation of the project V_t is a function of the agent's continuation utility W_t. The function is denoted as $V(W_t)$ and is referred to as firm value, since the principal represents a firm in this contracting relationship. The principal must earn an instantaneous return r under the optimal contract through the expected cash flow and the expected change rate of her valuation of the project, that is

$$V(W_t) = E\left[dY_t + dV(W_t)\right] . \tag{2}$$

Applying Ito's lemma to (2) and imposing $e_t = \overline{e}$ on equation (1), $V(W_t)$ is characterized by the following Hamiltonian-Jacobian-Bellman (HJB) equation:

$$rV(W_t) = \max_{\beta_t \geq \lambda} \mu + \gamma W_t V'(W_t) + \frac{1}{2}\beta_t^2 \sigma^2 V''(W_t) . \tag{3}$$

The principal maximizes her HJB equation by choosing β_t optimally. The concavity of the value function implies $\beta_t = \lambda$ for all t as long as the principal induces working as the equilibrium effort. On the one hand, $\beta_t \geq \lambda$ is necessary for incentive reasons. On the other hand, it is not optimal to set β above λ since higher sensitivity increases the likelihood of contract termination which is also a costly action for the principal.

The principal's value function must also satisfy three additional conditions which define two boundaries for W_t: the termination boundary and the payment boundary. First, the contract is terminated when W_t hits R. The principal cannot provide further incentives by lowering the agent's continuation utility any more since the agent is protected by limited liability. The principal receives the liquidation value L at the time of contract termination,

so $V_s(R) = L$. Secondly, since the principal can always make a lump sum transfer of dI to the agent, the principal's valuation of the project satisfies $V(W) \geq V(W - dI) - dI$. Equivalently this means $V'(W) \geq -1$, that the shadow value of the agent's continuation utility to the principal should not be lower than the cost of an instant cash payment. Defining $\overline{W} = \inf \{W|V'(W) = -1\}$ as the *payment boundary*. $\overline{W}$ satisfies the "smooth pasting" condition $V'(\overline{W}) = -1$. The agent will receive instant cash payment of size $W_t - \overline{W}$ once $W_t > \overline{W}$, which brings W_t back to $\overline{W}$ immediately. Finally, the payment boundary $\overline{W}$ is optimally chosen by the principal, which implies the "super contact" condition: $V''_s(\overline{W}) = 0$.

The following lemma, which is also Proposition 1 in DeMarzo and Sannikov (2006), summarizes the optimal incentive compatible contract in this setting:

Lemma 1. *The optimal incentive compatible contract, which maximizes the principal's payoff subject to delivering value W to the agent and satisfying the agent's limited liability, can without loss of generality be characterized by a value function $V(W)$ and a payment boundary $\overline{W}$ such that*

$$rV(W) = \mu + \gamma W V'(W) + \frac{1}{2}\lambda^2\sigma^2 V''(W) ,$$

with boundary conditions $V(R) = L$; $V'(\overline{W}) = -1$; $V''(\overline{W}) = 0$. Immediate cash payment of size $W_t - \overline{W}$ is made for all $W_t > \overline{W}$. When $W_t = R$, the contract is terminated.

Under the optimal contract, the agent receives compensation $dI = \max(W_t - \overline{W}, 0)$, which is always non-negative because of his limited liability and strictly positive whenever his continuation utility W_t exceeds the payment boundary $\overline{W}$. This compensation scheme resembles cash bonuses that managers in practice receive for good performance, where $\overline{W}$ represents the "bonus hurdle" managers must clear before getting paid. Hence, this model is suitable for studying the compensation of financial firms, in which cash bonuses make up the bulk of total employee compensation. Note that Lemma 1 concludes that any optimal incentive compatible contract can without loss of generality be characterized by a contract involving incentive pay that resembles cash bonuses, a natural means of incentive provision in this setting. Thus, in the remainder of this paper, I refer to the cash payment dI as bonuses. Keep in mind that it is representative of incentive pay in general, which in the current model can be contracted on and hence to which the principal can commit.

The optimal contract is associated with welfare losses due to moral hazard. The efficient allocation calls for the principal and the agent to split the maximal surplus generated by running the project permanently; that is, the agent's compensation and the principal's payoff satisfy $V(W) + W = \mu/r$, which marks the Pareto frontier of this model in the absence of moral hazard. However, when moral hazard is present, the principal must design incentive compatible contracts featuring delayed payments to the agent. Substituting boundary conditions $V'(\overline{W}) = -1$ and $V''(\overline{W}) = 0$ into the principal's HJB equation yields $rV(\overline{W}) + \gamma\overline{W} = \mu$, which is the "second best" frontier below the Pareto efficient frontier. This is because the agent is more impatient and some surplus is lost in the optimal incentive compatible contract with delayed payment. The "second best" frontier becomes a critical boundary when considering the optimal contract with principal's limited commitment, which is introduced in the following subsection.

2.2 The Commitment Constraint

So far, the structure of the optimal contract characterized in Lemma 1 relies on the principal's commitment to all future payments once the contract is signed. However, before the agent's continuation utility W hits the payment boundary $\overline{W}$, the agent is not actually paid. His continuation utility measures the present value of the total amount of payment he expects to receive in the future, only if the principal honors the contract. Just as the agent is tempted to quit his job when W approaches his reservation utility R, the principal will likewise be tempted to exercise her outside option, which in this model is liquidating the project and receiving L, if the firm value V drops below the liquidation value before W reaches the payment boundary. If enforcement is not perfect and commitment becomes a binding constraint before the cash payment boundary is reached, the dynamics of the optimal contract will consequently be different.

To consider this impact, I assume that the principal can terminate the agent's contract anytime. As discussed earlier, this assumption of limited commitment on the part of the principal is more realistic, as firms generally are free to fire managers or liquidate projects at any time in practice. Once the contract is terminated, I assume both parties will receive the value of their outside options: L for the principal and R for the agent. This assumption

sets this model apart from the other models in the relational contract literature in that termination time is the only aspect of the contract to which the principal cannot commit. Conditional on the continuation of the contract, the principal can still commit to all payments once the payment boundary is reached, suggesting the existence of long-term contracts although subject to a participation constraint from the principal[11]. More discussions of the validity and generality of this specific assumption are presented after I derive the optimal contract under the current constraint.

I introduce here a heuristic approach that derives the optimal contract under principal's limited commitment by separating the commitment constraint from moral hazard, the other contractual friction in the model. First, suppose the agent's effort is observable to the principal and then the only contractual constraints are the principal's limited commitment and the agent's limited liability. Limited commitment implies a participation constraint for the principal:

$$V_t \geq L . \tag{4}$$

Combined with the agent's participation constraint $W_t \geq R$, they define a payoff space $\{(W,V)|W \geq R, V \geq L\}$ where, if the continuation value delivered by a contract falls into the space, the contract will not be terminated, i.e. the contract is *self-enforcing*.

Given the self-enforcing contracting space, consider now adding moral hazard. The martingale representation theorem implies the dynamics of W still follows equation (1), and the optimal contract still features a reflecting payment boundary $\overline{W}$. Combined with the principal and agent's participation constraints, these conditions define a space between $W \geq R$, $V \geq L$ and the "second best" frontier $rV + \gamma W = \mu$ in which the optimal contract value function must lie. The principal's valuation of the project V is a function $V(W)$ of the agent's continuation utility, where $V(W)$ satisfies the same HJB in equation (3).

The exact shape of the value function is pinned down by appropriate boundary conditions similar to those characterized in Lemma 1. Intuitively, at the termination boundary, firm value must match the liquidation value L. At the payment boundary, the marginal value of cash payment always equals the shadow value of the agent's continuation utility. This implies

---

[11]At this point the principal can also commit not to renegotiate the contract, even though renegotiation can be mutually beneficial. I discuss the renegotiation-proof contract under this setting in the Appendix

the "smooth-pasting" condition $V'(\overline{W}) = -1$ always holds. The last boundary condition depends on whether the value function crosses the "second best" frontier $rV(W) + \gamma W = \mu$ or the self-enforcing border $V(W) = L$ first. If the value function meets $rV(W) + \gamma W = \mu$ first, it immediately follows $V''(\overline{W}) = 0$ and the principal's commitment constraint is not binding. In contrast, if the value function reaches $V(W) = L$ first, then $V(\overline{W}) = L$ is the boundary condition that replaces $V''(\overline{W}) = 0$.

These boundary conditions are intuitive. By concavity of the value functions, if the payment boundary is such that both the limited commitment constraint and the "super contact" condition are slack, in other words $V(\overline{W}) > L$ and $V''(\overline{W}) < 0$, the principal can always achieve a higher value by postponing the payment further, until either condition becomes binding. If the commitment constraint is binding with a lower $\overline{W}$, the payment boundary is no longer optimally chosen, and the "super contact" condition is replaced with a physical boundary condition $V(\overline{W}) = L$. The reason why only firm value at the payment boundary turns out to matter under limited commitment is the combination of a concave value function, $V(R) = L$ on the left boundary, and $\overline{W}$ as a reflecting right boundary.

Let variables with a superscript L represent variables in the limited commitment environment, the following proposition summarizes the optimal contract. A formal verification theorem of the optimality of this contract is provided in the appendix.

Proposition 1. *The optimal contract under the principal's limited commitment constraint is characterized by a value function $V^L(W)$ and a payment boundary $\overline{W}^L$, such that*

$$rV^L(W) = \mu + \gamma W V^{L\prime}(W) + \frac{1}{2}\lambda^2\sigma^2 V^{L\prime\prime}(W) \,,$$

subject to boundary conditions $V^L(R) = L$; $V^{L\prime}(\overline{W}^L) = -1$; and

$$V^{L\prime\prime}(\overline{W}^L) = 0, \ if \ V^L(\overline{W}^L) \geq L \,,$$

$$V^L(\overline{W}^L) = L, \ otherwise.$$

As in the full commitment benchmark case, the principal's value functions $V^L(W)$ under limited commitment is also a concave function with an interior maximal point. The prin-

cipal's commitment constraint becomes binding when W is high, in other words when the manager has accumulated adequate performance history. This is not counter-intuitive as W is a measure of the amount of future bonuses managers are owed, a form of debt of the firm induced by the labor contract. Higher debt in the form of bonuses lowers the share of profit investors can earn from investing in the firm. When it becomes large enough, investors will find it optimal to default like they would with any other form of debt. In that sense, the commitment constraint can also be motivated as an upper bound for the operating leverage of the firm.

The specific assumption used here, that firms can default on W any time, is a simple yet without loss generality characterization of the important commitment constraint prevalent in practice and has been commonly used in previous studies[12]. It may appear extreme that firms can default on their promised future payments W anytime without being penalized. One can think of certain mechanisms that would seemingly resolve the commitment issue. For instance a "severance pay" that equals W should firms ever terminate the contract, or an "escrow account" whose balance corresponds to W. However as briefly argued in the introduction, these payments are difficult to be made fully continent and perfectly enforceable in practice. In the case of bankruptcy, which takes place when exactly when firm value is lower than its liquidation value (L in this model), neither severance or other type of payments are fully guaranteed. More importantly, the nature of a incentive compatible contract requires firms to adjust the balance of W constantly according to their managers' performance. This in turn requires firms to have access and make adjustment to any hypothetical account where managers' deferred compensation is stored and therefore, cannot prevent firms from drawing down its balance right before defaulting. In other words, it is intrinsically difficult to distinguish contract termination due to poor performance, and termination due to firm commitment.

Furthermore, there are simple modification of the limited commitment assumption here that can address some of the aforementioned concern but do not change the qualitative results of this paper. For instance, one can introduce real costs associated with firm defaulting by lowering firms' residual value to $L^{def} < L$ in the event of a default. Such modification relaxes

---

[12]See, for instance, Levin (2003), Berk et al. (2010), Ai and Li (2013), Rampini and Viswanathan (2013)

the commitment constraint but still make default a plausible threat. Similarly, even if firms can credibly pledge certain payments when terminating the agent's contract, as long as some W is lost at the time of contract termination, the same concern of limited commitment still applies.

The assumption that both the principal and agent receive constant outside value whenever the contract is terminated is critical to obtaining a close-form solution. Existing models of limited commitment sometimes consider endogenous outside options, and common dynamic programming techniques usually do not apply directly. In my model the constant outside option assumption greatly simplifies the analysis by allowing the application of the Sannikov (2008) method to solve the moral hazard problem. Moreover, since the cash flow in this model follows an arithmetic Brownian motion, constant outside option value rules out strategic control of effort choices or firm size whenever the participation constraints are binding or close to binding.

Several immediate implications can be made by comparing the limited commitment contract and the full commitment contract. The following conclusions can be shown straightforwardly:

Corollary 1. *Under the same parameters,*

$$\overline{W}^L \leq \overline{W} \ .$$

The inequality is strict whenever $V(\overline{W}) < L$ and $V^L(\overline{W}^L) = L$. For all $W \in [R, \overline{W}^L]$,

$$V^L(W) \leq V(W) \ ,$$

$$\overline{W}^L - W \leq \overline{W} - W \ .$$

Corollary 1 states if the principal's participation constraint binds under the limited commitment constraint, the principal can no longer defer cash payment to the agent as much as she would like to under full commitment. This also implies $V^{L\prime\prime}(\overline{W}) < 0$, that is, the "second best" frontier $rV(W) + \gamma W = \mu$ cannot be reached, suggesting a further welfare loss associated with the limited commitment contract. Since lower payment boundary implies

higher likelihood of contract termination, the total surplus generated by the contract is lower when the commitment constraint binds. This provides the explanation for why firm value is always lower for any given W under the limited commitment contract: conditional on delivering the same utility to the agent, the continuation utility for the principal is lower due to earlier termination. This result is not surprising given that the self-enforcing contracting space is a strict subset of the contracting space under full commitment and $V(W)$ measures the highest value for the principal under any incentive compatible contract.

The third conclusion of Corollary 1 leads to some intuitive implications of comparing limited commitment contracts with full commitment contracts. Under the same parameter value, being closer to the payment boundary implies a higher likelihood of reaching the boundary given a certain period of time. In other words, compare two managers with the same level of continuation utility, the one under the limited commitment contract is more likely to receive bonuses in a short period of time. Meanwhile, managers under the limited commitment contract also face a higher turnover rate, because they will not be able to build a large continuation value as a result of the early payment. In all, whenever limited commitment is a binding constraint, firms' capacity to promise future payments is correlated with firm value. In normal times, firm value is high and compensation should be more "front-loaded"; whereas in crisis times, firm value is low and compensation should be more "back-loaded". Although intuitive in light of Corollary 1, these statements are so far only heuristic–I will establish them formally in the next section where I introduce the mathematical concepts needed to conduct the analysis.

