

Managerial Short-Termism, Turnover Policy, and the Dynamics of Incentives

Felipe Varas

Fuqua School of Business, Duke University

I study managerial short-termism in a dynamic model of project development with hidden effort and imperfect observability of quality. The manager can complete the project faster by reducing quality. To preempt this behavior, the principal makes payments contingent on long-term outcomes. I analyze the dynamics of the optimal contract and its implications for the level of managerial turnover. I show that optimal contracts might be stationary and entail no termination. In general, I show that the principal reduces the manager's temptation to behave myopically by reducing the likelihood of termination and deferring compensation. The model predicts a negative relation between the rate of managerial turnover and the use of deferred compensation that is consistent with evidence of managerial compensation contracts. (*JEL* C73, D86, G39, J33, M52)

Received May 23, 2016; editorial decision May 27, 2017 by Editor Francesca Cornelli.

A substantial amount of the literature starting with Stiglitz and Weiss (1983) and Bolton and Scharfstein (1990) shows that the threat of termination can be effective at providing incentives. Budget constraints, short-term financing, and deadlines are powerful tools to incentivize effort. However, the evidence shows us that this is not a panacea because these incentives encourage myopic behavior: a CEO can launch a product before it is ready, in order to increase short-term profits, and a research team can take shortcuts to complete a project in time and on budget. The main purpose of this paper is to study the effect of short-termism on the time structure of incentives, with special emphasis on the role of termination and turnover.

I consider a dynamic model of project development in which a manager exerts effort to complete a project that can be finished faster by reducing its quality. Reducing quality allows the manager to increase short-term performance, so there is a trade-off between the maximization of short- and long-term performance. For example, a struggling CEO may accelerate the development

This paper was previously circulated as "Contracting Timely Delivery with Hard-to-Verify Quality." I am extremely grateful to my advisers Peter DeMarzo and Andy Skrzypacz for numerous discussions and suggestions and two anonymous referees. I also thank Darrell Duffie, Felipe Aldunate, Manuel Amador, Simon Gervais, Paul Pfleiderer, Kristoffer Laursen, Dirk Jenter, Sebastian Infante, Ivan Marinovic, Monika Piazzesi, Martin Schneider, and Jeffrey Zwiebel for their helpful comments. Send correspondence to Felipe Varas, Duke University, 100 Fuqua Rd, Durham, NC 27708; telephone (919) 660 2920. E-mail: felipe.varas@duke.edu.

© The Author 2017. Published by Oxford University Press on behalf of The Society for Financial Studies.

All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.

doi:10.1093/rfs/hhx088

Advance Access publication August 2, 2017

of a new product or project to increase profits, as was the case in Ford's Pinto scandal in the 1960s.¹ A similar incentive problem is common in capital budgeting: managers with tight budgets have less incentive to waste resources but also more incentive to cut corners to finish the project on time and on budget.^{2,3}

The analysis of managerial short-termism is involved because of the persistent effect of short-termism, and this makes the analysis of long-term contracts and turnover challenging. The project development problem analyzed here is particularly tractable: in the absence of short-termism, the principal punishes delays by reducing the payment to the manager upon project completion and terminates the project if the manager fails to deliver before a pre-specified deadline. However, punishing low performance in this way is suboptimal once we introduce the possibility of managerial short-termism because attempts to punish the manager for low performance increase the incentives to engage in short-termism. In fact, when we reduce compensation, we also reduce the manager's skin-in-the-game, thereby stimulating short-termism. Hence, the principal relies less on the use of dynamic incentives and the contract may become stationary. This last result is analogous to previous results on the linearity of incentives in static multitasking models (Holmström and Milgrom 1987); but, in the case of managerial short-termism, it is a form of linearity in time. This is consistent with the ideas in Jensen (2001, 2003), who has criticized the use of compensation systems and budgeting processes that introduce nonlinearities over performance and over time.

The optimal contract is a combination of a dynamic, nonstationary phase followed by a stationary phase: the principal relies on dynamic incentives when the manager's rents are high—namely, when the manager has more skin-in-the-game—and the contract is stationary when their rents are low, in which case punishing the manager for low performance would induce myopic behavior. In the dynamic phase, the manager's payment is reduced after delays in completing the project, while in the stationary phase, the manager is no longer punished for delays and the terms of the contract remain constant over time. In this latter phase, the contract is stationary and incentives are provided by the threat of termination. In fact, I show that sometimes the optimal contract is completely stationary: in this extreme case, the manager is neither terminated nor punished

¹ In the late 1960s, Ford Motor Company faced strong competition from foreign producers selling small, fuel-efficient cars. The CEO for Ford Motor Company announced the challenging goal of producing a new car that would be competitive in this market and rushed the Pinto into production in less than the usual time. In doing so, they neglected many safety checks, a misstep resulting in a defective fuel system that could ignite on collision (see Dowie 1977 for more information on this case).

² For example, after the tragic *Columbia* space shuttle accident, the Department of Energy concluded that evaluation systems intended to measure worker performance against deadlines may have pressured workers who resorted to using shortcuts to complete the work more quickly US Department of Energy (2005, p. 9).

³ Similar examples can be found in Paté-Cornell (1990) and Austin (2001), who document that quality problems in software development projects are often associated with time pressure and tight development schedules.

for low performance and the optimal contract is given by the repetition of a static contract.

In the stationary phase, the contract is asymmetric: it rewards success but does not punish failure. A naive observer could interpret this as evidence that the manager is entrenched like in Bebchuk (2009). Indeed, following a sequence of periods with low performance, a manager remains in the company and their long-term compensation plan is not affected negatively by past performance; this feature is a natural response to the possibility of managerial short-termism.

I also study the effect of short-termism on the evolution of effort. In the absence of short-termism, the optimal contract frontloads effort, an effort that decreases over time; this is the natural consequence of punishing the manager for low performance and decreasing their reward over time. However, this is not always the case with managerial short-termism: in the stationary region, effort is constant, whereas in the nonstationary region, effort front-loading decreases when short-term manipulation is more difficult to detect.

The principal relies on deferred compensation because the negative effects of a bad project take a long time to materialize: to prevent short-termism, compensation is subject to clawbacks. Because the manager's incentive to undertake low-quality projects is stronger if the promised compensation is very low, pay duration is negatively correlated with the value of the compensation plan, and so the compensation of high-performing managers vests sooner. This happens because a manager with a valuable compensation plan has more skin-in-the-game, and so has less incentive to behave myopically. This suggests that vesting of long-term compensation plans should be contingent on long-term performance measures and positively correlated with the level of the overall compensation plan, which resembles some aspects of performance shares that are used in many compensation plans (performance shares are restricted shares in which vesting is contingent on long-term goals).

The optimal contract features random termination. The randomness in the contract represents the uncertainty in the mind of the manager about the termination date of the project. For example, random termination can be interpreted as a form of soft-budget constraint: the manager is allocated a minimum amount of funds, but the total amount of funds available is not fully communicated to them, and termination is random from their perspective as long they are uncertain about the available financing. This situation contrasts with the use of a fixed deadline—a hard budget constraint—in which the manager is provided funding only for a specific amount of time. A different implementation exists when the scale of the project can be adjusted. In this case, the project is gradually downsized rather than terminated outright, a low rate of termination is analogous to a low rate of downsizing in this case. Finally, the probability of termination attempts to capture in a reduced form the difficulty in terminating a manager or liquidating a project.

This paper contributes explaining some aspects of real-life contracts. Evidence shows that tolerance of failure combined with long-term compensation

induces CEOs to adopt longer-term policies (Baranchuk, Kieschnick, and Moussawi 2014; Tian and Wang 2014). Compensation schemes in R&D-intensive companies show a negative correlation between pay duration and managerial turnover, and these features are more pronounced in firms with growth opportunities, high R&D, and long-term assets. These are firms with intangible assets for which short-term manipulation that hurts the firm in the long-run is more difficult to detect, making them more prone to the type of incentive-related problems analyzed here.

1. Related Literature

This paper belongs to the literature studying optimal contracts with managerial short-termism. Edmans et al. (2012) consider a similar problem in which a manager can increase performance today by reducing performance in the long-run. However, they consider an exogenous retirement date for the manager; moreover, because of the absence of limited liability, termination is not necessary. Zhu (Forthcoming) and Sannikov (2012) also consider models in which the manager's actions can be inferred only in the long-run, but they do not consider a multitasking problem, so there is no tension between high-powered incentives and performance manipulation. The problem of designing deferred compensation is similar to the one in Hartman-Glaser, Piskorski, and Tchisty (2012) and Malamud, Rui, and Whinston (2013), who study the design of incentives to screen loans. In addition to the difference in their question and their focus, these papers consider static settings that do not incorporate the dynamic aspects, such as turnover, that are the main focus here.

This paper builds on the the extensive literature on moral hazard with multitasking (Holmström and Milgrom 1991) and imperfect performance measures (Baker 1992).⁴ Some of these trade-offs also arise in models with ex ante moral hazard and ex post asymmetric information. Benmelech, Kandel, and Veronesi (2010) study a related model of managerial short-termism focusing on the use of stock based compensation, and the interrelation between stock prices and managerial incentives, instead of turnover and the optimal contracting. Inderst and Mueller (2010) study the optimal replacement policy in the presence of ex ante moral hazard and interim asymmetric information.⁵ This paper shares some of the predictions about turnover with this previous literature; but, by considering a dynamic model of repeated effort and short-termism, I show that managerial short-termism might render the provision of dynamic incentives ineffective with the implication that the principal may forbear low performance,

⁴ Sinclair-Desgagné (1999) shows how high-powered incentives can be restored by combining performance-based compensation with a scheme of selective audits.

⁵ Levitt and Snyder (1997), Inderst and Ottaviani (2009), and Heider and Inderst (2012) are also examples of models with ex ante moral hazard and interim asymmetric information.

and this generates contracts that are more stationary than we would otherwise predict.

Dynamic multitasking problems also arise in experimentation problems like the one in Manso (2011). To induce experimentation, the optimal contract has excessive continuation and requires the use of a severance payment. Klein (2016) considers what happens when the agent replicates the results using a known technology. However, this is inconsequential in his setting because there is no cost of deferring compensation (the principal and agent share the same discount rate).⁶

Some of the same incentive problems arise in venture capital. Stage financing is valuable because the threat of abandonment creates incentives for entrepreneurs to work hard, but it can also induce entrepreneurs to focus on short-term goals. Several papers look at the optimal allocation of control rights, such as the authority to replace the entrepreneur or terminate a project (Hellmann 1998; Bergemann and Hege 1998; Cornelli and Yosha 2003). Cornelli and Yosha (2003) study the role of convertibles in the context of stage financing when entrepreneurs can bias short-term performance. They analyze the role of convertibles and how their use discourages window dressing due to the entrepreneur's fear of conversion by the VC.

Finally, this paper belongs to a broad literature that uses recursive methods to study dynamic moral hazard problems in models with risk-neutral managers protected by limited liability (DeMarzo and Sannikov 2006; DeMarzo and Fishman 2007).⁷ Biais et al. (2010) and Myerson (2015) consider an optimal contract in Poisson models with bad news: the arrival of a Poisson shock corresponds to a loss, which is “bad news” because the manager's effort reduces the probability of arrivals. Random termination/downsizing is also required in these papers but is driven by different economic considerations. In Biais et al. (2010) and Myerson (2015) it is impossible to provide incentives to exert effort if the continuation value is low, and the only way to provide incentives is to rely on termination or downsizing. That is not the case here; in my setting, it is always possible to incentivize effort—no matter how low the manager's continuation value is—but downsizing/randomization is optimal because of the presence of short-termism.

2. Main Setting

The principal hires the manager to develop a project that can be of good or bad quality, $q \in \{g, b\}$. A good project arrives at a rate $\lambda + \Delta e_t$, where $e = \{e_t\}_{t \geq 0}$,

⁶ Fong (2009) considers a dynamic model with moral hazard and asymmetric information in which bad types can manipulate performance in order to pool with the high types.

⁷ In particular, dynamics models with Poisson arrival of news, like in Hopenhayn and Nicolini (1997) and He (2012). It is also related to papers analyzing optimal contracts for exponential bandit problems; examples in this literature include Bergemann and Hege (2005), Bonatti and Horner (2011), and Gerardi and Maestri (2012).

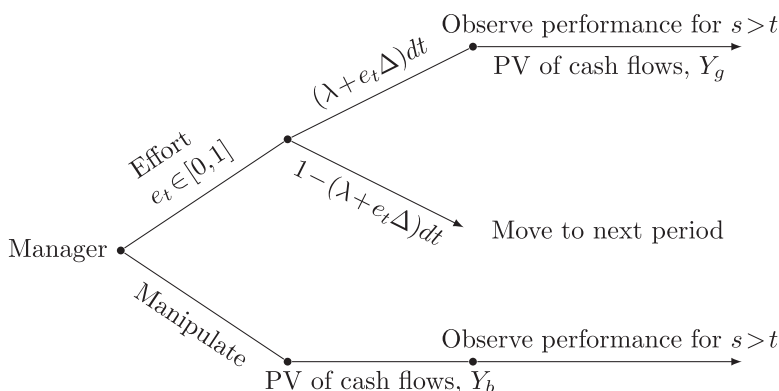


Figure 1
Time line of events at time t

$e_t \in [0, 1]$, is the manager’s unobservable effort. The manager can also produce a bad project—which looks like a good project in the short-term, but can generate losses in the long-term—at any time: once the project arrives, it generates a stream of cash flows $y > 0$ until a random failure time that is exponentially distributed with parameter ζ_q , where $\zeta_g < \zeta_b$. If the project fails, the principal suffers a loss $\ell > 0$, and so the expected value of the project is $Y_q \equiv (y - \zeta_q \ell) / (r + \zeta_q)$, where $Y_g > 0$ and $Y_b < 0$. In other words, a good project creates value, while a bad project destroys it, and is worse than no project at all.⁸ Figure 1 illustrates the timing of events.

The manager is risk neutral, has limited liability, and has a discount rate of γ . The manager’s cost of effort is C_{e_t} . The principal is also risk neutral and has a discount rate $r < \gamma$: because the manager is more impatient than the principal, the principal finds deferring payments to be costly.

A contract specifies the manager’s compensation, and the probability of termination as a function of (1) the time spent by the manager developing the project and (2) the project’s subsequent performance. Because the quality of the project is only revealed over time, the contract must specify payments for the manager subsequent to the completion date of the project. If the manager is terminated before being able to deliver the project, then the principal receives a liquidation payoff L .

We can summarize the relevant information available to the principal using the current date t the completion time τ , and the failure time $\bar{\tau}$. The contract specifies the manager’s cumulative compensation $U = \{U_t\}_{t \geq 0}$ and the liquidation date T , as a function of these variables. Limited liability requires the function U_t to be nondecreasing. Because the optimal contract requires

⁸ Alternatively, if we consider an initial investment I_0 (made at the time the project is completed), then the model also can accommodate the case with $\ell = 0$, in which case we have $Y_g = y / (r + \zeta_g) - I_0 > 0$ and $Y_b = y / (r + \zeta_b) - I_0 < 0$.

termination to be random, it is useful to specify the mean arrival of termination $\theta = \{\theta_t\}_{t \geq 0}$ as part of the contract. In summary, a contract is specified by the pair $C = (U, \theta)$.

Most of the paper focuses on the optimal contract that implements full effort $e_t = 1$ and during which time the manager does not generate a bad project. Section 4.4 considers the case with time-varying effort. Focusing on contracts implementing no manipulation is without loss of generality because the principal would never want to implement manipulation: implementing no effort is better than implementing manipulation. So the only other possibility for an optimal contract is that the principal wishes to implement no effort after some time. Intuitively, it is optimal to implement full effort if λ is sufficiently small compared to Δ or if the principal's outside option is sufficiently high. I provide sufficient conditions for full effort to be optimal in Appendix B.

Throughout the paper, I will make the following standing assumptions over the parameters.

Assumption 1.

1. Full effort is efficient:

$$\Delta Y_g > C.$$

2. The arrival rate with effort is high relative to the difference in discount rates:

$$\lambda + \Delta \geq \gamma - r.$$

The first condition ensures that the benefit of exerting effort is greater than its cost. The second condition is more technical in nature; it is required in the verification step for optimality.

