# Dynamic Capital Allocation and Managerial Compensation

Shiming Fu *

June 2016

**ABSTRACT**

This paper proposes a dynamic theory of capital budgeting and managerial compensation when the manager has private information about the arrival time and quality of investment projects and the manager can obtain private benefits from inefficient allocation of capital. The firm may forgo projects with positive net present value, and distort pay-performance sensitivity and capital allocation downwards. The paper shows that agency costs and firm policies vary monotonically with financial slack. The distortions are severe when manager performs poorly in the past. As the firm accumulates more financial slack, fewer projects will be forgone, and the optimal contract will provide steeper incentives and allocate capital more efficiently to the projects selected. All the distortions vanish when the firm has enough financial slack. Finally, the optimal project selection, capital allocation, and compensation can be implemented by a simple budgeting account mechanism.

JEL classification: D82, D86, G31, G35.

*Keywords:* Dynamic agency; Capital budgeting; Project choice; Pay-performance sensitivity; Liquidity.

---

*Simon School of Business, University of Rochester. Email: shiming.fu@simon.rochester.edu. This paper is part of my dissertation at Duke University. I am very grateful to S. Viswanathan for his generous guidance and insightful suggestions. I also thank R. Vijay Krishna, Barney Hartman-Glaser, Adriano Rampini, Felipe Varas, Ming Yang, and audience members at the Econometric Society World Congress 2015 for helpful comments. The most recent version of the paper is available here.

Corporate investment is financed overwhelmingly out of internal funds[1]. The most pervasive and important frictions affecting capital allocation within firms include information asymmetries and agency problems (Stein [2003]). In large firms, investment is usually delegated to division managers (DM) who have private information about projects[2]. Headquarters (HQ) or CEOs potentially face adverse selection problems in allocating capital. In principle, HQs could gather information, but when firms are large or complex this becomes quite costly [3]. Moreover, in large firms it is very costly to monitor the way DMs deploy capital. They may be able to deploy capital inefficiently and gain private benefit,engendering moral hazard in corporate investment. To ensure long-term health, firms must provide proper incentives for truthful information about investment projects and deter inefficient utilization of resources.

Another important aspect of capital budgeting and the associated investment process is their dynamic nature. While in practice HQ and DM always interact repeatedly over a sequence of investment opportunities, the literature has mainly focused on static environments. For instance, Harris and Raviv [1996] study the capital allocation with manager having "empire building" preference; Bernardo et al. [2004] analyze compensation and capital budgeting mechanisms when managers have private information on project quality. This paper instead explores the optimal capital allocation and compensation in a dynamic setting. In this dynamic model agency costs are endogenous and the past performance (or financial slack) of a division determines the optimal mechanism. In particular, the paper analyzes how (i) project choice, (ii) capital allocation rules, and (iii) performance-based pay vary with investment quality and the division's past performance.

To study these questions, this paper posits a simple environment with one risk-neutral principal (HQ) and one risk-neutral agent (DM). The HQ has unlimited access to capital, and investment opportunities arise stochastically over time. When the firm has a project, it

[1]Internal funds accounted for 70%-110% of total investment by U.S. nonfinancial corporations between 1994 and 2008 Brealey et al. [2011].

[2]Colom and Delmastro [2004] shows that in a sample of 438 Italian metalworking firms capital spending decisions are mostly delegated to divisional managers, especially when the task is urgent. Graham et al. [2014] survey more than 1000 CEOs and find that they delegate investment decisions more than any other major corporate policies.

[3]For example, in the model of Aghion and Tirole [1997], more decision-making powers are delegated down the corporate ladder when the principal is overloaded (such as when he manages a large firm).

may invest in a value-enhancing technology that boosts the return. However, the HQ has no information on whether or not the firm has a project to invest, or on project quality either. The firm has to rely on the DM's information to discover projects and their quality. When a project is reported, the HQ allocates capital according to the report. But the deployment of capital is under the DM's control, and may be diverted for the latter's personal consumption.

With asymmetric information and moral hazard, the capital budgeting process faces the trade-off between raising investment efficiency and reducing compensation (information rent). Since project quality is private information, the DM can always misreport it downward and then appropriate part of the capital. Hence, greater investment in any project requires increasing the compensation of the DM who runs all the higher quality projects. Compared with the frictionless benchmark, the optimal contract provides flatter incentives and induces lower investment level. This intuition mimics the classic agency problem in Laffont and Tirole [1986][4].

When the budgeting process is repeated, the agency conflicts will be endogenous and will vary over time. The intuition is as follows. In the dynamic setting, the HQ has flexibility to pay over time. The contract compensates the DM for information rents by promised future payments until his continuation value is sufficiently high. Continuation value is the present value of promised future payments to the DM which summarizes the division's past performance. Though the DM has to be compensated for information rents at the time when investment takes place, the HQ can form expectations of future information rents and extract them from the DM's continuation value. Only the unexpected part is compensated. In this sense, compensation over time relaxes the constraint on investment and incentive provision imposed by information rents.

However, adjusting compensation over time increases the risk of the division being liquidated. The DM is "punished" during the no-investment period by deducting his continuation value. And since the DM can only be "punished" to the extent that limited liability is binding, the agency conflicts will not disappear. Significantly, limited liability implies that the continuation value determines the severity of the agency issues, and hence the optimal policy.

[4]The agent in Laffont and Tirole [1986] exerts unobservable but costly effort. The incentive to shirk in their paper is analogous to the incentive to divert capital here.

After periods of good performance, the continuation value will be accumulated to a high level and the division is far from liquidation. It is optimal to design contracts with high-powered incentives, since the liquidation risk is low and a steeper contract induces more efficient investment both in the extensive and the intensive margin. The extreme case is at the payout boundary, where investment distortions disappear for all project types. After periods of bad performance (e.g. no investment return), the continuation value will be significantly lowered and the division is close to liquidation. It is optimal to design a contract with low-powered incentives, which determines that investment will be severely distorted. The extreme case is that low quality projects will be completely excluded from capital allocation.

The optimal contract can be implemented by a capital budgeting account mechanism with observable balance. Moral hazard means that the DM can either use the funds to invest or divert funds for personal consumption. The mechanism provides incentives by replenishing and depleting this account at designed rates so that its balance mimics the continuation value. Then the DM won't divert funds and will reveal truthful project information. In this sense, the continuation value that shapes the optimal contract can be measured by either financial slack or past performance.

The key feature of this model is that the agency cost varies with project types and financial slack. Consequently, the optimal capital allocation and pay-performance sensitivity exhibit monotone properties over these two dimensions under the technology and distribution assumptions. The paper has the following empirical implications for capital budgeting and incentive provisions. First, the DM is given a steeper incentive scheme and allocated more capital either when she reports a higher quality project or when the division has more financial slack. Second, with poor performance in the past and little financial slack, the division may forgo low-quality projects even when they have positive NPV. Third, when the DM receives cash compensation, capital allocations resume efficient levels for all projects. Using 4080 DM compensation contracts from ExecuComp, Alok and Gopalan [2013] finds that the DM pay is less sensitive to the performance of his division during periods of industry distress. This empirical finding is consistent with the implications of our model.

The paper proceeds as follows. Section 2 reviews the literature. Section 3 sets up the model. Section 4 summarizes the optimal contract in a frictionless environment. Section 5 provides necessary conditions for incentive compatibility (IC). Section 6 first derives the optimal contract heuristically under the obtained necessary conditions and then verifies that the proposed contract does maximize the HQ's expected payoff and satisfies IC. Section 7 characterizes the evolution of the DM's continuation value and the dynamic properties of optimal policies. Section 8 implements the contract using budgeting account mechanism. Section 9 concludes the paper.

# 1 The Literature

The static trade-off between private information and moral hazard is studied in Laffont and Tirole [1986]. They consider ways of providing incentives to regulate a monopoly contractor with unobserved cost efficiency and unobserved effort. They show that inducing greater effort in order to lower production cost (solving moral hazard problem) will make it more costly to induce information revelation (aggravating adverse selection). Effort in Laffont and Tirole [1986] is analogous to physical investment in this model. Adopting a dynamic framework, we show that this basic trade-off is endogenously determined and dynamically evolving, since the severity of the agency conflict is affected by the limited liability constraint.

Laffont and Tirole [1988] apply a two period auction model to analyze the regulation of an incumbent monopoly that could be replaced by another in the second period. Investment efficiencies are private information of the firm and are independent and identically distributed draws. Laffont and Tirole [1988] focus on how to design the selection rule and the incumbent's intertemporal incentive scheme in order to lower regulation costs. Depending on whether investments are transferable between the two firms, the optimal selection rule could favor either the incumbent or the entrant, and the optimal slope of the incumbent's incentive scheme could either be front-loaded or time invariant. Although this paper adopts similar information frictions and is also an i.i.d. model, we focus on how the agent's accumulative past performance changes optimal project selection, capital allocation and incentive slope.

More importantly, this paper shows that delaying payments, which is not considered in Laffont and Tirole [1988], is an important mechanism for easing information frictions.

This paper relates to the continuous-time dynamic contracting literature. While previous works mostly focus on moral hazard, this paper broadens the framework by considering both hidden action and hidden information. The key friction in the existing literature,, moral hazard, is exogenous and constant. For instance, the agent's ability to steal cash flows in DeMarzo and Sannikov [2006], or the private benefit of shirking in Biais et al. [2010] is an exogenous parameter. And these models mainly consider contracts implementing the first best action [5]. By contrast, in the present model the agency problem is endogenously determined by the level of the agent's continuation value. Therefore, optimal incentive and investment are time-varying, and investment is distorted below the first best level. The hidden information adds another dimension of dynamics over project quality, and generates interesting interactions with the dynamics of investment over time. Using the martingale representation approach over time dimension and the mechanism design approach (as in Myerson [1981]) over quality dimension, we have a rich setting to explore the dynamic implications.

This paper also relates to the literature on dynamic mechanism design. Malenko [2013] studies a dynamic capital budgeting model with costly state verification. The DM in his model has private investment information and empire-building preference. The HQ can verify the DM's reports at a fixed cost. High-quality projects are monitored and financed by the HQ, while the low quality projects are not monitored and financed out of the division's own budgeting account. The present model differs from Malenko [2013] chiefly in the possibility that the DM can also divert capital. Instead of monitoring, the HQ will design a compensation scheme to resolve the agency problems. This paper, unlike Malenko [2013], can characterize how capital allocation policy and the slope of compensation vary over time.

Eso and Szentes [2013] analyze a discrete-time dynamic auction where agent or buyer type is private information and types are correlated over time. They show that any fea-

[5] Exceptions include Zhu [2012] where the optimal contract can implement shirking either as a reward or punishment mechanism, and the moral hazard models with risk averse agent, for example, Sannikov [2008] or Gryglewicz and Hartman-Glaser [2013].

sible allocation gives the seller the same expected revenue as where the seller can observe orthogonalized agent types beyond the first period. So from the perspective of revenue, what matters is only the initial hidden information and its persistence; any subsequent orthogonal information is not compensated with information rents. Garrett and Pavan [2012] analyze a dynamic contracting model where a firm's cash flow is determined by its manager's hidden type (productivity), hidden effort, and an i.i.d. noise. The manager's various levels of productivity are correlated over time. In this setting, the dynamics of optimal policy are driven entirely by the persistence of the manager's initial private information, which is characterized by an impulse response function.

Although in this paper investment information is not persistent, the model clearly shows that all the orthogonalized future information must be remunerated by information rents. This key difference derives from the assumption of limited liability. In the mechanism posited by Eso and Szentes [2013], the buyer must make a large payment equal to all future expected payments in the case in which all future orthogonalized information can be observed. The implicit assumption is that the buyer has deep enough pockets to make this payment upfront. In the optimal contract of Garrett and Pavan [2012], the conditional information rents in all future periods are subtracted from the current-period compensation. If the manager is not patient enough or his outside option is not large enough, he could actually get negative cash payments. This paper makes the natural assumption that the DM is subject to limited liability. That is, negative cash payments are never possible. On this hypothesis, financial slack is crucial and investment and compensation policies are monotonically varying with the division's financial slack.

# 2 The Model

The model studies a large firm that consists of a HQ and a DM. The firms' investment decisions are delegated to the DM. Time is continuous and the horizon is infinite. Both the HQ and the DM are risk neutral. The HQ discounts the future cash flows at rate $r$. The DM is more impatient and he discounts future consumptions at rate $\gamma > r$.

## 2.1 Investment Opportunity

A distinct feature in this model is that the division's investment opportunities are sparse, arriving stochastically over time [6]. In particular, the projects arrive according to a Poisson process $\{\pi_t : t \geq 0\}$ with intensity $\lambda$, where $\pi_t$ is the total number of projects arrived before time $t$. The division only obtains an opportunity to invest when a project arrives. Another feature is that projects are heterogenous in quality which determines investment returns. We use the random variable $J_n$ ($n \in \mathbb{N}$) to characterize the quality of the $n$th arrived project. The qualities $\{J_n, n \in \mathbb{N}\}$ are i.i.d and uniformly distributed over the interval $\Theta = [\underline{\theta}, \bar{\theta}]$ with $\underline{\theta} > 0$. Moreover, project arrival is independent of project quality. The accumulated project quality $X_t = \sum_{n=1}^{\pi_t} J_n$ is therefore a Compound Poisson process with $dX_t \in \{0\} \times \Theta$. In short, the evolution of investment opportunities is summarized by $dX_t$: if no project arrives at time $t$, then $dX_t = 0$; and if a project with quality $\theta$ arrives at time $t$, then $dX_t = \theta$.

When a project of quality $\theta$ arrives at time $t$, the firm will have the opportunity to invest $k_t \geq 0$. This investment will increase the project return through a value enhancing technology $R : \mathbb{R}_+ \to \mathbb{R}_+$. The total return obtained from this project is $\theta + R(k_t)$. Capital $k_t$ depreciates completely at time $t$. After time $t$, the firm has to wait until next project arrives to make another investment. In short, the investment return during the infinitesimal time interval $(t, t+dt]$, denoted by $dY_t$, is determined by project arrival, project quality, and investment level in the relation of

$$dY_t = (dX_t + R(k_t))d\pi_t \tag{1}$$

The value enhancing technology $R$ exhibits decreasing return to scale, that is, $R' > 0$, $R'' < 0$. Moreover, to guarantee an interior optimal investment in the frictionless case, we also assume $R(0) = 0$, $\lim_{k\downarrow 0} R'(k) = \infty$, and $\lim_{k\uparrow\infty} R'(k) = 0$. It is easy to see that the first best investment $k^*$ satisfies $R'(k^*) = 1$.

[6] In discrete time models, for example Clementi and Hopenhayn [2006], firms are assumed to have investment opportunity every period.

## 2.2 Information Frictions and Mechanism Design

The agency issues arise due to two reasons. First, the DM has private information about the arrival time of projects and their qualities. The information about project arrival is crucial because the firm can discover an investment opportunity only if the DM reveals any arrived project. If the DM does not reveal an arrived project to the HQ, then firm won't get any return from such project and has to wait for future projects. The information about project quality is also important because the DM can potentially obtain private benefit from misreporting project quality. Second, the HQ can not observe the investment level in each project. Once the capital is allocated to the DM , he can either invest it in projects or divert it for personal consumption. These agency issues are modeled in a principal-agent contracting environment.

Because the HQ can use more accounting or real data to review the performance of investments, we adopt the assumption that all investment returns are observable [7]. By the revelation principle, it suffices to restrict attention to the truth-telling direct mechanisms. Because investment information can be backed out from the reported quality and the observed return, it is sufficient to consider only the mechanisms in which the DM reports whether a project arrives or not and the project quality if there is one. And the optimal mechanism is designed to induce truthful report.