Assuming limited commitment actually expands the space of parameters in which the optimal contract exists: the case when $r = \gamma$. In the full commitment model, $r = \gamma$ means the principal can costlessly delay payments to the agent. The payment boundary is therefore infinity, and the optimal contract does not exist. In contrast, the limited commitment constraint puts a physical bound on the payment boundary such that the payment boundary is where the limited commitment constraint binds.

3   Optimal Compensation with Uncertainty Shocks

In this section I introduce uncertainty shocks. I allow the volatility of cash flows to be stochastic, representing the transition between normal and crisis times, and derive the optimal contract under both full and limited commitment. I then show the implications of different contracts for the agent's compensation first through numerical examples from simulations and then with formal analytical arguments.

3.1   Volatility Regime Switching

The comparative statics above offer some intuition over the expected compensation and contract length regarding different levels of uncertainty when firms cannot commit to the contract termination time. While interesting, these comparative statics alone are not sufficient to make a compelling argument for the high level of compensation observed in crisis times. In practice, firms can make state-contingent payments. In other words, the principal can deploy the agent's continuation utility W_t optimally across different economic regimes. Comparative statics derived by holding W_t constant cannot capture the full dynamics of compensation in the presence of state transition.

To characterize the transitional dynamics of compensation under uncertainty shocks I extend the model by introducing regime switching. Specifically, I assume there are two states of the economy: σ_l and σ_h, with $\sigma_l < \sigma_h$, representing "normal" and "crisis" times respectively. Importantly, the state $s \in l, h$ is verifiable and can thus be contracted on. Given this assumption, the cash flow Y_t of the project follows

$$dY_t = \mu(e_t)dt + \sigma_s dZ_t$$

If the current state is s, in any given time interval $(t, t + dt)$, the transition probability to the other state $\widehat{s}$ is $\pi_s dt$. In the remainder of this paper, I further simplify the model by assuming that $\pi_h = 0$, so the state h is absorbing. This means the economy starts with low volatility σ_l and experiences a one-time transition into the high volatility state with probability $\pi_h dt$ within any time interval dt. I will refer to this one-time change in volatility

as the "uncertainty shock" to the economy. Although this may sound restrictive, most of the analytical results do carry through when I allow the states to be recurring, i.e. when $\pi_h > 0$. Discussion of the optimal contract under recurring states is given in the appendix[13].

Again, to offer a benchmark, assume for a moment that the principal has full commitment power. The same martingale method used to solve the single state optimal contract can be applied here but with the inclusion of an extra term in the dynamics of W_t to account for the state transition. Let N_t denote the total number of state transitions before time t. The martingale representation theorem implies

$$dW_t = \gamma W_t - dI_t - \lambda C \mathbb{1}_{\{e_t = \underline{e}\}} dt + \beta_t (dY_t - (\mu - C \mathbb{1}_{\{e_t = \underline{e}\}}) dt) + \delta(W_t)(dN_t - \pi_t dt) \quad (5)$$

The contract is still incentive compatible as long as $\beta_t \geq \lambda$. There are now two value functions for the principal $V_s(W_t)$, one for each state $s \in (l, h)$, that satisfy:

$$
\begin{aligned}
rV_s(W_t) &= \max_{\beta_t \geq \lambda, \delta_s} \mu + (\gamma W_t - \pi_t \delta_s(W_t)) V_s'(W_t) + \frac{1}{2}\beta_t^2 \sigma_s^2 V_s''(W_t) \\
&\quad + \pi_s \left(V_{\hat{s}}(W_t + \delta_s(W_t)) - V_s(W_t) \right) .
\end{aligned}
\quad (6)
$$

The principal now chooses β_t and $\delta_s(W_t)$ optimally. The choice of β_t remains the same as when there is only one state since the moral hazard problem is unchanged, that is $\beta_t = \lambda$ is optimal for a incentive compatible contract. The variable $\delta_s(W_t)$ denotes the discontinuous adjustment in the agent's continuation utility at the time of regime switching. Such adjustment exists because the shadow value of the agent's continuation utility is different for the principal in different states.The principal can promise future compensation conditional on the state of the economy and substitute immediate payments with more future payments if the value of the agent's continuation utility is higher in one state. Therefore, the choice of $\delta_s(W_t)$ is determined by matching the first order derivatives of the principal's value functions

---

[13]I also assume that π_l is a small number to ensure that states l and h have their proper definitions. If π_l is too large, the value function (derived later) in the low volatility state converges to the value function in the high volatility state. To keep them sufficiently distant, π_l must be small enough. See Appendix A for more a detailed discussion.

before and after the regime switching, that is

$$V_{\tilde{s}}'(W_t + \delta_s(W_t)) = V_s'(W_t), \text{ if } W + \delta_s > R , \tag{7}$$

$$\delta_s(W_t) = R - W_t, \text{ otherwise } , \tag{8}$$

In other words, the principal optimally deploys the agent's continuation utility until its marginal value to the principal is equalized across states. In the case where the first order derivatives cannot be matched for any δ_s that keeps the agent's continuation value in the high volatility state above his reservation utility, the contract is simply terminated.

The optimal contract still features a termination boundary R and payment boundaries $\overline{W}_s$ for each state. Boundary conditions under the full commitment setting are $V_s(R) = L$ ("value matching"), $V_s'(\overline{W}_s) = -1$ ("smooth pasting"), and $V_s''(\overline{W}_s) = 0$ ("super contact"). Payment boundaries are on the "second best" frontier $rV(W) + \gamma W = \mu$ for both the low and high volatility state. Numerical examples of the principal's value functions are illustrated in Panel A of Figure 1. This figure shows that firm value is always lower in the high volatility state for any given level of agent's continuation utility except at the termination boundary, that is, $V_h(W) < V_l(W)$ for every $W > R$. Intuitively, since cash flow serves as a signal for the principal to infer the agent's private effort, a more noisy signal increases the likelihood of contract termination which is a necessary but costly action for the principal to provide proper incentives. That is why the regime switching from the low to the high variance state is referred to as a negative shock in this paper.

When the principal has only limited commitment, an argument similar to Section 3.2 applies. The optimal contract features firm value functions $V_s^L(W)$ satisfying the same system of ODEs and same boundary conditions except for the "super contact" condition, which now follows the condition for the limited commitment contract proposed in Proposition 1. Specifically, for each state, if the firm value at the payment boundary is sufficiently high, then $V_s^{L\prime\prime}(\overline{W}_s^L) = 0$ is true. Otherwise, if in any state, firm value becomes too low when W approaches $\overline{W}_s^L$, then the limited commitment constraint $V_s^L(\overline{W}_s^L) = L$ binds in that state.

To summarize, the incentive compatible optimal contract under uncertainty shocks is characterized by the following proposition:

Proposition 2. *The optimal contract under volatility regime switching with full commitment defines a pair of value functions $V_s(W)$ and payment boundaries $\overline{W}_s$, $s \in \{l, h\}$ such that*

$$
\begin{aligned}
rV_s(W) &= \mu + (\gamma W - \pi_s \delta_s(W)) V_s'(W) + \frac{1}{2}\lambda^2 \sigma_s^2 V_s''(W) \\
&\quad + \pi_s \left(V_{\widehat{s}}(W + \delta_s(W)) - V_s(W) \right) ,
\end{aligned}
\tag{9}
$$

subject to boundary conditions $V_s(R) = L$; $V_s'(\overline{W}_s) = -1$; and $V_s''(\overline{W}_s) = 0$. $\delta_s(W)$ is determined by (7) and (8).

If the principal has only limited commitment, the optimal contract defines a pair of value functions $V_s^L(W)$ and payment boundaries $\overline{W}_s^L$, $s \in \{l, h\}$, such that $V_s^L(W)$ satisfies the same system of ODE (10) and boundary conditions $V_s^L(R) = L$; $V_s^{L\prime}(\overline{W}_s^L) = -1$, and

$$
V_s^{L\prime\prime}(\overline{W}_s^L) = 0, \text{ if } V_s^L(\overline{W}_s^L) \geq L ,
$$

$$
V_s^L(\overline{W}_s^L) = L, \text{ otherwise .}
$$

The boundary conditions specified in Proposition 2 imply that under limited commitment, the optimal contract takes three different forms depending on whether the principal's participation constraint (4) is binding at the payment boundary in each state: first, if (4) is not binding for either $s = l$ or $s = h$, this contract is simply identical to the one characterized in Proposition 1. Whether or not the principal can fully commit does not affect the contract. Secondly, the limited commitment constraint can be binding in the high volatility state but not the low volatility state. Third, the constraint can be binding in both states.[14]

Of the three types of contracts, the first type is obviously the least interesting since it is identical to the contract with full commitment. The second type can resemble contracts of either the first or the third type, depending on specific parameter values. I leave the details of this type of contracts to the appendix. The third type of contract produces the most distinct implications for the dynamics of compensation between the full commitment and the limited commitment case. In the remainder of this paper, I will concentrate discussion on this type of contract. That is, unless stated otherwise, I assume the parameter space is

---

[14]It is impossible for the constraint to be binding in the low volatility state but not in the high volatility state, since firm value is always lower when volatility is higher. See Lemma A2.1 in the Appendix for details.

such that under full commitment, $V_s(\overline{W}) < L$ for both $s = l$ and $s = h$[15].

Numerical examples shown in Figure 1 contrast contracts under different commitment assumptions given this parameter space. Recall that panel A shows the value functions for the full commitment contract. Noticing that in both states firm value is below the liquidation value L and therefore payment boundaries in neither state can be sustained without principal's full commitment. Panel B shows the value functions under the same parameters as Panel A but after imposing limited commitment.

Panel A. Full Commitment	Panel B. Limited Commitment
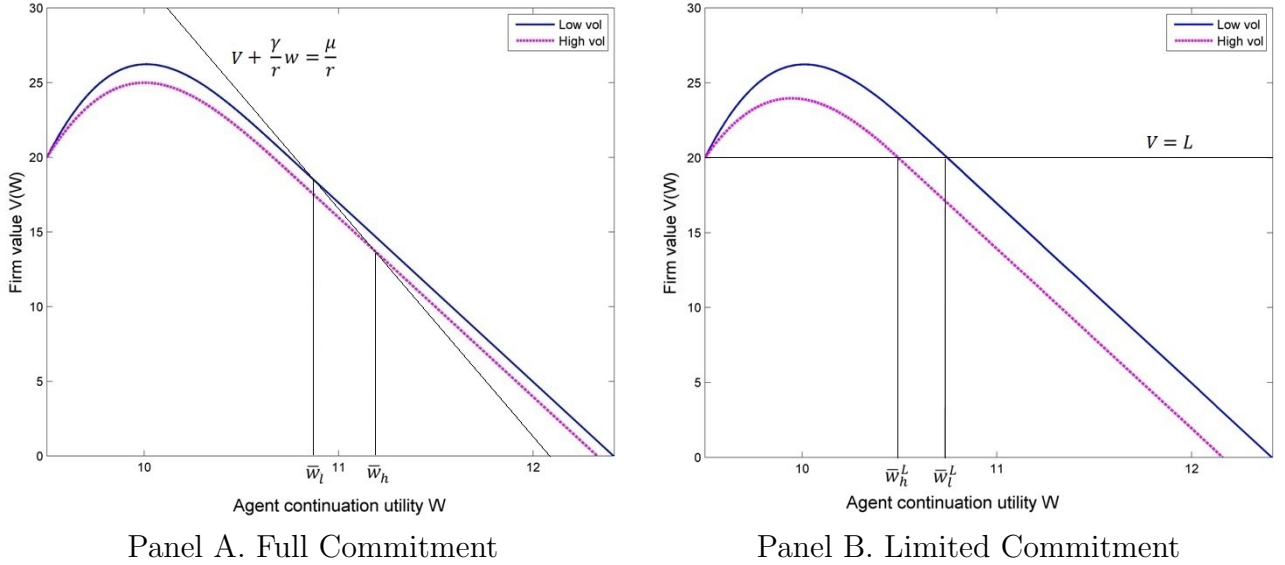
Figure 1 – Value Functions for the Optimal Contracts

This figure plots firm value functions under regime switching. The full commitment case is shown in the left panel. The limited commitment case is shown in the right panel. Parameter values are $L = 20$, $R = 0$, $\gamma = 0.04$, $r = 0.02$, $\mu = 1$, $\lambda = 0.1$, $\sigma_l = 5.9$, $\sigma_h = 6.5$, $\pi_l = 0.001$, $\sigma_h = 0$

The most crucial difference made by imposing the limited commitment condition is the position of the payment boundary. The following result highlights the point:

Corollary 2. *If $V_s(\overline{W}) < L$ and $V_s^L(\overline{W}^L) = L$ for both $s = l$ and $s = h$, then $\overline{W}_h > \overline{W}_l$; $\overline{W}_h^L < \overline{W}_l^L$. That is, the payment boundary under high volatility is higher for the full commitment contract but lower for the limited commitment contract.*

---

[15]The exact space of parameters satisfying such condition is difficult to characterize. However, $\overline{W}$ is larger and $V(\overline{W})$ smaller whenever γ is closer to r, holding other parameters constant. This implies if the principal and the agent have similar patience level, there is a potentially large parameter space in which the limited commitment constraint will be binding in both states once it is imposed.

Corollary 2 states that the principal defers payments to the agent when volatility is high under the full commitment contract. Since the cost of providing incentives to the agent is the possibility of early termination after sufficiently poor performance, it is higher when volatility is higher, as rising uncertainty of cash flows increases the likelihood of sufficiently poor performance and the subsequent early termination. The principal adjusts the contract optimally by giving the agent more financial slack. Here financial slack, defined as $\overline{W}_s - R$, measures how much loss the principal is willing to take before terminating the agent's contract. Greater flexibility to the agent regarding his performance lowers the possibility of costly early termination and is thus optimal for the principal under higher volatility.

In contrast, if the principal has only limited commitment ability, payments to the agent are expedited. Limited commitment implies that the principal's participation constraint $V_s^L(W) \geq L$ must be satisfied at any given time. In other words, the contract must guarantee firm value of at least L, which restricts the amount of future cash flow generated by continuing the project that can be used as compensation to the agent. When uncertainty becomes higher, the total value of the project is lower. A principal lacking the ability to commit to future payments when firm value is too low is forced to pay the agent earlier because the principal can now credibly promise less compensation in the future. These relative positions of the payment boundaries under each volatility state determine the timing of the cash payment to the agent, the expected length of the contract, as well as the concavity of the principal's value function, all of which are essential in studying the compensation structure in the next section.