3. Manager's Incentive Compatibility Constraint

In this section, I consider the agent incentive compatibility constraint. As is usual in the dynamic contracting literature, I use the manager's continuation value as the main state variable. The manager's continuation value given a contract C is

$$W_t = E_t \left[\int_t^\infty e^{-\gamma(s-t)} (dU_s - \mathbf{1}_{\{s < T \wedge \tau\}} e_s C ds) \right] \mathbf{1}_{\{t < T\}}. \tag{1}$$

Manipulation has a persistent effect on the output process – this captures the notion of short-termism; hence, like in Fernandes and Phelan (2000), I have to distinguish between the continuation value on-the-equilibrium path and the continuation value off-the-equilibrium path that follows a deviation. I denote

these continuation values by

$$\bar{W}_t^g \equiv E_t^g \left[\int_t^\infty e^{-\gamma(s-t)} dU_s \right] \mathbf{1}_{\{t < T\}}, \tag{2}$$

$$\bar{W}_t^b \equiv E_t^b \left[\int_t^\infty e^{-\gamma(s-t)} dU_s \right] \mathbf{1}_{\{t < T\}}, \tag{3}$$

where $E_t^q(\cdot)$ is the expected value conditional on quality q . The value \bar{W}_t^g corresponds to the manager's expected payoff from a good project (on-the-equilibrium-path), and the value \bar{W}_t^b corresponds to the expected payoff from a bad project (off-the-equilibrium-path).

Problems with persistent private information are usually difficult to analyze. However, I can analyze the model using standard recursive techniques because, after the project is completed, the manager is no longer working and is just waiting for the payment. This allows to separate the problem after the project is completed from the problem in the employment stage (before the project is completed), and analyze the problem using standard recursive methods and backward induction. First, I look at the incentives to manipulate performance, and then—given that the manager does not manipulate—I look at the incentives to exert effort. Let's consider the incentives to generate a bad project at time t . The value that the manager obtains by not generating a bad project (and continue work on the good project) is W_t , and the value of generating a bad project is \bar{W}_t^b : thus, not completing a bad project is incentive compatible if and only if

$$W_t \geq \bar{W}_t^b.$$

Note that W_t is the manager's expected payoff *immediately before* completing the project, while \bar{W}_t^g is the manager's expected payoff *immediately after* completing the good project.

Next, I consider the incentives to exert effort in the good project. Because it is never optimal to compensate a risk-neutral manager before they complete the project, I have that $dU_t = 0$ for $t < \tau$. If the manager chooses not to complete a bad project, then their continuation payoff at time $t < \tau$ is

$$W_t = \int_t^\infty e^{-(\lambda+\gamma)(s-t) - \int_t^s (e_u \Delta + \theta_u) du} ((\lambda + e_s \Delta) \bar{W}_s^g - C e_s) ds.$$

I can differentiate the previous expression with respect to t and obtain

$$\dot{W}_t = (\gamma + \theta_t) W_t + C e_t - (\lambda + e_t \Delta) (\bar{W}_t^g - W_t). \tag{4}$$

The previous equation implies that the manager's effort is

$$e_t = \operatorname{argmax}_e [(\bar{W}_t^g - W_t) \Delta - C] e.$$

If I let $c \equiv C/\Delta$ be the marginal cost of effort measured in units of arrival intensity, then I can write the incentive compatibility constraints as

Lemma 1. Full effort, $e_t = 1$, and no manipulation are incentive compatible if and only if

$$\overline{W}_t^g - W_t \geq c \tag{5}$$

$$W_t \geq \overline{W}_t^b. \tag{6}$$

Section C.1 of the appendix provides the formal proof of Lemma 1. Next, I provide an intuition for the incentive compatibility constraint: Equation (5) says that to induce the manager to exert effort, the marginal benefit of effort $\overline{W}_t^g - W_t$ must be greater than its marginal cost c . Equation (6) says that, because the manager can always secure an immediate payoff of \overline{W}_t^b by delivering a bad project, the continuation value must be greater or equal than \overline{W}_t^b . I can provide an alternative interpretation of the incentive compatibility constraint (5) by comparing the payoffs of effort and shirking between time t and time $t + dt$: the manager payoff of shirking is

$$\text{Payoff Shirking} = \lambda dt \overline{W}_t^g + (1 - \lambda dt)e^{-\gamma dt} W_{t+dt} + o(dt),$$

whereas the payoff of exerting effort is

$$\text{Payoff Full Work} = \text{Payoff Shirking} + \Delta dt \left(\overline{W}_t^g - e^{-\gamma dt} W_{t+dt} \right) - C dt + o(dt).$$

The previous equations show that if the principal increases the payoff for failure today, W_{t+dt} , this requires that the principal also increases the reward for success \overline{W}_t^g . In a sense, the constraint (5) becomes more stringent when the current continuation value is very high. In contrast, inequality (6) is more stringent if the current continuation value because in this case the manager has little to lose by manipulating performance. This tension between both constraints, (5) and (6), captures the tension between the incentives to exert effort and the incentives to manipulate performance.

4. Principal Contracting Problem

After deriving the manager's incentive compatibility constraint, I can proceed to solving the principal's optimization problem. The expected payoff for the principal from a contract \mathcal{C} given effort $e = \{e_t\}_{t \geq 0}$ and no manipulation is

$$P_0 = E \left[e^{-r\tau} \mathbf{1}_{\{\tau \leq T\}} Y_g + e^{-rT} \mathbf{1}_{\{\tau > T\}} L - \int_0^\infty e^{-rt} dU_t \right]. \tag{7}$$

The principal's problem is to design an incentive-compatible contract \mathcal{C} that maximizes the principal's profits. This problem can be separated into two parts: (1) the design of the deferred compensation plan for $t \geq \tau$ and (2) the contracting problem in the employment state for $t < \tau$. Thus, I can solve for the optimal contract using backward induction: I solve for the payment at $t \geq \tau$, which I will denote by $U_t^+ \equiv \{U_s\}_{s \geq t}$, and then I solve for the optimal contract in the employment state $t < \tau$, which will determine the termination rate θ_t .

4.1 Optimal deferred compensation

Next, I solve for the optimal payment for $t \geq \tau$: this amounts to finding the least expensive way of delivering a payoff w , while inducing effort and deterring a bad project. I find the deferred payment by solving the following optimization problem:

$$\Pi(w) \equiv \sup_{U^+} Y_g - E_\tau^g \left[\int_\tau^\infty e^{-r(t-\tau)} dU_t^+ \right],$$

subject to

$$\bar{W}^g \geq w + c,$$

$$\bar{W}^b \leq w.$$

This problem is similar to that analyzed by Hartman-Glaser, Piskorski, and Tchisty (2012) in the context of securitization.⁹ Because this is a linear optimization problem with a convex set of constraints, it is natural to look for an extremal solution; so, I can conjecture and then verify that the optimal payment takes the form of a single deferred bonus that is paid only if the project does not fail before the payment date. The probability that a project of quality q does not fail before the bonus is paid is $e^{-\zeta_q \delta}$, and so the manager's (expected) payoff from a good project is $e^{-(\gamma+\zeta_g)\delta} \bar{U}$, where \bar{U} is the bonus and δ is the deferral, while the expected payoff from a bad project is $e^{-(\gamma+\zeta_b)\delta} \bar{U}$. Hence, the incentive compatibility constraints can be written as

$$e^{-(\gamma+\zeta_g)\delta} \bar{U} \geq w + c,$$

$$e^{-(\gamma+\zeta_b)\delta} \bar{U} \leq w.$$

If the two incentive compatibility constraints are binding, finding the optimal payment reduces to solving the system of equations for the bonus \bar{U} and the deferment δ . In the proof of the following lemma, I verify that both constraints are binding, and so the optimal payment is given by the solution to the system of equations.

Lemma 2. The optimal contract has a payment U^+ given by

$$dU_t^+ = \bar{U}_\tau \mathbf{1}_{\{t=\tau+\delta_\tau, \bar{\tau} > \tau+\delta_\tau\}}, \tag{8}$$

where

$$\delta_\tau = \frac{1}{\zeta_b - \zeta_g} \log \left(\frac{c + W_\tau}{W_\tau} \right), \tag{9}$$

$$\bar{U}_\tau = e^{(\gamma+\zeta_g)\delta_\tau} (c + W_\tau). \tag{10}$$

⁹ Malamud, Rui, and Whinston (2013) extend the analysis to more general distributions.

In the optimal contract, both incentive compatibility constraints are binding. The principal expected payoff under this contract is

$$\Pi(w) = Y_g - (c + W_\tau)^{\phi+1} W_\tau^{-\phi}, \tag{11}$$

where $\phi \equiv \frac{\gamma-r}{\zeta_b - \zeta_g} > 0$, and Π is a concave function.

When $\gamma = r$, the profit function reduces to $\Pi(w) = Y_g - w - c$. In this case, the principal profits are the same as those in the case with observable quality.

4.1.1 Costly monitoring. In many situations, it might be difficult to use deferred payments that are contingent on subsequent performance. For example, it may be difficult to determine ex post the quality of an article of equipment if failures can arise due to misuse by the buyer. In fact, this is one of the reasons why many procurement contracts have warranty clauses with limited coverage (Burt 1984, p. 194).

We can capture the main economic mechanism in many of these situations by considering the case in which the principal can implement costly monitoring once the project is completed. For example, if we consider a simple monitoring technology that allows to discover a bad project with probability m_τ , where m_τ is the intensity of monitoring chosen by the principal.¹⁰ If I let \bar{U}_t be the manager's bonus conditional on a positive monitoring outcome, then I can show that the incentive compatibility constraints becomes

$$\begin{aligned} \bar{U}_t - W_t &\geq c, \\ W_t &\geq (1 - m_t)\bar{U}_t, \end{aligned}$$

and the principal profits $Y_g - w - c - h(m(w))$, share same qualitative features as the profit function in Equation (11); accordingly, the features of the contract are similar when the principal must rely on costly monitoring rather than deferred compensation: in some sense, this equivalence highlights the fact that deferred compensation is a way of costly monitoring.

4.2 Project termination

Given the optimal compensation design at $t \geq \tau$, I can now solve for the optimal contract in the employment stage at time $t < \tau$. Because the manager is risk neutral, it is never optimal to compensate the manager before they complete the project, and I can write the principal problem as

$$P(W_0) = \max_{e_t \in [0, 1], \theta_t \geq 0} \int_0^\infty e^{-(r+\lambda)t - \int_0^t (e_s \Delta + \theta_s) ds} \left((\lambda + e_t \Delta) \Pi(W_t) + \theta_t L \right) dt,$$

subject to the evolution of the continuation value in Equations (4). If the optimal contract implements maximum effort, then both incentive compatibility

¹⁰ To keep matters simple, I assume that the monitoring technology does not generate false positives and that the cost of monitoring is given by a cost function $h(\cdot)$. The function h is increasing, is convex, is continuously differentiable, and satisfies the conditions $\lim_{m \rightarrow 1} h(m) = \infty$ and $\lim_{m \rightarrow 1} h'(m) = \infty$.

constraints must be binding, so the evolution of the manager's continuation value is

$$\dot{W}_t = (\gamma + \theta_t)W_t - \lambda c. \tag{12}$$

This is a deterministic optimal control problem that can be solved using dynamic programming: the value function $P(w)$ satisfies the Hamilton-Jacobi-Bellman (HJB) equation

$$rP(w) = \max_{\theta \geq 0} \left\{ ((\gamma + \theta)w - \lambda c)P'(w) + (\lambda + \Delta)[\Pi(w) - P(w)] + \theta(L - P(w)) \right\}. \tag{13}$$

The first term in the HJB equation reflects the effect of changes in the continuation value, the second term captures the expected profits from the project, and the third term captures the effect of inefficient liquidation. The possibility of stochastic termination implies that $P(w) - wP'(w) \geq L$, and $\theta(w)$ is nonnegative only if this inequality holds with equality. It is optimal to defer compensation until the agent produces a project as long as $P'(w) \geq -1$. Limited liability (of the manager) implies that the project must be terminated as soon as $w = 0$, so the solution to (13) must satisfy the boundary condition $P(0) = L$. For values of w with no termination, the HJB equation (13) simplifies to

$$rP(w) = (\gamma w - \lambda c)P'(w) + (\lambda + \Delta)[\Pi(w) - P(w)]. \tag{14}$$

The termination rate is zero if the solution to the HJB equation is strictly concave; however, if the solution to Equation (14) is not strictly concave, then the termination rate must be positive, and in this case there is a threshold w_* such that for any $w \leq w_*$,

$$wP'(w) - P(w) + L = 0.$$

The value function is linear in this range and is continuously differentiable at w_* . The threshold w_* is determined by the super contact condition $P''(w_*) = 0$.¹¹

As soon as the continuation value reaches the threshold w_* , the contract becomes stationary. The termination intensity θ_t is set at a level consistent with a constant continuation value. Equation (12) implies that such a termination policy is given by

$$\theta_t = \mathbf{1}_{\{W_t = w_*\}} \left(\frac{\lambda c}{w_*} - \gamma \right). \tag{15}$$

The rate of termination is positive, and the contract becomes stationary. Notice that the termination rate is positive only if $\lambda > 0$. This happens because a manager who exerts no effort never terminates the project when $\lambda = 0$, which

¹¹ Lemma 4 in the appendix shows that the maximal solution to the HJB equation satisfies the super contact condition. The super contact condition arises because this is a singular optimal control problem.

implies that termination is not needed and the optimal contract is stationary. Also, notice that the termination intensity is decreasing in the threshold w_* as a lower turnover is needed to provide the manager with a higher continuation payoff. The following proposition provides a summary of the optimal contract.

Proposition 1. Suppose that

$$\frac{\lambda + \Delta}{r + \lambda + \Delta} \Pi\left(\frac{\lambda c}{\gamma}\right) - L > \frac{\lambda + \Delta}{r + \lambda + \Delta - \gamma} \frac{\lambda c}{\gamma} \Pi'\left(\frac{\lambda c}{\gamma}\right). \quad (16)$$

Then the HJB equation (13) has a maximal solution. The threshold $w_* \in (0, \lambda c / \gamma)$ is determined by the super contact condition $P''(w_*) = 0$ and is the unique solution to

$$\Pi'(w_*) = \frac{r + \lambda + \Delta - \gamma}{(r + \lambda + \Delta - \gamma)w_* + \lambda c} \Pi(w_*) - \frac{r + \lambda + \Delta - \gamma}{(r + \lambda + \Delta - \gamma)w_* + \lambda c} \frac{r + \lambda + \Delta}{\lambda + \Delta} L. \quad (17)$$

The optimal contract implementing effort and no manipulation is given by

1. A cumulative payment process U_t^+ described by (8)–(10)
2. A stochastic termination time T with hazard rate $\theta(W_t) = \mathbf{1}_{\{W_t = w_*\}} \left(\frac{\lambda c}{w_*} - \gamma\right)$

The expected payoff for the principal under the optimal contract is given by $P(W_0)$.

The value function characterizes the optimal contract for any continuation value at time zero: hence, it provides the solution for any division of the bargaining power between the principal and the manager. In the particular case in which the principal has all the bargaining power, then the contract is initialized at the promised W_0 that maximizes $P(W_0)$. It is not difficult to verify that there is some \bar{Y}_g large enough so the condition (16) is satisfied for any $Y_g > \bar{Y}_g$. Proposition 1 describes the optimal contract implementing effort; later, in Appendix B, I provide conditions for effort to be optimal and discuss the case in which it is not optimal to implement full effort all the time.

Figure 2 illustrates the optimal contract. The contract can be described as a function of time: letting T_* be the time that it takes for the continuation value to reach the threshold w_* , I find that for $t < T_*$, the contract is dynamic and the manager is punished for delays. This scenario helps to provide incentives to exert effort. The manager has incentives to exert effort before time T_* because the (present value) bonus they receive after completing the project decreases over time. However, the contract becomes stationary for $t \geq T_*$: the payment is no longer reduced, and incentives are provided by the possibility of being terminated.