A contract specifies capital allocation, compensation to the DM, and termination decisions as functions of the history of past reports and investment returns. We let a nonnegative and increasing process $K = \{K_t\}_{t\geq 0}$ describe the accumulative capital allocation [8]. Because of limited liability, the process $I = \{I_t\}_{t\geq 0}$, which describes the accumulative payments to the DM, also has to be nonnegative and increasing. In formulating and deriving the optimal contract, we assume that the DM cannot save. He will consume any cash immediately whenever it is paid. We show later that given the optimal contract, the DM optimally chooses not to save even if he is allowed to do so as long as the interest rate of saving is smaller than $\gamma$. The time at which liquidation of the division occurs is denoted as $\tau$. At any time $t$

[7] Return observability rules out the case where the DM lies about project arrival and invests by himself.

[8] To ease notation, we denote the increment of the process $\{K_t\}_{t\geq 0}$ by $k_t$, i.e., $k_t := dK_t$.

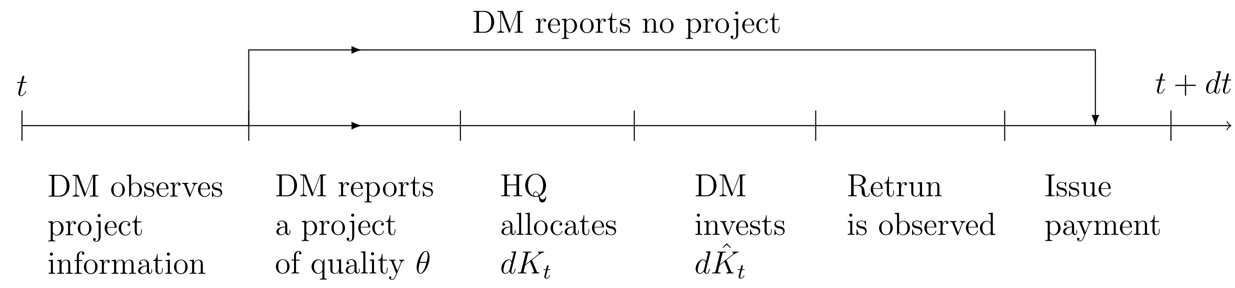


Figure 1: Timing

prior to liquidation, the sequence of events during the infinitesimal period $[t, t + dt]$ can be described as the following.

- The DM observes the project information $dX_t$, and makes a report $d\hat{X}_t \in \{0\} \times \Theta$.
- If no project is reported, no investment action will be taken, and no return will be generated.
- If any project is reported, then
  - the HQ decides the capital allocation $dK_t$ to the division;
  - the DM makes an investment $d\hat{K}_t \in [0, dK_t]$ [9];
  - the investment return $[dX_t + R(d\hat{K}_t)]dN_t$ is observed.
- The DM receives a nonnegative payment $dI_t$.
- The division is either liquidated or continued.

Formally, let us denote the probability space as $(\Omega, \mathcal{F}, P)$, and denote the filtration generated by $\{X_t\}_{t\geq 0}$ as $\{\mathcal{F}_t\}_{t\geq 0}$. Then the process of accumulative payments $I$ is a $\mathcal{F}$-adapted process and $\tau$ is an $\mathcal{F}$-stopping time. The process $K$ is $\mathcal{F}$-predictable. In particular, the capital allocation policy is a function $K_t : \Omega \times R_+ \times \Theta \to R_+$ which is $\mathcal{F}$-predictable with its first two arguments. Because the capital allocation affects the DM's deviation

[9]No saving directly implies that $d\hat{K}_t \leq dK_t$.

payoff, it has to be specified before the DM reports project information. So the process $K$ and its increments $dK$ is designed as $\mathcal{F}$-predictable. The DM's reporting strategy $\hat{X}$ and investment strategy $\hat{K}$ are $\mathcal{F}$-adapted, since the DM makes these decisions after obtaining project information.

Given a history of the DM's report, the capital allocation, and the realized return, the HQ can detect if the DM misreported project qualities and if the DM reported a project but none actually arrived. If the DM has never been detected a lie at termination, he receives an outside value zero. However, if the DM has ever been detected a lie, his outside option at termination will be dropped by $c > k^*$. This punishment reflects the fact that any detected lie will be a public record for the DM, making it harder for him to find a new job if the current contract is terminated. Let $C$ be a random variable that equals $c$ if a lie is ever detected and 0 otherwise. Once the contract is terminated, the HQ receives a liquidation value $L \geq 0$.

Given a contract $\{K, I, \tau\}$, a reporting strategy $\hat{X}$, and an investment strategy $\hat{K}$, the DM's expected payoff at time 0 is:

$$W_0 = E\left[\int_0^\tau e^{-\gamma t}\big(dI_t + dK_t - d\hat{K}_t\big) - e^{-r\tau}C\right]$$

The HQ's expected payoff at time 0 is:

$$P_0 = E\left[\int_0^\tau e^{-rt}(dY_t - dK_t - dI_t) + e^{-r\tau}L\right]$$

The expectation $E$ is taken with respect to the measure $P$. Moreover, we assume that terminating the contract is inefficient. This means $rL < \lambda[(\underline{\theta} + \bar{\theta})/2 + R(k^*) - k^*]$.

# 3 Frictionless Benchmark

Before investigating the optimal policies in the frictional environment, let us consider the first best case with no information friction. That is the HQ knows the arrival of projects and their qualities. Also, the HQ is able to implement any investment level. The frictionless contract has a simple form and is described in the following result.

**PROPOSITION 1.** *In the frictionless contract $\{K, I, \tau\}$ that delivers $W_0 \geq 0$ to the DM:*

*(a) investment is constant: $dK_t = k^*$ when there is a project, where $R'(k^*) = 1$.*

*(b) all compensation is paid out at time zero: $dI_0 = W_0$, $dI_t = 0$ for $t > 0$.*

*(c) the division of the firm will run without liquidation: $\tau = \infty$.*

# 4 Continuation Value and Incentive Compatibility

The technology and information structures imply that incentives in this model have the following features. First, no misreport of project information will be detected in equilibrium. To provide maximum punishment and relaxes the incentive constraints, the contract will be immediately terminated with zero payment issued to the DM if the HQ ever detects any misreport. And if his misreport is ever detected, the DM also has to bear the cost $c$, drop in his outside option. Since the largest private benefit from any possible misreporting cannot exceed $k^*$ which is smaller than $c$, it is never optimal for the DM to choose a reporting strategy in which any deviation can be detected.

Second, the DM tends to misreport project quality downward. Because the HQ does not observe either the project quality or the investment level, the DM can potentially report a lower quality and divert part (or all) of the allocated capital without being detected. Consider the case in which the DM reports a project of quality $\theta$ to be $\theta'$. According to the contract, capital $k_t(\theta')$ is allocated to the DM. The misreport won't be detected if the DM chooses an investment $d\hat{K}_t$ that satisfies $\theta + R(d\hat{K}_t) = \theta' + R(k_t(\theta'))$. Denote the investment level in this deviation as

$$k_t(\theta'; \theta) := d\hat{K}_t = R^{-1}[\theta' + R(k_t(\theta') - \theta] \tag{2}$$

For a downward misreport ($\theta' < \theta$), we have $k_t(\theta'; \theta) < k_t(\theta')$. Hence the DM obtains a private benefit of $k_t(\theta') - k_t(\theta'; \theta)$ from this deviation. Third, the DM always obtains nonnegative information rent from any report. This is because the DM can always report no project and not be detected.

The challenge in the dynamic setting is the complexity of the contract space. The contract can depend on the entire path of reported project information and observed returns. As noted in the literature of dynamic contracting, the DM's continuation value is a sufficient statistic. It can help analyze the agency issues in a tractable way. In this section, we first use the martingale techniques to characterize the continuation value. Then we identify the necessary conditions for any incentive compatible mechanism.

When making a report at time $t < \tau$, the DM considers how his decision will affect his utility promised by the contract. Define this continuation value $W_t(\hat{X}, Y)$ after a history of reports and observed returns $\{(\hat{X}_s, Y_s), 0 \le s \le t\}$ to be the total expected payoff that the DM receives if he tells the truth after time $t$:

$$W_t(\hat{X}, Y) = E\left[\int_t^\tau e^{-\gamma(s-t)} dI_s \middle| \mathcal{F}_t\right] \tag{3}$$

Recall from (2) that if the DM truthfully reports project information, he will invest all the allocated capital ($d\hat{K}_t = dK_t$) and get zero private benefit. So the DM's continuation value is only determined by payments from the contract. To characterize how the DM's continuation value evolves over time, it is useful to consider his lifetime expected utility, evaluated conditional on the information available at time $t \le \tau$, if the reports reveal true project information, i.e., $\hat{X} = X$:

$$\begin{aligned} V_t(X, Y) &= E\left[\int_0^\tau e^{-\gamma s} dI_s \middle| \mathcal{F}_t\right] \\ &= \int_0^t e^{-\gamma s} dI_s + e^{-\gamma t} W_t(X, Y) \end{aligned} \tag{4}$$

Because $V_t$ is the expectation of a given random variable conditional on $\mathcal{F}_t$, the process $V = \{V_t(X, Y)\}_{t \ge 0}$ is an $\mathcal{F}$-martingale. Using the martingale property, we now seek an alternative way to represent $V_t(X, Y)$. Define the number of projects that arrived before time $t$ and have quality located in the interval $U = [a, b]$ as

$$N(t, U) = \sum_{0 \le s \le t} \mathbb{1}_U(dX_s) \tag{5}$$

where $a, b$ are arbitrary values satisfying $\underline{\theta} \le a \le b \le \bar{\theta}$. The differential term $N(dt, d\theta)$ then indicates the number of projects that arrived in the time interval $[t, t+dt]$ and have quality within $[\theta, \theta + d\theta]$. Correspondingly, $\hat{N}(dt, d\theta)$ denotes the reported number. The martingale representation theorem for marked point process then implies the following result.

**LEMMA 1.** *There exists a function $\beta : \Omega \times R^+ \times [\underline{\theta}, \bar{\theta}] \to R^+$, which is $\mathcal{F}$-predictable with its first two arguments, such that at any moment $t \leq \tau$*

$$V_t = V_0 + \int_0^t \int_{\underline{\theta}}^{\bar{\theta}} e^{-\gamma s} \beta(\omega, s, \theta) \left[ N(ds, d\theta) - \frac{\lambda}{\Delta} ds d\theta \right] \tag{6}$$

Note that $W_t(\hat{X}, Y)$ is also the DM's continuation value if $\hat{X}_s, 0 \leq s \leq t$, were the true information and the DM reports truthfully. Therefore, without loss of generality we can derive the evolution of continuation value for the case in which the DM reports truthfully, i.e., $\hat{X} = X$, and hence, $N(dt, d\theta) = \hat{N}(dt, d\theta)$. To ease notation, we use $\beta_t(\theta)$ to denote the predictable function $\beta(\omega, t, \theta)$, and use $\beta_t'(\theta)$ to denote the partial derivative of $\beta(\omega, t, \theta)$ with respect to $\theta$ if it exists. Equations (4) and (6) together imply that the continuation value evolves as

$$dW_t = \gamma W_t dt - dI_t + \int_{\underline{\theta}}^{\bar{\theta}} \beta_t(\theta) \left[ \hat{N}(dt, d\theta) - \frac{\lambda}{\Delta} dt d\theta \right] \tag{7}$$

According to (7), $\beta_t(\theta)$ is the sensitivity of the DM's continuation value to project quality. If no project arrives during the time interval $[t, t+dt]$, this continuation value will have zero jump. If a project of quality $\theta$ is reported, this continuation value will jump by the magnitude of $\beta_t(\theta)$. Because project qualities reflect the performance of the division, we interpret $\beta_t$ as the DM's pay-performance sensitivity. A key feature of this model is that the optimal pay-performance sensitivity is not constant but contingent on project type and the performance history of the division. The following Lemma characterizes the properties of pay-performance sensitivity over the dimension of project qualities. It also characterizes the trade-off between investments and information rents.

**LEMMA 2.** *In any incentive compatible contract, the pay for performance sensitivity $\beta_t$ and the capital allocation $k_t$ satisfy ($t \leq \tau$):*

*(a) $\beta_t(\theta)$ is strictly increasing at $\hat{\theta} < \bar{\theta}$, if $k_t(\hat{\theta}) > 0$.*

*(b)* $\beta_t'(\hat{\theta}) \geq 1/R'(k_t(\hat{\theta}))$ *at* $\hat{\theta} < \bar{\theta}$*, if* $k_t(\hat{\theta}) > 0$ *and* $\beta_t(\theta)$ *is differentiable at* $\hat{\theta}$*.*

*(c)* $\beta_t(\theta) \geq 0$*.*

Part $(a)$ of Lemma 2 states that the pay-performance sensitivity is monotone over project qualities. This is because a higher quality project can generate the same return as a lower quality project with less investment. The DM who is reporting a project with higher quality must obtain a larger compensation. Part $(b)$ states that the marginal increase in pay-performance sensitivity has a lower bound which is determined by the capital allocation. This condition corresponds to the envelope condition that usually arises in mechanism design problems. By reporting a marginally lower type the DM can divert the amount $1/R'(k_t(\theta))$ from the allocated capital. So any incentive compatible contract must compensate the DM at least this amount in order to induce truth-telling. Part $(c)$ is implied by the fact that the DM can always get zero compensation by reporting no project.

Recall that if the DM misreports project quality $\theta$ to be $\theta'$, he has to invest $k_t(\theta'; \theta)$ as in (2). Because the DM has no savings, the investment $k_t(\theta'; \theta)$ has to be smaller than the allocated capital $k_t(\theta')$. Equation (2) then implies that $\theta' \leq \theta$, meaning only downward reporting is feasible. Moreover, since investment $k_t(\theta'; \theta)$ is nonnegative, we must have $\theta' \geq \theta - R(k_t(\theta'))$ from (2). Accordingly, the feasible set of misreports is $\Gamma(\theta, k_t) =: \{\theta' \in \Theta : \theta - R(k_t(\theta')) \leq \theta' \leq \theta\}$. It is easy to see from this feasible set that the DM can possibly report a lower quality only if the lower quality project receives positive capital allocation. Therefore, the monotonicity of $\beta_t(\theta)$ and the envelope condition hold only when capital allocation is positive.

# 5 Optimal Contract

In this section, we derive the firm policy that maximizes the HQ's value when the DM has private information about investment opportunities. The HQ's value is a function of the the DM's continuation value which is denoted by $P(W)$. We use the dynamic programming approach to determine the most profitable way to deliver the promised value to the DM.

## 5.1 Optimal Payment

Since the HQ can always provide the DM with a lump-sum cash, the marginal cost of compensating the DM can never exceed the marginal cost of immediate cash payment. So the value function must satisfy $P'(W) \geq -1$ at any $W$. The cash payment boundary $\bar{W}$ is the smallest value such that $P(\bar{W}) = -1$. In deriving the optimal contract, we assume that $P(.)$ is concave and strictly concave when $W < \bar{W}$, which we will show in the verification section.

The twin assumptions that (i) DM is risk neutral, and (ii) terminating the division is inefficient jointly determine that cash payments are postponed until the continuation value reaches the threshold $\bar{W}$. The DM is compensated purely through promised future values before $\bar{W}$.