It is worth noting that the discontinuity in the agent's continuation utility δ has non-trivial solutions even when the transition probability π_l approaches zero, that is when the pair of value functions $V_s(W)$ converge to two independent functions with different values of variance. The effect of π_l on determining δ is small when π_l is close to zero because $V_s(W)$ moves relatively little. This implies analyses of δ can be made almost independently of π_l for small π_l which greatly simplifies the mathematics.

3.2 Numerical Illustration

The optimal contracts under full and limited commitment differ in how the payment boundary is determined. Their implications for compensation thus also differ, as the agent only receives payments in the form of cash bonuses once his continuation utility W exceeds the payment boundary. In this section I show how considering the optimal contract under limited commitment generates conclusions about compensation that standard contracts with full commitment cannot explain. Specifically, I argue that the large compensation observed during the recent crisis can result from optimal contracts under limited commitment, and that managers who receive large cash bonuses also face a shorter expected contract length. Contract termination can be equivalently interpreted as managerial turnover or firm liquidation in the model. I will focus on the former interpretation when discussing the results.

I begin the analysis with numerical simulations, in order to provide a transparent view of contract dynamics. In the simulations I segment the continuous-time model into discrete time intervals. The economy starts with low volatility, and the agent's initial wealth W_0 is drawn uniformly from the interval $(R, \overline{W}_l)$. I simulate N different paths of cash flows. Each path can be interpreted as one manager running an independent project. I then allow the state to switch to σ_h following a poisson arrival process, representing the transition into the crisis time. I calculate W for each of the realized cash flows and, given the payment boundaries, record the timing as well as the size of the cash bonuses. Finally, for each period before and after the uncertainty shock, I calculate the frequency of cash payments by taking the average number of recorded payments among all firms still surviving after the crisis.

I repeat this simulation procedure for both the full commitment and limited commitment contract. Results are shown in Figure 2, with $N = 5000$.

Figure 2 plots the frequency of payments in Panel A plots and the fraction of active projects (managers) at each given time in Panel B. Both contracts share exactly the same parameter values. They differ only in whether or not the commitment constraint is imposed, which alters their payment boundaries. I choose the parameters such that once the limited commitment constraint is imposed, it will be binding in both low and high volatility states, which most clearly manifests the implications of limited commitment.

Two observations emerge from the frequency of cash bonuses shown in Figure 2. On the

Panel A. Frequency of Payments   Panel B. Survival Rate
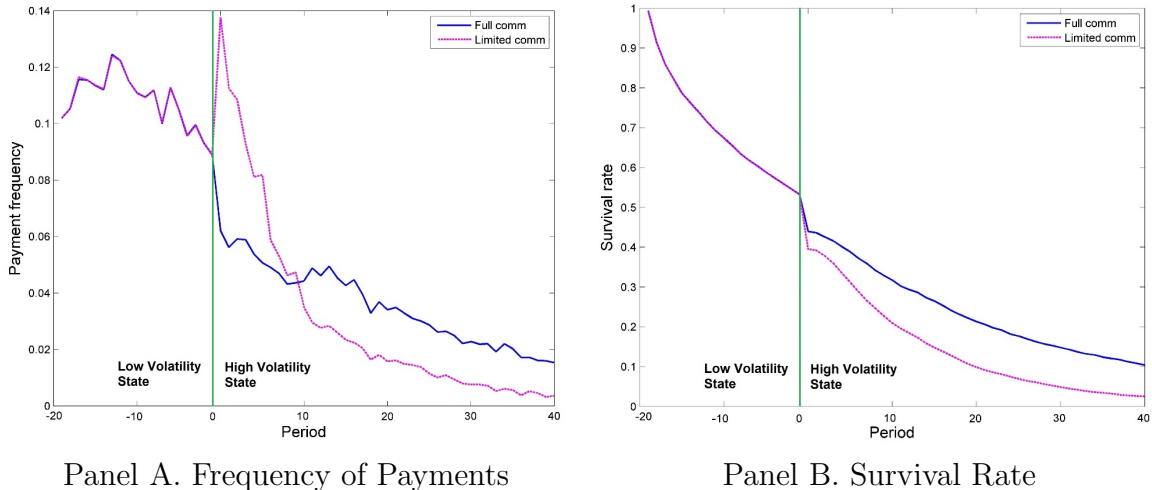
Figure 2 – Simulation Results

This figure plots the frequency of cash compensation (bonuses) and the fraction of active projects from simulating 5000 paths of cash flows, Parameter values are the same as those in Figure 2.

one hand, under the limited commitment contract, the frequency of payments in the few periods immediately after the uncertainty shock is much higher than the frequency under the full commitment contract and the frequency in the low volatility state. On the other hand, payment frequency under the limited commitment contract quickly diminishes to zero due to a higher rate of contract termination, while it is much more persistent under the full commitment contract. These two observations can be summarized into two theoretical predictions:

Predictions: *Managers of firms with limited commitment (1) receive more cash bonuses immediately after entering the crisis time; and (2) face a higher expected turnover rate during the crisis time, compared to managers during normal times and managers of firms with full commitment*

Both results can be formalized using mathematical concepts in stochastic calculus and are rigorously proven in the next subsection. Here I offer readers with a general interest a heuristic derivation and an intuitive explanation of the mechanism behind these results.

Frequency of cash payments can be rationalized when considered jointly with the likelihood of contract termination. When uncertainty is higher, firms with full commitment power optimally set higher bonus hurdles so managers are able to build large continuation

utility, reducing the likelihood of early contract termination. In contrast, without full commitment power, large deferred payments are no longer credible. The higher the cash flow uncertainty, the lower the value from running the firm and the more likely it is for firms to terminate managers' contracts before any bonuses are realized. Managers thus have to be compensated with bonuses early for the increased likelihood of turnover. Put differently, during crisis when firm value is low, firms have to make higher payments to retain their managers, who are more worried about loosing their jobs in the near future. A similar argument applies to the comparison between normal and crisis times for the limited commitment contract, under which the capacity of credibly deferring payments is correlated with firm value in each state.

It is important to clarify here that, despite the above description of immediate payment as a result of shorter expected tenure, the two are not fruit and tree to each other but rather two sides of the same coin. Both compensation and contract length are endogenously determined by the dynamics of the state variable W. In the high volatility state, the dynamics of the agent's continuation utility are given by $dW_t = \gamma W_t - dI_t + \lambda dZ_t$. Payment dI_t reduces W_t and thus increases the likelihood of termination. As shorter length stimulates more front-loaded contracts, a front-loaded contract also implies more aggressive managerial replacement following negative performance.

Both the result regarding compensation and the result regarding turnover are supported by empirical evidence. Besides the level of compensation during the recent crisis which motivates this paper, it has also been suggested that the high level of cash bonuses can be attributed to firm setting lower bonus hurdles[16]. The prediction that managers face higher turnover rate during market downturns is also consistent with empirical studies such as Jenter and Kanaan (2010), Kaplan and Minton (2012), Eisfeldt and Kuhnen (2013). The empirical finding of significant correlation between managerial turnover and poor market returns is puzzling, given that market returns are beyond managers' direct control. Note that those studies usually focus on CEOs while this paper applies to a broader range of employees; they also do not usually control for managerial payment in their regressions and

---

[16]See the Deloitte Directors' Remuneration Report (2010) and related articles on The Times and on Management Today

are therefore unable to directly test the theoretical hypotheses proposed here. Empirical work that simultaneously studies managerial compensation and turnover during market downturns is a potentially interesting direction for further research.

The dynamics of compensation revealed in this section also explain the variety of contracts used in practice. If investors cannot promise to refrain from withdrawing their investment when firm value drops below a certain level, managers will be hesitant to agree to a contract with *back-loaded* payments, that is, a contract that postpones most payments until satisfactory performance is reached. Notice that here satisfactory performance does not necessarily increase investors' valuation of the firm, because the firm's labor bill grows larger as well and, in the model where $V'(W_t) < 0$, better performance from the agent implies less value to the principal because the agent is paid a higher share of the profit. The manager's concern is greatest precisely during crisis, when total value from the firm's projects is the lowest, and firms are more likely to close in the near future. This leads managers to demand *front-loaded* contracts instead, where they are paid sooner rather than later as suggested by the simulation results. On the contrary, if the the principal can fully commit to retain managers they expect a longer tenure and may agree to postpone their payment further to achieve a higher total payoff from the contract.

In the aftermath of the recent financial crisis many economists and politicians blamed the current managerial compensation scheme for not aligning managerial incentives with long-term investor benefit. Consequently, policy recommendations to propagate the use of delayed payment as a solution to that problem were suggested. For instance, the Troubled Asset Relief Program (TARP) limits the ability of executives of TARP firms to cash out their restricted stock until the government is repaid in full [17]. However, the effectiveness of such recommendation hinges on the credibility that future payment promised to the executives will be delivered at full value. If executives believe that when firms are in distress, investors will withdraw by selling their shares, then the value of their stock holdings is less the longer they have to wait to cash them out. As a result executives may require even higher and more immediate compensation at the time of distress.

---

[17]See TARP Standards for Compensation and Corporate Governance, 74 Fed. Reg. 28,394, 28,410 (June 15, 2009)

I also calculate the average size of cash payments for each period before and after the uncertainty shock, which produces a pattern almost identical to that observed in Figure 2. This is not surprising given that conditional on receiving payments, the size of payments depends only on the variance σ which is a constant once the state is fixed. Moreover, instead of drawing the agent's initial utility randomly and uniformly from $(R, \overline{W}_l)$, I also conduct similar simulations but fix the manager's initial utility to be $W_l^* \equiv \arg\max_W V_l(W)$, which is the optimal level of continuation utility promised by the principal if she were to offer the contract. Results of this exercise are again very similar to those in Figure 2 and are thus omitted here.

3.3   Formal Analysis of the Regime Switching Model

While numerical simulations provide intuitive and transparent stories, I now formally state and prove the results. In addition to being mathematically rigorous, the formal argument also provides new insights into some important and controversial topics in the research on executive compensation.

The argument for more immediate payments under the limited commitment contract consists of two major steps, both of which are stated relative to the full commitment contract: (1) since the payment boundary is lower when volatility is higher, at the time of regime switching, the agent's continuation utility is adjusted downward, closer to the payment boundary; (2) when the drift of the agent's continuation utility process is low, being closer to the payment boundary implies a higher likelihood of reaching that boundary and incurring cash compensation in a shorter period of time, suggesting more frequent payments immediately after the crisis hits. The argument for a higher turnover rate is relatively simpler: the agent's adjusted continuation utility after the volatility increase is also closer to the termination boundary, because the limited commitment contract gives managers less financial slack when the cash flow signal is noisier.

3.3.1 Adjustment in the Agent's Wealth

The first step in formally showing the result of crisis time compensation is to derive the adjustment of the agent's continuation utility $\delta(W)$. Consider the full commitment contract first. The optimal contract characterized in Proposition 2 suggests that W is discontinuous at the time of the uncertainty shock. Such discontinuity arises because the principal adjusts the shadow value of the agent's wealth optimally to keep incentives the same before and after the shock, which is reflected by the slope matching procedure introduce in Propositions 2. However, such equalization of incentive is not always viable as the two value functions corresponding to each state have different ranges of slopes. In particular at the termination boundary, because $V_s(R) = L$ for both s but $V_l(W) > V_h(W)$ for all W imply $V_l'(R) > V_h'(R)$. By concavity, $V_h'(R) > V_h'(W)$ for every $W > R$. Therefore, there is some cut-off level W_F such that if the agent's wealth before the uncertainty shock, denoted by W_{t^-}, falls below W_F, the principal simply cannot keep the same marginal value of agent's wealth after the shock. Consequently the contract is terminated as soon as the shock occurs.

This same argument applies to limited commitment contracts, but the difference lies in the sign of $\delta_l(W)$ (full commitment) and $\delta_l^L(W)$ (limited commitment) given each W. Figure 3 Panel A illustrates the change in $\delta_l(W)$ and $\delta_l^L(W)$ as functions of the agent's wealth W before the state transition. Both the full commitment and limited commitment contracts are presented to offer comparison. Note that there is a kink point, before which $\delta_l(W)$ and $\delta_l^L(W)$ both first starts from 0 and then descends until the kink. This represents the region in which contracts are terminated once the uncertainty shock arrives. However, the full commitment and limited commitment contracts behave very differently thereafter: while $\delta_l(W)$ for the full commitment contract continues to grow until it becomes positive, $\delta_l^L(W)$ remains negative all the way to the payment boundary for the limited commitment case. This suggests that if the agent has accumulated sufficiently good performance before the shock, the size of $\delta(W)$ takes different values depending on the different payment boundaries specified according to the type of the contract.

The observation from Panel A of Figure 3 can be formally summarized in the following proposition:
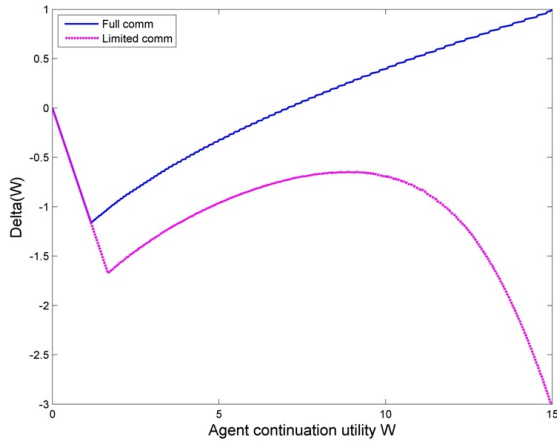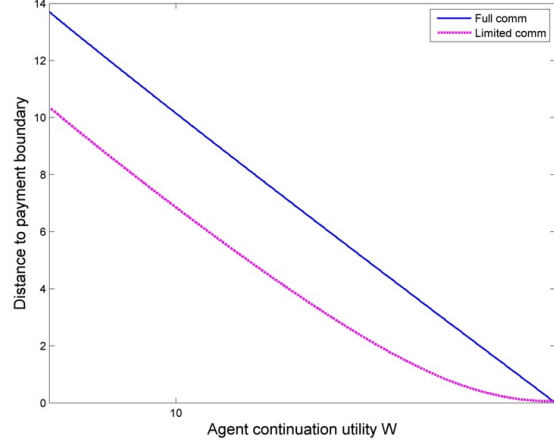
| Panel A. Size of $\delta_l(W)$ | Panel B. Distance to the Payment Boundary |

Figure 3 – Allocation of Agent's Continuation Utility

This figure plots the size of δ_l (left panel) and $\overline{W}_h - W_{t^+}$ (right panel), the distance between agent's continuation utility and the payment boundary after the uncertainty shock.