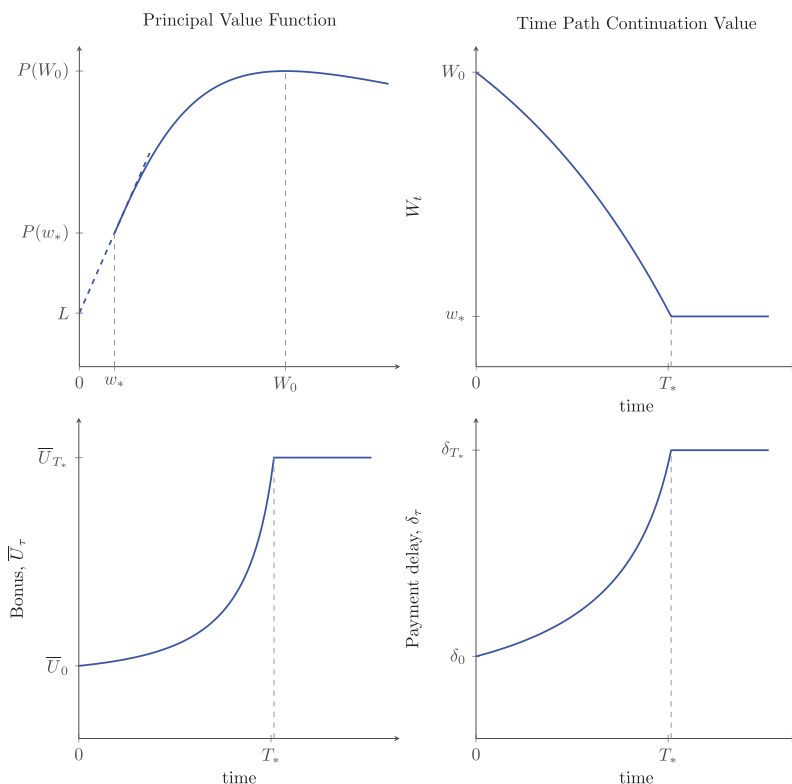


Figure 2
Optimal contract with contractible randomization

The contract is initialized at the value W_0 ; before time $T_* = \min\{t > 0: W_t = w_*\}$, incentives are provided through front-loaded payments. After time T_* , payments remain constant and incentives are provided through probabilistic termination. Parameters: $r=0.05$, $\gamma=0.1$, $\lambda=0.05$, $\Delta=0.5$, $\zeta_g=0$, $\zeta_b=0.2$, $Y_g=500$, $L=150$.

Why does the optimal contract becomes stationary when w is low? The economic intuition is that there are two possible ways to provide incentives to exert effort:

1. the principal can punish the manager for delays by reducing the promised payment (i.e., the continuation value), and
2. The principal can use a stationary contract in which compensation is constant but the project is terminated with positive probability if the manager fails to deliver.

If we ignore the possibility of manager short-termism, using (1) is always more efficient than using (2). Because offering a high continuation value tomorrow makes it more difficult to satisfy the incentive compatibility constraint today. The manager can always make little effort today and work tomorrow, suffering little cost from delays. In contrast, if the principal uses stochastic termination,

the manager risks being terminated if the project is not completed today; however, stochastic termination is costly because the principal suffers the risk of terminating the manager early, which is suboptimal. When the principal and the manager are equally patient, the problems with and without managerial short-termism are equivalent because deferring compensation is costless. This last observation highlights that the crux of the problem is that the manager is more focused on the short-term than is the principal.

If the manager is more impatient than the principal, it is costly to defer compensation, which means that it is also costly to reduce the manager's promised payment. Limited liability constrains the punishment the principal can inflict on the manager *ex post* if a bad project fails, so the incentive compatibility constraint $W_t \geq e^{-(\gamma+\zeta_b)\delta t} \bar{U}$ becomes more difficult to satisfy when W_t is close to zero. As a consequence, it is suboptimal to punish the manager if the continuation value is low, and the contract becomes stationary (conditional on retaining employment). However, a stationary contract may require the use of (random) liquidation to provide incentives. In other words, rather than reducing the manager's promised payment, the principal keeps the compensation constant but terminates the contract with positive probability if there are further delays.

4.2.1 Project downsizing. The specific way in which the optimal contract is implemented will depend on the precise context that we are considering. One common situation arises when the scale of the project can be adjusted over time; in that case, the principal can gradually downsize the project rather than terminating it outright. So the question in this case is whether the principal prefers to use an investment with gradual downsizing or a policy with a deadline at which the project is terminated. A well-known feature in the dynamic contracting literature is that stochastic liquidation shares many features with downsizing. In fact, when production technology has constant returns to scale, downsizing and random termination are mathematically equivalent, like in Biais et al. (2010) and Myerson (2015). In particular, the project starts at the maximum scale of 1 but at any point in time can be downsized to any scale $K_t \in [0, 1]$ (the liquidation value of the assets is L). Because the project technology has linear returns to scale, both the cash-flow Y_q and the cost of effort C are proportional to the scale of the project K , and so if I interpret the continuation value and the manager's payments as per unit of capital K , then the optimal contracting problem looks exactly the same as before. The main difference now is that the principal gradually downsizes the project at a rate $\theta(w_*)$ when the continuation value (per unit of capital) reaches the lower threshold w_* rather than terminating the manager.

Figure 3 shows the evolution of the project scale when quality is difficult to observe. When we interpret this result, we must keep in mind that when quality is observable, the project is always at full scale, there is no intermediate downsizing, and the project is operated at full scale before being fully liquidated at the deadline.

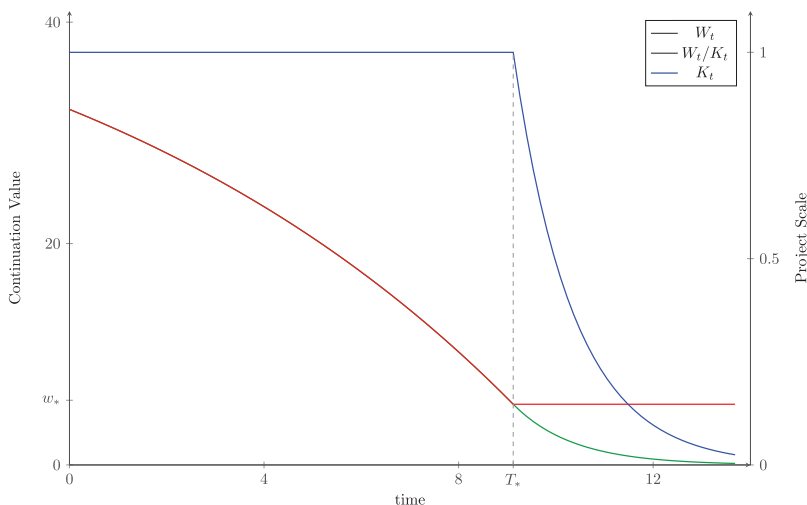


Figure 3
Time path of project scale and continuation value

In the presence of quality concerns, the project is operated at full scale up to time T_* . The project is gradually downsized after that point. Parameters: $r=0.05$, $\gamma=0.1$, $\lambda=0.05$, $\Delta=0.5$, $\zeta_g=0$, $\zeta_b=0.2$, $Y_g=500$, $L=150$.

4.3 Turnover, compensation, and noisy quality

In this section, I analyze the implications of the optimal contract for turnover and how turnover is related to the difficulty of detecting short-termism, that is, how noisy is quality. Proposition 1 specifies the optimal contract for any division of the surplus between the principal and the manager. In this section, I discuss the comparative statics for any division of the surplus between the manager and the principal: that is, I consider the case in which W_0 is given and the case in which the principal has all the bargaining power; so W_0 is chosen as the maximizer of $P(w)$. The cost of deterring manipulation depends on the informativeness of the signal $\bar{\tau}$ (recall that $\bar{\tau}$ is the time when the project fails). The log-likelihood ratio between the failure time of a good and a bad project is $(\zeta_b - \zeta_g)t$ and measures the precision of the information about quality.

We start by looking at the extreme case in which the optimal contract is completely stationary and the manager is never terminated: this is optimal if quality is very noisy. Using Equation (12), I find that $w^s \equiv \lambda c / \gamma$ is the steady state of the continuation value when no termination is used. If $W_0 = w^s$, then the contract is completely stationary: using termination is so costly that the manager is never terminated after low performance. The following proposition shows that this is the case when $\zeta_b - \zeta_g$ is sufficiently low.

Proposition 2. Under the assumptions in Proposition 1, a stationary contract is optimal if and only if

$$\zeta_b - \zeta_g \leq \frac{\gamma(\gamma - r)}{\lambda}. \tag{18}$$

In this case, the contract:

1. has no deadline to complete the project, that is $T = \infty$ ($\theta_t = 0$ for all t);
2. promises a single payment $dU = e^{\gamma\delta(w^s)}w^s$ to be paid at time $\tau + \delta(w^s)$, where τ is the date of project completion and δ is given by (9).

The main takeaway of Proposition 2 is that a manager is never terminated when it is either difficult to differentiate high- and low-quality projects (low $\zeta_b - \zeta_g$) or costly to defer compensation (high $\gamma - r$). In this case, the manager is motivated to exert effort because that allows them to receive the payment as soon as possible. However, because the manager is never terminated, the compensation required to induce effort is very high. In other words, only a carrot is being used to provide incentives, and this would never be optimal in the absence of managerial short-termism. With hidden effort, there is a substitution between the incentives to exert effort today and the incentives to exert effort tomorrow. By exerting effort today, the manager increases the probability of finishing now; yet, if the project is finished today, the manager gives up the possibility of finishing the project tomorrow with the associated reward. Thus, a higher reward tomorrow makes it harder to incentivize the manager today. This intuition in the standard case with pure hidden effort indicates that rewards should decrease over time. As has been already mentioned in the previous section, eventually, limited liability will make it impossible to reduce the reward further: at this point, the project must be terminated. This is the deadline common to the previous literature. But this intuition ignores the effect that reducing the reward has on the incentive to accelerate the project by taking shortcuts: the optimal contract balances these two incentives. When quality is too difficult to observe, the second effect dominates, and the principal does not reduce the reward, and the manager is never terminated.

Now, I can discuss the more general case in which the contract consists of a nonstationary phase followed by a stationary phase. I derive comparative statics that relate the difficulty of detecting short-termism to the manager's compensation and turnover. Later, in Section 5, I discuss the empirical implications and compare the prediction of the model with the evidence. Turnover is determined by two numbers: the threshold w_* , where the contract becomes stationary, and the initial continuation value W_0 . First, I show that w_* is higher when quality is more noisy. This, in turn, implies that, for any given fixed continuation value W_0 , the expected duration is decreasing in $\zeta_b - \zeta_g$.

Proposition 3. The random termination threshold is decreasing in the precision of the signal $\bar{\tau}$. That is, w_* is a decreasing function of $\zeta_b - \zeta_g$. This means that for any fixed $W_0 > w_*$ the expected termination date $E(T|\tau > T)$ is a decreasing function of $\zeta_b - \zeta_g$.

Next, I consider the case in which the principal has all the bargaining power, so $W_0 = \operatorname{argmax} P(w)$, and show that W_0 is also higher when quality is noisier.

In this case, the rents that the manager receives are directly linked to the punishments for delays. In addition, because deferring payment is costly, the principal reduces the manager's incentive to manipulate performance by reducing the punishment for delays. This implies that the previous result about the duration of the contract extends to the case in which the principal has all the bargaining power and that in this case the manager's rents are higher in the presence of quality concerns.

Proposition 4. Let $W_0 = \operatorname{argmax}_w P(w)$, and suppose that

$$\zeta_b - \zeta_g > \frac{\gamma(\gamma - r)}{\lambda};$$

that is, $W_0 < w^s$, where w^s is the manager's payoff in the stationary contract. Then, the manager's payoff is decreasing in the precision of the signal $\bar{\tau}$. That is, W_0 is a decreasing function of $\zeta_b - \zeta_g$.

Recalling that T_* is the time at which the manager is fired with positive probability, the expected termination date is

$$\begin{aligned} E(T|\tau > T) &= T_* + E(T - T_*|\tau > T) \\ &= \frac{1}{\gamma} \left[\log \left(\frac{w^s - W_0}{w^s - w_*} \right) + \frac{w_*}{w^s - w_*} \right]. \end{aligned} \tag{19}$$

The previous equation, together with Propositions 3 and 4, implies that the duration of the contract is decreasing in the informativeness of the failure time regarding quality.

Proposition 5. Suppose that $W_0 = \operatorname{argmax} P(w)$; under the assumptions in Proposition 1, the expected termination date $E(T|\tau > T)$ is a decreasing function of $\zeta_b - \zeta_g$.

Proposition 2 states that the manager is never terminated if $\zeta_b - \zeta_g$ is sufficiently low; now, I conclude that even if the manager is sometimes terminated, the expected duration of the contract is decreasing in the precision of the information about quality. In addition, it is also the case that $E(T|\tau > T) \rightarrow \infty$ as $\zeta_b - \zeta_g \downarrow \gamma(\gamma - r)/\lambda$.

4.4 Convex cost of effort

The previous analysis largely relies on the assumption that it is optimal for the principal to implement effort all the time. I provide sufficient conditions for full effort to be optimal in the appendix. In this section, I consider the case in which the cost of effort is a strictly convex function. This case allows us to see the effect of short-termism on the evolution of effort. It has been highlighted that managerial short-termism makes the optimal contract more stationary; this

stationarity becomes even more apparent when we look at the time evolution of effort. One standard result in models without short-termism is that effort is frontloaded, meaning that the power of incentives (and so effort) decreases over time. This is not necessarily the case in the presence of managerial short-termism. In the stationary region, effort is constant, and so the slope of incentives is constant even after low performance. Moreover, even in the nonstationary region, effort becomes less sensitive to performance: effort decreases at a lower speed when quality is more noisy.