**LEMMA 3.** *When $W_t < \bar{W}$, no cash payment is issued, i.e., $dI_t = 0$. When $W_t \geq \bar{W}$, cash payment $dI_t = W_t - \bar{W}$ is immediately issued, and $P(W_t) = P(\bar{W}) - (W_t - \bar{W})$.*

Different from the Brownian models, e.g. DeMarzo and Sannikov [2006], the payment issuance in this model is determined by the jumps in DM's continuation value. Equation (7) has shown that the DM's continuation value immediately jumps to a new level when a project arrives. According to Lemma 3, cash payment will be issued to the DM if this new level of continuation value achieves the threshold $\bar{W}$.

## 5.2 A Heuristic Derivation

In the interior region $W_t \in (0, \bar{W})$, the HQ holds all the investment returns. The HQ's flow payoff consists of two parts: the net expected return of investment and the expected change in value function induced by the variation in $W_t$. The net expected return of investment at time $t$ is:

$$E[(dY_t - k_t)d\pi_t] = \frac{\lambda dt}{\Delta} \int_{\underline{\theta}}^{\bar{\theta}} [\theta + R(k_t(\theta)) - k_t(\theta)] d\theta$$

The left limit $W_{t-}(\hat{X}, Y) = \lim_{s \uparrow t} W_s(\hat{X}, Y)$ is the DM's continuation utility evaluated before project information is reported. So the process $\{W_{t-}(\hat{X}, Y)\}_{t \geq 0}$ is $\mathcal{F}$-predictable. To ease notation, $W_t$ is generally used to denote $W_{t-}$ throughout the paper. The HQ's expected change in contract value is given by:

$$E[dP(W_t)] = \left[\gamma W_t dt - \frac{\lambda dt}{\Delta}\int_{\underline{\theta}}^{\bar{\theta}} \beta_t(\theta) d\theta\right] P'(W_t)$$
$$+ \frac{\lambda dt}{\Delta}\int_{\underline{\theta}}^{\bar{\theta}} [P(W_t + \beta_t(\theta)) - P(W_t)] d\theta$$

Because at the optimum the HQ should earn an instantaneous return of $rP(W_t)$, the Hamilton-Jacobi-Bellman(HJB) equation has the form:

$$rP(W_t) = \max_{k_t \geq 0, \beta_t \geq 0} \frac{\lambda}{\Delta}\int_{\underline{\theta}}^{\bar{\theta}} [\theta + R(k_t(\theta)) - k_t(\theta)] d\theta \tag{HJB}$$
$$+ \left(\gamma W_t - \frac{\lambda}{\Delta}\int_{\underline{\theta}}^{\bar{\theta}} \beta_t(\theta) d\theta\right) P'(W_t) + \frac{\lambda}{\Delta}\int_{\underline{\theta}}^{\bar{\theta}} \left[P(W_t + \beta_t(\theta)) - P(W_t)\right] d\theta$$

The optimization in (HJB) has to satisfy the incentive compatibility constraint

$$\beta_t(\theta) \geq \beta_t(\hat{\theta}) + k_t(\hat{\theta}) - k_t(\hat{\theta}; \theta), \quad \forall \hat{\theta} \in \Gamma(\theta, k_t) \tag{IC}$$

The conditions that pin down the value function $P$ and the payout boundary $\bar{W}$ are

$$P(0) = L, P'(\bar{W}) = -1 \tag{BC}$$

The incentive compatible constraint and the constraint of $\beta_t \geq 0$ guarantee that the DM truthfully reveals project information. To solve this problem, we use the first order approach that is typically applied in mechanism design issues. In particular, we consider a relaxed problem by replacing (IC) with the envelope condition. We then show that the policy derived from this relaxed problem actually satisfies (IC).

Recall from Lemma 2 that the envelope condition holds only when capital allocation is positive. However, as we will show later the HQ may exclude the project from capital allocation if the agency cost is too high. The difficulty in relaxing the (HJB) is to find out which projects are excluded from capital allocation. We handle this issue in two steps. First, we consider the case where no project is excluded, i.e., $k_t(\theta) > 0$ for all $\theta$, for any fixed value of $W_t$. This is true if the spread of project qualities ($\Delta$) is sufficiently small, because the information asymmetry is not severe. In the case of $\Delta \approx 0$, the HQ almost knows the true project quality even without DM's report. So the DM is paid little information rent. The infinite marginal return of value enhancing technology at zero ($R'(0) = \infty$) then implies that no project will be excluded from capital allocation. Second, we show that as the spread of project quality increases, projects with bottom qualities are possibly forgone.

## 5.3 Policy without Project Exclusion

When no project is excluded, we can get a relaxed problem by replacing (IC) by the envelope condition $\beta_t'(\theta) \geq 1/R'[k_t(\theta)]$. From the optimality of the relaxed problem we can derive the following necessary condition that characterizes the policy of capital allocation and pay-performance sensitivity.

**LEMMA 4.** *In the interior region ($W_t \in (0, \bar{W})$), if no project is excluded from capital allocation, then the optimal policy satisfies:*

$$R'(k_t(\theta)) - 1 = \frac{R''(k_t(\theta))}{[R'(k_t(\theta))]^2} \int_{\theta}^{\bar{\theta}} \left[ P'(W_t + \beta_t(u)) - P'(W_t) \right] du \tag{8}$$

*Moreover, capital allocation $k_t(\theta)$ increases in project quality $\theta$.*

Lemma 4 characterizes the distortion in investment due to agency issues. The left-hand side of (8) is the net return of marginally raising investment in type $\theta$ project. The strict concavity of the value function $P$ and the technology $R$ together implies that the right-hand side of (8) is positive and represents the agency cost. It is easy to see that investment in any project is lower than the first best level $k^*$. Investment distortion in any project arises from the information rents paid to the DM reporting projects of higher qualities. As project quality decreases, more information rent is paid to the DM, and therefore investment is more downward distorted.

An important feature is that the agency cost in capital budgeting is endogenous. The information rents are costly to the firm because they induce positive jumps in DM's continuation value when projects arrive, and induce downward drifts when no project arrives. These variations in continuation value exacerbate liquidation of the division. In particular, the increase in agency cost from a marginal information rent i n project $\theta$ is measured by $P'(W_t + \beta_t(\theta)) - P'(W_t)$. So the curvature of the value function and the level of the continuation value together determine the magnitude of the agency cost. And both the curvature of the value function and the continuation value are endogenous elements. In this respect, this model is very different from static problems such as Laffont and Tirole [1986], and Myerson [1981]. In those static models, the agency costs are always exogenously related to the inverse hazard ratio.

### 5.4 Policy with Project Exclusion

The optimal contract distorts investment downward to economize on information rents. As the spread of project quality increases, investments are severely distorted. When the agency cost is sufficiently high, the extreme strategy is to forgo positive NPV investment opportunities. If type $\theta$ project is reported but gets no capital allocation, then the DM observing higher types will not be able to misreport project quality as $\theta$. So excluding projects from capital allocation relaxes the agency problems. We now examine the general case where project exclusion possibly occurs.

**PROPOSITION 2.** *There exists threshold project quality $\theta_t^c = \theta^c(W_t) \leq \bar{\theta}$ such that*

*(a) Projects with reported quality above $\theta_t^c$ receive positive capital allocation and have positive pay-performance sensitivity: $k_t(\theta) > 0$, $\beta_t(\theta) > 0$, and $\beta_t'(\theta) = 1/R'(k_t(\theta))$ for any $\theta > \theta_t^c$.*

*(b) Projects with reported quality below $\theta_t^c$ receives no capital allocation and has zero pay-performance sensitivity: $k_t(\theta) = \beta_t(\theta) = 0$, for any $\theta < \theta_t^c$, if $\theta_t^c > \underline{\theta}$.*

Proposition 2 shows that the optimal exclusion strategy exhibits threshold property. If project exclusion ever occurs, it is the projects with bottom qualities that gets no capital allocation. Recall from Lemma 4 that the lowest quality project receives the smallest amount of capital allocation when no project is excluded. If we keep the state variable $W_t$ constant but increase the spread of project quality $\Delta$, then the capital allocation to the lowest quality project will reach zero first.

This project exclusion policy implies zero pay-performance sensitivity for projects with quality lower than the threshold value $\theta_t^c$. Because all projects with quality lower than $\theta_t^c$ gets no capital allocation, it is not feasible for the DM observing a $\theta \leq \theta_t^c$ project to report a lower quality. In other words, the only feasible report is the project's true quality. So it is optimal to offer no information rent to the DM. A positive pay-performance sensitivity only adds variation to the DM's continuation value and lowers the HQ's expected value.

With this threshold exclusion policy, we can rewrite the relaxed problem as:

$$\begin{aligned} rP(W_t) = \max_{k_t,\beta_t,\theta_t^c} \frac{\lambda}{\Delta}\left[\int_{\underline{\theta}}^{\theta_t^c} \theta d\theta + \int_{\theta_t^c}^{\bar{\theta}} (\theta + R(k_t(\theta)) - k_t(\theta))d\theta\right] \\ + \left(\gamma W_t - \frac{\lambda}{\Delta}\int_{\theta_t^c}^{\bar{\theta}} \beta_t(\theta)d\theta\right) P'(W_t) + \frac{\lambda}{\Delta}\int_{\theta_t^c}^{\bar{\theta}} \left[P(W_t + \beta_t(\theta)) - P(W_t)\right] d\theta \end{aligned} \tag{HJB'}$$

subject to the condition of $\beta(\theta_t^c) \geq 0$ and the envelope condition of

$$\beta_t'(\theta) = 1/R'(k_t(\theta)), \quad \forall \theta \geq \theta_t^c \tag{EN}$$

Note that (HJB') has $\theta_t^c$ as an additional control. The objective of the (HJB') is simply obtained by setting $k_t(\theta) = \beta_t(\theta) = 0$ for $\theta \leq \theta_t^c$ in the objective of the (HJB). The threshold $\theta_t^c$ is an endogenous object that is moving with the state $W_t$. This is exactly because the agency cost itself is endogenous. When the endogenous agency cost is high, then only marginal investment in higher quality project can balance this cost. So the threshold $\theta_t^c$ takes a higher value. Similarly, when the endogenous agency cost is low, the threshold $\theta_t^c$ takes a lower value.

Also note that for a given value of $\theta_t^c$ problem (HJB') has the same structure as the relaxed problem considered in Lemma 4. Hence, the optimal policy above the threshold quality $\theta_t^c$ must satisfy (8) and the envelope condition $\beta_t'(\theta) = 1/R'(k_t(\theta))$. Using these conditions, we can obtain the following characterizations of the optimal policy.

**PROPOSITION 3.** *The optimal capital allocation and pay-performance sensitivity satisfy:*

*(a) The highest quality project has first best investment before liquidation:*

$$k_t(\bar{\theta}) = k^* \quad \textit{when } W_t > 0$$

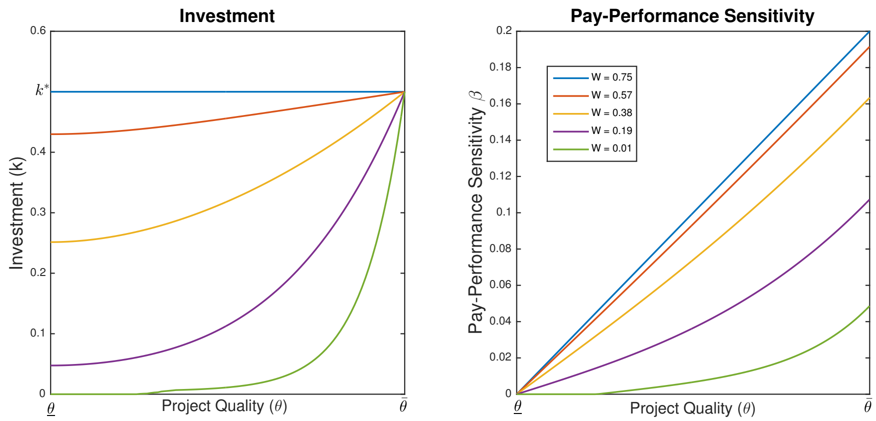


Figure 2: Policy Dynamics Over Project Qualities

*(b) Underinvest in all other projects before payout:*

$$k_t(\theta) < k^*, \forall \ \theta < \bar{\theta} \quad \textit{when } 0 < W_t < \bar{W}$$

*(c) At the payout boundary, investments in all projects resume first best:*

$$\theta_t^c = \underline{\theta}, \ k_t(\theta) = k^*, \forall \ \theta \quad \textit{when } W_t \geq \bar{W}$$

*(d) Investment and pay-performance sensitivity are both increasing in project quality*

Figure 2 plots the optimal $k_t$ and $\beta_t$ through a numerical example in which $\bar{W} = 0.75$, and $k^* = 0.5$. It reflects all the properties in Proposition 3. First, the right panel shows at the payout boundary $\bar{W}$ investment in all projects resumes the first best level $k^*$. Second, along any curve (fix continuation value) the policy is increasing in project quality. Third, when fixing a level of $\theta$ and considering a higher curve, we can see that the policy is increasing in continuation value.

The "no distortion at top" result in this model coincides with that in Mussa and Rosen [1978]. Surprisingly, this model characterizes an additional no-distortion dimension: "no distortion for all types" at the highest continuation value level ($\bar{W}$). In other words, the policy reaches first best at either the highest project quality or the highest continuation value. This is due to the endogenous agency cost in this model.

The level of continuation value indicates the tightness of the limited liability constraint. When the continuation value is close to the liquidation boundary, the limited liability constraint is very tight. As the continuation value goes up, the limited liability constraint relaxes and the liquidation concern decreases. At the payout boundary $\bar{W}$, the liquidation concern disappears. Moreover, the liquidation concern determines the trade-off between investment efficiency and information rents. To induce larger investment, the contract has to design larger pay-performance sensitivities, lowering the drift of DM's continuation value. When the liquidation concern is sever, it is optimal to lower pay-performance sensitivities which severely distorts capital allocation. When the liquidation concern is low, it is not very costly to design large pay-performance sensitivities. So the optimal contract induces more efficient investments.

Proposition 2 concludes that if the HQ decides to exclude any positive NPV investment, the exclusion must occur to the bottom quality projects. Now we explore under what condition project exclusions will ever happen in the optimal contract. It depends on two driving forces: (i) the marginal return of investment at $k = 0$, and (ii) the agency cost induced by this marginal investment. According to the envelope condition, a marginal investment in project $\theta$ will increase the slope of the pay-performance sensitivity at $\theta$ by $\lim_{k\downarrow 0} \frac{R''(k)}{[R'(k)]^2}$, which makes the (IC) constraints tighter for all types above $\theta$. This effect will induce the agency cost to increase by $\lim_{k\downarrow 0} \frac{\mu_t(\theta)R''(k)}{[R'(k)]^2}$, where $\mu_t(\theta)$ is the multiplier of (IC) at type $\theta$. The optimal contract trades off these two competing forces in determining project exclusion.

**LEMMA 5.** *The optimal project exclusion has the following property:*

*(a) If the investment technology satisfies* $\lim_{k\downarrow 0} \frac{R''(k)}{[R'(k)]^3} = 0$*, then no project will be excluded* $(\theta_t^c = \underline{\theta})$ *at any state before liquidation* $(W_t > 0)$*.*

*(b) If the investment technology satisfies* $\lim_{k\downarrow 0} \frac{R''(k)}{[R'(k)]^3} < 0$*, then depending on the value of the state variable* $W_t$*, projects with bottom qualities will possibly be excluded, i.e.,* $\theta_t^c > \underline{\theta}$*.*

Lemma 5 illustrates that if the marginal return dominates the agency cost induced by the marginal investment as capital allocation converges to zero, then exclusion will not happen at any state. However, if the agency cost induced by the marginal investment dominates the marginal return as capital allocation converges to zero, then exclusion of investment possibly occurs, depending on the state $W_t$.