Proposition 3. *There exist cut-off levels of the agent's continuation utility $\widehat{W}$ such that if $W_{t^-} > \widehat{W}$:*

$$\delta_l(W_{t^-}) > 0$$

$$\delta_l^L(W_{t^-}) < 0$$

This proposition links the type of contract to different predictions of the "pay-for-luck" phenomenon documented by empirical works such as Bertrand and Mullainathan (2001) and Axelson and Baliga (2009). These studies find that manager's compensation is related to market performance beyond their control, which is contradictory to earlier standard contract theories such as Holmstrom (1982) which argues that the principal should filter out any signals unrelated to the manager's own performance. Many theories motivated by this contradiction have been developed in recent years, most of which feature some kind of managerial entrenchment or hidden talent that is partly related to market signals. See Bebchuk and Fried (2006) and Jenter and Kanaan (2010) for a more detailed survey.

In contrast to existing theories, pay-for-luck is a natural feature of the regime switching model used in this paper. As the agent is not responsible for the occurrence of uncertainty shocks, his continuation utility from the contract is indeed adjusted to keep its marginal value to the principal unchanged. However, whether that adjustment corresponds to a "reward"

or a "punishment" depends on the principal's commitment power. The full commitment contract rewards the agent with higher continuation value when uncertainty is higher. Considering higher uncertainty naturally as "bad luck" since firm value is lower in this state, this prediction is less intuitive. On the contrary, when the principal has limited commitment, agents are "punished" for higher uncertainty as their continuation value is brought down. In fact, if the model allows recurring state transitions, the direction of δ flips signs when the state switches from "crisis" to "normal", and the limited commitment contract predicts a "reward" for "good luck", which is largely consistent with the empirical findings.

It should be noted nonetheless that here neither the "reward" nor "punishment" involves any instant cash transfer. As W measures the agent's present value of all future payments, the adjustment of W is merely a reflection of the different payment boundaries and termination likelihood. The actual payments are related not only to the shift in boundaries but also the distance between the boundaries and the agent's continuation value after the adjustment. The next result based on Proposition 3 sheds light on this point:

Corollary 3. *Let W_{t-} be the agent's continuation utility before the uncertainty shock, and $W_{t+} \equiv W_{t-} + \delta_l(W_{t-})$ and $W_{t+}^L \equiv W_{t-} + \delta_l^L(W_{t-})$ be the agent's continuation utility after the uncertainty shock under full and limited commitment contract, respectively, then*

$$\overline{W}_h^L - W_{t+}^L < \overline{W}_h - W_{t+}, \textit{if } W_{t-} > \widehat{W} \ .$$

This result is illustrated in Panel B of Figure 3. It shows that while the payment boundary of the limited commitment contract is lower, the agent's adjusted continuation utility after the uncertainty shock is also closer to the payment boundary. This conclusion plays a leading role in the analysis of compensation later, as being close to the payment boundary implies a larger probability of receiving more cash payments in the near future. At the same time, a lower payment boundary suggests a higher likelihood of contract termination following a series of poor performances. This trade-off between immediate cash payments and likelihood of termination is the central mechanism behind the dynamics of compensation.

After establishing the direction of $\delta(W)$ and the position of W_{t+} relative to the payment boundary, I can formalize the observations from the simulation example using standard

methods in stochastic calculus. The argument is presented in the next subsection

3.3.2   Analytical Characterization

Here I characterize the dynamics of compensation following uncertainty shocks. Given any W_{t+}, the agent's continuation utility after the volatility increase, the goal is to characterize the distribution of the agent's wealth after a certain amount of time elapses. Following Cox and Miller (1977), given the dynamics of W, the transition density function $f(t, W; W_{t+})$ for a process starts with W_{t+} and satisfies the Kolmogorov forward equation:

$$\frac{\partial}{\partial t} f(t, W; W_{t+}) = \frac{1}{2} \frac{\partial^2}{\partial W^2} \left[\lambda^2 \sigma_h^2 f(t, W; W_{t+}) \right] - \frac{\partial}{\partial W} \left[\gamma W f(t, W; W_{t+}) \right]$$

subject to boundary conditions:

$$f(t, R; W_{t+}) = 0$$

$$\frac{1}{2} \frac{\partial}{\partial W} \left[\lambda^2 \sigma_h^2 f(t, W; W_{t+}) \right] |_{W = \overline{W}_h} - \gamma \overline{W}_h f(t, \overline{W}_h; W_{t+}) = 0$$

Unfortunately, this partial differential equation is generally intractable. However, when γ is small, the dynamics of W can be approximated by a standard Brownian motion with one absorbing boundary R and one reflecting boundary $\overline{W}_h$, whose transition density has an explicit form[18]. Details on the approximation and the derivation of the transition density are shown in the Appendix by virtue of the method developed in Ward and Glynn (2003)

After obtaining the transition density, I can measure the likelihood of cash payments given a certain time period T after the shock using the concept of local time in stochastic processes. Given a time period T and initial point W_{t+}, define local time $\mathcal{L}$

$$\mathcal{L}_h(T; W_{t+}) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^T \mathbb{1}_{\left\{ \overline{W}_h - \varepsilon < W_t < \overline{W}_h + \varepsilon \right\}} dt | W_0 = W_{t+}$$

where $\mathbb{1}_{\{\cdot\}}$ is the indicator function. This local time is a random variable that measures the amount of time W spends in the neighborhood of the payment boundary. Since being at the payment boundary implies cash payments, this can be interpreted as the frequency of

---

[18]The assumption $\gamma > r$ is still needed for the benchmark full commitment contract to exist. It is not necessary, though, for the limited commitment contract. See the end of Section 3.2. for the discussion

payments an agent with initial wealth W_{t+} receives within time T after the economy enters crisis mode.

The interesting value is the expectation of local time given by

$$E\left[\mathcal{L}_h(T; W_{t+})\right] = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^T dt \int_{\overline{W}_h - \varepsilon}^{\overline{W}_h + \varepsilon} f(t, W; W_{t+}) dW$$

This value measures the expected frequency of cash payments given initial wealth W_{t+}. The higher this value, the more frequently the agent can expect to receive cash payments before time T during the crisis. As the numerical simulations show, under the limited commitment contract, agents on average receive cash payments more frequently for a short period immediately after the crisis begins. This observation can now be stated formally using the expected local time defined above. Last but not least, the numerical simulations begin with a random agent's wealth in the low variance state. In order for the analytical results to better match those from simulations as well as what is seen in practice, I need to replace the fixed initial wealth W_{t+} in the high variance state with the agent's wealth W_{t-} in the low variance state. Thanks to Proposition 3 and Corollary 3 there is a one to one mapping between the two variables which allows me to summarize the analytical finding in the next proposition.

Proposition 4. *Assume γ is small. There exists $\widehat{T}$ and $\widehat{W}_{t-}$ such that if $T < \widehat{T}$:*

$$E^L\left[\mathcal{L}_h(T; W_{t-})\right] > E\left[\mathcal{L}_h(T; W_{t-})\right]$$

$$E^L\left[\mathcal{L}_h(T; W_{t-})\right] > E^L\left[\mathcal{L}_l(T; W_{t-})\right]$$

for all $W_{t-} > \widehat{W}_{t-}$, where E^L represents expectation under the limited commitment contract

Proposition 4 provides the formal conditions under which the frequency of payments is higher under the limited commitment contract, in the high volatility state. Despite the mathematical complexity, its basic intuition is quite simple: first, compare the limited and full commitment contract, Corollary 3 shows W is closer to the payment boundary after the uncertainty shock under the limited commitment contract. When γ is small, the process of W behaves similarly to a standard Brownian motion and thus spends more time at the

payment boundary whenever the starting point is closer to the boundary. In more intuitive terms, the agent should expect more frequent payments in the near future if his cumulative performance is closer to the target bonus hurdle set by his contract. The similar argument applies to the comparison between the limited commitment contract in low and high volatility states: W_{t^+} is closer to $\overline{W}_h^L$ than W_{t^-} is to $\overline{W}_l^L$.

Why does Proposition 4 hold only when T is small? This is because while W_{t^+} is closer to the payment boundary after the shock under the full commitment contract, it is also closer to the termination boundary because the agent is overall punished according to Proposition 3. As T increases, the likelihood of contract termination rises faster for the limited commitment contract. That is, agents now operate under tighter financial slack. The longer into a crisis, the more likely is termination, as the possibility of realizing a series of losses becomes more real. The conclusion in Proposition 4 thus holds only for T small enough, when the probability of termination is negligible. As shown by the numerical simulations, this pertains to the second observation that cash payment vanishes very quickly under the high volatility state under limited commitment. The notion of termination likelihood can be formally described using the concept of stopping time, as the next proposition shows:

Proposition 5. *Define* $\tau_s = \inf\left\{t : W_t = R | \overline{W}_s\right\}$ *as the termination time given payment threshold* $\overline{W}_s$ *then:*

$$E^L(\tau_h) < E^L(\tau_l)$$

$$E^L(\tau_h) < E(\tau_h)$$

When the commitment constraint is binding, the agent's expected termination time is shorter under high volatility

Intuitively, given the absorbing boundary R and reflecting boundary $\overline{W}_h$, a process with initial value W_{t^+} is in expectation stopped earlier whenever W_{t^+} is closer to R and $\overline{W}_h$ is smaller. The limited commitment contract satisfies both conditions. Further, it should be noted that this proposition does not require the assumption of a small γ, as the expected speed of growth for W is lower when W_{t^+} is lower, which the limited commitment contract again satisfies. Nevertheless the proof of Proposition 5 still imposes the restriction on γ for the sole purpose of analytical tractability.

The results of this subsection imply that the recipients of crisis time bonuses are those who perform relatively well before the crisis. Proposition 4 states that more frequent cash compensation is conditional on the agent's wealth before the shock W_{t-} surpassing a certain threshold, and higher W_{t-} represents better before-shock performance. Those who perform relatively poorly ex ante are no longer around after the crises as a result of either replacement or firm liquidation. Combined with Proposition 3, this suggests that those who produce the largest profits before the crisis are being criticized the most for receiving bonuses during the crisis. One should keep in mind, however, that the huge loss of firm wealth is primarily due to the risky aggregate environment and, despite receiving bonuses for a short period into the crisis, managers are being harmed overall.

The optimal contract derived in this paper is not renegotiation-proof, which may raise a legitimate concern but does not affect the main results. Renegotiation-proof contracts require the value function to be downward sloping everywhere. In the Appendix I derive the renegation-proof contract and show that the main results carry through. Despite the principal having only limited power of commitment, renegotiation-proofness is not a necessary feature of the resulting equilibrium contract, because it is assumed in the model that the principal can commit not to renegotiate the contract but just cannot commit to when to terminate the contracting relationship.

The assumption of constant liquidation value L and reservation utility R for both high and low volatility states is for the sake of simplicity but could be extended to better match reality. It is quite plausible that both outside options could have state-contingent values. Managers could have difficulty finding another job if laid off during crisis times. Similarly, if L is interpreted as the fixed cost for firms to replace an incumbent manager, both L and R should be lower in the high volatility state.

Nevertheless, allowing reservation value to be state-contingent does not affect the validity of the main results. The renegotiation-proof contracts discussed in the Appendix feature endogenous renegotiation boundaries, which are lower when volatility is higher, similar to assuming a lower R in the high volatility state. The main results of this paper still hold under the renegotiation-proof contracts because they depend mainly on the relative positions of the payment boundaries. As for the liquidation value L, a lower value implies a lower value

for the principal's outside option, thus reducing the tightness of the commitment constraint imposed on the contract. Therefore, whether more frequent cash bonuses are paid in the high volatility state simply depends on the decrease in L relative to the increase in σ.

4   Extensions

In this section I discuss two extensions of my main model. One, the implementation of the limited commitment contract, which justifies the limited commitment constraint by revealing a similarity between a firm's commitment to a contract termination time and its commitment to a capital structure. Two, I explore the equilibrium in which shirking is optimal and its implications for both empirical studies and policy recommendations

4.1   Contract Implementation, Capital Structure and the Commitment Constraint

Results from the previous section highlight different dynamics of compensation generated by full commitment contracts and limited commitment contracts. In this section I further explore the difference between those contracts via contract implementation and establish a novel equivalence between firms' commitment to compensation contracts and commitment to capital structure. The equivalence provides a justification for the limited commitment assumption made throughout this paper, as firms' commitment to capital structure is known to be implausible.

Implementing the full commitment contract involves the use of debt and equity, and thus creating a conflict between debt and equity holders that leads to potential commitment issues. When the regime switches from low to high volatility, the face value of debt must be brought down at the expense of equity holders. Such implementation imposes an implicit assumption that equity holders must commit to maintain a certain capital structure, which is generally implausible since equity holders do not always act for the benefit of the entire firm. In contrast, contracts with limited commitment do not require firms to make such commitment to capital structure and should thus be more prevalent in practice.

The implementation in this paper follows the standard literature in using a set of common securities with limited liability: equity, long-term debt, and credit line. Equity can be held by both the manager as well as outside investors who receive dividend payments and can decide the firm's capital structure. The long-term debt is a callable consol bond that pays a fixed rate and has a fixed face value. The firm can issue more long-term debt or call it back at its face value [19]. Finally, the credit line provides the manager with limited liquidity. The manager decides both the dividend and the credit line balance, but incentive compatibility under the optimal contract renders irrelevant who makes dividend and credit line decisions.

There is more than one implementation of the optimal contract. The following proposition provides a standard result:

Proposition 6. *Both the full and limited commitment contract can be implemented by*
  (a) manager holding inside equity share λ;
  (b) face value of the callable debt satisfying $D_s = V_s(\overline{W}_s)$;
  (c) credit line balance M_t and credit limit C^* satisfying $W_t = \lambda(C_s^* - M_t)$ and $\lambda C_s^* = \overline{W}_s$.
  Dividend is paid when $M_t = 0$. Liquidation occurs when M_t reaches C_s^*.

The implementation is intuitive and hence the explanation here concise. Since λ measures the portion of private benefits the manager can derive from shirking, its value represents the least degree of sensitivity to cash flow to which the manager is exposed. Managers can draw down the line of credit for operating liquidity. Dividends serve as reward to the manager as well as returns to outside investors. Since the manager has a higher discount rate, dividends will not be saved inside the firm as long as the credit line is paid in full. Finally, the amount of long-term debt is used to adjust the profit rate of the firm such that incentives remain unchanged.