We generalize the model to a strictly convex cost of effort: the manager continuously chooses a level of effort $e_t \in [0, \bar{e}]$ at an instantaneous cost $c(e_t)$. The cost function is assumed to be strictly increasing, convex, and twice continuously differentiable. Given any effort level e_t , I assume that the good project is completed with intensity $\lambda + e_t$. The equation for the evolution of the continuation value in this case is

$$\dot{W}_t = (\gamma + \theta_t)W_t + c(e_t) - (\lambda + e_t)(\bar{W}_t^g - W_t), \tag{20}$$

and the incentive compatibility constraint now is given by following maximization problem:

$$e_t = \operatorname{argmax}_e e(\bar{W}_t^g - W) - c(e).$$

This optimization problem yields the incentive compatibility constraint $\bar{W}_t^g - W_t = c'(e_t)$. The appendix provides the formal proof. In addition, the non-manipulation incentive constraint is $W_t \geq \bar{W}_t^b$. As I did before, I first look at the principal's problem at time $t \geq \tau$ and then solve for $t < \tau$. Noting that this optimization problem is the same as the optimization problem in Lemma 2, with the minor difference that I replace the marginal cost c with $c'(e_t)$, I obtain the principal's profit as a function of the promised value and the effort level

$$\Pi(w, e) = Y_g - (c'(e) + w)^{\phi+1} w^{-\phi}. \tag{21}$$

When the cost of effort is strictly convex, it is simpler to solve the model using the Pontryagin maximum principal rather than by using dynamic programming. The optimization problem for the principal in the first stage, before the project is completed, is

$$\max_{e_t \in [0, \bar{e}], \theta_t \geq 0} \int_0^\infty e^{-(r+\lambda)t - \int_0^t (e_s + \theta_s) ds} \left((\lambda + e_t)\Pi(W_t, e_t) + \theta_t L \right) dt,$$

where optimization is subject to the evolution of the continuation value in (20) and the incentive compatibility constraint. If I replace the incentive compatibility constraint in the evolution of the continuation value, and I use the auxiliary state variable $\Lambda_t = \int_0^t (e_s + \theta_s) ds$, then I can write this optimization problem in a form that is more convenient for an application of optimal control techniques:

$$\max_{e_t \in [0, \bar{e}], \theta_t \geq 0} \int_0^\infty e^{-(r+\lambda)t - \Lambda_t} \left((\lambda + e_t)\Pi(W_t, e_t) + \theta_t L \right) dt,$$

subject to

$$\begin{aligned} \dot{W}_t &= (\gamma + \theta_t)W_t + c(e_t) - (\lambda + e_t)c'(e_t) \\ \dot{\Lambda}_t &= e_t + \theta_t, \Lambda_0 = 0. \end{aligned}$$

When the cost of effort is a strictly convex function, I cannot solve the previous optimization problem in closed form; however, I can address it numerically, and I can also obtain a reasonable amount of intuition from its first-order conditions. For simplicity, I relegate the analysis of the necessary and sufficient conditions to the appendix. Just like in the case with a linear cost of effort, the qualitative nature of the results will depend on the liquidation value L ; if the liquidation value is relatively high, then termination is better than low effort, and if the liquidation value is sufficiently low, no effort is better than liquidating the project. In this latter case, the manager is never fired. I focus in the case with a relatively high liquidation value, so it is not optimal to implement zero effort. The first-order condition for effort is given by

$$P'(W_t)(\lambda + e_t)c''(e_t) - (\lambda + e_t)\Pi_e(W_t, e_t) = \Pi(W_t, e_t) - P(W_t). \quad (22)$$

The left-hand side in (22) represents the cost of increasing effort; this cost consists of two terms: the first term captures the impact of reducing the continuation value over time—this is the punishment for low performance—which has an effect on the principal expected payoff of $P'(W_t)$. The second term reflects the effect of increasing the power of incentives, which makes short-termism more attractive and requires more deferred compensation. The right-hand side captures the benefit given by the difference between the profits of a complete project and an incomplete project. The termination threshold is pinned down by the condition

$$\frac{P'(w^*)\left(c(e^*) - (\lambda + e^*)c'(e^*)\right) + (\lambda + e^*)\Pi(w^*, e^*)}{r + \lambda + e^*} = \frac{\lambda + e^*}{r + \lambda + e^*} \Pi_w(w^*, e^*)w^* + L. \quad (23)$$

If the cost of effort is linear, then the previous condition reduces to the same condition in the baseline model (equation Equation (17)). I find the level of effort in the stationary phase by evaluating Equation (22) at (w^*, e^*) and solving the system of equations (22)–(23).¹²

Figure 4a shows the evolution of the continuation value and effort for two different values of $\zeta_b - \zeta_g$. The optimal contract implements lower effort when it is more difficult to distinguish a good project from a bad one. This difference is particularly important at the beginning of the contract. Over time, the level of effort converges to a similar level. Like in the case with the linear cost of

¹² As it was also the case in the baseline model, I can solve for $P(w^*)$ and $P'(w^*)$ without solving the HJB equation.

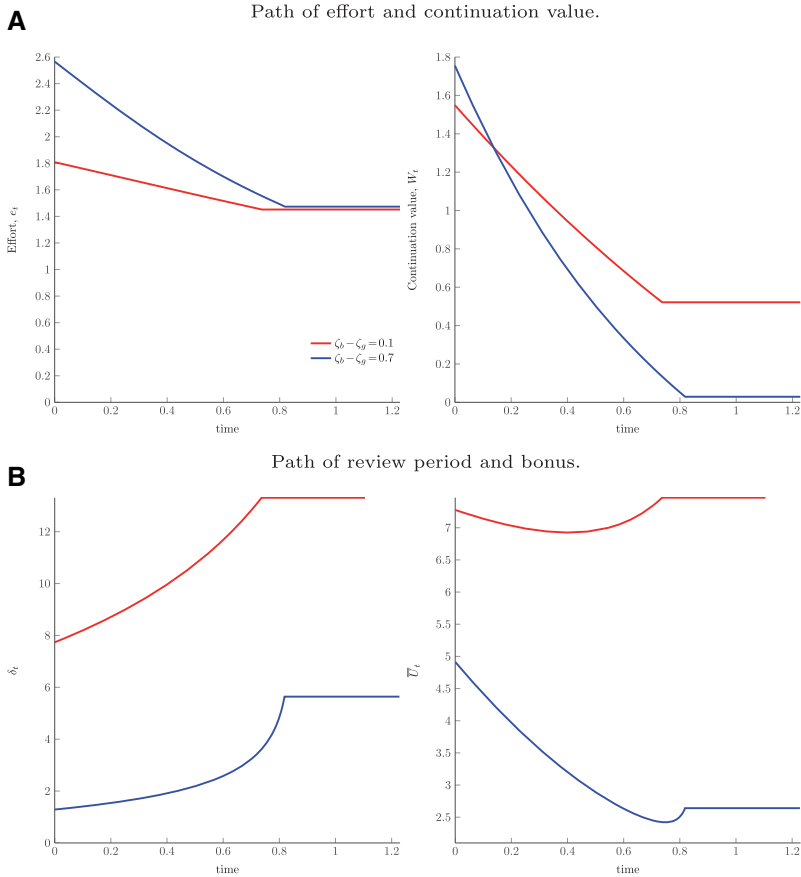


Figure 4
Path in the optimal contract
 Parameters: $r=0.05, \gamma=0.1, \lambda=0.1, Y=170, L=160, c(e)=0.5 \cdot e^2$.

effort, the punishment for delay is used less when the information about quality is noisier, so it is more difficult to detect deviations in expected quality. This is reflected in the fact that the continuation value falls faster when $\zeta_b - \zeta_g$ is relatively high. The dynamics of the continuation value are similar to those with the linear cost of effort. At date T_* , the continuation value is decreasing for $t < T_*$. In this first phase, incentives are provided mainly by reducing the manager’s compensation and effort decreases over time as it becomes increasingly costly to incentivize the manager. After time T_* , compensation and effort remain constant in a second phase. From here on, the possibility of termination provides the incentives.

Figure 4a highlights the effect of managerial short-termism on the time evolution of incentives. The evolution of effort – as well as the evolution of

Downloaded from https://academic.oup.com/ft/article-abstract/31/9/3409/4060545 by Acq/Serials Dept-Periodicals user on 17 February 2020

the continuation value – becomes more flat when quality is more noisy. This captures the idea that effort is not front-loaded as much and the contract becomes more stationary. The principal does not rely as much on the dynamic provision of incentives because this makes preventing managerial short-termism more difficult. One difference between the case with linear cost and the case with convex cost of effort is that while the manager's rent at time zero W_0 is always decreasing in $\zeta_b - \zeta_g$, in the linear case, this is not always the case when the cost of effort is convex. This difference should not come as a surprise; in the case of a convex cost of effort, we have two forces working in opposite directions. On the one hand, the temptation to deviate and work on the bad project is lower when the rents from work on the good project are high; this was the effect identified in previous sections, and this means that the principal might want to increase the manager's payoff. On the other hand, because high effort is more costly to implement, the principal might want to reduce the power of incentives, and that implies that the manager's rent is lower – this is the traditional effect on effort in the multitasking literature. Then, depending on which of these effects dominates, the manager's payment may go up or down. We should expect that the distortion in effort will be low if effort is very productive and if the cost of effort is not too convex; when this the case, the first effect is likely to dominate. For example, if the project is large enough, the benefit of effort greatly surpasses the cost of effort, so maximal effort $e_t = \bar{e}$ is optimal; this is the argument made by Edmans et al. (2012) to focus on contracts implementing high effort in the study of CEO compensation. Formally, this will happen if $c'(\bar{e})$ is low relative to Y_g – in this case, the analysis in the previous sections applies. The overall effect on the expected duration of the contract is presented in Figure 5, and we find that the duration of the contract becomes longer when the signal about quality becomes more noisy.

5. Applications and Empirical Implications

The purpose of this section is to discuss the different implications of the model for managerial short-termism, and other applications, in particular, venture capital contracts. I begin by discussing the existent empirical evidence related to managerial short-termism and the different implications of the model. Later, I discuss the implications of the model in the context of venture capital contracts and the relation to some stylized facts in the empirical literature on venture capital contracting.

An extensive empirical literature analyzes the perverse effects of ill-designed high-powered incentive schemes. For example, Burns and Kedia (2006) study the effect of CEO compensation contracts on misreporting and find that stock options are associated with stronger incentives to misreport. Similarly, Larkin (2014) shows that high-powered incentives lead salespeople to distort the timing, quantity, and price of sales in order to game the system. In a different context, Agarwal and Ben-David (Forthcoming) and Gee and Tzioumis (2013)

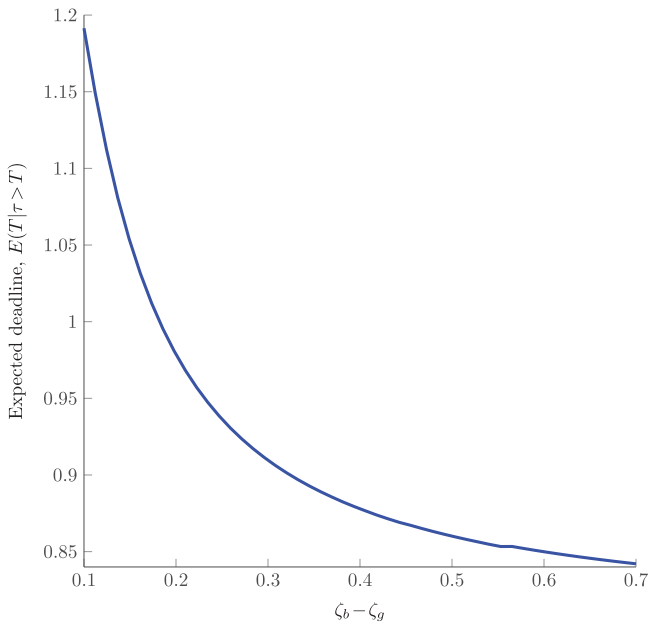


Figure 5
Expected deadline

Parameters: $r=0.05$, $\gamma=0.1$, $\lambda=0.1$, $Y=170$, $L=160$, $c(e)=0.5 \cdot e^2$.

find that loan officers who are compensated based on the volume of loans increase origination at the expense of quality. In an experimental setting, Schweitzer, Ordonez, and Douma (2004) find that people with unmet short-term goals are more likely eventually to engage in unethical behavior.

The analysis predicts that companies should become more lenient with a manager's performance when short-termism is an important concern. The project development setting in the paper is particularly well suited to study managerial compensation in research-intensive industries, but the general economic mechanism should extend to other situations. The model predicts that long-term contract should have a low turnover and high level of compensation that is deferred over time. These predictions are consistent with recent evidence on the duration of executive compensation in innovative firms. Baranchuk, Kieschnick, and Moussawi (2014) find that a combination of tolerance to failure and long-term compensation induces CEOs to adopt more innovative policies: firms with high R&D encourage innovation by combining deferred compensation and short-term protection. In fact, this pattern appears to be more pronounced in innovative firms, and the combination of these contractual features is different in firms that pursue innovation from that in the ones that do not. Moreover, the level of compensation is positively correlated with the degree of takeover protection (entrenchment) and the length of vesting

periods. Taken together, all these stylized facts are consistent with the idea that firms wishing to pursue innovation provide CEOs with more incentives, longer vesting periods, and more protection from termination (lower turnover). Ederer and Manso (2013) find (in a controlled laboratory setting) that tolerance for early failure and reward for long-term success are effective in motivating innovation and that termination undermines incentives to innovate.

Additional evidence on the duration of incentives is provided by Gopalan et al. (2014), who develop a measure of executive pay duration and quantify the mix of short- and long-term compensation.¹³ The comparative statics in Section 4.3 predict that the level and the duration of compensation should be positively correlated. This prediction is consistent with the evidence in Gopalan et al. (2014) who look at the correlation between pay duration and firm characteristics. They find that the duration of payments is positively correlated with growth opportunities, long-term assets, and R&D intensity: these are firms with intangible assets where the possibility of short-termism considered here is more likely to be severe and short-term manipulation more difficult to detect (small $\zeta_b - \zeta_g$ in the context of the model). They find that pay duration is positively correlated with managerial entrenchment and total compensation. Alternative theories of managerial entrenchment, based on CEO bargaining and rent seeking, can explain the positive correlation between entrenchment and compensation but cannot explain the positive correlation with pay duration: a manager who has bargaining power over the board tends to prefer a compensation package that is not deferred as much.

The same underlying problems of short-term manipulation appear in the context of venture capital financing. To secure financing, entrepreneurs may have incentives to sacrifice long-term value to increase short-term performance. Kaplan and Strömberg (2003) document that many venture capital contracts make the vesting of the entrepreneur's shares contingent on long-term measures of consumer satisfaction or patent approvals: these contingencies are similar to the deferred compensation in the model. In addition, venture capital contracts specify the allocation of cash-flow and control rights in different states of the world and commonly specify state-contingent control rights that allows for removal of the entrepreneur for performance. For example, many venture capitalist (VC) contracts incorporate provisions under which the VC can only vote for all owned shares if some performance measure, such as EBIT, is below some threshold. Other contracts specify that VCs obtain additional board members if the net worth falls below some prespecified value.

The main idea behind all the previous mechanisms is to increase the ability of the VC to remove the entrepreneur or terminate the project after low performance. I can explicitly incorporate the distinction between control and cash-flow rights by considering the case in which termination is not contractible

¹³ Their measure is related to the traditional measure of duration used in bond markets. They measure pay duration as a value-weighted average of the vesting period of the different components of the compensation package.

ex ante. In this context, I can reinterpret the termination of the project as a state-contingent allocation of control. However, if termination arises through an allocation of control rights, then it is not clearly reasonable to assume that the agent can commit ex ante to replace the manager (entrepreneur) unless it is ex post optimal to do it. I can easily extend the model to the case in which termination is optimal ex post: so I can interpret random termination as the outcome of an allocation of control rights to the principal.¹⁴ If I assume that randomization is not contractible, so randomization arises only through the principal equilibrium strategy then the contract only specifies the payments to the manager (the allocation of cash-flow rights) and the right to terminate the manager (the allocation of control rights). Even if the principal cannot commit to terminate the manager ex ante, the main qualitative features of the contract remain the same as those in the baseline model. In this case, the principal solves an optimal stopping problem and the liquidation threshold is pinned down by the traditional value-matching and smooth-pasting conditions.

Whether randomization is contractible or not does not affect the qualitative aspects of the contracts; moreover, the termination threshold is increasing in the difficulty to determine short-term manipulation, and this decreases the probability of termination and can be interpreted as more entrepreneurial control rights. In addition, I find that the randomization threshold is higher when randomization is contractible, which implies that the probability of liquidation in the stationary region of the contract is higher when random liquidation strategies are noncontractible. In general, this means that the VC would like to commit to a higher duration of the contract, which commitment can be partially achieved by increasing the difficulty of terminating the project after low performance.

Kaplan and Strömberg (2003) distinguish between rights that are contingent on performance (performance vesting) and rights that are contingent on the entrepreneur staying at the company (time vesting). In the case of time vesting, the entrepreneur's compensation is contingent on the board's decision to retain them instead of explicit benchmarks. Although highly stylized, the stationary region in the optimal contract captures many of the qualitative features of the time vesting contract: expected compensation is constant, and incentives are partly driven by the decision to terminate the manager (entrepreneur). In addition, Kaplan and Strömberg (2003) also find that contracts in industries characterized by high volatility, R&D, and small size rely more on the replacement of the entrepreneur by the board (time vesting) to induce pay performance sensitivity rather than on explicit performance benchmarks (performance vesting). This is consistent with the predictions of the model, as these are industries where short-termism might be more difficult to detect and long-term performance more difficult to assess. In terms of control

¹⁴ The appendix provides the formal analysis.

rights, a positive but low probability of termination (a low θ_t) can be interpreted as a situation in which the VC has some but not all the required control of the board to terminate the entrepreneur. In fact, Kaplan and Strömberg (2003) find that state-contingent control – where neither the VC nor the entrepreneur has control and outside directors are pivotal – are common in pre-revenue R&D ventures, and the allocation of control requires that less successful ventures transfer the control from the entrepreneur to the VC.