## 5.5 Verification

The results so far only relies on the necessary conditions of (HJB'), i.e., condition (8) and $\beta_t'(\theta) = 1/R'(k_t(\theta))$. However, applying these conditions may not allow us to pin down a unique policy of $k_t$ and $\beta_t$. To guarantee (HJB') has a unique solution, we impose a technology assumption: $R'''R' \leq 3(R'')^2$. In the case of Cobb-Douglas technology, i.e., $R(k) = ak^\alpha$, this assumption is equivalent to $\alpha \leq 0.5$. It's shown in the Appendix that this technology assumption is sufficient to guarantee a unique optimal contract.

Now we verify in the following two steps that the relaxed problem from the heuristic characterization does correspond to the optimal contract. First, we verify that the optimal policy derived from the relaxed problem actually satisfies (IC).

**LEMMA 6.** *The policy $(k_t(\theta), \beta_t(\theta))$ that satisfies*

1. *condition* (8) *and the envelope condition $\beta_t'(\theta) = 1/R'(k_t(\theta))$, when $\theta \geq \theta_t^c$;*
2. *$\beta_t(\theta) = 0$, when $\theta \leq \theta_t^c$;*
3. *$k_t(\theta) = 0$, when $\theta \leq \theta_t^c$ and $\theta_t^c > \underline{\theta}$*

*is the optimal solution of* (HJB).

In the no exclusion region ($\theta > \theta_t^c$), the monotonicity of capital allocation implies that the envelope condition is sufficient for incentive compatibility. In the exclusion region ($\theta \leq \theta_t^c$), because the DM cannot misreport as lower types, zero pay-performance sensitivity is sufficient.

Second, following the standard argument in optimal control, we show that HQ's value from any incentive compatible mechanism that delivers the DM continuation value $W_0$ is at most $P(W_0)$, which is the HQ's expected payoff from the conjectured mechanism.

**PROPOSITION 4.** *The contract $(K, I, \tau)$ that maximizes the HQ's expected profit and delivers value $W_0 \in [0, \bar{W}]$ to the DM has the following form:*

1. *$W_t$ evolves according to (7), where $\beta_t(\theta) = \int_{\underline{\theta}}^{\theta} \frac{1}{R'(k_t(u))} du$.*
2. *When $W_t \in [0, \bar{W})$, $dI_t = 0$; When $W_t = \bar{W}$, payments $dI_t$ cause $W_t$ to reflect at $\bar{W}$.*
3. *Liquidation of the division occurs when $W_t$ reaches 0.*
4. *The HQ's expected payoff $P(W_t)$: matches the objective of the (HJB) evaluated at $(k_t(\theta), \beta_t(\theta))$ on interval $[0, \bar{W}]$; satisfies $P'(W_t) = -1$ when $W_t \geq \bar{W}$, and $P(0) = L$.*
5. *The value function $P(W_t)$ is globally concave and strictly concave when $W_t < \bar{W}$.*

## 5.6 Private Savings

So far, we restrict the DM from saving, hence he cannot overreport project qualities. We now relax this restriction. In particular, we assumes that the DM can save in a private account at interest rate $\rho \leq r$. But the balance of this account has to be nonnegative. The contract $(K, I, \tau)$ in Proposition 4 remains incentive compatible even if we allow the DM to save. The intuition is that the envelope condition and the monotonicity of $k_t(\theta)$ imply that the DM has no incentive to over-report project qualities. So the marginal benefit of consuming cash is constant over time. Given that private savings grow at rate $\rho < \gamma$, there is no incentive to save.

**PROPOSITION 5.** *Suppose a contract $(K, I, \tau)$ satisfies $k_t(\theta)$ increases in $\theta$. Also suppose that the process $W_t \geq 0$ evolves according to (7) with $\beta_t(\theta) = \int_{\underline{\theta}}^{\theta} \frac{1}{R'(k_t(u))} du$ until the stopping time $\tau = \min\{t | W_t = 0\}$. Then the DM earns a payoff of at most $W_0$ from any feasible strategy. The expected payoff $W_0$ is obtained if the DM reports project information truthfully and maintains zero savings.*

# 6 Dynamics of the Optimal Contract

The previous analysis shows that the optimal investment and compensation exhibit different properties at different levels of continuation value. Investments are distorted downward

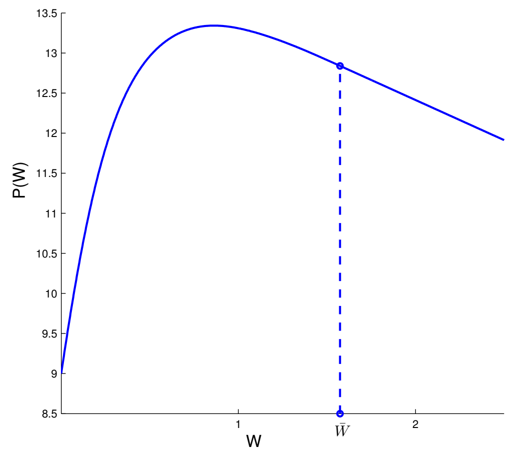


Figure 3: Value Function

in general, but the distortions disappear when the DM's continuation value reaches its upper bound. This clearly shows that the severity of the agency issues are endogenous and therefore the optimal policy should be designed to vary with the state of the contract. In this section, we will characterize how the continuation value evolves under the optimal compensation. We will also show that the optimal policy exhibit monotone properties. These properties sharpen the intuition that as continuation value goes up, the agency cost falls. Accordingly, the HQ allocates capital more efficiently and the DM gets higher compensation when the division accumulates more financial slack in the past.

## 6.1 Liquidity Evolution

The evolution of the agent's continuation value reflects how the firm's liquidity varies over time. Because of agency issues, the firm has to pay information rent to the DM if it decides to invest in the value-enhancing technology. When a project above the exclusion threshold arrives, the DM's continuation value will immediately jump up to reflect this positive compensation. However, the expected amount of future compensations will be taken out from the DM's continuation value at other times, as long as the division is not liquidated. This mechanism poses a downward force to the continuation value when no project arrives

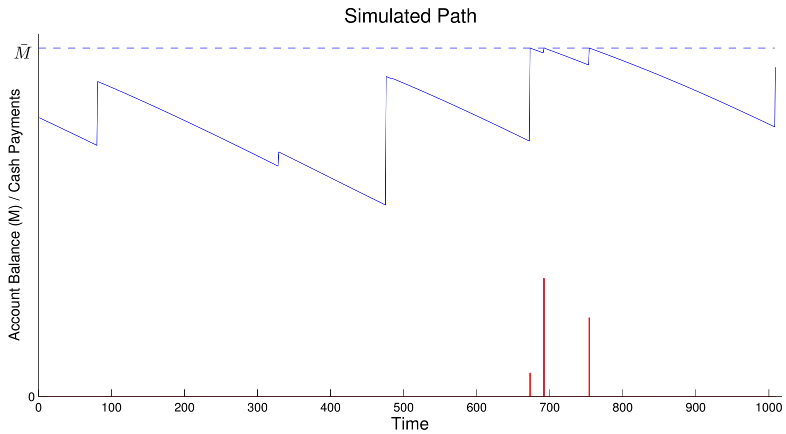


Figure 4: Simulated Path of Continuation Value and Payout

or a project below the exclusion threshold arrives. In this sense the contract "punishes" the DM at the time of no capital allocation. This mechanism is feasible under limited liability simply because the DM is "punished" by the promised payments not immediate transfers. The following result formally shows this intuition.

**PROPOSITION 6.** *Firm liquidity (continuation value) jumps up when projects above the exclusion threshold arrives and it drifts downward during other times.*

Figure 4 shows a simulated path of continuation value. The continuation value keeps drifting down in the no project periods and jumps up when a project arrives. We can also see that the jump size becomes larger as the continuation value level increases. Cash payments will be issued when the continuation value jumps beyond the payout boundary. In Figure 4, the red bars on the horizontal axis reflect cash payments which drive the continuation level immediately back to the payout boundary.

## 6.2 Monotone Policies

As the DM's continuation value varies, there is a feedback mechanism at play. On one hand, the pay-performance sensitivity affects the agency cost and hence the optimal capital allocation. On the other hand, the capital allocation imposes constraint on pay-performance sensitivity through the incentive compatibility condition. So in general, it is hard to pin down the complete policy dynamics. We proceed with the characterization in two steps. First, we examine investment dynamics around the payout boundary. Second, we restrict attention to the Cobb-Douglas technology, but analyze the policy dynamics in more general regions.

**PROPOSITION 7.** *The capital allocation $k_t(\theta)$ and the pay-performance sensitivity $\beta_t(\theta)$ are both increasing in firm liquidity ($W_t$) for all $\theta$. The exclusion threshold $\theta_t^c$ is decreasing in liquidity ($W_t$).*

Proposition 7 implies that there is a uniform lower bound of continuation value for all projects beyond which capital allocation becomes more efficient (at both the extensive and the intensive margins) and compensation becomes larger as continuation value increases. More importantly, Proposition 7 offers sufficient conditions that allow us to quantify the cutoff value $W^1$. When either the investment efficiency is high or the project quality spread is low, the cutoff $W^1$ becomes very small, meaning the optimal policy is monotone over a very large range of the domain $(0, \bar{W}]$.

# 7 Implementation

The foregoing results characterize the properties of the optimal contract in a principal-agent setting. To understand the capital budgeting process in practice, in this section we show in this section how the optimal contract so derived can be implemented using a simple budgeting account mechanism.

To facilitate the investments and to deal with the agency problems, the HQ fixes a budgeting account for the division. Given moral hazard, the HQ cannot determine the actual use of the funds by the DM. In other words, at any time the DM can withdraw funds at his discretion to either invest or issue compensation. What the HQ can observe and control is the evolution of the account balance. Based on project reports, the HQ can either deplete or replenish the budgeting account. If it sets the right depleting and replenishment rates, it can give the DM the incentive to invest as in the optimal contract. Let $(k_t, \beta_t, \theta_t^c)$

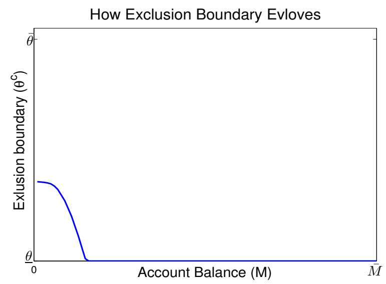


Figure 5: How project exclusion boundary evolves

be the optimal policy derived as above. The budgeting account mechanism includes the following key elements:

- The budgeting account is set up with an initial balance $M_0$.
- The account is constantly depleted by the rate $q(M_t) = \frac{\lambda}{\Delta M_t} \int_{\underline{\theta}}^{\bar{\theta}} \beta_t(\theta) d\theta - \gamma$.
- If a project of quality $\theta > \theta_t^c$ is reported, the account is replenished by $k_t(\theta)$, and then by $\beta_t(\theta)$ if investment return $\theta + R(k_t(\theta))$ is futher observed.
- Investment and cash compensation are paid from the budgeting account balance, and are at the discretion of the DM.
- The division is liquidated when the account balance reaches 0.

By this mechanism, the budgeting account balance is raised by $\beta_t(\theta)$ when a project of quality $\theta \geq \theta_t^c$ arrives at time $t$, and evolves as $dM_t = -q(M_t)M_t dt$ at other times. The HQ can adjust the initial balance and cash payout boundary appropriately to induce truth-telling of project information and implement the optimal contract.

**PROPOSITION 8.** *In the capital budgeting mechanism with initial account balance* $M_0 = W_0$*, it is incentive compatible for the DM to report project information truthfully, and not to steal funds. The division issues cash compensation* $\{M_t - \bar{W}, 0\}$*. Moreover, this mechanism implements the optimal contract.*

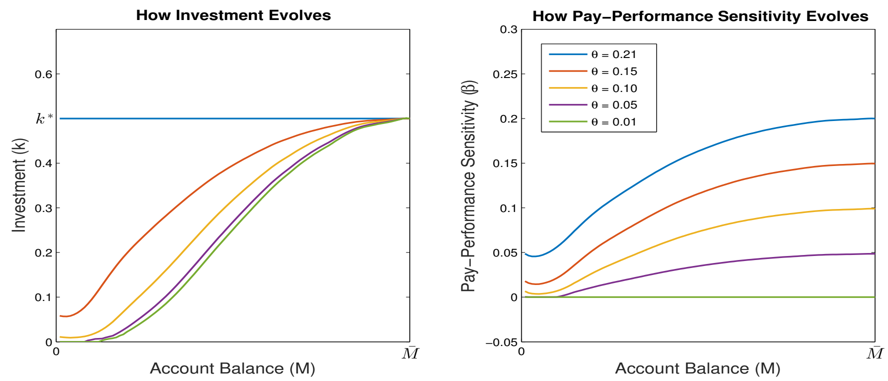


Figure 6: How investment and compensation evolve

Under this mechanism, the DM has the freedom to pay out the entire budgeting account balance $M_{t-}$ as cash compensation and thus have the division liquidated at any time $t$. But it is obviously not optimal for the HQ. How can we ensure that the DM does not opt for this deviation? This payoff from waiting until the account balance reaches $\bar{M}$ to receive payment is also $M_{t-}$. This is because by construction the evolution of the budgeting account balance mimics that of the continuation value in the optimal contract. So the DM has no incentive to take cash compensation before the account balance reaches $\bar{M}$. Moreover, when a project $\theta$ arrives at time $t$, the funds replenished to the budgeting account raise the balance by $\beta_t(\theta)$ and the DM's expected payoff increases by the same amount, which satisfies (IC). So the DM will report project information truthfully.

Through an numerical example, we now illustrate how optimal policies vary with the budgeting account balance. In figure 5, we can see the exclusion boundary is above zero at a low account balance but it decreases as account balance becomes large which means that the fewer projects are excluded. And the exclusion boundary stays at zero, i.e., no project is excluded, when account balance is above some threshold. In figure 6, we fix project quality at various levels and plot how investment and pay-performance sensitivity vary with account balance. We can see that they are both increasing in account balance at all levels of project quality.

# 8 Conclusion

Using a dynamic principal-agent model with continuous time and continuous project quality levels, this paper shows how capital budgeting and managerial compensation contracts can be jointly designed to mitigate two important agency problems in large firms: asymmetric investment information and capital diversion. The trade-off for the headquarter between getting truthful information from the division manager and inducing efficient investment is shown to be time-varying and endogenously determined by the financial slack of the division. When there is little financial slack, the model predicts that the optimal contract will have low pay-performance sensitivity and allocate much less than efficient capital to the DM. It may also be optimal to exclude positive NPV projects when the division is in financial distress. The high liquidation risk constrains the extent of the compensation contract to provide incentive. So extracting information will be costly and the problem of underinvestment will be severe. When financial slack is substantial, the model predicts that the optimal contract will have high pay-performance sensitivity and will allocate capital more efficiently in both extensive and intensive margins. These policies are monotonically increasing as financial slack accumulates. Investments in every type of project are shown to reach the first best level when the DM starts getting cash compensation. Finally, the HQ can implement the optimal policies by appropriately adjusting the balance of a budgeting account through which the DM gets funds for investment and compensation.

# A Appendix

In our notation, $k_t(\theta)$ and $k(\theta, W_t)$ are equivalent. Also, $\beta_t(\theta)$ and $\beta(\theta, W_t)$ are equivalent. We will use them interchangeably in different contexts.