Security implementation of the optimal contract implies a certain capital structure which can potentially raise questions under the regime switching environment: the boundary conditions of the full commitment contract implies $V_h(\overline{W}_h) < V_l(\overline{W}_l)$. Since $V_h(\overline{W}_h)$ and $V_l(\overline{W}_l)$ also correspond to the face value of the callable debt in the high and low volatility state, the implementation of the full commitment contract requires that the face value of long-term

---

[19]Although callable debt is usually redeemed at a premium, the specific value of the premium does not play a role in this model and is thus without loss of generality assumed to be zero.

debt be brought down when volatility increases. In other words, some portion of the long-term debt must be called back, hence the usage of callable debt here. These callbacks induce a transfer of wealth out of equity holders' pockets while debt holders are paid in full.

To further investigate this problem, I compute the value of the aforementioned securities, in particular that of equity. To simplify the analysis, I assume that $L = 0$, so that there will be no residual value in the event of firm liquidation, eliminating the need to specify the priority of residual claims among equity holders and debt holders. Let V_s^E denote the equity value, which is defined by

$$V^E(W_t) = E\left(\int_t^\tau e^{-rs} dDiv_s | W_t\right)$$

where W_t, the manager's continuation utility, can be transferred to M_t, the credit line balance, through the relationship defined in Proposition 6.

The value function of equity value can be characterized by the following differential equation:

$$rV_s^E(W) = (\gamma W - \pi_s \delta_s(W)) V_s^{E\prime} + \frac{1}{2}\lambda^2\sigma_s^2 V_s^{E\prime\prime} + \pi_s(V_{\widehat{s}}^E(W + \delta_s(W)) - V_s^E(W))$$

subject to boundary conditions

$$V_s^E(0) = 0$$

$$V_s^{E\prime}(\overline{W}_s) = 1$$

where $V_{\widehat{s}}^E$ is the value of equity in state $\widehat{s}$ The implementation requires equity holders to commit to the particular capital structure specified in the optimal contract by redeeming the outstanding debt at the time of the uncertainty shock.

Do equity holders always find it preferable to recall debt when uncertainty is high? The answer is hardly yes, as equity holders can usually withdraw investment in practice and default on any debt obligation. In this model, let $D_{\widehat{s}} - D_s$ measure the value of debt redemption. Equity holders will find it optimal to default when $V_{\widehat{s}}^E$, the value from maintaining the firm, is lower than $(D_{\widehat{s}} - D_s)$, the cost of doing so. On the contrary, under the limited commitment contract, $V_s(\overline{W}) = L$ for both $s = l$ and h implies an identical face value of long-term debt

before and after regime switching. That is, the capital structure of the limited commitment contract can be maintained without a tendency on the part of equity holders to default ex post.

Equity holders' making ex post default decisions is common in both financial research and in practice. There is a large body of literature studying the endogenous default decision of equity holders and the conflict with debt holders, notably Leland (1994), Leland and Toft (1996), He and Xiong (2012). In this paper, I do not characterize the exact default boundary of equity holders; rather, the point I want to make is that equity holders cannot (credibly) commit to not defaulting for the sake of the entire firm when there is a chance of their finding default preferable ex post.

In additional to justifying the prevalence of limited commitment contracts in practice, the equivalence between the commitment to contract termination time and the commitment to capital structure also offers an empirically testable hypothesis: investors of more distressed firms are more likely to withdraw their investment, default on firms' debt and alter firms' capital structure. The degree of distress can be a potential proxy for the commitment power firms have over their labor contracts which is difficult to observe.

4.2 Optimal Compensation with Shirking

Throughout previous analyses, it has been assumed that working is always preferred by the principal regardless of the level of uncertainty. This section relaxes this assumption and examines when the optimal contract allows shirking in equilibrium. The results carry both policy and empirical implications. When the contract allows shirking during crisis times, no bonuses are paid. This may appear agreeable to policymakers and to public sentiment, but it is actually worse because the average productivity of the economy is lower as a result of lower managerial effort. Empirically, studies of compensation and performance, such as those examining pay-performance sensitivity, could be confounded by the endogeneity between return and volatility driven by unobservable changes in managerial effort.

Which effort level is optimal in equilibrium depends on the cost of allowing shirking. Working is preferred as long as C, the social cost of shirking measured by the reduction in average cash flow, is high. This section explores this assumption in more detail. Consider

a contract that involves no payment but simply allows the agent to shirk forever. Define (W^S, V^S) as the pair of payoffs for the agent and the principal respectively if the agent exerts effort $e_t = \underline{e}$ for all t, then

$$W^S = \frac{\lambda C}{\gamma} \; ;$$

$$V^S = \frac{1}{r}\left(\mu - \frac{\lambda C}{\gamma}\right) \; . \tag{10}$$

Notice that this payoff is a function of A, which is an irrelevant variable in the incentive compatible contract characterized in Proposition 1 and 2. Therefore, whether the incentive compatible contract is optimal for the principal depends on the level of V^S. When C is sufficiently low, $V^S > V(W^*)$, where $V(W^*) \equiv \max V(W)$ is the maximal value the principal can derive from an incentive compatible contract, the principal is better off stopping incentive provision[20]. The agent will choose to shirk, receive no payment from the principal and instead be compensated by his private benefit from shirking. The optimal contract is static, unrelated to the agent's performance, and therefore involves no termination.

The different maximal firm value under low and high volatility states raises the possibility that working is not always optimal for both states. If $V_l^* > V^S > V_h^*$, the optimal contract will induce working as long as $s = l$ and switches to the static contract at times when $s = h$. In the model where only one state transition occurs, the dynamics of the optimal contract follow the ODEs described in Proposition 2, except the value function V_h is replaced by the static payoff given by equation (10).

If shirking is optimal in the high volatility state, the procedure that pins down $\delta_l(W)$ is slightly different. Under the static contract that allows shirking when $s = h$, the agent's continuation utility is a singleton W^S. Jumps of W from the low to the high volatility state is simply $\delta_l(W) = W^S - W$. Note that W^S measures not the discounted future income but the present value of private shirking benefit to the agent. This value is automatically achieved as long as the principal immediately ceases any payment. Moreover, because W^S is no longer sensitive to the agent's performance, there is no contract termination after the

---

[20]Strictly speaking, the contract that allows shirking forever is optimal only when $V^S > B(W)$, where $B(W)$ is a V-shaped function that extends above V_h^*. See Zhu (2012) for the details. Here I avoid the complicated situations where V^S lies above b^* but below $B(W)$ by assuming that C is either high enough or low enough so that either working or shirking permanently is the optimal effort.

regime switches to high volatility. All firms survive regardless of their agent's performance history up to the regime switching time. The next proposition summarizes these findings:

Proposition 7. *Suppose C is low or σ_h is sufficiently high, the optimal contract induces $e_t = \bar{e}$ under σ_l but $e_t = \underline{e}$ under σ_h. The principal's value function $V_l(W)$ and payment boundaries $\overline{W}_l$ satisfy:*

$$
\begin{aligned}
rV_l(W) &= \mu + (\gamma W - \pi_l \delta_l(W))\, V_l'(W) + \frac{1}{2}\lambda^2 \sigma_L^2 V_l''(W) \\
&\quad + \pi_l \left(V_h - V_l(W) \right)
\end{aligned}
$$

subject to boundary conditions $V_l(R) = L$; $V_l'(\overline{W}_l) = -1$; and $V_l''(\overline{W}_l) = 0$. Furthermore, $\delta_l(W) = W^S - W = \frac{\lambda C}{\gamma} - W$, and V_h is given by

$$
V_h = V^S = \frac{1}{r}\left(\mu - \frac{\lambda C}{\gamma} \right) .
$$

The existence of an optimal contract that involves shirking in the equilibrium has important policy implications. Since the manager is compensated through the private benefit of shirking when uncertainty is high, no cash payment is made under that regime. This implies the possibility of observing little or no bonuses during a recession. However, though much to the media or public's liking, this equilibrium is actually worse in terms of total welfare, because productivity, measured by mean cash flow, is now lower due to less effort from managers. This is true as long as $\lambda < 1$ so that there is deadweight loss associated with managerial shirking. This result highlights the importance of compensation in keeping mangers properly incentivized, even though the exact timing of their compensation may not match their overall performance at the time when a large negative shock occurs.

The shirking equilibrium also reveals a potential endogeneity problem between profitability and volatility. Existing empirical work that studies compensation often considers profitability and volatility as independent factors. However, fluctuation in profitability can be driven by the change in volatility through the channel of managerial effort, raising empirical challenges since such effort is normally difficult to measure. It also provides further evidence in addition to previous work that uncertainty is the key to understanding the recent

financial crisis.

It is worth noting that the change in the agent's equilibrium effort is a feature of increasing volatility but not necessarily of decreasing profitability. This sets this paper apart from those with similar regime switching techniques such as Hoffmann and Pfeil (2010). While lower average cash flow μ does bring down firm value under an incentive compatible contract, it also lowers V^S, firm value under a static contract that allows shirking. As a result, working can still be the optimal effort to induce if $V^S < V_h^*$. In contrast, V^S does not depend on σ, but V_h^* does. When cash flow volatility becomes higher, V_h^* becomes lower until falling below V^S, and the incentive compatible contract is dominated by the static contract, a unique feature of stochastic volatility.

5   Conclusion

This paper studies the optimal compensation contract under the twin assumptions of limited commitment by the principal and regime switching of cash flow volatility. Sudden and dramatic increases in market uncertainty have been argued as the most critical aspect of financial crises. When uncertainty is high, investment becomes more risky, and the value of continuing the firm is correspondingly low. Principals without full commitment cannot credibly pledge sufficient amounts of future payments and must provide agents with more immediate compensation. This offers an explanation for the highly controversial large compensation–especially cash bonuses–paid by many financial firms during the recent crisis. At the same time, managers are subject to a higher turnover rate. Therefore, despite the bonuses, managers are still worse off in crisis times as their overall present value from the contract is lower.

The model provides a new perspective on the current compensation structure not only in regard to bonuses. While equity-based compensation contracts provide a solution to the problem of aligning managers' incentives with those of investors, their effectiveness depends on the ability of both parties to commit to the contract. Furthermore, payment and expected tenure are two counter weights both endogenously determined in the optimal contract. Any measure that intends to provide better long-term incentives to managers must take into

account both payment and the expected contract length.

There are several directions in which this model can be fruitfully extended. In the model there is only one representative firm and one representative agent, so the uncertainty shock can be interpreted as either aggregate or idiosyncratic. A model that allows firm heterogeneity and differentiates idiosyncratic shocks from aggregate shocks may generate interesting results, such as the cross-sectional predictions regarding the response of compensation to firm level investment risks. This model also potentially speaks to the the important issue of liquidity management in response to market downturns. Recent studies such as Campello et al. (2011) examine cross-sectional liquidity management along different dimensions. A slightly modified dynamic model à la Bolton et al. (2011) would be readily equipped to provide theoretical insights for these observations.

Appendix A. Proofs

Proof of Proposition 1: This optimality condition for a single state environment is identical to the baseline model in DeMarzo and Sannikov (2006) if the commitment constraint $V''(\overline{W}) \geq L$ is not binding, so I will only focus on the case when such condition is violated. I use superscript L to denote variables and functions for the limited commitment case.

Define the social benefit function as $F(W) = W + V(W)$, which satisfies

$$F''(\overline{W}) = \frac{rF(\overline{W}) + (\gamma - r)\overline{W} - \mu}{\frac{1}{2}\lambda^2\sigma_s^2} \; .$$

When the principal's participation constraint is binding, $F^L(\overline{W}) = L + \overline{W}^L$ implying

$$F^{L''}(\overline{W}^L) = \frac{rL + \gamma\overline{W}^L - \mu}{\frac{1}{2}\lambda^2\sigma_s^2} \; .$$

Suppose $F^{L''}(\overline{W}^L) > 0$, that is, $rL + \gamma\overline{W}^L > \mu$, this implies that $V^{L''}(\overline{W}^L) > 0$. Since $V^L(\overline{W}^L) = L$, $rV^L(\overline{W}) + \gamma\overline{W}^L > \mu$. Compare this result to the case of full commitment, where $rV(\overline{W}) + \gamma\overline{W} = \mu$. If $\overline{W}^L < \overline{W}$, since $rV(W) + \gamma W < \mu$ for all $W < \overline{W}$, it must be that $rV(\overline{W}^L) + \gamma\overline{W}^L < \mu$, which implies $V(\overline{W}^L) < V^L(\overline{W}^L)$. However this is a contradiction since $V(W) \geq V^L(W)$ for every W. If, on the other hand, $\overline{W}^L > \overline{W}$, but $V^L(\overline{W}^L) = L$ and $V(\overline{W}) < L$, which implies that $V^L(\overline{W}^L) > V(\overline{W}^L)$ again, contradiction. Therefore $F^{L''}(W) < 0$ in the neighbourhood of $\overline{W}^L$.

The rest of the argument about F^L being also concave besides the neighbourhood of $\overline{W}^L$ follows the standard argument. The proof also implies immediately that $rV^L(W) + \gamma W \leq \mu$ for all W if the boundary condition $V^{L''}(\overline{W}^L) = L$ is true. This conclusion is used in the following verification theorem.