6. Conclusion

The main purpose of this paper has been to analyze the effect of managerial short-termism on the dynamic provision of incentives and its effect on turnover. Like in previous multitasking models, high-powered incentives, though necessary to stimulate effort, also generate incentives for a manager to manipulate performance. When managers can manipulate performance over time—that is, they affect the timing of cash flow by increasing short-term performance at the expense of long-term performance—the optimal contract relies less on the dynamic provision of incentives and becomes more stationary: this has implications for turnover and the role of termination in dynamic settings.

The main assumption is that quality only can be assessed by observing the performance of the project over time. The principal considers the trade-off between the rents they provide to the manager and the amount of deferred compensation necessary to prevent manipulation. The optimal contract keeps the manager's continuation value high, thereby increasing their skin-in-the-game. Doing so reduces the amount of deferred compensation. This trade-off between monitoring (more deferred compensation) and the level of compensation is reminiscent of the literature on efficiency wages. In the efficiency wage literature, workers receive an above-market wage to make layoffs more costly for them, thereby reducing the amount of monitoring necessary to increase effort. Similarly, in my model, the only way to provide incentives to exert effort, while still giving the manager high rents (not punishing them by reducing the continuation value) is to use random termination. The problem is that the incentives for the manager to manipulate performance are too high when termination is predictable. One way of sidestepping this problem is to make termination unpredictable, and this is optimal.

The analysis has implications for the dynamic provision of incentives and, in particular, the duration of employment relationships and worker turnover. The expected duration is an increasing function of the difficulty of assessing quality. We should observe longer contracts (or lower turnover rates) in jobs or projects in which workers can easily increase performance measures by reducing quality and quality is more difficult to observe. The model predicts a negative correlation between turnover rates and pay duration that is consistent

with patterns observed in managerial compensation contracts in innovative firms.

As mentioned before, the model implies that contracts are more stationary in the presence of managerial short-termism; this is a form of linearity over time that is analogous to the linearity over outcomes in Holmström and Milgrom (1987). Most dynamic principal-agent models (particularly models with limited liability) predict that contracts should be highly nonstationary and should depend on the history of performance in a complicated way. However, we observe that contracts are often much simpler than that, and one of the messages of this paper is that one reason for this is that highly dynamic contracts increase incentives to game the system and engage in managerial short-termism. This is the point that Jensen (2001, 2003) has informally made, calling for the elimination of several nonlinearities in the budgeting process. Introducing managerial short-termism in a dynamic contracting model (and in particular in a model with limited liability) is challenging because of the persistent effect of managerial myopia: the project development model setting that has been analyzed here is tractable and has allowed us to obtain a clean characterization of the optimal contract. Although stark, this project development setting captures some of the main incentive problems that we face in many managerial situations and highlights economic mechanisms that should be relevant for other, more complex settings.

Appendix

A. Solution Optimal Contract

Proof of Lemma 2. I prove the proposition using the saddle point theorem (Luenberger 1968, theorem 2, p. 221). Let U^{**} be the payment process characterized by (δ, \bar{U}) in Lemma 2. Let the Lagrangian be defined by

$$\mathcal{L} \equiv \int_0^\infty \left(-e^{-(r+\zeta_g)s} + (\bar{\mu} - P'(x))e^{-(\gamma+\zeta_g)s} - \eta e^{-(\gamma+\zeta_b)s} \right) dU_s^+ - \bar{\mu}(c+w) + \eta w. \tag{A1}$$

Defining $\mu \equiv \bar{\mu} - P'(x)$, I obtain

$$\mathcal{L} = \int_0^\infty \left(-e^{-(r+\zeta_g)s} + \mu e^{-(\gamma+\zeta_g)s} - \eta e^{-(\gamma+\zeta_b)s} \right) dU_s^+ - \bar{\mu}(c+w) + \eta w. \tag{A2}$$

For fixed multipliers (μ, η) , the gradient of \mathcal{L} with respect to U^+ in direction H is

$$\nabla \mathcal{L}(U^+; H) = \int_0^\infty \left(-e^{-(r+\zeta_g)s} + \mu e^{-(\gamma+\zeta_g)s} - \eta e^{-(\gamma+\zeta_b)s} \right) dH_s. \tag{A3}$$

By construction, both constraints are binding under the conjectured contract U^{**} . Hence, if I can find $(\bar{\mu}^*, \eta^*) > 0$ such that $\nabla \mathcal{L}(U^{**}; H) \leq 0$ in all feasible directions H (that is, for all H such that the process $U^{**} + \epsilon H$ is nondecreasing for ϵ sufficiently small), then $(U^{**}, \bar{\mu}^*, \eta^*)$ is a saddle point of \mathcal{L} . Noting that H must be nondecreasing for any $t \neq \delta$, I have $\nabla \mathcal{L}(U^{**}; H) \leq 0$ if and only if

$$-e^{-(r+\zeta_g)t} + \mu e^{-(\gamma+\zeta_g)t} - \eta e^{-(\gamma+\zeta_b)t} \leq 0, \tag{A4}$$

$$-e^{-(r+\zeta_g)\delta} + \mu e^{-(\gamma+\zeta_b)\delta} - \eta e^{-(\gamma+\zeta_g)\delta} = 0. \tag{A5}$$

Let $G(t, \mu, \eta) \equiv -e^{-(r+\zeta_g)t} + \mu e^{-(\gamma+\zeta_g)t} - \eta e^{-(\gamma+\zeta_b)t}$ and $\Delta\zeta \equiv \zeta_b - \zeta_g$. I can find multipliers (μ^*, η^*) that solve the system of equations $G(\delta, \mu^*, \eta^*)=0$ and $G_t(\delta, \mu^*, \eta^*)=0$.

$$\eta^* = \frac{\gamma - r}{\Delta\zeta} e^{(\gamma-r+\Delta\zeta)\delta} = \frac{\gamma - r}{\Delta\zeta} \left(\frac{c+w}{w} \right)^{\frac{\gamma-r+\Delta\zeta}{\Delta\zeta}} \tag{A6}$$

$$\mu^* = \frac{\gamma - r + \Delta\zeta}{\Delta\zeta} e^{(\gamma-r)\delta} = \frac{\gamma - r + \Delta\zeta}{\Delta\zeta} \left(\frac{c+w}{w} \right)^{\frac{\gamma-r}{\Delta\zeta}} \tag{A7}$$

We can see from (A6) and (A7) that $\eta^* > 0$ and $\mu^* > 1$. Hence, $\tilde{\mu}^* = \mu^* + P'(x) > 0$ given the hypothesis $P'(x) \geq -1$.

Replacing in G , I obtain

$$G(t, \mu^*, \eta^*) = e^{-rt} \left[\frac{\gamma - r + \Delta\zeta}{\Delta\zeta} e^{-(\gamma-r)(t-\delta)} - \frac{\gamma - r}{\Delta\zeta} e^{-(\gamma-r+\Delta\zeta)(t-\delta)} - 1 \right]. \tag{A8}$$

From (A8), it suffices to show that for all $x \in \mathbb{R}$

$$\frac{\gamma - r + \Delta\zeta}{\Delta\zeta} e^{-(\gamma-r)x} - \frac{\gamma - r}{\Delta\zeta} e^{-(\gamma-r+\Delta\zeta)x} - 1 \leq 0.$$

Rearranging terms, I obtain the condition

$$\frac{\gamma - r + \Delta\zeta}{\Delta\zeta} - \frac{\gamma - r}{\Delta\zeta} e^{-\Delta\zeta x} - e^{(\gamma-r)x} \leq 0. \tag{A9}$$

Using the inequality $e^{ax} \geq 1 + ax$ and (A9), I obtain

$$\frac{\gamma - r + \Delta\zeta}{\Delta\zeta} - \frac{\gamma - r}{\Delta\zeta} e^{-\Delta\zeta x} - e^{(\gamma-r)x} \leq \frac{\gamma - r + \Delta\zeta}{\Delta\zeta} - \frac{\gamma - r}{\Delta\zeta} - 1 = 0,$$

which means that $G(t, \mu^*, \eta^*) \leq 0$ for all $t \geq 0$. Moreover, by construction $G(\delta, \mu^*, \eta^*)=0$. Thus, conditions (A4) and (A5) are satisfied. Finally, I obtain the expected payoff by replacing the optimal policy in the objective function, and I verify concavity by simple differentiation.

A.1 Verification of optimality

Lemma 3. Let V be any solution to $\mathcal{D}V - rV = 0$. If some $\hat{w} \in [0, \frac{\lambda c}{\gamma})$ such that $V''(\hat{w}) \leq 0$, then $V''(w) \leq 0$ for all $w \in [\hat{w}, \frac{\lambda c}{\gamma})$.

Proof. Looking for a contradiction, suppose some $w^\dagger > \hat{w}$ such that $V''(w^\dagger) > 0$. By continuity of V'' , there exist some $y \in (\hat{w}, w^\dagger)$ such that $V''(y) = 0$ and $V^{(3)}(y) > 0$. The third derivative of V is given by

$$V^{(3)}(w) = \frac{\gamma}{\lambda c - \gamma w} V''(w) + \frac{1}{\lambda c - \gamma w} \left\{ (\gamma - r - \lambda - \Delta) V''(w) + (\lambda + \Delta) \Pi''(w) \right\}. \tag{A10}$$

Using concavity of Π and (A10), I obtain that $V^{(3)}(y) = \frac{(\lambda + \Delta) \Pi''(y)}{\lambda c - \gamma y} < 0$. This is a contradiction. The case with $V''(\hat{w}) = 0$ follows as $V^{(3)}(\hat{w}) < 0$ implies that $V''(\hat{w} + \epsilon) < 0$ for $\epsilon > 0$ sufficiently close to zero. ■

Let $V(w, z)$ be the solution to the initial value problem $rV(x) = \mathcal{D}V(x)$, $V(z) = P_*(z)$ where

$$P(w_*) = E \left[e^{-r\tau} \mathbf{1}_{\{\tau \leq T\}} \Pi(w_*) + e^{-rT} \mathbf{1}_{\{\tau > T\}} L \mid W_\tau = w_* \right] \tag{A11}$$

$$= \frac{(\lambda + \Delta)w_* \Pi(w_*) + (\lambda c - \gamma w_*)L}{(r + \lambda + \Delta - \gamma)w_* + \lambda c}.$$

I can solve for V in closed form

$$V(w, z) = (\lambda + \Delta)(\lambda c - \gamma w)^\psi \int_z^w (\lambda c - \gamma x)^{-(\psi+1)} \Pi(x) dx + P_*(z) \left(\frac{\lambda c - \gamma w}{\lambda c - \gamma z} \right)^\psi, \tag{A12}$$

where $\psi \equiv \frac{r+\lambda+\Delta}{\gamma} > 0$. It turns out that, if I maximize (A12) with respect to z , I obtain that smooth fit ($P'(w_*) = 0$) is just the right condition I need to find the threshold w_* .

Lemma 4. Let $V(w, z)$ be given by (A12). For all $w \in [0, \lambda c/\gamma]$, $w_* = \arg \max_z V(w, z)$ if and only if $V_{ww}(w_*, w_*) = 0$. Moreover, under the assumptions in Proposition 1 such $w_* \in [0, \lambda c/\gamma]$ exists and it is the unique solution to

$$\Pi'(w_*) = \frac{r + \lambda + \Delta - \gamma}{(r + \lambda + \Delta - \gamma)w_* + \lambda c} \Pi(w_*) - \frac{r + \lambda + \Delta - \gamma}{(r + \lambda + \Delta - \gamma)w_* + \lambda c} \frac{r + \lambda + \Delta}{\lambda + \Delta} L. \tag{A13}$$

Proof. I first solve for $V_{ww}(w_*, w_*) = 0$.

$$V_{ww}(w, z) = -\frac{r + \lambda + \Delta}{(\lambda c - \gamma w)^2} [\gamma V(w, z) + (\lambda c - \gamma w)V_w(w, z)]$$

$$+ \frac{\lambda + \Delta}{(\lambda c - \gamma w)^2} [\gamma \Pi(w) + (\lambda c - \gamma w)\Pi'(w)].$$

Replacing $V_w(w, z)$, I obtain

$$(\lambda c - \gamma w)^2 V_{ww}(w, z) = (r + \lambda + \Delta)[(r + \lambda + \Delta - \gamma)V(w, z) - (\lambda + \Delta)\Pi(w)]$$

$$+ (\lambda + \Delta)[\gamma \Pi(w) + (\lambda c - \gamma w)\Pi'(w)].$$

Evaluating at $(w, z) = (w_*, w_*)$, I obtain

$$(\lambda c - \gamma w)^2 V_{ww}(w_*, w_*) = (r + \lambda + \Delta) \frac{(r + \lambda + \Delta - \gamma)(\lambda c - \gamma w_*)L - \lambda c(\lambda + \Delta)\Pi(w_*)}{(r + \lambda + \Delta - \gamma)w_* + \lambda c}$$

$$+ (\lambda + \Delta)[\gamma \Pi(w_*) + (\lambda c - \gamma w_*)\Pi'(w_*)].$$

Hence, after some straightforward algebra, $V_{ww}(w_*, w_*) = 0$ if and only if

$$\Pi'(w_*) = \frac{r + \lambda + \Delta - \gamma}{(r + \lambda + \Delta - \gamma)w_* + \lambda c} \Pi(w_*) - \frac{r + \lambda + \Delta - \gamma}{(r + \lambda + \Delta - \gamma)w_* + \lambda c} \frac{r + \lambda + \Delta}{\lambda + \Delta} L. \tag{A14}$$

Next, for any w , I maximize $V(w, z)$ with respect to z . The first-order condition is

$$(\lambda c - \gamma w)^\psi (\lambda c - \gamma w_*)^{-\psi-1} \left[-(\lambda + \Delta)\Pi(w_*) + P'_*(w_*)(\lambda c - \gamma w_*) + (r + \lambda + \Delta)V(w_*, w_*) \right] = 0, \tag{A15}$$

where

$$P'_*(w_*) = \frac{(\lambda + \Delta)(\Pi'(w_*)w_* + \Pi(w_*)) - \gamma L}{(r + \lambda + \Delta - \gamma)w_* + \lambda c} - \frac{r + \lambda + \Delta - \gamma}{(r + \lambda + \Delta - \gamma)w_* + \lambda c} V(w_*, w_*)$$

Using this expression, I can write (A15) as

$$(\lambda + \Delta)w_*[(\lambda c - \gamma w_*)\Pi'(w_*) - (r + \lambda + \Delta)\Pi(w_*)] + [\gamma(\lambda c - \gamma w_*) + (r + \lambda + \Delta)^2 w_*]V(w_*, w_*) - (\lambda c - \gamma w_*)\gamma L = 0$$

Replacing $V(w_*, w_*)$ and after some algebra, I obtain the condition

$$\Pi'(w_*) = \frac{r + \lambda + \Delta - \gamma}{(r + \lambda + \Delta - \gamma)w_* + \lambda c} \Pi(w_*) - \frac{r + \lambda + \Delta - \gamma}{(r + \lambda + \Delta - \gamma)w_* + \lambda c} \frac{r + \lambda + \Delta}{\lambda + \Delta} L. \tag{A16}$$

Comparing Equations (A14) and (A16), I obtain the desired conclusion. Finally, I can verify that w_* is indeed the maximizer of $V(w, z)$. Following the same computations I used to solve the first-order conditions, I obtain

$$\text{sign } V_z(w, z) = \text{sign } H(z),$$

where

$$H(z) \equiv (r + \lambda + \Delta - \gamma)[z\Pi'(z) - \Pi(z)] + \lambda c\Pi'(z) + (r + \lambda + \Delta - \gamma) \frac{r + \lambda + \Delta}{\lambda + \Delta} L.$$

Differentiating, I obtain that

$$H'(z) = [(r + \lambda + \Delta - \gamma)z + \lambda c]\Pi''(z) < 0. \tag{A17}$$