*Proof of Proposition 1.* Since both the project information and the investment amount are observable, there will be no information rent paid to the DM. So it is always optimal to invest the first best level for any project. And $k^*$ can be implemented by a forcing contract. Given any mechanism, DM's discounted future compensation will be constant over time and, therefore, the division of the conglomerate will never be liquidated. Finally, no liquidation means delay payment to the DM is not optimal, since the DM is more impatient. So the optimal contract will payout $W_0$ immediately and always enforce the first best investment. □

*Proof of Lemma 1.* A basic result from the point process theory is that the compensated point process defined by

$$Z(t, U) = N(t, U) - \frac{\lambda t(b-a)}{\Delta} \tag{A.9}$$

is an $\mathcal{F}$-martingale. The differential form of this martingale is $dZ = N(dt, d\theta) - \frac{\lambda}{\Delta} dt d\theta$.

The predictable representation (6) of the martingale $V$ follows from Last and Brant [1995] (Chapter 1, page 25-26). The factor $e^{-\gamma s}$ is just a convenient rescaling. □

*Proof of Lemma 2. Part (a)*: Take any $\hat{\theta} \in [\underline{\theta}, \bar{\theta})$ such that $k_t(\hat{\theta}) > 0$. There exists sufficiently small $\varepsilon > 0$ such that $\hat{\theta} + \varepsilon - R(k_t(\hat{\theta})) \leq \hat{\theta}$ and $\hat{\theta} + \varepsilon \leq \bar{\theta}$. This means $\hat{\theta} \in \Gamma(\hat{\theta} + \varepsilon, k_t)$. By misreporting project quality of $\hat{\theta} + \varepsilon$ as $\hat{\theta}$, the DM gets $\beta_t(\hat{\theta})$ in promised utility which is implied by (7) and diverts capital $k_t(\hat{\theta}) - k_t(\hat{\theta}; \hat{\theta} + \varepsilon)$. If the DM reports truthfully, he gets only $\beta_t(\hat{\theta})$ in promised utility. Incentive compatibility then implies:

$$\beta_t(\hat{\theta} + \epsilon) \geq \beta_t(\hat{\theta}) + k_t(\hat{\theta}) - k_t(\hat{\theta}; \hat{\theta} + \varepsilon)$$

So we get

$$\beta_t(\hat{\theta} + \epsilon) - \beta_t(\hat{\theta}) \geq k_t(\hat{\theta}) - R^{-1}[R(k_t(\hat{\theta})) - \epsilon] > 0 \tag{A.10}$$

The first inequality is from (2). Therefore, $\beta_t(\theta)$ is strictly increasing at $\hat{\theta}$.

*Part (b)*: Take any $\hat{\theta} \in [\underline{\theta}, \bar{\theta})$ such that $k_t(\hat{\theta}) > 0$ and $\beta_t'(\hat{\theta})$ exists. Equation (A.10) implies that

$$\begin{aligned}\beta_t'(\hat{\theta}) &= \lim_{\epsilon \downarrow 0} \frac{\beta_t(\hat{\theta}+\epsilon) - \beta_t(\hat{\theta})}{\epsilon} \\ &\geq \lim_{\epsilon \downarrow 0} \frac{k_t(\hat{\theta}) - R^{-1}[R(k_t(\hat{\theta})) - \epsilon]}{\epsilon} = 1/R'(k_t(\hat{\theta}))\end{aligned}$$

*Part (c)*: If $\beta_t(\theta) < 0$ for any $\theta \in \Theta$, then the DM's continuation value will jump downward when he reports a project of quality $\theta$. However, the DM is better off by not revealing such project, because his continuation value will have a zero jump if he reports no project. To induce truthful information about project arrival, it is necessary to set $\beta_t(\theta) \geq 0$ in any contract. □

*Proof of Lemma 3.* The concavity of the value function and the definition of $\bar{W}$ implies: $P'(W_t) > -1$ when $W_t < \bar{W}$ and $P'(W_t) = -1$ when $W_t \geq \bar{W}$. In the region $W_t < \bar{W}$, the marginal cost of compensating the DM through continuation value is lower than the immediate payment. So $dI_t = 0$. In the region $W_t \geq \bar{W}$, the marginal costs of compensating the DM are the same through continuation value or immediate payment. Since the DM is more impatient, an immediate payment $dI_t = W_t - \bar{W}$ will be mad. The immediate payment will cause $W_t$ to reflect at $\bar{W}$. So $P(W_t) = P(\bar{W}) - (W_t - \bar{W})$. □

*Proof of Lemma 4.* Since $k_t(\theta) > 0$ for all $\theta$, Lemma 2 implies that $\beta_t(\theta)$ is strictly increasing over $[\underline{\theta}, \bar{\theta})$. So $\beta_t(\theta)$ is differentiable a.e. Then Lemma 2 future implies $\beta_t'(\theta) \geq 1/R'(k_t(\theta))$ a.e. Using this envelope condition to replace (IC) in Problem (HJB), we know the optimal policy $(k_t, \beta_t)$ solves

$$\begin{aligned}\max_{k_t,\beta_t} &\frac{\lambda}{\Delta} \int_{\underline{\theta}}^{\bar{\theta}} [\theta + R(k_t(\theta)) - k_t(\theta)] d\theta \\ &+ \frac{\lambda}{\Delta} \int_{\underline{\theta}}^{\bar{\theta}} [P(W_t + \beta_t(\theta)) - P(W_t) - \beta_t(\theta) P'(W_t)] d\theta \\ &s.t. \quad \beta_t'(\theta) \geq 1/R'(k_t(\theta)), \quad \beta_t(\underline{\theta}) \geq 0\end{aligned} \tag{P1}$$

Differentiate the objective in Problem (P1) with respect to $\beta_t(\theta)$ to get

$$\lambda/\Delta[P'(W_t+\beta_t(\theta))-P'(W_t)]\leq 0 \tag{A.11}$$

Since $P$ is strictly concave when $W_t<\bar{W}$, (A.11) holds as strict inequality when $\beta_t(\theta)>0$. So the constraints in Problem (P1) must hold as equality at optimal, i.e., $\beta_t'(\theta)=1/R'(k_t(\theta))$ and $\beta_t(\underline{\theta})=0$.

Then (P1) can be viewed as an optimal control problem. From the view of optimal control theory, $k_t(\theta)$ is the "control" and $\beta_t(\theta)$ is the "state" of this problem. Moreover, the initial state $\beta_t(\underline{\theta})$ is given as zero; and the terminal state $\beta_t(\bar{\theta})$ is a free value. To apply the Maximum Principal in optimal control theory, we define the Hamiltonian at $W_t$ as:

$$\begin{aligned}H_t(\theta,k_t,\beta_t,\mu_t)=&\lambda/\Delta[\theta+R(k_t(\theta))-k_t(\theta)]\\&+\lambda/\Delta[P(W_t+\beta_t(\theta))-P(W_t)-\beta_t(\theta)P'(W_t)]+\frac{\mu_t(\theta)}{R'[k_t(\theta)]}\end{aligned} \tag{A.12}$$

By the Maximum Principal, the necessary conditions of problem (P1) are:

$$\mu_t'(\theta)=-\frac{\partial H_t}{\partial \beta_t}=\frac{\lambda}{\Delta}[P'(W_t)-P'(W_t+\beta_t(\theta))] \tag{A.13}$$

$$k_t(\theta)=\arg\max_{k\geq 0}\{g_t(k;\theta)\} \tag{A.14}$$

where $g_t(k;\theta):=\lambda/\Delta[R(k)-k]+\frac{\mu_t(\theta)}{R'(k)}$. Moreover, $\mu_t(\bar{\theta})=0$ by the fact $\beta_t(\bar{\theta})$ is free. Integrate (A.13) to get

$$\mu_t(\theta)=\frac{\lambda}{\Delta}\int_{\theta}^{\bar{\theta}}[P'(W_t+\beta_t(u))-P'(W_t)]du \tag{A.15}$$

Since $k_t(\theta)>0$ by assumption, the maximizer of (A.14) must satisfy

$$\frac{\lambda}{\Delta}[R'(k_t(\theta))-1]-\frac{\mu_t(\theta)R''(k_t(\theta))}{[R'(k_t(\theta))]^2}=0 \tag{A.16}$$

Combining (A.15) and (A.16), we conclude that the optimal policy satisfies (8).

Now we show the monotonicity of $k_t(\theta)$. Note that $g_t(k;\theta)$ satisfies the single crossing property. Since $\mu_t'(\theta)>0$ and $R''(k)<0$, we know that $-\frac{\mu_t'(\theta)R''(k)}{[R'(k)]^2}>0$. Hence, $g_t'(k;\theta)$ is strictly increasing in $\theta$. Then (A.14) implies that the optimal capital allocation $k_t(\theta)$ is increasing in $\theta$. $\square$

**LEMMA A.1.** *If the capital allocation policy satisfies $k_t(\theta) > 0$ at some $\theta < \bar{\theta}$, then for any arbitrarily small $\varepsilon > 0$ there must exist $\hat{\theta} \in (\theta, \theta + \varepsilon)$ such that $k_t(\hat{\theta}) > 0$.*

*Proof of Lemma A.1.* Let $(k_t, \beta_t)$ be the optimal policy at time $t$, and let $\theta_2$ be any type that satisfies $k_t(\theta_2) > 0$. Suppose there exists some sufficiently small $\varepsilon > 0$ such that $k_t(\theta) = 0$ for any $\theta \in (\theta_2, \theta_2 + \varepsilon)$. We will show in the following steps that we can find an alternative capital allocation $k'_t$ that satisfies $k'_t(\hat{\theta}) = \hat{k} > 0$ for some $\hat{\theta} \in (\theta_2, \theta_2 + \varepsilon)$ and $k'_t(\theta) = k_t(\theta)$ for $\theta \neq \hat{\theta}$ such that $(k'_t, \beta_t)$ is incentive compatible. Hence, we reach a contradiction because $(k'_t, \beta_t)$ gives a larger objective in (HJB) than the proposed optimal policy $(k_t, \beta_t)$.

Let $\hat{\theta} \in (\theta_2, \theta_2 + \varepsilon)$ be any project type that satisfies $\hat{\theta} - \theta_2 < R[k_t(\theta_2)]$. Let $cl[\Gamma(\theta, k_t)]$ denote the closure of the feasible set $\Gamma(\theta, k_t)$. Let $k_t(\theta^+)$ and $k_t(\theta^-)$ denote the right and left limit of $k_t(\theta)$ if they exist. From (HJB) we get

$$\beta_t(\hat{\theta}) = \sup_{\theta \in \Gamma(\hat{\theta}, k_t)} \left\{ \beta_t(\theta) + k_t(\theta) - R^{-1}[\theta + R(k_t(\theta)) - \hat{\theta}] \right\}$$

So there exist $\theta_1 \in cl[\Gamma(\hat{\theta}, k_t)]$ and $k_1 := \max\{k_t(\theta_1), k_t(\theta_1^+), k_t(\theta_1^-)\}$ such that $\beta_t(\hat{\theta}) = \beta_t(\theta_1) + k_1 - R^{-1}[\theta_1 + R(k_1) - \hat{\theta}]$.

First, we show that $\hat{\theta} < \theta_1 + R(k_1)$. Suppose not. Then $\theta_1 + R(k_1) \leq \hat{\theta} < \theta_2 + R(k_t(\theta_2))$. The second inequality is from the construction of $\hat{\theta}$. This relation plus the fact that $\hat{\theta} > \theta_2$ imply that $R(k_t(\theta_2)) > \max\{\theta_1 + R(k_1) - \theta_2, \theta_2 + R(k_t(\theta_2)) - \hat{\theta}\}$. Because $R^{-1}(.)$ is strict convex, we obtain

$$\begin{aligned} &\frac{1}{2}R^{-1}[R(k_t(\theta_2))] + \frac{1}{2}R^{-1}[\theta_1 + R(k_1) - \hat{\theta}] \\ &\quad > \frac{1}{2}R^{-1}[\theta_1 + R(k_1) - \theta_2] + \frac{1}{2}R^{-1}[\theta_2 + R(k_t(\theta_2)) - \hat{\theta}] \end{aligned} \tag{A.17}$$

Moreover, we know $\Gamma(\hat{\theta}, k_t) \subseteq \Gamma(\theta_2, k_t)$ because $k_t(\theta) = 0$ for $\theta \in (\theta_2, \hat{\theta}]$. So $\theta_1 \in cl[\Gamma(\theta_2, k_t)]$. From (IC) we get

$$\beta_t(\theta_2) \geq \beta_t(\theta_1) + k_1 - R^{-1}[\theta_1 + R(k_1) - \theta_2] \tag{A.18}$$

And because $\theta_2 \in \Gamma(\hat{\theta}, k_t)$, we get

$$\begin{aligned}\beta_t(\hat{\theta}) &\geq \beta_t(\theta_2) + k_t(\theta_2) - R^{-1}[\theta_2 + R(k_t(\theta_2)) - \hat{\theta}] \\ &\geq \beta_t(\theta_1) + k_1 - R^{-1}[\theta_1 + R(k_1) - \theta_2] + k_t(\theta_2) - R^{-1}[\theta_2 + R(k_t(\theta_2)) - \hat{\theta}] \\ &> \beta_t(\theta_1) + k_1 - R^{-1}[\theta_1 + R(k_1) - \hat{\theta}]\end{aligned}$$

which is a contradition. The second line is from plugging in (A.18). The third line is from plugging in (A.17). Hence, we must have $\hat{\theta} < \theta_1 + R(k_1)$. Then there exists $\hat{k} > 0$ such that $\hat{\theta} + R(\hat{k}) < \theta_1 + R(k_1)$. To show that $(k'_t, \beta_t)$ satisfies (IC), we only need to consider types in the range $(\hat{\theta}, \max\{\hat{\theta} + R(\hat{k}), \bar{\theta}\}]$ because the incentive compatible constraints of all other types will be the same as under $(k_t, \beta_t)$. Let $\tilde{\theta}$ be any type in this range.

Second, we show that the following relation holds

$$\beta(\tilde{\theta}) \geq \beta(\theta_1) + k_1 - R^{-1}[\theta_1 + R(k_1) - \tilde{\theta}] \tag{A.19}$$

By construction, we know that $\tilde{\theta} \leq \hat{\theta} + R(\hat{k}) < \theta_1 + R(k_1)$. If $k_1 = k_t(\theta_1)$, then $\tilde{\theta} < \theta_1 + R(k_t(\theta_1))$, which implies that $\theta_1 \in \Gamma(\tilde{\theta}, k_t)$. So (A.19) holds by (IC) under $(k_t, \beta_t)$. If $k_1 = k_t(\theta_1^+)$, then there must exist an sufficiently small $\eta > 0$ such that $\tilde{\theta} < \theta + R(k_t(\theta))$ for any $\theta \in (\theta_1, \theta_1 + \eta)$. Then (IC) under $(k_t, \beta_t)$ implies $\beta(\tilde{\theta}) \geq \beta(\theta) + k_t(\theta) - R^{-1}[\theta + R(k_t(\theta)) - \tilde{\theta}]$ for any $\theta \in (\theta_1, \theta_1 + \eta)$. Continuity then implies (A.19) holds. Similar argument shows that (A.19) holds if $k_1 = k_t(\theta_1^-)$.