Verification Theorem: for any incentive compatible contract, define an auxiliary gain process G as

$$G_t = \int_0^t e^{-rs}(dY_s - dI_s) + e^{-rt}V(W_t) \; ,$$

where W_t evolves according to dW_t. By Ito's lemma

$$e^{rt}G_t = \left(\mu + \gamma W_t V'(W_t) + \frac{1}{2}\beta_t^2\sigma^2 V''(W_t) - rV(W_t)\right)dt - (1 + V'(W_t))dI_t + (1 + \beta_t V(W_t))\sigma dZ_t \; .$$
$$(11)$$

The first two terms are negative and therefore G_t is a supermartingale. Now evaluating the

principal's payoff for this contract

$$E\left[\int_0^\tau e^{-rs}(dY_s - dI_s) + e^{-r\tau}L\right] =$$

$$E\left(G_{t^\wedge\tau}\right) + e^{-r\tau}E\left[1_{\{t\le\tau\}}\left(E_t\left(\int_t^\tau e^{-r(s-t)}(dY_s - dI_s) + e^{-r(\tau-t)}L\right) - V(W_t)\right)\right].$$

First, $E\left(G_{t^\wedge\tau}\right) \le G_0$ since G_t is a supermartingale. Then, $E_t\left(\int_t^\tau e^{-rs}(dY_s - dI_s) + e^{-rt}L\right) \le \frac{\mu}{r} - W_t$, since by the argument above, $rV(W) + \gamma W \le \mu$ for all W. Letting $t \to \infty$ implies that

$$E\left(\int_0^\tau e^{-rs}(dY_s - dI_s) + e^{-r\tau}L\right) \le V_0(W).$$

$\square$

Proof of Corollary 1: The relationship between $V(W)$ and $V^L(W)$ is fairly straightforward: if $V^L(W) > V(W)$ instead, then $V(W)$ cannot be the optimal value function for the principal since the contracting space with the commitment constraint is a strict subset of the contracting space without the constraint. The relationship between $\overline{W}$ and $\overline{W}^L$ follows that $rV(\overline{W}) + \gamma\overline{W} = \mu$ and $rV(\overline{W}^L) + \gamma\overline{W}^L \le \mu$ and the inequality is strict whenever $V(\overline{W}) < L\square$

Proof of Proposition 2: Since the high volatility state is assumed a absorbing state, the value function in such state follows directly from Proposition 1. The optimality conditions for the low volatility state can be proved in a very similar manner as that in Proposition 1. Differentiate the corresponding social benefit function with respect to W, substituting in the boundary conditions and evaluating the equation at the payment boundary $\overline{W}_L$ implies

$$F_l'''(\overline{W}_l) = \frac{(\gamma - r) + \left(\gamma\overline{W}_l - \pi_l\delta_l(\overline{W}_L)\right)F_l''(\overline{W}_l)}{\frac{1}{2}\lambda^2\sigma_l^2},$$

where $F_l''(W_l)$ is given by

$$F_l''(\overline{W}_l) = \frac{rF_l(\overline{W}_l) + (\gamma - r)\overline{W}_l - \mu + \pi_l\left(F_h(\overline{W}_l + \delta_l(\overline{W}_l)) - F_l(\overline{W}_l)\right)}{\frac{1}{2}\lambda^2\sigma_l^2}.$$

Piskorski and Tchistyi (2010) show the optimality conditions for the full commitment case. Under limited commitment, if the commitment constraint is not binding in the low volatility state, the proof is identical to theirs. Now suppose it is binding, which implies that it must also be binding in the high volatility state. Given the fact that $V_l^{L\prime}(\overline{W}_l^L) = V_h^{L\prime}(\overline{W}_h^L) = -1$, the slope matching procedure that pins down δ implies $\delta_l^L(\overline{W}_l^L) = \overline{W}_h^L - \overline{W}_l^L$. Given that $rF_l^L(\overline{W}_l^L) + \gamma\overline{W}_l^L < \mu$ from Corollary 1, if $\overline{W}_h^L < \overline{W}_l^L$, then $\delta_l^L(\overline{W}_l^L) < 0$,

and $F_l^{L'''}(\overline{W}_l^L) > 0$. If $\overline{W}_h^L > \overline{W}_l^L$, then $\gamma\overline{W}_l^L - \pi_l\delta_l^L(\overline{W}_l^L) > 0$ as $\pi_l < \frac{\gamma\overline{W}_l^L}{\delta_l^L(\overline{W}_l^L)}$. Since $\delta_l^L(\overline{W}_l^L) < \overline{W}_h^L < \frac{\mu-rL}{\gamma}$, $\gamma\overline{W}_l^L - \pi_l\delta_l^L(\overline{W}_l^L) > 0$ as long as $\pi_l < \frac{\overline{W}_l^L}{\mu-rL}$. Note that for a non trivial contract, $\overline{W}_l^L > R = 0$, there is always π_l small enough such that $\pi_l < \frac{R}{\mu-rL}$ is satisfied. The subsequent verification is similar to that used in Piskorski and Tchistyi (2010) thus is omitted here. $\square$

Proof of Corollary 2: From Corollary 1 and Proposition 2, $V_h(W) < V_l(W)$ and $V_h^L(W) < V_l^L(W)$ for all $W > R$. For the full commitment contract, $V_l''(\overline{W}_l) = V_h''(\overline{W}_h) = 0$ implies $rV_s(\overline{W}_s) + \gamma\overline{W}_s$ for $s = l, h$, then Corollary 1 implies $\overline{W}_l < \overline{W}_h$. For the limited commitment contract, $V_l^L(\overline{W}_l^L) = V_h^L(\overline{W}_h^L) = L$ and $V_s'(W) < 0$ near the payment boundary implies $\overline{W}_l^L > \overline{W}_h^L$ $\square$

Proof of Proposition 3: This proposition is proved in two steps. First, I show that both V_l' and V_h' are convex functions at the payment boundary. This conclusion utilizes the concavity of the value function which is true for both full commitment and limited commitment so only the former is shown. Differentiate the principal's HJB equation with respect to W, and substituting in $V_h'(W+\delta_l(W)) = V_h'(W)$, a condition that is always satisfied around the neighbourhood of the payment boundary because $V_s'(\overline{W}) = -1$ regardless of state and contract type. This yields

$$rV_s'(W) = (\gamma W - \pi_s\delta_s(W))V_s''(W) + \frac{1}{2}\lambda^2\sigma_s^2V_s'''(W) + (\gamma - \pi_s\delta_s'(W))V_s'(W).$$

Evaluating this at the payment boundary in the high volatility state yields

$$V_h'''(\overline{W}_h) = \frac{(\gamma - r) - \gamma\overline{W}_hV_h''(\overline{W}_h)}{\frac{1}{2}\lambda^2\sigma_h^2} > 0,$$

since $V_h''(\overline{W}_h) \leq 0$. Similarly,

$$V_l'''(\overline{W}_l) = \frac{(\gamma - r) - \gamma\overline{W}_lV_l''(\overline{W}_l) + \pi_lX(\overline{W}_l)}{\frac{1}{2}\lambda^2\sigma_l^2},$$

where

$$X(\overline{W}_l) = \delta_l'(\overline{W}_l)V_l'(\overline{W}_l) + \delta_l(\overline{W}_l)V_l''(\overline{W}_l).$$

Letting $\pi_L \to 0$ yields

$$V_l'''(\overline{W}_l) = \frac{(\gamma - r) - \gamma\overline{W}_lV_l''(\overline{W}_l)}{\frac{1}{2}\lambda^2\sigma_l^2} > 0.$$

Therefore $V_l'''(\overline{W}_l) > 0$ for small enough π_l, and both V_l' and V_h' are convex functions at the payment boundary. The same argument applies for the limited commitment contract.

The second step compares the variation of δ near the payment boundary. Consider the full commitment contract: by Corollary 1, $V_h(W) < V_l(W)$ and $V_h(R) = V_l(R) = L$ implies $V_h'(R) < V_l'(R)$. Since $V_h'(\overline{W}_h) = V_l'(\overline{W}_l) = -1$, there must exist $\widehat{W}$ such that $V_h'(\widehat{W}) = V_l'(\widehat{W})$. Moreover, $\overline{W}_h > \overline{W}_l$ by Corollary 2 and, because V_s' are convex functions, there is a unique $\widehat{W}$ after which $\delta_l(W) > 0$ for all $W > \widehat{W}$.

For the limited commitment contract, $V_h^{L\prime}(R) < V_l^{L\prime}(R)$, $V_h^{L\prime}(\overline{W}_h^L) = V_l^{L\prime}(\overline{W}_l^L) = -1$ and $\overline{W}_h^L < \overline{W}_l^L$ by Corollary 2 implies there exists $\widehat{W}^L$ such that $\delta_l^L(W) > 0$ for all $W > \widehat{W}^L$ Let $\widehat{W}$ be the largest between the two cut-offs for the full and limited commitment contract, and note that $\widehat{W} < \overline{W}_l^L$ since $\delta_l^L(\overline{W}_l^L) < 0$ and $\delta_l(\overline{W}_l) > 0$ proves this proposition.□

Proof of Corollary 3

Define $\Delta(W) = \left(\overline{W}_h - (W + \delta_l(W))\right) - \left(\overline{W}_l - W\right)$ as the difference between the distances to the payment boundary before and after the uncertainty shock for the full commitment contract, and $\Delta^L(W)$ as the same distance but for the limited commitment contract. Then $\Delta^L(W) - \Delta(W) = \left(\overline{W}_h^L - \overline{W}_h\right) - \left(\overline{W}_l^L - \overline{W}_l\right) - \left(\delta_l^L(W) - \delta_l(W)\right)$. For small π_l, $\overline{W}_l^L - \overline{W}_l$ is small. Therefore $\Delta^L(W) - \Delta(W) < 0$ as long as $\overline{W}_h^L - \delta_l^L(W) < \overline{W}_h - \delta_L(W)$. Notice that $\overline{W}_h - \overline{W}_h^L = \delta_L(\overline{W}_l) - \delta_l^L(\overline{W}_l^L)$, and $\delta_l'(\overline{W}_l) > 0$ while $\delta_l^{L\prime}(\overline{W}_l^L) < 0$ by Proposition 3. Therefore $\delta_l(W) - \delta_l^L(W) < \delta_l(\overline{W}_l) - \delta_l^L(\overline{W}_l^L) = \overline{W}_h - \overline{W}_h^L$ for any $W > \widehat{W}$. That is, $\Delta^L(W) - \Delta(W) < 0$ for all $W > \widehat{W}$.□

Proof of Proposition 4: Following Cox and Miller (1977), the transition density of the process W in the high variance state given initial value W_{t+} follows the Kolmogorov forward equation:

$$\frac{\partial}{\partial t} f(t, W; W_{t+}) = \frac{1}{2}\frac{\partial^2}{\partial W^2}\left[\lambda^2 \sigma_h^2 f(t, W; W_{t+})\right] - \frac{\partial}{\partial W}\left[\gamma W f(t, W; W_{t+})\right],$$

subject to the boundary conditions

$$f(t, 0; W_{t+}) = 0$$

$$\frac{1}{2}\frac{\partial}{\partial W}\left[\lambda^2 \sigma_h^2 f(t, W; W_{t+})\right]|_{W=\overline{W}_h} - \gamma \overline{W}_h f(t, \overline{W}_h; W_{t+}) = 0,$$

where f is a density function conditional on $W_{t+} = W$.

Define $\sigma^2 = \lambda^2 \sigma_h^2$ as the overall variance of the W process. Let f_γ be the solution to this boundary value problem for a particular γ. According to Ward and Glynn (2003), when γ

is closer to zero, f_γ can be approximated by

$$f_\gamma(t, W; W_{t+}) = k(\gamma)g(t, W; W_{t+}) + o(\gamma) , \tag{12}$$

where $k(\gamma) = \left(1 - \frac{\gamma}{2\sigma^2}W_{t+}^2 + \frac{\gamma}{2\sigma^2}W^2 + \frac{\gamma}{2}t\right)$ and g is the corresponding transition density function for the same process but with $\gamma = 0$.

Now the problem becomes a Brownian motion between an absorbing and a reflecting barrier. In particular, $g(t, W; W_{t+})$ satisfies the differential equation:

$$\frac{\partial}{\partial t}g(t, W; W_{t+}) = \frac{1}{2}\frac{\partial^2}{\partial W^2}\left[\sigma^2 g(t, W; W_{t+})\right] ,$$

subject to boundary conditions $g(t, R; W_{t+}) = 0$, $\frac{1}{2}\sigma^2 \frac{\partial}{\partial W}[g(t, W; W_{t+})]|_{W=\overline{W}_h} = 0$.

The solution to this problem has been derived by Schwarz (1992) as

$$g(W, t) = \sum_{n=1}^{\infty} A_n \exp\left(-\alpha_n^2 \frac{1}{2}\sigma^2 t\right) \cos(\alpha_n W) ,$$

where $\alpha_n = \frac{(2n-1)\pi}{2\overline{W}_h}$ and $A_n = \frac{\cos(\alpha_n W_{t+})}{\overline{W}_h}$.

Substituting this into the approximation function (12) yields $f(W, t)$ which can be used in the definition of the expected local time at the payment boundary

$$E\left[L_h(T; W_{t+})\right] = \lim_{\varepsilon \to 0}\frac{1}{2\varepsilon}\int_0^T dt \int_{\overline{W}_h-\varepsilon}^{\overline{W}_h+\varepsilon} f(t, W; W_{t+})dW$$

Fixed some $W_{t+} < \overline{W}_h^L$, Let

$$E[\mathcal{L}_h(T; W_{t+})] \equiv E[\mathcal{L}_h(T; W_{t+})|\overline{W}_h]$$

be the expected local time given the full commitment value functions and payment boundaries, and define

$$E^L[\mathcal{L}_h(T; W_{t+})] \equiv E[\mathcal{L}_h(T; W_{t+})|\overline{W}_h^L]$$

be the expectation of local time at the payment boundary under the limited commitment contract. First, $\frac{\partial}{\partial T}E[\mathcal{L}_h(T; W_{t+})]|_{T=0} > 0$, that is, the expected time spend at one point is longer whenever the time interval is longer, in particular when the time interval expands a small amount from 0. Secondly, such derivative is larger for smaller $\overline{W}_h$ because for a fixed W, $f(W, t)$ is decreasing in $\overline{W}_h$. The effect of expanding the time interval is bigger, the shorter distance between W_{t+} and the reflecting boundary is. Note that in the case of $\sigma \gg \gamma$, the approximation adjustment term $h(\gamma)$ is close to one if W and W_{t+} are near each

other, this implies the most precise approximation is around the payment boundary, exactly the target of the analysis given here.