Hence, from (A16) and (A17), I have $V_z(w, z) > 0$ for $z < w_*$ and $V_z(w, z) < 0$ for $z > w_*$. Thus, $V(w, z)$ attains its maximum at $z = w_*$. Moreover, (A17) implies that w_* is the unique solution to $H(z) = 0$. The only step left is to show that a solution to $H(z) = 0$ exists. First, noting that $\lim_{z \downarrow 0} \Pi(z) = -\infty$ and $\lim_{z \downarrow 0} \Pi'(z) = \infty$, I can verify that $\lim_{z \downarrow 0} H(z) > 0$. As $H(z)$ is a continuous function of z and $H'(z) < 0$ I find a unique solution if and only if $H(\lambda c/\gamma) < 0$, which corresponds to condition (16) in Proposition 1. ■

Lemma 5. Assume that $w_* \in (0, \lambda c/\gamma)$ satisfies Equation (A13). Then the function P satisfies the variational inequality

$$\max \left(wP'(w) - P(w) + L, \mathcal{D}P(w) - rP(w), -P'(w) - 1 \right) = 0. \tag{A18}$$

Proof. By construction, $wP'(w) - P(w) + L = 0$ for $w \leq w_*$, $\mathcal{D}P(w) - rP(w) = 0$ for $w \in (w_*, w^*)$, and $P'(w) = -1$ for $w \geq w^*$. From lemma 3 P is concave, so $P'(w) \geq -1$ and $wP'(w) - w + L \leq 0$ for all w . Hence, it only remains to show that $\mathcal{D}P(w) - rP(w) \leq 0$. Let $\Phi(w) := \mathcal{D}P(w) - rP(w)$. As P is C^2 at w_* , I can differentiate Φ and obtain

$$\Phi'(w) = (\gamma - r - \lambda - \Delta)P'(w) + (\gamma w - \lambda c)P''(w) + (\lambda + \Delta)\Pi'(w).$$

Case 1. For $w \leq w_*$

$$\begin{aligned} \Phi'(w) &= (\gamma - r - \lambda - \Delta) \frac{P(w_*) - L}{w_*} + (\lambda + \Delta)\Pi'(w) \\ &= (\lambda + \Delta) \left[\Pi'(w) - \frac{r - \gamma + \lambda + \Delta}{(r - \gamma + \lambda + \Delta)w_* + \lambda c} \Pi(w_*) \right] + (r + \lambda + \Delta - \gamma) \frac{L}{w_*} \\ &= (\lambda + \Delta) [\Pi'(w) - \Pi'(w_*)] + (r + \lambda + \Delta - \gamma) \left[\frac{1}{w_*} - \frac{r + \lambda + \Delta}{(r + \lambda + \Delta - \gamma)w_* + \lambda c} \right] L \geq 0. \tag{A19} \end{aligned}$$

Where (A19) follows from Π concavity and $w_* \leq \lambda c/\gamma$. Therefore, as $\mathcal{D}P(w_*) - rP(w_*) = 0$ I have $\mathcal{D}P(w) - rP(w) \leq 0$ for all $w \leq w_*$.

Case 2. For $w > w_*$

$$\begin{aligned} \Phi(w) &= \Phi(w) - \Phi(w^*) \\ &= (w - w^*) \left[r + \lambda + \Delta - \gamma + (\lambda + \Delta) \frac{\Pi(w) - \Pi(w^*)}{w - w^*} \right] \\ &\leq (w - w^*) [r + \lambda + \Delta - \gamma + (\lambda + \Delta) \Pi'(w^*)]. \end{aligned} \tag{A20}$$

Where (A20) follows from the concavity of Π . $\Phi(w) = 0$ for all $w \in [w_*, w^*]$ imply that $\Phi'(w) = 0$ for $w \in (w_*, w^*)$. Hence,

$$\lim_{w \uparrow w^*} \Phi'(w) = -(\gamma - r - \lambda - \Delta) + (\gamma w - \lambda c) P''(w^* -) + (\lambda + \Delta) \Pi'(w^*) = 0$$

So

$$(\lambda + \Delta) \Pi'(w^*) = \gamma - r - \lambda - \Delta + (\lambda c - \gamma w^*) P''(w^* -) \tag{A21}$$

Replacing (A21) in (A20), I obtain that for $w > w^*$

$$\Phi(w) \leq (w - w^*) [r + \lambda + \Delta - \gamma + (\lambda + \Delta) \Pi'(w^*)] = (w - w^*) (\lambda c - \gamma w^*) P''(w^* -) \leq 0.$$

Proof of Proposition 1 For any termination policy θ , let $\Theta_t = \int_0^t \theta_s ds$. I can write the principal expected payoff as

$$P_0 = \int_0^\infty e^{-(r+\lambda+\Delta)t - \Theta_t} ((\lambda + \Delta)\Pi(W_t) + \theta_t L) dt - \int_0^\infty e^{-(r+\lambda+\Delta)t - \Theta_t} dU_t^-. \tag{A22}$$

Using the HJB equation, I obtain

$$\begin{aligned} e^{-(r+\lambda+\Delta)t - \Theta_t} P(W_t) &= P(W_0) + \int_0^t e^{-(r+\lambda+\Delta)s - \Theta_s} [\mathcal{D}P(W_s) - rP(W_s) \\ &\quad + \theta_s W_s P(W_s) - \theta_s P(W_s) + \theta_s L - (\lambda + \Delta)\Pi(W_s)] ds \\ &\quad - \int_0^t e^{-(r+\lambda+\Delta)s - \Theta_s} P'(W_s) dU_s^- \\ &\leq P(W_0) - \int_0^t e^{-(r+\lambda+\Delta)s - \Theta_s} (\lambda + \Delta)\Pi(W_s) ds \end{aligned} \tag{A23}$$

Where inequality (A23) follows from lemma 5. Because P is bounded on $[0, w^*]$, linear on (w^*, ∞) , and $\gamma \leq r + \lambda + \Delta$, I can conclude that $\lim_{t \rightarrow \infty} e^{-(r+\lambda+\Delta)t - \Theta_t} P(W_t) = 0$. It follows that

$$P(W_0) \geq \int_0^\infty e^{-(r+\lambda+\Delta)s - \Theta_s} ((\lambda + \Delta)\Pi(W_s) + \theta_s L) ds.$$

Thus, P is an upper bound for the principal expected payoff under any admissible contract. In the case of the conjectured optimal contract I have

$$\mathcal{D}P(W_s) - rP(W_s) + \theta_s W_s P(W_s) - \theta_s P(W_s) + \theta_s L = 0.$$

and

$$\int_0^t e^{-(r+\lambda+\Delta)s - \Theta_s} P'(W_s) dU_s^- = \int_0^t e^{-(r+\lambda+\Delta)s - \Theta_s} dU_s^-,$$

Hence, the conjectured optimal contract attains the upper bound. ■

Proof of Proposition 2

Proof. I prove the proposition showing that whenever the conditions in the proposition are satisfied I have $P'_-(w^s) \geq 0$, where $P'_-(w^s)$ is the left derivative of P evaluated at w^s .

Differentiating Equation (14), I obtain

$$(r + \lambda + \Delta - \gamma)P'(w) = (\gamma w - \lambda c)P''(w) + \lambda \Pi'(w). \tag{A24}$$

Evaluating in w^s , I obtain

$$(r + \lambda + \Delta - \gamma)P'_-(w^s) = \lambda \Pi'(w^s). \tag{A25}$$

Given that $(r + \lambda + \Delta) > 0$, I have that a necessary and sufficient condition for $P'_-(w^s) \geq 0$ is that $\Pi'(w^s) \geq 0$. Replacing w^s , I obtain that the latter inequality is satisfied iff

$$\phi \left(\frac{\lambda + \gamma}{\gamma} \right)^{\phi+1} \geq (\phi + 1) \left(\frac{\lambda + \gamma}{\gamma} \right)^\phi.$$

We arrive to inequality (18) by replacing ϕ and rearranging terms. ■

Proof of Proposition 6

Proof. We can write condition (A32) as

$$(\lambda + \Delta)(\Pi(w) - P(w)) - \lambda(Y_g - w) \geq \lambda c P'(w) - \lambda P(w), \tag{A26}$$

where I have just subtracted $\lambda P(w)$ at both sides. From the HJB equation, I have that, for all $w \in [w_*, W_0]$,

$$\Pi'(w) - P'(w) = \frac{-(\gamma - r)P'(w) + (\lambda c - \gamma w)P''(w)}{\lambda + \Delta} < 0.$$

Thus, I have that

$$(\lambda + \Delta)(\Pi(w) - P(w)) - \lambda(Y_g - w) \geq (\lambda + \Delta)(\Pi(w_*) - P(w_*)) - \lambda(Y_g - w_*). \tag{A27}$$

We also have that

$$\lambda c P''(w) - \lambda P'(w) < 0.$$

So

$$\lambda c P'(w) - \lambda P(w) \leq \lambda c P'(w_*) - \lambda P(w_*). \tag{A28}$$

Combining (A26)–(A27), I arrive at the sufficient condition

$$(\lambda + \Delta)(\Pi(w_*) - P(w_*)) - \lambda(Y_g - w_*) \geq \lambda c P'(w_*) - \lambda P(w_*), \tag{A29}$$

which after rearranging terms give us

$$\Delta[\Pi(w_*) - P(w_*)] - \lambda[Y_g - w_* - \Pi(w_*)] \geq \lambda c P'(w_*)$$

■

Proof of Lemma 3 From Equation (A16), w_* is given by the unique solution to $H(z, \phi) = 0$ where

$$H(z, \phi) \equiv (r + \lambda + \Delta - \gamma)[z\Pi'(z, \phi) - \Pi(z, \phi)] + \lambda c \Pi'(z, \phi) + (r + \lambda + \Delta - \gamma) \frac{r + \lambda + \Delta}{\lambda + \Delta} L.$$

Where by definition $\phi = (\gamma - r) / \Delta \zeta$. I have the following derivative

$$\begin{aligned} \Pi_{\Delta \zeta}(z, \Delta \zeta) &= -\log\left(1 + \frac{b}{w}\right) (b+w)^{\phi+1} w^{-\phi} \frac{\partial \phi}{\partial \Delta \zeta} \\ \Pi_{w \Delta \zeta}(z, \Delta \zeta) &= \left(\frac{b+w}{w}\right)^\phi \left[\frac{b}{w} - \log\left(1 + \frac{b}{w}\right) + \phi \frac{b}{w} \log\left(1 + \frac{b}{w}\right) \right] \frac{\partial \phi}{\partial \Delta \zeta}. \end{aligned}$$

Hence, $\Pi_{\Delta \zeta}(z, \phi) > 0$ and using the inequality $x > \log(1+x)$, $\Pi_{w \Delta \zeta}(z, \phi) < 0$. Accordingly, $H_{\Delta \zeta}(z, \phi) > 0$ so $w_*(\Delta \zeta)$ is decreasing in $\Delta \zeta$.

Proof of Proposition 4 Let's define $\Delta\zeta \equiv \zeta_g - \zeta_b$. Given the parametric restriction, $W_0 < w^s$ the optimal W_0 is interior and $P'(W_0)=0$. Thus, using the HJB equation I have that

$$P(W_0) = \frac{\lambda + \Delta}{r + \lambda + \Delta} \Pi(W_0).$$

Using implicit differentiation and using $P'(W_0)=0$, I obtain

$$W'_0(\Delta\zeta) = \frac{P_{\Delta\zeta}(W_0) - \frac{\lambda + \Delta}{r + \lambda + \Delta} \Pi_{\Delta\zeta}(W_0)}{\frac{\lambda + \Delta}{r + \lambda + \Delta} \Pi'(W_0)}.$$

Differentiating the HJB equation with respect to w and evaluating at W_0 , I obtain

$$(\lambda + \Delta)\Pi'(W_0) = \gamma(w^s - W_0)P''(W_0) < 0.$$

Thus, W_0 is decreasing in $\Delta\zeta$ iff

$$P_{\Delta\zeta}(W_0) > \frac{\lambda + \Delta}{r + \lambda + \Delta} \Pi_{\Delta\zeta}(W_0). \tag{A30}$$

From Lemma 4 I have that

$$P(W_0) = \max_{w_* \geq 0} V(W_0, w_*),$$

where $V(w, w_*)$ is given by Equation (A12). By the envelope theorem,

$$P_{\Delta\zeta}(W_0) = \frac{\lambda + \Delta}{\gamma} (w^s - W_0)^\psi \int_{w_*}^{W_0} (w^s - x)^{-(\psi+1)} \Pi_{\Delta\zeta}(x) dx + \left(\frac{w^s - W_0}{w^s - w_*} \right)^\psi \frac{\lambda + \Delta}{r + \lambda + \Delta} \Pi_{\Delta\zeta}(w_*).$$

Finally, given that $\Pi_{\Delta\zeta} > 0$ and $\Pi_{\Delta\zeta} w < 0$ (proof Proposition 3), I have that

$$P_{\Delta\zeta}(W_0) > \frac{\lambda + \Delta}{\gamma} (w^s - W_0)^\psi \Pi_{\Delta\zeta}(W_0) \int_{w_*}^{W_0} (w^s - x)^{-(\psi+1)} dx + \left(\frac{w^s - W_0}{w^s - w_*} \right)^\psi \frac{\lambda + \Delta}{r + \lambda + \Delta} \Pi_{\Delta\zeta}(W_0) = \frac{\lambda + \Delta}{r + \lambda + \Delta} \Pi_{\Delta\zeta}(W_0),$$

which yields the desired result.

Proof of Proposition 5 The expected contract duration is

$$E(T|\tau > T) = \frac{1}{\gamma} \left[\log \left(\frac{w^s - w_*}{w^s - W_0} \right) + \frac{w_*}{w^s - w_*} \right]$$

Take $\Delta\zeta' > \Delta\zeta$ and let w'_*, W'_0, T' and w_*, W_0, T be the respective solutions. From Lemma 3 and Proposition 4 I have that $w'_* < w_*$ and $W'_0 \leq W_0$ (with equality only when $T = \infty$). Then,

$$\begin{aligned} E(T|\tau > T) &= \frac{1}{\gamma} \left[\log \left(\frac{w^s - w'_*}{w^s - W_0} \right) + \log \left(\frac{w^s - w_*}{w^s - w'_*} \right) + \frac{w_*}{w^s - w_*} \right] \\ &= \frac{1}{\gamma} \left[\log \left(\frac{w^s - w'_*}{w^s - W_0} \right) + \frac{w'_*}{w^s - w'_*} + \log \left(\frac{w^s - w_*}{w^s - w'_*} \right) + \frac{(w_* - w'_*)w^s}{(w^s - w_*)(w^s - w'_*)} \right] \\ &> \frac{1}{\gamma} \left[\log \left(\frac{w^s - w'_*}{w^s - W'_0} \right) + \frac{w'_*}{w^s - w'_*} + \log \left(\frac{w^s - w_*}{w^s - w'_*} \right) + \frac{(w_* - w'_*)w^s}{(w^s - w_*)(w^s - w'_*)} \right] \\ &= E(T'|\tau > T') + \frac{1}{\gamma} \left[\log \left(\frac{w^s - w_*}{w^s - w'_*} \right) + \frac{(w_* - w'_*)w^s}{(w^s - w_*)(w^s - w'_*)} \right] \end{aligned}$$

Finally, noting that $\log\left(\frac{w^s - w_*}{w^s - w'_*}\right)$ is convex as a function of w'_* , I have that

$$\begin{aligned} \log\left(\frac{w^s - w_*}{w^s - w'_*}\right) + \frac{(w_* - w'_*)w^s}{(w^s - w_*)(w^s - w'_*)} &\geq \frac{w'_* - w_*}{w^s - w_*} + \frac{(w_* - w'_*)w^s}{(w^s - w_*)(w^s - w'_*)} \\ &= \frac{(w_* - w'_*)w'_*}{(w^s - w_*)(w^s - w'_*)} > 0, \end{aligned}$$

so $E(T|\tau > T) > E(T'|\tau > T')$.