Third, the strict convexity of $R^{-1}(.)$ and the relation $\theta_1 + R(k_1) - \hat{\theta} > \max\{R(\hat{k}), \theta_1 + R(k_1) - \tilde{\theta}\}$ implies

$$\begin{aligned}&\frac{1}{2}R^{-1}[\theta_1 + R(k_1) - \hat{\theta}] + \frac{1}{2}R^{-1}[\hat{\theta} + R(\hat{k}) - \tilde{\theta}] \\ &\quad > \frac{1}{2}R^{-1}[R(\hat{k})] + \frac{1}{2}R^{-1}[\theta_1 + R(k_1) - \tilde{\theta}]\end{aligned} \tag{A.20}$$

Combining (A.19) and (A.20), we get

$$\begin{aligned}\beta_t(\tilde{\theta}) &> \beta_t(\theta_1) + k_1 - R^{-1}[\theta_1 + R(k_1) - \hat{\theta}] + \hat{k} - R^{-1}[\hat{\theta} + R(\hat{k}) - \tilde{\theta}] \\ &= \beta_t(\hat{\theta}) + \hat{k} - R^{-1}[\hat{\theta} + R(\hat{k}) - \tilde{\theta}]\end{aligned} \tag{A.21}$$

Hence, $(k'_t, \beta_t)$ satisfies (IC). Then we get a contradiction because $(k_t, \beta_t)$ is not optimal. □

*Proof of Proposition 2. Part (a)*: Let $(k_t^*,\ \beta_t^*)$ be the optimal policy at time $t$. Define the lowest quality project with positive capital allocation as $\theta_t^c = \inf\{\underline{\theta} \leq \theta \leq \bar{\theta} : k_t^*(\theta) > 0\}$. Let us consider $\theta_t^c < \bar{\theta}$ since conclusion only applies to this case. If $k^*(\theta_t^c) > 0$, then Lema A.1 implies that there must exist an interval $(\theta_t^c, \theta_t^c + \varepsilon)$ over which $k_t^*(\theta) > 0$. If $k^*(\theta_t^c) = 0$, then the definition of $\theta_t^c$ also implies that such an interval of positive capital allocation exists. Suppose there are $\theta_1, \theta_2$ such that $k_t^*(\theta) > 0$ for $\theta \in (\theta_t^c, \theta_1)$ and $k_t^*(\theta) = 0$ for $\theta \in (\theta_1, \theta_2)$. Then by Lemma 2, $\beta_t^{*'}(\theta) \geq 1/R'(\hat{k}_t(\theta))$ for $\theta \in (\theta_t^c, \theta_1)$. Then $(k_t^*(\theta), \beta_t^*(\theta))$ must solve the following problem for $\theta \in (\theta_t^c, \theta_1)$:

$$
\begin{aligned}
&\max_{k_t,\beta_t} \frac{\lambda}{\Delta} \int_{\theta_t^c}^{\theta_1} [\theta + R(k_t(\theta)) - k_t(\theta)] d\theta \\
&+ \frac{\lambda}{\Delta} \int_{\theta_t^c}^{\theta_1} [P(W_t + \beta_t(\theta)) - P(W_t) - \beta_t(\theta) P'(W_t)] d\theta \\
&\quad s.t. \quad \beta_t'(\theta) = 1/R'(k_t(\theta)), \quad \beta_t(\theta_t^c) = \hat{\beta}_t(\theta_t^c)
\end{aligned}
\tag{P2}
$$

This is because the objective of (P2) is part of the objective of (HJB). Note that problem (P2) has the same structure as problem (P1). If we consider (P2) and (P1) as optimal control problems, then they differ only by initial states, and the starting and ending points. The same procedure as in Lemma 4 shows that $k_t^*(\theta)$ increases over $(\theta_t^c, \theta_1)$. Hence, we get $k_t^*(\theta_1) > 0$ and $k_t^*(\theta) = 0$ for all $\theta \in (\theta_1, \theta_2)$. This is a contradiction with Lemma A.1. So we must have $k_t^*(\theta) > 0$ for $\theta > \theta_t^c$. Hence, $k_t^*(\theta) > 0$, $\beta_t^*(\theta) > 0$, and $\beta_t^{*'}(\theta) = 1/R'(k_t^*(\theta))$ for $\theta > \theta_t^c$.

*Part (b)*: The conclusion applies only to the case $\theta_t^c > \underline{\theta}$. By the definition of $\theta_t^c$, we know $k_t^*(\theta) = 0$ for $\theta < \theta_t^c$. Since projects below $\theta_t^c$ do not get capital allocation, $\Gamma(\theta, k_t^*) = \emptyset$ if $\theta < \theta_t^c$. Therefore, the optimal contract will set $\beta^*(\theta) = 0$ for $\theta < \theta_t^c$. □

With the threshold property in Proposition 2, we can simplify (HJB) to be (HJB'). To characterize the optimal policy, we define the following two auxiliary problems:

$$Q_t(\theta_t^c) = \max_{k_t,\beta_t} \frac{\lambda}{\Delta} \int_{\theta_t^c}^{\bar{\theta}} [\theta + R(k_t(\theta)) - k_t(\theta)] d\theta \tag{P3}$$

$$+ \frac{\lambda}{\Delta} \int_{\theta_t^c}^{\bar{\theta}} [P(W_t + \beta_t(\theta)) - P(W_t) - \beta_t(\theta) P'(W_t)] d\theta$$

$$s.t. \quad \beta_t'(\theta) = 1/R'(k_t(\theta)), \quad \beta_t(\theta_t^c) = 0$$

and

$$\max_{\underline{\theta} \le \theta_t^c \le \bar{\theta}} \frac{\lambda}{\Delta} \int_{\underline{\theta}}^{\theta_t^c} \theta d\theta + Q_t(\theta_t^c) \tag{P4}$$

We can solve (HJB') in two steps: (i) given $\theta_t^c$, we find the optimal $k_t$ and $\beta_t$ that solves (P3); and (ii) given the value function $Q_t$ from (P3), we find the optimal threshold $\theta_t^c$ that solves (P4). Note that the first step simply repeats the procedure of solving (P1). We will use the same notations as before and focus on deriving the optimal $\theta_t^c$. Let $\eta_t \ge 0$ and $\delta_t \ge 0$ be the Lagrangian multipliers of the constraints $\theta_t^c \ge \underline{\theta}$ and $\theta_t^c \le \bar{\theta}$ respectively in (P4). The necessary condition of (P4) is:

$$\frac{\lambda}{\Delta} \theta_t^c + Q'(\theta_t^c) = -\eta_t + \delta_t \tag{A.22}$$

Moreover,

$$Q'(\theta_t^c) = -H_t(\theta_t^c, k_t, \beta_t, \mu_t) \tag{A.23}$$

Combine (A.23) and (A.22), and apply the constraint of $\beta_t(\theta_t^c) = 0$ to obtain:

$$\eta_t - \delta_t = \frac{\lambda}{\Delta} [R(k_t(\theta_t^c)) - k_t(\theta_t^c)] + \frac{\mu_t(\theta_t^c)}{R'(k_t(\theta_t^c))}$$

$$= g_t(k_t(\theta_t^c); \theta_t^c) \tag{A.24}$$

*Proof of Lemma 5. Part (a)*: if $\lim_{k \downarrow 0} \frac{R''(k)}{[R'(k)]^3} = 0$, then we get

$$\lim_{k \downarrow 0} g_t'(k; \theta_t^c) = \lim_{k \downarrow 0} \frac{\lambda R'(k)}{\Delta} \left[ 1 - \frac{1}{R'(k)} - \frac{\Delta \mu_t(\theta_t^c) R''(k)}{\lambda [R'(k)]^3} \right] = \infty$$

since $\mu_t(\theta_t^c)$ is finite. From (A.14), we know $g_t(k_t(\theta_t^c); \theta_t^c) > \lim_{k \downarrow 0} g_t(k; \theta_t^c) = 0$. Then (A.24) implies $\eta_t > \delta_t \geq 0$. Hence, $\theta_t^c = \underline{\theta}$ by complementary slackness.

*Part (b)*: if $\lim_{k\downarrow 0} \frac{R''(k)}{[R'(k)]^3} < 0$, then we have

$$\lim_{k\downarrow 0} \frac{\lambda[R'(k)-1][R'(k)]^2}{\Delta R''(k)} = \lim_{k\downarrow 0} \frac{\lambda[R'(k)]^3}{\Delta R''(k)} > -\infty$$

By continuity, there exists $a < 0$ such that

$$0 \geq \frac{\lambda[R'(k)-1][R'(k)]^2}{\Delta R''(k)} \geq a \tag{A.25}$$

for any $k \in (0, k^*]$. Suppose that $\theta_t^c = \underline{\theta}$. Then the optimal $k_t(\theta) > 0$ for any $\theta > \underline{\theta}$. Depending on the level of $W_t$, we possibly have $\mu_t(\underline{\theta}) < a$. In that case, there exists $\hat{\theta} > \underline{\theta}$ such that $\mu_t(\hat{\theta}) < a$. Then from (A.25),

$$\mu_t(\hat{\theta}) < \frac{\lambda[R'(k)-1][R'(k)]^2}{\Delta R''(k)}$$

for any $x \in (0, K^*]$. This implies the problem $\max_{k\geq 0} g_t(k;\hat{\theta})$ has no interior solution. Hence, we must have $k_t(\hat{\theta}) = 0$, a contradiction. Therefore, when $|\mu_t(\underline{\theta})|$ is sufficiently large, exclusion possibly occurs. □

*Proof of Proposition 3. Part (a)*: From equation (8), it is easy to get $R'(k_t(\bar{\theta})) - 1 = 0$.

*Part (b)*: Take any $\theta_t^c \leq \theta < \bar{\theta}$ and $0 < W_t < \bar{W}$. Since $P(W_t)$ and $R$ are strictly concave, and $\beta_t(\hat{\theta}) > 0$ for all $\hat{\theta} > \theta$, we know the right-hand side of (8) is positive. This implies $R'(k_t(\theta)) < 1$.

*Part (c)*: Take any $\theta$ and $W_t \geq \bar{W}$. Since $P'(W_t) = P'(W_t + \beta_t(\hat{\theta})) = -1$ for any $\hat{\theta} \geq \theta$, we know the right-hand side of (8) is zero. This implies $R'(k_t(\theta)) = 1$. Since $\theta$ is arbitrary and $k_t(\theta) > 0$, we must have $\theta_t^c = \underline{\theta}$. □

**LEMMA A.2.** *Under the assumption $R'R''' \leq 3(R'')^2$, equation $g_t'(k;\theta) = 0$ has at most one solution of $k$ given any value of $\mu_t(\theta) < 0$.*

*Proof of Lemma A.2.* Define $h(k) = \frac{(R'(k)-1)R'(k)^2}{R''(k)}$ and we have,

$$\begin{aligned} h'(k) &= \frac{R'(k)}{R''(k)^2}\left[(3R'(k)-2)R''(k)^2 - (R'(k)-1)R'(k)R'''(k)\right] \\ &> \frac{R'(k)(R'(k)-1)}{R''(k)^2}\left[3R''(k)^2 - R'(k)R'''(k)\right] \geq 0 \end{aligned}$$

when $k < k^*$. The last inequality is because $R'(k) > 1$ and $R'R''' \leq 3(R'')^2$. Note that $g'_t(k;\theta) = 0$ if and only if $h(k) = \frac{\Delta\mu_t(\theta)}{\lambda}$. For a given $\mu_t(\theta) < 0$, $g'_t(\hat{k};\theta) = 0$ implies that $\hat{k} < k^*$. Since $h(\cdot)$ is strictly monotone over $[0, k^*)$, the equation $h(k) = \frac{\Delta\mu_t(\theta)}{\lambda}$ has at most one solution of $k$. □

**Corollary A.1.** *Take any $\hat{W}, \tilde{W} \leq \bar{W}$ and $\hat{W} \neq \tilde{W}$. In the optimal contract, $k(\theta, \hat{W}) > k(\theta, \tilde{W})$ if and only if $\mu(\theta, \hat{W}) > \mu(\theta, \tilde{W})$. Moreover, if $k(\theta, \hat{W}) = k(\theta, \tilde{W})$ and $k_\theta(\theta, \hat{W}) > k_\theta(\theta, \tilde{W})$, then $\mu_\theta(\theta, \hat{W}) > \mu_\theta(\theta, \tilde{W})$*

*Proof.* From Lemma A.2, we know in the optimal contract $h[k(\theta, W)] = \frac{\Delta\mu(\theta,W)}{\lambda}$. Hence, the first part of the result is directly implied by the strict monotonicity of $h(\cdot)$. Moreover, since $h'[k(\theta, W)]k_\theta(\theta, W) = \frac{\Delta\mu_\theta(\theta,W)}{\lambda}$, $h'(\cdot) > 0$ also implies the second part of the result. □

**LEMMA A.3.** *Under the technology assumption $R'''R' \leq 3(R'')^2$, the problem* (HJB') *has a unique solution $(k_t, \beta_t, \theta^c_t)$.*

*Proof. Part(a)*: we show that the capital allocation policy $k_t(\theta)$ is continuous in $\theta$. When $\theta > \theta^c_t$, $k_t(\theta) > 0$ by Proposition 2. So $k_t(\theta)$ satisfies (A.16). From Lemma A.2, $k_t(\theta)$ is the unique solution of $g'_t(k;\theta) = 0$. Since $g'_t(k;\theta)$ is continuous in $k$ and $\theta$, we must have $k_t(\theta)$ is continuous in $\theta$.

Now let us consider the continuity at $\theta^c_t$. Let $\hat{k}_t = \lim_{\theta\downarrow\theta^c_t} k_t(\theta)$. We first show that $k_t(\theta^c_t) = \hat{k}_t$ meaning $k_t(\theta)$ is right continuous at $\theta^c_t$. It obviously holds if $k_t(\theta^c_t) > 0$ by the argument above. Suppose $k_t(\theta^c_t) = 0 < \hat{k}_t$. By continuity from $\theta > \theta^c_t$, we know that $g'_t(\hat{k}_t;\theta^c_t) = 0$. By Lemma A.2, either $g'_t(k;\theta^c_t) > 0$ for all $k < \hat{k}_t$ or $g'_t(k;\theta^c_t) < 0$ for all $k < \hat{k}_t$. The former case implies that $g_t(\hat{k}_t;\theta^c_t) > 0 = g_t(k_t(\theta^c_t);\theta^c_t)$, contradicting with $k_t(\theta^c_t)$ being optimal. The latter case implies that $g_t(\hat{k}_t;\theta^c_t) < 0$. However, we know from (A.14) that $g_t(k_t(\theta);\theta) \geq \lim_{k\downarrow 0} g_t(k;\theta) = 0$ for all $\theta$. So by continuity $g_t(\hat{k}_t;\theta^c_t) \geq 0$, a contradiction.

If $\theta_t^c = \underline{\theta}$, then right continuity means that $k_t(\theta)$ is continuous for all $\theta$. If $\theta_t^c > \underline{\theta}$, then complementary slackness means that $\eta_t = 0$. From (A.24), $g_t(k_t(\theta_t^c);\theta_t^c) = -\delta_t \leq 0$. So we must have $g_t(k_t(\theta_t^c);\theta_t^c) = 0$. Suppose $k_t(\theta_t^c) > 0$, then $g_t'(k_t(\theta_t^c);\theta_t^c) = 0$. Lemma A.2 implies that either $g_t'(k;\theta_t^c) > 0$ for all $k < k_t(\theta_t^c)$ or $g_t'(k;\theta_t^c) < 0$ for all $k < k_t(\theta_t^c)$. So we have either $g_t(k_t(\theta_t^c);\theta_t^c) > 0$ or $g_t(k_t(\theta_t^c);\theta_t^c) < 0$, a contradiction. Therefore $k_t(\theta_t^c) = 0$, implying that $k_t(\theta)$ is continuous at $\theta_t^c$ when $\theta_t^c > \underline{\theta}$.