From Corollary 2, $\overline{W}_h > \overline{W}_h^L$, and $E^L[\mathcal{L}_h(0; W_{t+})] = E[\mathcal{L}_h(0; W_{t+})] = 0$ implies

$$E^L\left[\mathcal{L}_h(T; W_{t+})\right] > E\left[\mathcal{L}_h(T; W_{t+})\right] \ , \ as \ T \to 0 \ .$$

The expected local time grows faster for closer reflecting boundary near $T = 0$. Also

$$E^L\left[\mathcal{L}_h(T; W_{t+})\right] < E\left[\mathcal{L}_h(T; W_{t+})\right] \ , \ as \ T \to \infty \ ,$$

which implies there is some $\widehat{T}$ such that

$$E^L\left[\mathcal{L}_h(\widehat{T}; W_{t+})\right] = E\left[\mathcal{L}_h(\widehat{T}; W_{t+})\right] \ ,$$

and

$$E^L\left[\mathcal{L}_h(T; W_{t+})\right] > E\left[\mathcal{L}_h(T; W_{t+})\right] \ , for \ all \ 0 < T < \widehat{T} \ .$$

Finally, notice that given $\overline{W}_h$, $E\left[L_h(T; W_{t+})\right]$ is decreasing in W_{t+}, that is, the further W_{t+} is from the reflecting barrier, the less time it spends there within a certain time. Therefore $E^L\left[\mathcal{L}_h(T; W_{t+}^L)\right] > E\left[\mathcal{L}_h(T; W_{t+})\right]$ as long as $\overline{W}_h^L - W_{t+}^L < \overline{W}_h - W_{t+}$. By Corollary 3 $\overline{W}_l^L - W_{t+}^L < \overline{W}_h - W_{t+}$, if $W_{t-} > \widehat{W}$, therefore $E^L\left[\mathcal{L}_h(T; W_{t-})\right] > E\left[\mathcal{L}_h(T; W_{t-})\right]$ for all $0 < T < \widehat{T}$ as long as $W_{t-} > \widehat{W}\square$

Proof of Proposition 5: Consider the process of W in the high volatility state with initial position W_{t+}. Let N be the number of times W reaches the reflecting boundary $\overline{W}_h$ before it is stopped. Then

$$E\left[\tau\right] = \sum_{i=0}^{\infty} E\left[\tau, N = i\right] \ .$$

First consider $N \geq 1$, if W reaches $\overline{W}_h$ at least once before it is stopped, then starting from $\overline{W}_h$, the expected stopping time is smaller whenever $\overline{W}_h - R$ is a shorter interval. Next consider the case $M = 0$, the expected stopping time is smaller whenever W_{t+} is closer to R. Finally, the average speed of growth of W, γW, is slower for smaller W. From Corollary 2 and 3 it can be concluded that $E^L\left[\tau_h\right] < E\left[\tau_h\right]$ because $\overline{W}_h^L < \overline{W}_h$ and $W_{t+}^L < W_{t+}$ for the same W_{t-}.

Same comparison can be conducted between $E^L\left[\tau_h\right]$ and $E^L\left[\tau_l\right]$. The expected stopping time is smaller when $\overline{W}$ and the initial W is closer to R, and when σ is larger.

The exact value of $E\left[\tau\right]$ is difficult to compute due to the irregular process W follows. However, when γ is small, the same approximation method used in the proof of Proposition 4 can be applied here as well. The problem thus becomes a standard absorbing time question

for a Brownian motion between an absorbing and a reflecting barrier, whose solution is given by Cox and Miller (1977) as

$$E\left[\tau\right] = \frac{W_{t^+}(2\overline{W}_h - W_{t^+})}{\sigma^2}$$

This solution confirms that $E\left[\tau\right]$ is positively related to $\overline{W}_h$ and W_{t^+} while negatively related to σ. Since $\overline{W}_h^L < \overline{W}_l^L < \overline{W}_h$, $W_{t^+}^L < W_{t^-}^L < W_{t^+}$, and $\sigma_h > \sigma_l$, $E^L\left[\tau_h\right]$ must be the smallest compare to $E\left[\tau_h\right]$ and $E^L\left[\tau_l\right]$ $\square$

Proof of Proposition 6:

Without loss of generality, assume that the interest rate of the credit line is γ. Begin with the high volatility state σ_h, Then the credit line balance evolves according to:

$$dM_t = \gamma M_t dt + xdt + dDiv_t - dY_t \tag{13}$$

where Div_t represents the cumulative dividends paid by the firm and x is the consol bond rate. Using the fact that $x = rD_t$ and substituting that into equation (13) implies:

$$dWt = -\lambda dM_t = -\lambda\gamma Mtdt - \lambda xdt - \lambda dDiv_t + \lambda dY_t$$

$$= \gamma W_t dt - \lambda dI_t + \lambda(dY_t - \mu dt)$$

satisfying incentive compatibility. The argument for state σ_L can be made analogously subject to a jump δ_L, whose value is pinned by the matching first order derivatives procedure$\square$

Proof of Proposition 7:

This Proposition is a natural extension given the proofs of Proposition 1 and 2. The exact conditions under which shirking forever is optimal can be found in Zhu (2013)$\square$

Appendix B. Recurring States

In the main body of the paper I assume that the transition probability from high to low uncertainty state π_h is zero, that is the crisis state is absorbing. This assumption greatly simplifies the verification of the optimality of the contract provided by Proposition 2, but is unnecessarily for the results of this paper to hold. In this appendix I provide a full characterization of the optimal contract when I relax such assumption. That is, when $\pi_h > 0$ and the economy switches between normal and crisis times stochastically. The following proposition summarizes the result:

Proposition 8. *Suppose $\pi_l > 0$ and $\pi_h > 0$. Let N_t be the total number of state transitions at time t. The agent's continuation utility W_t follows*

$$dW_t = \gamma W_t - dI_t + \lambda(dY_t - \mu dt) + \delta_t(dN_t - \pi_t dt); . \tag{14}$$

The optimal contract is a pair of value functions $V_s(W)$ and payment boundaries $\overline{W}_s$, $s \in \{l, h\}$ such that

$$
\begin{aligned}
rV_s(W) &= \mu + (\gamma W - \pi_s \delta_s(W)) V_s'(W) + \frac{1}{2}\lambda^2 \sigma_s^2 V_s''(W) \\
&\quad + \pi_s \left(V_{\hat{s}}(W + \delta_s(W)) - V_s(W) \right) ,
\end{aligned} \tag{15}
$$

subject to boundary conditions $V_s(R) = L$; $V_s'(\overline{W}_s) = -1$; and

$$V_s''(\overline{W}_s) = 0 ,$$

where $\delta_s(W)$ follows (7) and (8). If the principal has only limited commitment, the optimal contract is a pair of value functions $V_s^L(W)$ and payment boundaries $\overline{W}_s^L$, $s \in \{l, h\}$, such that $V_s^L(W)$ satisfies the same system of ODE (15) and boundary conditions $V_s^L(R) = L$; $V_s^{L\prime}(\overline{W}_s^L) = -1$, and

$$V_s^{L\prime\prime}(\overline{W}_s^L) = 0, \text{ if } V_s^L(\overline{W}_s^L) \geq L ,$$

$$V_s^L(\overline{W}_s^L) = L, \text{ otherwise} .$$

Proof: the proof builds on iteration procedure described in Li (2012). I therefore only sketch the argument here in the interest of space. Consider first the case of full commitment. Applying the martingale method of Sannikov (2008), the agent's continuation utility follows (14). Ito's lemma implies that the principal's HJB equation satisfies (6). Let $\widetilde{V}_s(W)$ be a solution to (6). The concavity of $\widetilde{V}_s(W)$ can be shown using the method similar to Proposition 2. Take $\widetilde{V}_l(W)$ as given, define an auxiliary value function U_h^S as the payoff assuming the principal ceases to provide any incentive to the agent in the high volatility state until the next volatility shock arrives. The concavity of $\widetilde{V}_s(W)$ implies that $\widetilde{V}_h(W) > U_h^S$. Apply the similar argument to $\widetilde{V}_l(W)$ but take $\widetilde{V}_h(W)$ as given, Li (2012) shows that the procedure converges to a pair of function $V_s(W)$ satisfying equation (15). Finally, the same procedure also applied to the limited commitment contract as long as $V_s(W)$ remain concave, which is shown in Proposition 1 by replacing the $V_h(W)$ with $\widetilde{V}_h(W)$ in its proof.□

The optimal contract characterized under recurring state is qualitatively identical to the one summarized in Proposition 2 under a one-time shock. In fact, principal's value functions of the contract under recurring states converge to value functions under a one-time shock when $\pi_h \to 0$. Given π_s are assumed to be small numbers the case of a one-time shock

provides a good approximation for the general case of recurring states and does not lose any important result.

All the remaining results discussed in the main body regarding the position of payment boundaries, the frequency of cash payment and expected termination time are preserved in the recurring state contract, as long as the parameters are that once the limited commitment constraint is imposed, it is binding in both states. The discussion of "pay-for-luck" can be expanded to not only negative shocks but also positive shocks. The following result can be inferred from Proposition 3: Under the full commitment contract, managers whose accumulated performance is well enough receive less utility when volatility decreases; meanwhile managers under the limited commitment contract receive higher utility. The conclusion for limited commitment contract is consistent with empirical findings of "pay-for-luck" which further reinforce the importance of taking firms' commitment ability into account when understanding compensation under shocks.

Appendix C. Contracts with Limited Commitment Binding in One State Only

Section 3.1 introduced three types of contracts based on when the limited commitment constraint is binding. While the main text focuses on the third types, here I also provide some discussions of the second type: the contract where the limited commitment constraint is binding only in the high volatility state. In general, this type of contract can behave like contracts with either full commitment or those with limited commitment but the commitment constraint is binding in both states, depending on the parameter value σ in each state.

The main goal of this appendix is to establish conditions under which the main Propositions in Section 4 are still valid for the optimal contract when the limited commitment constraint is imposed. The proofs of Proposition 4 reveal that the key variable driving the dynamics of compensation is the distance between W_{t+} and payment boundary $\overline{W}_h$. This leads to the conjecture that the dynamics of compensation of the type of contract discussed in this appendix section will be similar to the dynamics of the limited commitment contracts described in Section 4 as long as when the commitment constraint is imposed, the agent's continuation utility W_{t+} is closer to the payment boundary $\overline{W}_h^L$ compared to the full commitment case. Due to the implicit form of the value function, it is analytically difficult to characterize the exact range of parameters under which this conjecture is true. Nevertheless the following proposition gives one sufficient condition for it.

Proposition 9. *If the commitment constraint is binding only in the high variance state, then there exist $\widehat{W}$ such that $W_{t+}^L - \overline{W}_h^L < W_{t+} - \overline{W}_h$ for all $W_{t-} > \widehat{W}$ as long as $\overline{W}_h < \overline{W}_l$.*

Proof: Similar to the proof of Proposition 3 and Corollary 3, consider the full commitment contract first. $V_h(W) < V_L(W)$ implies $\overline{W}_h > \overline{W}_l$ and $V_h'(R) < V_l'(R)$. Since $V_h'(\overline{W}_h) = V_l'(\overline{W}_l) = -1$, there must exist $\widehat{W}$ such that $V_h'(\widehat{W}) = V_l'(\widehat{W})$ and $\delta_l(W) > 0$ for all $W > \widehat{W}$.

Next, if the limited commitment constraint is imposed and binding, $V_h^{L\prime}(R) < V_l^{L\prime}(R)$, $V_h^{L\prime}(\overline{W}_h^L) = V_l^{L\prime}(\overline{W}_l^L) = -1$. If $\overline{W}_h^L < \overline{W}_l^L$, then there exists $\widehat{W}^L$ such that $\delta_h^L(W) > 0$ for all $W > \widehat{W}$. Let $\widehat{W} = \max\left\{\widehat{W}, \widehat{W}^L\right\}$, then $\delta_l'(\overline{W}_l) > 0$ while $\delta_l^{L\prime}(\overline{W}_l^L) < 0$ for all $W > \widehat{W}$. Finally, define $\Delta^L(W_{t-}) - \Delta(W_{t-}) = \left(\overline{W}_h^L - \overline{W}_h\right) - \left(\overline{W}_l^L - \overline{W}_l\right) - \left(\delta_l^L(W_{t-}) - \delta_l(W_{t-})\right)$. For small π_l, $\overline{W}_l^L - \overline{W}_l$ is small. Notice that $\overline{W}_h - \overline{W}_h^L = \delta_l(\overline{W}_l) - \delta_l^L(\overline{W}_l^L)$, $\delta_l'(\overline{W}_l) > 0$ and $\delta_l^{L\prime}(\overline{W}_l^L) < 0$ implies $\delta_l = (W_{t-}) - \delta_l^L(W_{t-}) < \delta_l = (\overline{W}_l) - \delta_l^L(\overline{W}_l^L) = \overline{W}_h = -\overline{W}_h^L$ for any $W_{t-} > \widehat{W}$. Therefore, $\Delta^L(W_{t-}) - \Delta(W_{t-}) < 0$ for all $W_{t-} > \widehat{W}$.□

Given the sufficient condition above, the rest of the analysis follows exactly the one shown in the main text. Figure 4 demonstrate the difference between two levels of volatility in the high volatility state. For the same level of σ_l, the relative position of $\overline{W}_l$ and $\overline{W}_h^L$ are similar to the full commitment case when σ_h is moderate, but converge to the case in which the commitment constraint is binding in both states when σ_h becomes high enough.

Panel A. Low σ_h Panel B. High σ_h
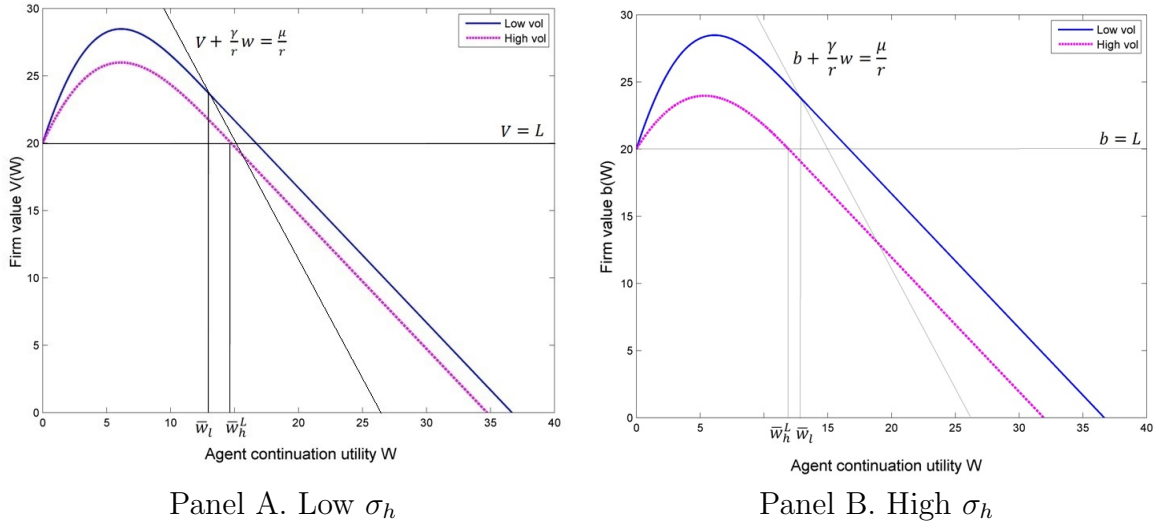
Figure 4 – Contracts with the Commitment Constraint Binding in One State Only
This figure plots firm value functions when the limited commitment constraint is binding only in the high volatility state. Parameter values are the same as those in Figure 1 except $\sigma_l = 5$ and $\sigma_h = 6$ for the left panel, and $\sigma_h = 6.5$ for the right panel

The finding of this section greatly expands the domain of contracts to which Propositions 4 and 5 can apply. Large bonuses in crisis times could be possible if the abrupt volatility increase is substantial enough that many firms that operate smoothly during normal times suddenly become constrained in the amount they can credibly pledge to pay their managers

in the long-run. The greater increase of market risks during the crises, the more severe is this concern. Future research that calibrates or empirically investigates the real scope of this commitment constraint will be helpful in determining the proportion of firms that are subject to limited commitment contracts and firms whose dynamics of bonuses follow the predictions in this paper.