B. Optimality of Effort

In this appendix, I provide sufficient conditions for full effort to be optimal. In the absence of any effort, it is not necessary to provide incentives to the manager; this means that the manager's continuation value evolves according to

$$\dot{W}_t = \gamma W_t.$$

The fact that the manager is not being incentivized to exert effort also implies that the manager has no incentives to manipulate performance, so it is not necessary to defer compensation. To verify the optimality of effort, I have to compare the principal expected payoff of implementing effort to the expected payoff of no effort at all. The HJB equation implies that the principal finds it optimal to implement effort only if the following condition is satisfied:

$$rP(w) \geq \gamma w P'(w) + \lambda[Y_g - w - P(w)]. \tag{A31}$$

We can use the HJB equation to replace $rP(w)$ and simplify the previous condition, and I arrive at the following condition

$$\Delta[\Pi(w) - P(w)] \geq \lambda[Y_g - w - \Pi(w)] + \lambda c P'(w), \quad w \in [w_*, W_0]. \tag{A32}$$

One important property of Equation (A32) is that the inequality becomes tighter when the continuation value is low: a low continuation value makes inducing effort more costly, and accordingly it is sufficient to check the previous condition only at the threshold w_* . From the previous argument, I find the following sufficient condition for effort optimality:

Proposition 6. Under the assumptions in Proposition 1, a sufficient condition for optimality of effort is

$$\Delta[\Pi(w_*) - P(w_*)] \geq \lambda[Y_g - w_* - \Pi(w_*)] + \lambda c P'(w_*). \tag{A33}$$

First, note that Condition (A33) can be computed directly from $\Pi(w_*)$ – I can easily verify this assertion by computing $P(w_*)$ and $P'(w_*)$ in terms of $\Pi(w_*)$ and its derivatives. Thus, Condition (A33) imposes a direct condition on the primitive parameters, and this condition can be verified numerically without the need to solve the differential equation for $P(w)$.

If the liquidation value is low enough, then Condition (A33) may be violated when the continuation value is sufficiently low; in this case, the optimal contract may require zero effort. The problem now is that new technical complications arise because the evolution of the continuation value is given by

$$\dot{W}_t = \gamma W_t - \lambda c \mathbf{1}_{\{e_t > 0\}}. \tag{A34}$$

The right-hand side in Equation (A34) is not convex in effort. For example, suppose that at time t it is optimal to implement no effort (that is, $e_t = 0$), and that this is optimal when the continuation value reaches some lower boundary $\underline{w} > 0$. If this is true, then as soon as the continuation value hits \underline{w} , its derivative would be $\dot{W}_t > 0$, which means that at time $t + dt$ the continuation value would

be $W_{t+dt} > \underline{w}$; but this also implies that at time $t + dt$, the derivative would be $\dot{W}_{t+dt} < 0$, which again would bring the value of W_t back to \underline{w} . As soon as the continuation value reaches the lower threshold, the level of effort starts chattering between no effort and full effort. Mathematically, this means that an optimal control (in the traditional sense) fails to exist, and this happens because the evolution of the continuation value fails to be convex. One possibility for dealing with this technical problem would be to consider a larger set of admissible controls, namely, the set of relaxed controls (Davis 1993).¹⁵ The way of interpreting these controls is that instead of implementing a fixed level of effort $e_t \in [0, 1]$ at time t , the optimal contract randomizes over the set $[0, 1]$ according to some distribution $v_t(de)$.¹⁶ The optimal contract mixes between $e_t = 0$ and $e_t = 1$ with probability v_t as soon as the continuation value reaches \underline{w} ; the mixing probability is chosen such that $\dot{W}_t = 0$. The previous approach is probably unnecessarily technical. Rather than looking at relaxed effort policies, I sidestep the previous issue by considering a strictly convex cost of effort.

C. Convex Cost of Effort

C.1 Manager Incentive Compatibility

I start deriving the incentive compatibility constraint for effort. As before, let \bar{W}_t^i , $i \in \{g, b\}$ be the expected payoff from a good and bad project. The manager's expected payoff if chooses effort $\tilde{e} = \{\tilde{e}_t\}_{t \geq 0}$ and delivers a bad project at time τ^b is

$$\bar{W}_t = \int_t^{\tau^b} e^{-(\lambda+\gamma)(s-t) - \int_t^s (\tilde{e}_u + \theta_u) du} ((\lambda + \tilde{e}_s) \bar{W}_s^g - c(\tilde{e}_s)) ds + e^{-(\lambda+\gamma)(\tau^b-t) - \int_t^{\tau^b} (\tilde{e}_u + \theta_u) du} \bar{W}_{\tau^b}^b.$$

Differentiating with respect to time, I obtain that the continuation value follows the differential equation

$$\frac{d}{dt} \bar{W}_t = (\gamma + \lambda + \tilde{e}_t + \theta_t) \bar{W}_t - (\lambda + \tilde{e}_t) \bar{W}_t^g + c(\tilde{e}_t)$$

Similarly, the expected payoff if the manager follows the recommended level of effort e_t (and does not deliver a bad project) evolves according to

$$\frac{d}{dt} W_t = (\gamma + \lambda + e_t + \theta_t) W_t - (\lambda + e_t) \bar{W}_t^g + c(e_t).$$

Let's define $Z_t = W_t - \bar{W}_t$. Then I have that

$$\frac{d}{dt} Z_t = (\gamma + \lambda + \theta_t) Z_t - e_t (\bar{W}_t^g - W_t) + \tilde{e}_t (\bar{W}_t^g - \bar{W}_t) + c(e_t) - c(\tilde{e}_t).$$

Adding and subtracting $\tilde{e}_t W_t$, I obtain

$$\frac{d}{dt} Z_t = (\gamma + \lambda + \tilde{e}_t + \theta_t) Z_t - e_t (\bar{W}_t^g - W_t) + \tilde{e}_t (\bar{W}_t^g - W_t) + c(e_t) - c(\tilde{e}_t).$$

Integrating this differential equation forward between time zero and time τ^b I find that

$$Z_t = \int_t^{\infty} e^{-(\lambda+\gamma)(s-t) - \int_t^s (\tilde{e}_u + \theta_u) du} \left\{ [e_s (\bar{W}_s^g - W_s) - c(e_s)] - [\tilde{e}_s (\bar{W}_s^g - W_s) - c(\tilde{e}_s)] \right\} ds + e^{-(\lambda+\gamma)(\tau^b-t) - \int_t^{\tau^b} (\tilde{e}_u + \theta_u) du} (W_{\tau^b} - \bar{W}_{\tau^b}^b).$$

From here I see that $Z_t \geq 0$ for all alternative strategies (\tilde{e}, τ^b) if and only if $e_t = \arg \max_e \left\{ (\bar{W}_t^g - W_t) e - c(e) \right\}$ and $W_t \geq \bar{W}_t^b$.

¹⁵ Another possibility is to consider ϵ -optimal controls that alternate between zero effort and full effort over short periods. These controls can be designed to approximate the relaxed control arbitrarily close.

¹⁶ The set of relaxed controls is the set of measurable functions $v: [0, \infty) \rightarrow \mathcal{P}([0, 1])$, where $\mathcal{P}([0, 1])$ is the set of probability measures on $[0, 1]$ (Davis 1993, Definition 43.2, p. 148).

C.2 Principal Problem

We can apply the results in Lemma 2 immediately by noting that I can replace the condition $\bar{W}^g = W + c$ by $\bar{W}^g = W + c'(e)$. The principal's profit is

$$\Pi(w, e) = Y_g - (c'(e) + w)^{\phi+1} w^{-\phi}, \tag{A35}$$

and the evolution of the continuation value is given by

$$\dot{W}_t = (\gamma + \theta_t)W_t + c(e_t) - (\lambda + e_t)c'(e_t). \tag{A36}$$

Hence, the optimal effort and termination probability solves the optimal control problem

$$\max_{e_t \geq 0, \theta_t \geq 0} \int_0^\infty e^{-(r+\lambda)t - \int_0^t (e_s + \theta_s) ds} \left((\lambda + e_t)\Pi(W_t, e_t) + \theta_t L \right) dt$$

subject to the evolution of the continuation value in (A36). Using the auxiliary variable $\Lambda_t = \int_0^t (e_s + \theta_s) ds$ I can write the optimization problem in the following form which is more suitable for an application of the maximum principle

$$\max_{e_t \geq 0, \theta_t \geq 0} \int_0^\infty e^{-(r+\lambda)t - \Lambda_t} \left((\lambda + e_t)\Pi(W_t, e_t) + \theta_t L \right) dt$$

subject to

$$\begin{aligned} \dot{W}_t &= (\gamma + \theta_t)W_t + c(e_t) - (\lambda + e_t)c'(e_t) \\ \dot{\Lambda}_t &= e_t + \theta_t, \Lambda_0 = 0. \end{aligned}$$

I formulate the problem as an optimal control problem in Mayer form so the Hamiltonian is concave (Cesari 1983). For this purpose, I introduce the state variable P_t given by

$$\dot{P}_t = (r + \lambda + e_t + \theta_t)P_t - (\lambda + e_t)\Pi(W_t, e_t) - \theta_t L,$$

where P_t is the principal payoff. The optimal control now is to maximize P_0 subject to the odes for W and P . The Hamiltonian for this problem is

$$H = \mu_0 \left((\gamma + \theta)w + c(e) - (\lambda + e)c'(e) \right) + \mu_1 \left((r + \lambda + e + \theta)p - (\lambda + e)\Pi(w, e) - \theta L \right), \tag{A37}$$

where μ_0 and μ_1 are the (present value) costate variables. Assuming an interior solution for e_t , the first-order condition is

$$\mu_1 \left(p - \Pi(w, e) - (\lambda + e)\Pi_e(w, e) \right) - \mu_0(\lambda + e)c''(e) = 0.$$

Similarly, the first-order condition for θ is

$$\theta = \begin{cases} 0 & \text{if } \mu_0 w - \mu_1 L + \mu_1 p < 0 \\ [0, \infty) & \text{if } \mu_0 w - \mu_1 L + \mu_1 p = 0 \\ \infty & \text{if } \mu_0 w - \mu_1 L + \mu_1 p > 0. \end{cases}$$

The evolution of the adjoint variables is

$$\dot{\mu}_{0t} = (r + \lambda + e_t - \gamma)\mu_{0t} - (\lambda + e_t)\Pi_w(W_t, e_t), \mu_{00} = 0 \tag{A38}$$

$$\dot{\mu}_{1t} = (r + \lambda + e_t + \theta_t)\mu_{1t} - (r + \lambda + e_t + \theta_t), \mu_{11} = -1. \tag{A39}$$

Accordingly, $\mu_{1t} = -1$, so I can write the previous first-order conditions as

$$\Pi(w, e) + (\lambda + e)\Pi_e(w, e) - p - \mu_0(\lambda + e)c''(e) = 0.$$

Similarly, the first-order condition for θ is

$$\theta = \begin{cases} 0 & \text{if } \mu_0 w + L - p < 0 \\ [0, \infty) & \text{if } \mu_0 w + L - p = 0 \\ \infty & \text{if } \mu_0 w + L - p > 0. \end{cases}$$

The evolution of the adjoint variables is The second order condition is satisfied if the Hamiltonian is jointly concave in (e, θ) . As the Hamiltonian is linear in θ , it is enough to verify $H_{ee} \leq 0$, where

$$H_{ee} = 2\Pi_e(w, e) + (\lambda + e)\Pi_{ee}(w, e) - \mu_0 c''(e) - \mu_0(\lambda + e)c'''(e)$$

and

$$\Pi_e(w, e) = -(\phi + 1)(c'(e) + w)^\phi w^{-\phi} c''(e)$$

$$\Pi_{ee}(w, e) = -(\phi + 1)\phi(c'(e) + w)^{\phi-1} w^{-\phi} (c''(e))^2 - (\phi + 1)(c'(e) + w)^\phi w^{-\phi} c'''(e).$$

A sufficient condition for $H_{ee} \leq 0$ is that $c'''(e) \geq 0$, which is satisfied for example if $c(e) = c_0 \cdot e + c_1 \cdot e^2$.

The intensity of termination θ_t enters linearly into the optimization problem and is unbounded above. Hence, if the probability of termination is positive, it must correspond to a singular arc. Let's define the switching function $\Gamma(t) \equiv L - P_t + \mu_{0t} W_t$; in any interval of time in which $\theta_t > 0$, the switching function must be constant. This means that $\dot{\Gamma}(t) = 0$ or

$$\dot{\mu}_{0t} W_t + \mu_{0t} \dot{W}_t - \dot{P}_t = 0$$

Replacing the differential equations for W , μ_0 , and μ_1 , I obtain

$$(r + \lambda + e_t + \theta_t)(\mu_{0t} W_t - P_t) - (\lambda + e_t)\Pi_w(W_t, e_t)W_t + \mu_{0t} \left(c(e_t) - (\lambda + e_t)c'(e_t) \right) + (\lambda + e_t)\Pi(W_t, e_t) + \theta_t L = 0.$$

Using the equality $L + \mu_{0t} W_t - P_t = 0$, I obtain the condition

$$-(r + \lambda + e_t)L - (\lambda + e_t)\Pi_w(W_t, e_t)W_t + \mu_{0t} \left(c(e_t) - (\lambda + e_t)c'(e_t) \right) + (\lambda + e_t)\Pi(W_t, e_t) = 0. \quad (A40)$$

Equation (A40) reduces to condition (16) if I consider the case with linear cost, $c(e_t) = c \cdot e_t$, and maximal effort $\bar{e} = \Delta$. Because it must be the case that $\dot{\Gamma}(t) = 0$ at all times in a singular arc it also must be the case that $\ddot{\Gamma}(t) = 0$. Differentiating the previous expression once again and replacing the first-order conditions, I obtain

$$\ddot{\Gamma}(t) = -\dot{e}_t \left[L - P_t + (\Pi_w(W_t, e_t) + (\lambda + e)\Pi_{we}(W_t, e_t))W_t \right] - (\lambda + e_t)\Pi_{ww}(W_t, e_t)W_t \dot{W}_t + \dot{\mu}_{0t} \left(c(e_t) - (\lambda + e_t)c'(e_t) \right) = 0$$

This condition can be satisfied by setting $\dot{W}_t = \dot{\mu}_{0t} = \dot{P}_t = 0$ (in which case $\dot{e}_t = 0$). Hence, I obtain that in a singular arc

$$\theta_t = \theta^* = \frac{(\lambda + e^*)c'(e^*) - c(e^*) - \gamma w^*}{w^*}$$

$$\mu_{0t} = \mu_0^* = \frac{(\lambda + e^*)\Pi_w(w^*, e^*)}{r + \lambda + e^* - \gamma}$$

$$P_t = P^* = \frac{(\lambda + e^*)\Pi(w^*, e^*) + \theta^* L}{r + \lambda + e^* + \theta^*},$$

where (w^*, e^*) solve

$$\begin{aligned} \mu_0^*(\lambda + e^*)c''(e^*) &= \Pi(w^*, e^*) + (\lambda + e^*)\Pi_e(w^*, e^*) - P^* \\ \mu_0^* \left(c(e^*) - (\lambda + e^*)c'(e^*) \right) + (\lambda + e^*)\Pi(w^*, e^*) &= (\lambda + e^*)\Pi_w(w^*, e^*)w^* + (r + \lambda + e^*)L \end{aligned}$$

To obtain the expressions in the text, I note that the costate variable μ_{0t} corresponds to the derivative $P'(W_t)$ evaluated at the optimal path W_t ; this is a standard result in optimal control theory connecting the maximum principle and dynamic programming.