*Part(b)*: we show that (HJB') has a unique optimal policy. Suppose $k_t(\theta), \beta_t(\theta)$ and $\hat{k}_t(\theta), \hat{\beta}_t(\theta)$ are both optimal and $\beta_t(\theta) \neq \hat{\beta}_t(\theta)$. Since $\beta_t(\underline{\theta}) = \hat{\beta}_t(\underline{\theta}) = 0$, we know there exists some $\tilde{\theta} < \bar{\theta}$ such that: (1) $\beta_t(\tilde{\theta}) = \hat{\beta}_t(\tilde{\theta})$, and (2) either $\beta_t(\theta) \geq \hat{\beta}_t(\theta)$ in the interval $[\tilde{\theta}, \bar{\theta}]$ or $\beta_t(\theta) \leq \hat{\beta}_t(\theta)$ in the interval and strict for some $\theta > \tilde{\theta}$. Without loss of generality, we assume $\beta_t(\theta) \geq \hat{\beta}_t(\theta)$ in the interval $[\tilde{\theta}, \bar{\theta}]$ and strict for some $\theta$. From (A.14) and (A.15), we can derive that $k_t(\theta) \leq \hat{k}_t(\theta)$ for all $\theta \in [\tilde{\theta}, \bar{\theta}]$. This is because larger $\beta$ leads to smaller multiplier $\mu$ and hence smaller $k$. However, the envelope condition implies that:

$$\begin{aligned}\beta_t(\theta) &= \beta_t(\tilde{\theta}) + \int_{\tilde{\theta}}^{\theta} \frac{1}{R'(k_t(u))} du \\ &\leq \hat{\beta}_t(\tilde{\theta}) + \int_{\tilde{\theta}}^{\theta} \frac{1}{R'(\hat{k}_t(u))} du = \hat{\beta}_t(\theta)\end{aligned}$$

for any $\theta > \tilde{\theta}$, a contradiction. So the optimal $\beta_t(\theta)$ is unique. Moreover, given any optimal $\beta_t(\theta)$, Lemma A.2 and the continuity of $k_t(\theta)$ in $\theta$ together imply that the optimal $k_t(\theta)$ is uniquely determined. Therefore, (HJB') has a unique optimal policy. □

*Proof of Lemma 6.* First, we need to show the necessary condition (8) for $\theta \geq \theta_t^c$ is sufficient for (P3). Let us define $H_t^0(\theta, \beta_t, \mu_t) = \max_{k_t \geq 0} H_t(\theta, \beta_t, k_t, \mu_t)$. Because $P(W)$ is concave, it is easy to see from (A.12) that $H_t$ is concave in $\beta_t$. Also from (A.12), the optimal $k_t$ that maximizes $H_t$ does not depend on the value of $\beta_t$. Hence, $H_t^0(\theta, \beta_t, \mu_t)$ is a concave function of $\beta_t$ for any given $\mu_t$. By Arrow Theorem (see P.222 of Kamien and Schwartz [1991]), the policy $k_t(\theta)$ and $\beta_t(\theta)$ will maximize (P3).

Second, we need to show the proposed policy $(k_t, \beta_t)$ satisfies (IC). Note that $\Gamma(\theta, k_t) \neq \varnothing$ only if $\theta > \theta_t^c$. Take any $\theta > \theta_t^c$ and $\hat{\theta} \in \Gamma(\theta, k_t)$. We know $\hat{\theta} \leq \theta$. Because $k_t(\theta)$ is increasing in $\theta$, we know $\hat{\theta} + R(k_t(\hat{\theta})) \leq R(k_t(x)) + x$ for any $x \in [\hat{\theta}, \theta]$. Let us define

$$b_t(x) = \int_{\hat{\theta}}^{x} 1/R'(k_t(u))du - k_t(\hat{\theta}) + k_t(\hat{\theta}; x)$$

It is easy to see that $b_t(\hat{\theta}) = 0$ and $b_t'(x) = R^{-1'}[R(k_t(x))] - R^{-1'}[\hat{\theta} + R(k_t(\hat{\theta})) - x] \geq 0$. The last inequality is because $R^{-1'}(.)$ is increasing which is implied by the concavity of $R$. Hence, $b_t(\theta) \geq 0$. Using the envelope condition $\beta_t'(x) = 1/R'(k_t(x))$, we can rewrite $b_t(\theta)$ as $b_t(\theta) = \beta_t(\theta) - \beta_t(\hat{\theta}) - k_t(\hat{\theta}) + k_t(\hat{\theta}; \theta) \geq 0$. Therefore, (IC) is satisfied. □

*Proof of Proposition 4.* Following the standard argument in optimal control theory, we show that HQ's value from any incentive compatible mechanism that delivers DM continuation value $W_0$ is at most $P(W_0)$, which is HQ's expected payoff from the conjectured mechanism. Let us define

$$G_t = \int_0^t e^{-rs}[(dX_s + R(k_s) - k_s)dN_s - dI_s] + e^{-rt}P(W_t)$$

In any incentive compatible contract, $W_t$ evolves according to (7) with $\hat{N}(dt, d\theta) = N(dt, d\theta)$. By Ito's Lemma,

$$\begin{aligned} e^{rt}E_t(dG_t) =& \frac{\lambda}{\Delta}\int_{\underline{\theta}}^{\bar{\theta}} (\theta + R(k_t(\theta)) - k_t(\theta))d\theta dt \\ &+ \left[\gamma W_t - \frac{\lambda}{\Delta}\int_{\underline{\theta}}^{\bar{\theta}} \beta_t(\theta)d\theta\right] P'(W_t)dt \\ &+ \frac{\lambda}{\Delta}\int_{\underline{\theta}}^{\bar{\theta}} \left[P(W_t + \beta_t(\theta)) - P(W_t)\right] d\theta dt - rP(W_t)dt \\ &- (1 + P'(W_t))dI_t \end{aligned}$$

Since $P(W_t)$ is constructed from the optimal policies $k_t^*(\theta), \beta_t^*(\theta)$, the first three lines must be less than or equal to zero. Moreover, $P'(W_t) \geq -1$ implies the last line is also less than or equal to zero. So, $G_t$ is a supermartingale. It is a martingale if and only if $k_t(\theta) = k_t^*(\theta), \beta_t(\theta) = \beta_t^*(\theta)$, and $dI_t \geq 0$ only when $W_t \geq \bar{W}$.

Now we evaluate the time zero expected payoff of the HQ from an arbitrary incentive compatible mechanism:

$$
\begin{aligned}
& E\left[\int_0^\tau e^{-rs}((dX_s + R(k_s) - k_s)dN_s - dI_s) + e^{-r\tau}L\right] \\
=& E\left[G_{t\wedge\tau} + \mathbb{1}_{t\le\tau}\left(\int_t^\tau e^{-rs}((dX_s + R(k_s) - k_s)dN_s - dI_s) + e^{-r\tau}L - e^{-rt}P(W_t)\right)\right] \\
=& E(G_{t\wedge\tau}) - e^{-rt}E[\mathbb{1}_{t\le\tau}P(W_t)] \\
& + e^{-rt}E\left\{\mathbb{1}_{t\le\tau}\left[E_t\left(\int_t^\tau e^{-r(s-t)}((dX_s + R(k_s) - k_s)dN_s - dI_s) + e^{-r(\tau-t)}L\right)\right]\right\} \\
\le& G_0 + e^{-rt}E\left[\mathbb{1}_{t\le\tau}\left(\frac{\lambda[E(\theta) + R(k^*) - k^*]}{r} - W_t - P(W_t)\right)\right] \\
\le& P(W_0) + e^{-rt}E\left[\mathbb{1}_{t\le\tau}\left(\frac{\lambda[E(\theta) + R(k^*) - k^*]}{r} - L\right)\right]
\end{aligned}
$$

The first inequality is from: (1) $G_t$ is a supermartingale; (2) HQ's expected payoff from time $t$ on is smaller than the first best surplus minus continuation value. The second inequality is from the fact that total surplus at any time is at least as large as liquidation value $L$. Therefore, as $t \to \infty$, the second piece in the last line converges to zero. And by definition, $G_0 = P(W_0)$. So we get

$$
E\left[\int_0^\tau e^{-rs}((dX_s + R(k_s) - k_s)dN_s - dI_s) + e^{-r\tau}L\right] \le P(W_0)
$$

Therefore, HQ's expected payoff under any incentive compatible mechanism is at most the expected payoff obtained from the mechanism in the proposition.

In the following, we first show that the value function $P$ is concave, and then show that there exists a continuation value $\bar{W} > 0$ such that $P(W_t)$ is strictly concave when $W_t < \bar{W}$ and $P'(W_t) = -1$ when $W_t \ge \bar{W}$. To ease notation, we define

$$
f(\beta_t, W_t) = \frac{\lambda}{\Delta}\int_{\underline{\theta}}^{\bar{\theta}} [P(W_t + \beta_t(\theta)) - P(W_t) - \beta_t(\theta)P'(W_t)]d\theta
$$

Let us consider any incentive compatible policy $(k_t, \beta_t)$ at $W_t$. From the objective of (HJB),

$$\begin{aligned}
&\frac{\lambda}{\Delta}\int_{\underline{\theta}}^{\bar{\theta}}[\theta + R(k_t(\theta)) - k_t(\theta)]d\theta + f(\beta_t, W_t) \\
\leq& rP(W_t) - \gamma W_t P'(W_t) \\
\leq& rP(W_t) + \gamma W_t \\
\leq& \frac{\lambda}{\Delta}\int_{\underline{\theta}}^{\bar{\theta}}[\theta + R(k_t(\theta)) - k_t(\theta)]d\theta
\end{aligned} \tag{A.26}$$

The second inequality is from $P'(W_t) \geq -1$. The last inequality is because the sum of HQ's flow profit and DM's flow utility should be less than the expected cash flow. Therefore, $f(\beta_t, W_t) \leq 0$. Note that for any sufficiently small $\varepsilon > 0$ the following policy is incentive compatible: $k_t(\theta) = 0$, $\beta_t(\theta) = \varepsilon$, for all $\theta$ at $W_t$. Under this policy,

$$f(\beta_t, W_t) = \lambda[P(W_t + \varepsilon) - P(W_t) - \varepsilon P'(W_t)] \leq 0 \tag{A.27}$$

which implies that $P(.)$ is concave at $W_t$ since $\varepsilon$ is arbitrarily small.

Since the DM is more impatient, the optimal contract cannot delay cash payment for ever. There must exist a $W$ such that $P'(W) = -1$. We define $\bar{W} = \min_{W \geq 0}\{P'(W) = -1\}$. Suppose that $\bar{W} = 0$. Since $P$ is concave and $P'(W) \geq -1$, $P$ has to be a straight line with slope $-1$. From the necessary condition (8), capital allocation is always first best. Also, the objective of (HJB) can be simplified as

$$rP(W_t) = -\gamma W_t + \lambda[E(\theta) + R(k^*) - k^*]$$

which contradicts with the assumption that $P'(W_t) = -1$, because $r < \gamma$. Hence, $\bar{W} > 0$. And by concavity of $P$, we know that $P'(W_t) = -1$ when $W_t \geq \bar{W}$ and $P'(W_t) < -1$ when $W_t < \bar{W}$.

We now show that $P(W_t)$ is strictly concave when $W_t < \bar{W}$. Suppose not. Then there exists $\hat{W}_t < \bar{W}$ and sufficiently small $\eta > 0$ such that $P'(W_t)$ is constant over $[\hat{W}_t, \hat{W}_t + \eta]$. Since $P'(\hat{W}_t) > -1$, (A.26) implies that $f(\beta_t, \hat{W}_t) < 0$ for any incentive compatible policy $(k_t, \beta_t)$. As before, we can find a incentive compatible policy such that (A.27) holds at $\hat{W}_t$ for any sufficiently small $\varepsilon > 0$. However, when $\varepsilon < \eta$, (A.26) implies that $f(\beta_t, \hat{W}_t) = 0$, a contradiction. $\square$

*Proof of Proposition 5.* Denote $dC_t$ as the DM's consumption and $S_t$ as the DM's saving account balance which evolves according to

$$dS_t = \rho S_t + dI_t + \int_{\underline{\theta}}^{\bar{\theta}} [k_t(\theta) - \hat{k}_t(\theta)]\hat{N}(dt, d\theta) - dC_t$$

We show that for any feasible strategy $(C, \hat{X}, \hat{K})$ of the DM, $\hat{V}_t = \int_0^t e^{-\gamma s} dC_s + e^{-\gamma t}(S_t + W_t)$ is a supermartingale. To see this,

$$\begin{aligned} e^{\gamma t} d\hat{V}_t &= dC_t - \gamma(S_t + W_t)dt + dS_t + dW_t \\ &= (\rho - \gamma)S_t dt + \int_{\underline{\theta}}^{\bar{\theta}} \beta_t(\theta) \left[\hat{N}(dt, d\theta) - \frac{\lambda}{\Delta} dt d\theta\right] + \int_{\underline{\theta}}^{\bar{\theta}} [k_t(\theta) - \hat{k}_t(\theta)]\hat{N}(dt, d\theta) \end{aligned}$$

The envelope condition and the monotonicity of $k_t(\theta)$ imply that $\beta_t(\theta) \geq \beta_t(\hat{\theta}) + k_t(\hat{\theta}) - R^{-1}[\hat{\theta} + R(k_t(\hat{\theta})) - \theta]$ for feasible $\theta'$ [10]. Hence,

$$\int_{\underline{\theta}}^{\bar{\theta}} [\beta_t(\theta) + k_t(\theta) - \hat{k}_t(\theta)]\hat{N}(dt, d\theta) \leq \int_{\underline{\theta}}^{\bar{\theta}} \beta_t(\theta) N(dt, d\theta) \tag{A.28}$$

which implies $\hat{V}_t$ is a supermartingale because $\rho < \gamma$ and $N(dt, d\theta) - \frac{\lambda}{\Delta} dt d\theta$ is a martingale. If the DM reports truthfully so that (A.28) holds as equality, and if there is no savings ($S_t = 0$), then $\hat{V}_t$ is a martingale. So we know,

$$W_0 = \hat{V}_0 \geq E(\hat{V}_\tau) = E\left[\int_0^\tau e^{-\gamma s} dC_s + e^{-\gamma \tau} S_\tau\right]$$

with equality only if the DM reports truthfully and does not save. $\square$

*Proof of Proposition 6.* Consider any $W \in (0, \bar{W})$. Suppose that $W_t = W$ and $\gamma W_t - \frac{\lambda}{\Delta} \int_{\underline{\theta}}^{\bar{\theta}} \beta_t(\theta) d\theta \geq 0$. Then $dI_t = 0$ by Lemma 3. Then $dW_t \geq 0$ by (7), because $\beta_t(\theta) \geq 0$ for all $\theta$. In addition, from (7) we know that $W_t$ cannot jump downward when $W_t \in (W, \bar{W})$. So if the contract starts at $W_0 \geq W$, then the continuation value will never drop below $W > 0$, which means liquidation never occurs (i.e., $\tau = \infty$). Then $k_t(\theta) = k^*$ for all $t$, a contradiction. $\square$

[10] Here $\hat{\theta}$ is possibly greater than $\theta$ since the DM has savings.

**LEMMA A.4.** *There exits $\hat{W} < \bar{W}$ such that $P''(\cdot)$ strictly increases over $[\hat{W}, \bar{W})$.*

*Proof.* Suppose such $\hat{W}$ does not exist. Then because of continuity, the value function determined by (HJB) will have negative second derivative for a small neighborhood above $\bar{W}$. Hence, $\int_{\underline{\theta}}^{\bar{\theta}} P(\bar{W}_- + \beta(\theta, \bar{W}_-)) - P(\bar{W}_-) - \beta(\theta, \bar{W}_-)P'(\bar{W}_-)d\theta < 0$, where $\bar{W}_-$ indicates the left limit of $\bar{W}$. Then (HJB) implies

$$\lambda[E(\theta) + R(k^*) - k^*] > rP(\bar{W}_-) - \gamma \bar{W} P'(\bar{W}_-) = rP(\bar{W}) + \gamma \bar{W}$$

The second equality is because $P'(\bar{W}_-) = -1$ by (BC). This forms a contradiction. □

Define $W^l := \inf_W \{P''(\cdot) \text{ strictly increases over } (W, \bar{W}]\}$. Now we show in the following two results that the optimal policies are monotone when continuation value is above $W^l$.