Appendix D. Renegotiation-Proof Contracts

A renegotiation-proof contract requires the slope of the principal's value function to be non-positive. Such condition is ruled out in the main context of this paper because $V^L(R) = V^L(\overline{W}^L) = L$ when the limited commitment constraint binds, hence a non-trivial contract must have a region where the principal's valuation is increasing in the agent's continuation value W. To allow renegotiation-proof contracts I modify the assumption about the principal's commitment ability. I assume that, instead of having the option to liquidate the project any time during the contract, the principal will only withdraw the investment when firm values is below zero. This assumption is similar to the one made by Ai and Li (2013). The corresponding constraint on the payment boundary is now:

$$V(\overline{W}) \geq 0 .$$

The dynamics of the agent's continuation value under renegotiation-proof contracts follow

$$dW_t = \gamma W_t - dI_t + \lambda(dY_t - \mu dt) + \delta_t(dN_t - \pi_t dt) + dP_t .$$

The new term dP_t defines a reflecting termination boundary $\underline{W}$ which satisfies the boundary condition $V(\underline{W}) = L$ and $V'(\underline{W}) = 0$. Termination is stochastic at this boundary, with probability $\frac{dP_t}{\underline{W} - R}$ to account for the extra term on the agent's continuation value and keep the contract incentive compatible.

For the main results of this paper to carry through, it is sufficient to prove the following proposition:

Proposition 10. *Under the renegotiation-proof contract, $\overline{W}_h > \overline{W}_l$ under full commitment and $\overline{W}_h^L < \overline{W}_l^L$ when the constraint is binding in both states.*

Proof: Clearly, Corollary 1 and 2 still apply to renegotiation-proof contracts. Therefore $\overline{W}_h^L < \overline{W}_l^L$ since $V_h(W) < V_l(W)$ and $V(\overline{W}) = 0$ if the commitment constraint is binding in both states. Without the constraint the contract is a standard continuous-time dynamic contract with regime switching, and the argument of the boundary positions can be found in Hoffmann and Pfeil (2010)□

Given the relative positions of the payment boundaries for the full commitment and limited commitment contracts, one can easily see that a statement similar to Corollary 3 can be made here as well. Figure 5 shows the value functions for the renegotiation-proof contracts, where both the endogenous renegotiation boundaries as well as the payment boundaries are displayed. The graphs confirms Propositions regarding the relative positions of payment boundaries for both the full and limited commitment contracts. Such conclusions leads to the same dynamics of bonuses payment described in Section 3 and the details are thus omitted here.

Panel A. Full Commitment   Panel B. Limited Commitment
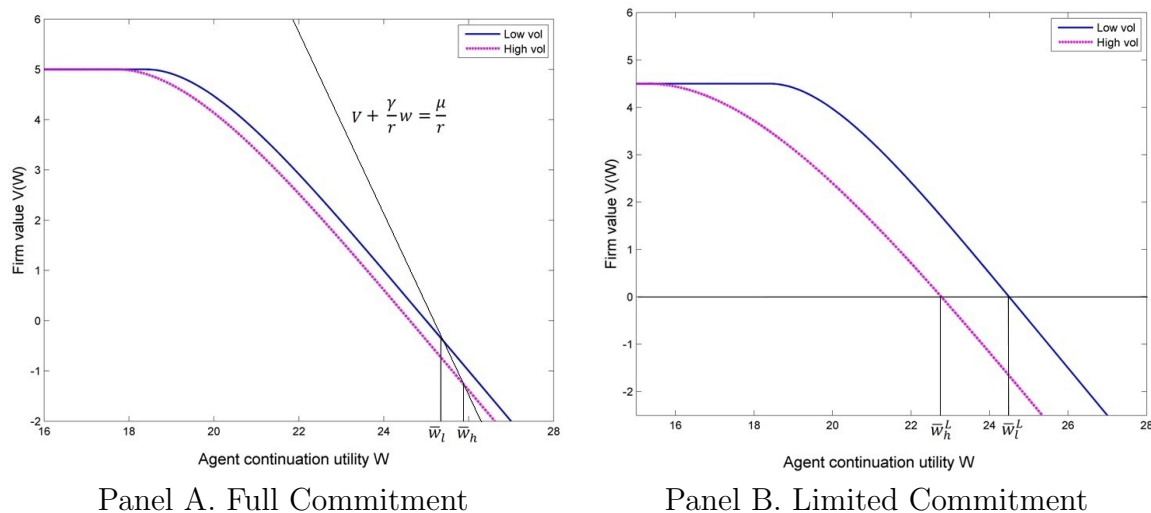
Figure 5 – Renegotiation-Proof Contracts

Last but not least, renegotiation-proofness is not a necessary feature for the contract to be optimal despite limited commitment. The principal is still able to rule out further renegotiation since the only action she cannot commit to is not to withdraw when firm value is negative. In particular, the principal can commit to the random termination schedule described above, which is crucial in keeping the manager's incentive properly. Further, the assumption of investors withdrawing their investment when firm value drops below zero replaces the earlier assumption of liquidation at any time, and therefore the value of the firm at the termination boundary is still the liquidation value since it is determined by the agent's effective limited liability constraint and the principal is able to commit to termination once that boundary is reached.

References

Abreu, D., Pearce, D., and Stacchetti, E. 1990. Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica*, 58:1041–1063.

Aggarwal, R. and Samwick, A. 1999. The other side of the trade-off: The impact of risk on executive compensation. *Journal of Political Economy*, 107:65–105.

Ai, H. and Li, R. 2013. Investment and CEO compensation under limited commitment. Working Paper. University of Minnesota and Purdue University.

Atkeson, A. 1991. International lending with moral hazard and risk of repudiation. *Econometrica*, 59:1069–1089.

Atkeson, A., Eisfeldt, A., and Weill, P.-O. 2013. Measuring the financial soundness of us firms 1926-2012. Working paper. University of California-Los Angeles.

Axelson, U. and Baliga, S. 2009. Liquidity and manipulation of executive compensation schemes. *Review of Financial Studies*, 22:3907–3939.

Axelson, U. and Bond, P. 2012. Wall street occupations: An equilibrium theory of overpaid jobs. Working paper. London School of Economics and University of Minnesota.

Bebchuk, L. A. and Fried, J. M. 2006. *Pay without performance: The unfulfilled promise of executive compensation*. Harvard University Press.

Bénabou, R. and Tirole, J. 2013. Bonus culture: Competitive pay, screening, and multitasking. Working paper. Princeton University and Toulouse School of Economics.

Berk, J., Stanton, R., and Zechnar, J. 2010. Human capital, bankruptcy, and capital structure. *Journal of Finance*, 65:891–926.

Bertrand, M. and Mullainathan, S. 2001. Are CEOs rewarded for luck? the ones without principals are. *Quarterly Journal of Economics*, 116:901–932.

Biais, B., Mariotti, T., Rochet, J.-C. R., and Villeneuve, S. 2010. Large risks, limited liability, and dynamic moral hazard. *Econometrica*, 78:73–118.

Bijlsma, M., Boone, J., and Zwart, G. 2012. Competition for traders and risk. Working paper, Tilburg University.

Bloom, N. 2009. The impact of uncertainty shocks. *Econometrica*, 77:623–685.

Bolton, P., Chen, H., and Wang, N. 2011. A unified theory of tobin's q, corporate investment, financing, and risk management. *Journal of Finance*, 66:1545–1578.

Bolton, P., Chen, H., and Wang, N. 2013. Market timing, investment, and risk management. *Journal of Financial Economics*, 109:40–62.

Bond, P. and Glode, V. 2013. Bankers and regulators. Working paper, University of Washington and University of Pennsylvania.

Brunnermeier, M. and Sannikov, Y. 2012. A macroeconomic model with a financial sector. Working paper, Princeton University.

Campello, M., Giambona, E., Graham, J., and Harvey, C. 2011. Liquidity management and corporate investment during a financial crisis. *Review of Financial Studies*, 24:1944–1979.

Core, J. and Guay, W. 2002. The other side of the trade-off: The impact of risk on executive compensation: A revised comment. Working Paper, University of Pennsylvania.

Cox, D. and Miller, H. 1977. *The theory of stochastic processes*. Chapman and Hall.

Cvitanić, J., Possamaï, D., and Touzi, N. 2014. Moral hazard in dynamic risk management. Working Paper, Caltech, University Paris-Dauphine, and École Polytechnique.

DeMarzo, P., Fishman, M., He, Z., and Wang, N. 2012. Dynamic agency and the q theory of investment. *Journal of Finance*, 67.

DeMarzo, P. and Sannikov, Y. 2006. Optimal security design and dynamic capital structure in a continuous-time agency model. *Journal of Finance*, 61:2681–2724.

Di Tella, S. 2013. Uncertainty shocks and balance sheet recessions. Working paper. Stanford University.

Eisfeldt, A. and Kuhnen, C. 2013. CEO turnover in a competitive assignment framework. *Journal of Financial Economics*, 109:351–372.

Ellul, A., Pagano, M., and Schivardi, F. 2014. Employment and wage insurance within firms: Worldwide evidence. Working paper. Indiana University, University of Naples Federico II, and LUISS.

Frydman, C. and Jenter, D. 2010. CEO compensation. *Annual Review of Financial Economics*, 2:75–102.

Glode, V., Green, R. C., and Lowery, R. 2012. Financial expertise as an arms race. *Journal of Finance*, 67:1723–1759.

Glode, V. and Lowery, R. 2013. Informed trading and high compensation in finance. Working Paper, University of Pennsylvania and University of Texas-Austin.

Grochulski, B. and Zhang, Y. 2011. Optimal risk sharing and borrowing constraints in a continuous-time model with limited commitment. *Journal of Economic Theory*, 146:2356–2388.

Grochulski, B. and Zhang, Y. 2013. Market-based incentives. Working paper. Federal Reserve Bank of Richmond and University of Texas-Austin.

He, Z. 2009. Optimal executive compensation when firm size follows geometric brownian motion. *Review of Financial Studies*, 22:859–892.

He, Z. and Krishnamurthy, A. 2013. Intermediary asset pricing. *American Economic Review*, 103:732–770.

He, Z. and Xiong, W. 2012. Rollover risk and credit risk. *Journal of Finance*, 67:391–430.

Hoffmann, F. and Pfeil, S. 2010. Reward for luck in a dynamic agency model. *Review of Financial Studies*, 23:3329–3345.

Holmstrom, B. 1982. Moral hazard in teams. *Bell Journal of Economics*, 13:324–340.

Jenter, D. and Kanaan, F. 2010. CEO turnover and relative performance evaluation. Working paper. Stanford University.

Kaplan, S. 2012. Executive compensation and corporate governance in the U.S.: Perceptions, facts and challenges. Working paper. University of Chicago.

Kaplan, S. and Minton, B. 2012. How has CEO turnover changed. *International Review of Finance*, 12:57–87.

Kaplan, S. and Rauh, J. 2010. Wall street and main street: What contributes to the rise in the highest incomes. *Review of Financial Studies*, 23:1004–1050.

Lambert, R. and Larcker, D. 1987. An analysis of the use of accounting and market measures of performance in executive compensation contracts. *Journal of Accounting Research*, 25:85–125.

Leland, H. E. 1994. Corporate debt value, bond covenants, and optimal capital structure. *Journal of Finance*, 49:1213–1252.

Leland, H. E. and Toft, K. B. 1996. Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *Journal of Finance*, 51:987–1019.

Levin, J. 2003. Relational incentive contracts. *American Economic Review*, 93:835–857.

Li, R. 2012. Optimal contract design over the business cycle. Working paper, Purdue University.

Lioui, A. and Poncet, P. 2014. Optimal benchmarking for active portfolio managers. *European Journal of Operational Research*, 226::268–276.

Miao, J. and Zhang, Y. 2013. A duality approach to continuous-time contracting problems with limited commitment. Working paper, Boston University and University of Texas-Austin.

Opp, M. and Zhu, J. 2013. Impatience vs. incentives. Working paper, University of California-Berkeley and University of Pennsylvania.

Ou-yang, H. 2003. Optimal contracts in a continuous-time delegated portfolio management problem. *Review of Financial Studies*, 16:173–208.

Palacios, M. and Stomper, A. 2014. Training the doubtful and timid. Working paper, Vanderbilt University and Humboldt University.

Peters, F. and Wagner, A. 2014. The executive turnover risk premium. *Journal of Finance*, 69(4):1529–1563.

Philippon, T. and Reshef, A. 2012. Wages and human capital in the u.s. financial industry: 1909-2006. *Quarterly Journal of Economics*, 127:1551–1609.

Piskorski, T. and Tchistyi, A. 2010. Optimal mortgage design. *Review of Financial Studies*, 23:3098–3140.

Prendergast, C. 2002. The tenuous trade-off between risk and incentives. *Journal of Political Economy*, 110:1071–1102.

Rampini, A. and Viswanathan, S. 2013. Collateral and capital structure. *Journal of Financial Economics*, 109:466–492.

Ray, D. 2002. The time structure of self-enforcing agreements. *Econometrica*, 70:547–582.

Sannikov, Y. 2008. A continuous-time version of the principal-agent problem. *Review of Economic Studies*, 75:957–984.

Schwarz, W. 1992. The wiener process between a reflecting and an absorbing barrier. *Journal of Applied Probability*, 65:597–604.

Thanassoulis, J. 2012. The case for intervening in bankers' pay. *Journal of Finance*, 67:849–895.

Thomas, J. and Worrall, T. 1988. Self-enforcing wage contracts. *Review of Economic Studies*, 61:541–554.

Ward, A. and Glynn, P. 2003. Properties of the reflected ornstein–uhlenbeck process. *Queueing Systems*, 44:109–123.

Zhu, J. Y. 2013. Optimal contracts with shirking. *Review of Economic Studies*, 80:812–839.