The maximized Hamiltonian is linear in (w, p) and so automatically concave; accordingly it satisfies Arrow's sufficient condition for optimality. In addition, because this is a singular control problem, the Legendre-Clebsch condition $\partial \ddot{\Gamma}(t)/\partial \theta \geq 0$ (Cesari 1983, p. 170) needs to be checked when I solve the model. Differentiating the first-order condition for e with respect to time and θ I obtain $\partial \dot{e}/\partial \theta$, and replacing in $\partial \ddot{\Gamma}(t)/\partial \theta$, I obtain

$$\begin{aligned} \frac{\partial}{\partial \theta} \ddot{\Gamma}(t) &= \frac{P^* - \Pi_w(w^*, e^*)w^*}{\mu_0^* - 2\Pi_e(w^*, e^*) - (\lambda + e^*)\Pi_{ee}(w^*, e^*)} \\ &\quad \left[L - P^* + (\Pi_w(w^*, e^*) + (\lambda + e^*)\Pi_{we}(w^*, e^*))w^* \right] - (\lambda + e^*)\Pi_{ww}(w^*, e^*)(w^*)^2. \quad (\text{A41}) \end{aligned}$$

The only step left is to determine the initial conditions W_0 and the principal payoff P_0 . If $\Gamma(t)$ is nondecreasing in time, then $T^* \equiv \inf\{t > 0 \mid (W_t, \mu_{0t}, P_t) = (w^*, \mu_0^*, P^*)\}$ such that $\theta_t = 0$ for $t < T^*$ and $\theta_t = \theta_*$ for $t \geq T^*$. Let $\tilde{W}(t, W_0, P_0)$, $\tilde{\mu}_0(t, W_0, P_0)$, and $\tilde{P}(t, W_0, P_0)$ be the solution of the differential equations at time t given the initial conditions W_0 and P_0 ; by construction, the solution must solve the system of equations

$$\begin{aligned} \tilde{W}(T^*, W_0, P_0) &= w^* \\ \tilde{\mu}_0(T^*, W_0, P_0) &= \mu_0^* \\ \tilde{P}(T^*, W_0, P_0) &= P^*. \end{aligned}$$

We can solve the previous system numerically using reverse shooting. For any conjectured T^* I can solve the differential equations backward in time starting at (w^*, μ_0^*, P^*) . I can find T^* iterating until I find T^* such that $|\mu_0(0)| \leq \epsilon$ for some stopping rule $\epsilon > 0$. Once I have found the candidate solution, I verify that condition (A41) is satisfied at $t \in [0, T^*]$ (If the condition is satisfied at T_* then it is trivially satisfied at all $t > T^*$).

D. Noncontractible Termination

This section provides the analysis of the case in which randomization is not contractible. The events leading to termination are not explicitly specified in the contract, and randomization arises only through the principal equilibrium strategy: the contract only specifies the payments to the manager (the allocation of cash-flow rights) and the right to terminate the manager (the allocation of control rights). Notice that this does not mean that contract is renegotiation-proof. In a renegotiation-proof contract, the principal has no commitment whatsoever and is not able to commit to renegotiate any aspect of the contract – neither payments nor termination.

If termination is not contractible, the liquidation threshold must satisfy the indifference condition $P(w_*) = L$ together with the traditional smooth-pasting condition $P'(w_*) = 0$ in optimal stopping problems. With noncontractible termination, I need to distinguish the manager's beliefs about the principal termination intensity, $\hat{\theta}_t$, from the termination intensity that the principal actually uses,

θ_t . In equilibrium both intensities coincide. Given the principal profit function $\Pi(\cdot)$, the principal value function and termination strategy is the maximal solution to the HJB equation

$$rP(w) = \max_{\theta \geq 0} \left\{ \mathcal{D}P(w) + \hat{\theta}(w)wP'(w) + \theta[L - P(w)] \right\}. \tag{A42}$$

From the manager’s perspective, the evolution of the continuation value depends on the expected intensity, $\hat{\theta}_t$, not the actual intensity, θ_t . Accordingly, the indirect effect of stochastic termination, coming from the drift of the continuation value, $W_t P'(W_t)$, is determined by $\hat{\theta}_t$. Only the direct effect, $P(W_t) - L$, depends on θ_t . In addition to the standard incentive compatibility constraints for the manager, the contract must satisfy the following incentive compatibility constraint for the principal:

$$\theta(w) = \begin{cases} 0 & \text{if } P(w) > L \\ [0, \infty) & \text{if } P(w) = L \\ \infty & \text{if } P(w) < L. \end{cases}$$

where $\theta_t = \infty$ means that the manager is fired immediately. This immediately yields the boundary condition $P(w_*) = L$, which corresponds to the standard indifference condition for mixed strategies. Moreover, if the value function is increasing, then the intensity of termination must be zero when $W_t > w_*$. Accordingly, the principal value function is the maximal solution to the initial value problem

$$rP(w) = \mathcal{D}P(w) + (\lambda + \Delta)[\Pi(w) - P(w)], \quad P(w_*) = L. \tag{A43}$$

The threshold w_* is pinned down using the smooth-pasting condition $P'(w_*) = 0$.¹⁷ The derivation of the smooth-pasting condition is similar to the derivation in the case with commitment. The termination intensity, θ_t , is such that the continuation value, W_t , has an absorbing barrier at w_* , which means that $\theta_t = \mathbf{1}_{\{W_t = w_*\}}(\lambda c/w_* - \gamma)$. If I compute $P(w_*)$ and combine it with the boundary condition $P(w_*) = L$, I obtain the condition

$$L = \frac{\lambda + \Delta}{r + \lambda + \Delta} \Pi(w_*). \tag{A44}$$

This indifference condition is intuitive. The right-hand side is the terminating the manager immediately, L , whereas the left-hand side is the benefit if the principal never terminate the manager and continue with a constant promised value of w_* . In equilibrium, both must be equal if the principal is using a mixed termination strategy.

In general, if a solution to Equation (A44) exists, then there are two solutions. A solution exists if $\max_{w \geq 0} \Pi(w) \geq (1 + \frac{r}{\lambda + \Delta})L$. The only case in which Equation (A44) has a unique solution is the knife-edge case, where $\max_{w \geq 0} \Pi(w) = (1 + \frac{r}{\lambda + \Delta})L$. The largest solution to Equation (A44) corresponds to the renegotiation proof contract.¹⁸ On the other hand, the smallest solution to Equation (A44), w_* , has the property that $\Pi'(w_*) > 0$, which implies that the value function P is convex in a neighborhood of w_* . Because of this convexity, the principal’s profit in this latter contract is strictly higher than the profits in the former renegotiation-proof contract, so it provides the maximal solution to Equation (A43).

¹⁷ Like before, the upper threshold is given by $w^* = \inf\{w > 0 : P'(w) = -1\}$.

¹⁸ Let w_{RP} be the largest solution. This solution has the property that $\Pi'(w_{RP}) < 0$. If I differentiate Equation (A43) and replace the smooth-pasting condition $P'(w_{RP}) = 0$, P is concave and attains its maximum at w_{RP} .

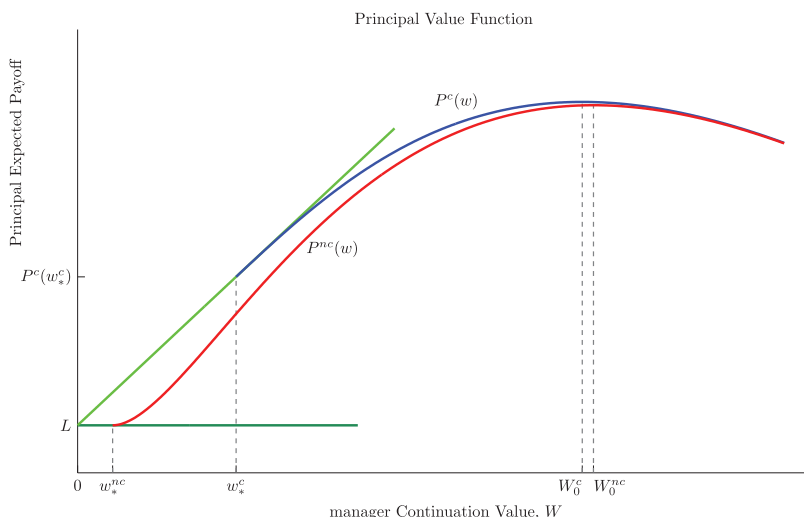


Figure D1

Optimal contract without contractible randomization

The superscript *c* indicates the solution when stochastic termination is contractible, whereas the superscript *nc* indicates the solution with noncontractible randomization. When the principal cannot commit to terminate the manager, the firing threshold w_*^{nc} is lower than the threshold with commitment w_*^c .

Proposition 7. Suppose that

$$\frac{(\lambda + \Delta) \max_{w \in [0, \lambda c / \gamma]} \Pi(w)}{r + \lambda + \Delta} > L. \tag{A45}$$

Let

$$w_* \equiv \min \left\{ w : L = \frac{\lambda + \Delta}{r + \lambda + \Delta} \Pi(w) \right\},$$

and let P be the solution to (A43). Then $W_0 = \arg \max_{w \geq 0} P(w)$, and the optimal contract when random termination is not contractible is given by

1. A cumulative payment process U_t^+ described by (8)–(10); and
2. A stochastic termination time T with intensity $\theta(W_t) = \mathbf{1}_{\{W_t = w_*\}} \left(\frac{\lambda c}{w_*} - \gamma \right)$.

The expected payoff for the principal under the optimal contract is given by $P(W_0)$.

Equation (A45) provides conditions for existence of a solution to Equation (A44) in the interval $[0, \lambda c / \gamma]$.

Figure D1 illustrates the differences between the value function in the cases with and without contractible randomization. Whether randomization is contractible or not does not affect the qualitative aspects of the contracts; moreover, the termination threshold is increasing in the difficulty to determine short-term manipulation, and this decreases the probability of termination. The randomization threshold is higher when randomization is contractible, which implies that the probability of liquidation in the stationary region of the contract is higher when random liquidation strategies are noncontractible.

References

- Agarwal, S. and I. Ben-David. Forthcoming. Loan prospecting and the loss of soft information. *Journal of Financial Economics*.
- Austin, R. D. 2001. The effect of time pressure on quality in software development: An agency model. *Information Systems Research* 12:195–207.
- Baker, G. P. 1992. Incentive contracts and performance measurement. *Journal of Political Economy* 100: 598–614.
- Baranchuk, N., R. Kieschnick and R. Moussawi. 2014. Motivating innovation in newly public firms. *Journal of Financial Economics* 111:578–88.
- Bebchuk, L. A. 2009. *Pay without performance: The unfulfilled promise of executive compensation*. Cambridge, MA: Harvard University Press.
- Benmelech, E., E. Kandel, and P. Veronesi. 2010. Stock-based compensation and ceo (dis)incentives. *Quarterly Journal of Economics* 125:1769–820.
- Bergemann, D. and U. Hege. 1998. Venture capital financing, moral hazard, and learning. *Journal of Banking & Finance* 22:703–35.
- . 2005. The financing of innovation: learning and stopping. *RAND Journal of Economics* 36: 719–52.
- Biais, B., T. Mariotti, J.-C. Rochet, and S. Villeneuve. 2010. Large risks, limited liability, and dynamic moral hazard. *Econometrica* 78:73–118.
- Bolton, P. and D. S. Scharfstein. 1990. A theory of predation based on agency problems in financial contracting. *American Economic Review* 80:93–106.
- Bonatti, A. and J. Horner. 2011. Collaborating. *American Economic Review* 101:632–63.
- Burns, N. and S. Kedia. 2006. The impact of performance-based compensation on misreporting. *Journal of Financial Economics* 79:35–67.
- Burt, D. N. 1984. *Proactive procurement: The key to increased profits, productivity, and quality*. Englewood Cliffs, NJ: Prentice-Hall.
- Cesari, L. 1983. *Optimization theory and applications*. New York, NY: Springer.
- Cornelli, F. and O. Yosha. 2003. Stage financing and the role of convertible securities. *Review of Economic Studies* 70:1–32.
- Davis, M. H. 1993. *Markov models and optimization*. London: Chapman & Hall.
- DeMarzo, P. M. and M. J. Fishman. 2007. Optimal long-term financial contracting. *Review of Financial Studies* 20:2079–128.
- DeMarzo, P. M. and Y. Sannikov. 2006. Optimal security design and dynamic capital structure in a continuous-time agency model. *Journal of Finance* 61:2681–724.
- Dowie, M. 1977. Pinto madness. *Mother Jones*, September and October.
- Ederer, F. and G. Manso. 2013. Is pay for performance detrimental to innovation? *Management Science* 59: 1496–513.
- Edmans, A., X. Gabaix, T. Sadzik, and Y. Sannikov. 2012. Dynamic CEO compensation. *Journal of Finance* 67:1603–47.
- Fernandes, A. and C. Phelan. 2000. A recursive formulation for repeated agency with history dependence. *Journal of Economic Theory* 91:223–47.
- Fong, K. 2009. Evaluating skilled experts: Optimal scoring rules for surgeons. Working Paper, Stanford University.

- Gee, M. and K. Tzioumis. 2013. Nonlinear incentives and mortgage officers' decisions. *Journal of Financial Economics* 107:436–53.
- Gerardi, D. and L. Maestri. 2012. A principal-agent model of sequential testing. *Theoretical Economics* 7:425–63.
- Gopalan, R., T. Milbourn, F. Song, and A. V. Thakor. 2014. Duration of executive compensation. *Journal of Finance* 69:2777–817.
- Hartman-Glaser, B., T. Piskorski, and A. Tchistyi. 2012. Optimal securitization with moral hazard. *Journal of Financial Economics* 104:186–202.
- He, Z. 2012. Dynamic compensation contracts with private savings. *Review of Financial Studies* 25:1494–549.
- Heider, F. and R. Inderst. 2012. Loan prospecting. *Review of Financial Studies* 25:2381–415.
- Hellmann, T. 1998. The allocation of control rights in venture capital contracts. *Rand Journal of Economics* 29:57–76.
- Holmström, B. and P. Milgrom. 1987. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica* 55:303–28.
- . 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization* 7:24–52.
- Hopenhayn, H. A. and J. P. Nicolini. 1997. Optimal unemployment insurance. *Journal of Political Economy* 105 412–38.
- Inderst, R. and H. M. Mueller. 2010. Ceo replacement under private information. *Review of Financial Studies* 23:2935–69.
- Inderst, R. and M. Ottaviani. 2009. Misselling through agents. *American Economic Review* 99:883–908.
- Jensen, M. C. 2001. Budgeting is broken - let's fix it. *Harvard Business Review* 79:94–101.
- . 2003. Paying people to lie: The truth about the budgeting process. *European Financial Management* 9:379–406.
- Kaplan, S. N. and P. Strömberg. 2003. Financial contracting theory meets the real world: An empirical analysis of venture capital contracts. *Review of Economic Studies* 70:281–315.
- Klein, N. 2016. The importance of being honest. *Theoretical Economics* 11:773–811.
- Larkin, I. 2014. The cost of high-powered incentives: Employee gaming in enterprise software sales. *Journal of Labor Economics* 32:199–227.
- Levitt, S. and C. Snyder. 1997. Is no news bad news? information transmission and the role of “early warning” in the principal-agent model. *RAND Journal of Economics* 28:641–61.
- Luenberger, D. G. 1968. *Optimization by vector space methods*. New York: John Wiley & Sons.
- Malamud, S., H. Rui, and A. B. Whinston. 2013. Optimal incentives and securitization of defaultable assets. *Journal of Financial Economics* 107:111–35.
- Manso, G. 2011. Motivating innovation. *Journal of Finance* 66:1823–60.
- Myerson, R. 2015. Moral hazard in high office and the dynamics of aristocracy. *Econometrica* 83:2083–126.
- Paté-Cornell, M. E. 1990. Organizational aspects of engineering system safety: The case of offshore platforms. *Science* 250:1210–16.
- Sannikov, Y. 2012. Moral hazard and long-run incentives. Working Paper, Princeton University.
- Schweitzer, M. E., L. Ordóñez, and B. Douma. 2004. Goal setting as a motivator of unethical behavior. *Academy of Management Journal* 47:422–32.
- Sinclair-Desgagné, B. 1999. How to restore higher-powered incentives in multitask agencies. *Journal of Law, Economics, and Organization* 15:418–33.

Stiglitz, J. E. and A. Weiss. 1983. Incentive effects of terminations: Applications to the credit and labor markets. *American Economic Review* 73:912–27.

Tian, X. and T. Y. Wang. 2014. Tolerance for failure and corporate innovation. *Review of Financial Studies* 27:211–55.

US Department of Energy. 2005. Department of Energy action plan lessons learned from the Columbia space shuttle accident and Davis-Besse reactor pressure-vessel head corrosion event, Technical Report, July.

Zhu, J. Y. Forthcoming. Myopic agency. *Review of Economic Studies*.