**LEMMA A.5.** *Take any $W^l \le W_1 < W_2$ and $\theta_1 < \theta_2$. Suppose $k(\theta, W_1) > k(\theta, W_2)$ for all $\theta \in (\theta_1, \theta_2)$, $k(\theta_2, W_1) = k(\theta_2, W_2)$, and $k_\theta(\theta_2, W_1) < k_\theta(\theta_2, W_2)$. Then $\mu_\theta(\theta_1, W_1) < \mu_\theta(\theta_1, W_2)$.*

*Proof.* From Lemma A.1, we know $\mu_\theta(\theta_2, W_1) < \mu_\theta(\theta_2, W_2)$. Since $P'(W_1) \ge P'(W_2)$, (A.15) implies taht $-P'(W_1 + \beta(\theta_2, W_1)) < -P'(W_2 + \beta(\theta_2, W_2))$. This further implies that $W_1 + \beta(\theta_2, W_1) < W_2 + \beta(\theta_2, W_2)$. Hence, we know $W_2 + \beta(\theta_2, W_2) > \max\{W_1 + \beta(\theta_2, W_1), W_1 + \beta(\theta_1, W_1), W_2 + \beta(\theta_1, W_2)\}$. Moreover, we know from the envelop condition (EN) that

$$\begin{aligned} &\Big[W_1 + \beta(\theta_1, W_1) + W_2 + \beta(\theta_2, W_2)\Big] - \Big[W_1 + \beta(\theta_2, W_1) + W_2 + \beta(\theta_1, W_2)\Big] \\ &= -\int_{\theta_1}^{\theta_2} \left( \frac{1}{R'(k(\theta, W_1))} - \frac{1}{R'(k(\theta, W_2))} \right) d\theta < 0 \end{aligned} \tag{A.29}$$

Since $P'(\cdot)$ is strictly convex and decreasing over $(W^l, \bar{W})$, we have

$$P'(W_1 + \beta(\theta_1, W_1)) + P'(W_2 + \beta(\theta_2, W_2)) > P'(W_1 + \beta(\theta_2, W_1)) + P'(W_2 + \beta(\theta_1, W_2))$$

which further implies that

$$
\begin{aligned}
&\mu_\theta(\theta_1, W_1) - \mu_\theta(\theta_1, W_2) \\
&\quad = \frac{\lambda}{\Delta}\Big[P'(W_1) - P'(W_2) + P'(W_2 + \beta(\theta_1, W_2)) - P'(W_1 + \beta(\theta_1, W_1)\Big] \\
&\quad < \frac{\lambda}{\Delta}\Big[P'(W_1) - P'(W_2) + P'(W_2 + \beta(\theta_2, W_2)) - P'(W_1 + \beta(\theta_2, W_1))\Big] \\
&\quad = \mu_\theta(\theta_2, W_1) - \mu_\theta(\theta_2, W_2) < 0 \qquad \square
\end{aligned}
$$

**LEMMA A.6.** *The optimal policies $k(\theta, W), \beta(\theta, W)$ are increasing in $W$ over $(W^l, \bar{W})$ for all $\theta$.*

*Proof.* Take $W_1 \geq W^l$ and $W_2 = W_1 + \varepsilon$. Suppose $k(\hat{\theta}, W_1) > k(\hat{\theta}, W_2)$ for some $\hat{\theta}$. Then there must exist $\theta_2 > \hat{\theta}$ such that $k(\theta_2, W_1) = k(\theta_2, W_2)$ and $k_\theta(\theta_2, W_1) < k_\theta(\theta_2, W_2)$. This is because $k(\bar{\theta}, W_1) = k(\bar{\theta}, W_2)$. There are two possible cases.

First, suppose there also exists $\theta_1 < \hat{\theta}$ such that $k(\theta_1, W_1) = k(\theta_1, W_2)$. Then we must have $k_\theta(\theta_1, W_1) > k_\theta(\theta_1, W_2)$. Without loss of generality, we can assume $k(\theta, W_1) > k(\theta, W_2)$ for all $\theta \in (\theta_1, \theta_2)$. Lemma A.5 implies $\mu_\theta(\theta_1, W_1) < \mu_\theta(\theta_1, W_2)$. However, by Lemma A.1 we must also have $\mu_\theta(\theta_1, W_1) > \mu_\theta(\theta_1, W_2)$, a contradiction.

Second, suppose $k(\theta, W_1) < k(\theta, W_2)$ for all $\theta < \hat{\theta}$. Then Lemma A.5 implies $\mu_\theta(\underline{\theta}, W_1) < \mu_\theta(\underline{\theta}, W_2)$. This also forms a contradiction because $\mu_\theta(\underline{\theta}, W_1) = \mu_\theta(\underline{\theta}, W_2) = 0$ by $\beta(\underline{\theta}, \cdot) = 0$.

Hence, $k(\theta, W_1) \leq k(\theta, W_2)$ for all $\theta$. By (EN) we know $\beta(\theta, W_1) \leq \beta(\theta, W_2)$ for all $\theta$. $\square$

To ease notation in the following proof, we define

$$
n(\beta, W) = P(W + \beta) - P(W) - \beta P'(W), \quad m(\theta, W) = n(\beta(\theta, W), W) \tag{A.30}
$$

**LEMMA A.7.** *If $W^l > 0$, then $\lim_{w \uparrow w^l} k_w(\underline{\theta}, W) \leq 0$.*

*Proof.* Since $W^l > 0$, we can define $\underline{W} = \inf_W\{P''(\cdot)$ strictly decreases over $(W, W^l)\}$. Take any sufficiently small $\varepsilon, \eta > 0$ such that $\underline{W} + \varepsilon + \eta < W^l$. Also, take any $W_1 \in (\underline{W}, W^l - \varepsilon - \eta)$. Let $W_2 = W_1 + \varepsilon$ and Let $\hat{\theta}$ be the project type that satisfies $\beta(\hat{\theta}, W_1) = \eta$. Consider the following program:

$$\max_{k,\beta} \frac{\lambda}{\Delta} \int_{\underline{\theta}}^{\hat{\theta}} \Big[\theta + R(k(\theta)) - k(\theta) + m(\theta, W_2)\Big] d\theta \tag{P4}$$
$$s.t. \quad \beta'(\theta) \geq 1/R'(k(\theta)), \quad \beta(\underline{\theta}) \geq 0, \quad \beta(\hat{\theta}) \leq \eta$$

Let $k(\theta), \beta(\theta)$ be the optimal policy of (P4). Suppose $k(\underline{\theta}) > k(\underline{\theta}, W_1)$. Then we must have $k(\theta) > k(\theta, W_1)$ for all $\theta \leq \hat{\theta}$. Suppose not. Then there must exist $\tilde{\theta} < \hat{\theta}$ such that $k(\theta) > k(\theta, W_1)$ for all $\theta \in (\underline{\theta}, \tilde{\theta})$, $k(\tilde{\theta}) = k(\tilde{\theta}, W_1)$, and $k_\theta(\tilde{\theta}) < k_\theta(\tilde{\theta}, W_1)$. By Corollary A.1, $\mu_\theta(\tilde{\theta}) < \mu_\theta(\tilde{\theta}, W_1)$.

From the envelop condition (EN),

$$\beta(\tilde{\theta}) - \beta(\tilde{\theta}, W_1) = \int_{\underline{\theta}}^{\tilde{\theta}} \left( \frac{1}{R'(k(\theta))} - \frac{1}{R'(k(\theta, W_1))} \right) d\theta > 0$$

Since $P'(\cdot)$ is strictly concave and decreasing over $(\underline{W}, W^l)$ and $W_2 + \beta(\tilde{\theta}) < W_2 + \eta < W^l$,

$$P'(W_1) + P'(W_2 + \beta(\tilde{\theta})) < P'(W_1 + \beta(\tilde{\theta}, W_1)) + P'(W_2)$$

Rearrange the above relation to get $\mu_\theta(\tilde{\theta}, W_1) < \mu_\theta(\tilde{\theta})$, a contradiction. Hence, $k(\theta) > k(\theta, W_1)$ for all $\theta \leq \hat{\theta}$. But this implies $\beta(\hat{\theta}) > \beta(\hat{\theta}, W_1) = \eta$, a contradiction. Therefore, $k(\underline{\theta}) \leq k(\underline{\theta}, W_1)$. The above argument also implies $k(\theta) < k(\theta, W_1)$ over some interval $(\underline{\theta}, \theta^h)$ with $\theta^h < \hat{\theta}$.

By continuity of $\beta(\cdot, \cdot)$ over $W$, the optimal policy $\beta(\hat{\theta}, W_2)$ is sufficiently close to $\eta = \beta(\hat{\theta}, W_1)$, since $\varepsilon$ is sufficiently small. This means $k(\theta)$, the solution of (P4), is also sufficiently close to $k(\theta, W_2)$ when $\theta \in (\underline{\theta}, \theta^h)$. Hence, we must have $k(\theta, W_1) > k(\theta, W_2)$ for $\theta \in (\underline{\theta}, \theta^h)$. This further means $k(\underline{\theta}, W_1) \geq k(\underline{\theta}, W_2)$ by continuity of $k(\cdot, \cdot)$ over $\theta$. Then the result follows from the arbitrary pick of $W_1$. □

*Proof of Proposition 7.* Let $\theta^c = \theta^c(W^l)$ and $\hat{W} = W^l + \varepsilon$ for any sufficiently small $\varepsilon > 0$. First, we show that $k(\theta^c, W^l) < k(\theta^c, \hat{W})$. Note by Lemma A.6 it suffices to show $k(\theta^c, W^l) \neq k(\theta^c, W^l + \varepsilon)$. Suppose not. By the convexity of $P'(\cdot)$ over $[W^l, \bar{W})$, we know $P'(W^l) + P'(\hat{W} + \eta) > P'(W^l + \eta) + P'(\hat{W})$ for any $\eta > 0$. Hence, there exists $\hat{\theta}$ (very close to $\theta^c$) such that $P'(W^l) + P'(\hat{W} + \beta(\theta, \hat{W})) > P'(W^l + \beta(\theta, W^l)) + P'(\hat{W})$ for $\theta \in (\theta^c, \hat{\theta})$.

This means $\mu_\theta(\theta, W^l) > \mu_\theta(\theta, \hat{W})$ for $\theta \in (\theta^c, \hat{\theta})$. Moreover, because $\mu(\theta^c, W^l) = \mu(\theta^c, \hat{W})$, we must have $\mu(\hat{\theta}, W^l) > \mu(\hat{\theta}, \hat{W})$ implying $k(\hat{\theta}, W^l) > k(\hat{\theta}, \hat{W})$. This forms a contradiction with Lemma A.6. Since $\varepsilon$ is arbitrarily small, we have $\lim_{w \downarrow w^l} k_w(\theta^c, W) > 0$.

If $W^l > 0$, then Lemma A.7 implies $\lim_{w \uparrow w^l} k_w(\theta^c, W) \leq 0$. This forms a contradiction with the fact that $k_w(\cdot, \cdot)$ is continuous in $W$. Hence, $W^l = 0$. Or in other words, $P''(\cdot)$ strictly increases over $(0, \bar{W})$. Therefore, Lemma A.6 implies $k(\theta, W)$ and $\beta(\theta, W)$ are increasing in $W$ for all $\theta$ and all interior $W$. □

*Proof of Proposition 8:* Let $Com_t$ be an increasing process denoting the cash compensation. The budgeting account balance evolves according to:

$$\begin{aligned} dM_t &= -e(M_t)M_t dt + \int_{\underline{\theta}}^{\bar{\theta}} \beta_t(\theta)\hat{N}(dt, d\theta)d\theta - dCom_t \\ &= \gamma M_t d_t + \int_{\underline{\theta}}^{\bar{\theta}} \beta_t(\theta) \left[\hat{N}(dt, d\theta) - \frac{\lambda}{\Delta} dt d\theta\right] - dCom_t \end{aligned}$$

Since the DM can issue all the account balance as cash compensation at any time, $M_t$ is the continuation value of the DM. Proposition 5 implies that DM always reports project information truthfully since the envelope and monotonicity conditions are satisfied. Moveover, no cash compensation is issued until $M_t$ reaches $\bar{M}$, because Proposition 5 shows that DM at most earns payoff $M_0$ under any $Com$ process. □

# References


Philippe Aghion and Jean Tirole. Formal and real authority in organizations. *The Journal of Political Economy*, 105:1–29, 1997.

Shashwat Alok and Radhakrishnan Gopalan. Managerial compensation in multi-division firms. *Working paper*, 2013.

Antonio E. Bernardo, Hongbin Cai, and Jiang Luo. Capital budgeting in multidivision firms: Information, agency, and incentives. *The Review of Financial Studies*, 17:739–767, 2004.

Bruno Biais, Thomas Mariotti, Jean-Charles Rochet, and Stephane Villeneuve. Large risks, limited liability, and dynamic moral hazard. *Econometrica*, 78:73–118, 2010.

Richard A. Brealey, Stewart C. Myers, and Franklin Allen. *Principles of Corporate Finance*. McGraw-Hill/Irwin, New York, NY, 2011.

Gian Luca Clementi and Hugo A. Hopenhayn. A theory of financing constraints and firm dynamics. *Quarterly Journal of Economics*, 121:229–265, 2006.

Massimo G. Colom and Marco Delmastro. Delegation of authority in business organizations: an empirical test. *Journal of Industrial Economics*, LII:53–80, 2004.

Peter M. DeMarzo and Yuliy Sannikov. Optimal security design and dynamic capital structure in a continuous-time agency model. *Journal of Finance*, 6:2681–2724, 2006.

Peter Eso and Balazs Szentes. Dynamic contracting with adverse selection: An irrelevance result. *Working Paper*, 2013.

Daniel F. Garrett and Alessandro Pavan. Managerial turnover in a changing world. *Journal of Political Economy*, 120:879–925, 2012.

John R. Graham, Campbell R. Harvey, and Manju Puri. Capital allocation and delegation of decision-making authority within firms. *Working Paper, Duke University*, 2014.

Sebastian Gryglewicz and Barney Hartman-Glaser. Dynamic agency and real options. *Working paper*, 2013.

Milton Harris and Artur Raviv. The capital budgeting process: Incentives and information. *Journal of Finance*, 4:1139–1174, 1996.

Morton I. Kamien and Nancy L. Schwartz. *Dynamic Optimization The Calculus of Variation and Optimal Control in Economics and Management*. Elsevier North Holland, Amsterdam, The Netherlands, 1991.

Jean-Jacques Laffont and Jean Tirole. Using cost observation to regulate firms. *Journal of Political Economy*, 94:614–641, 1986.

Jean-Jacques Laffont and Jean Tirole. Repeated auctions of incentive contracts, investment, and bidding parity with an application to takeovers. *The RAND Journal of Economics*, 19(4):516–537, 1988.

Günter Last and Andreas Brant. *Marked Point Processes on the Real Line: The Dynamical Approach*. Springer, NY, USA, 1995.

Andrey Malenko. Optimal dynamic capital budgeting. *Working Paper*, 2013.

Michael Mussa and Sherwin Rosen. Monopoly and product quality. *Journal of Economic Theory*, 18:301–317, 1978.

Roger B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6:58–73, 1981.

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

Jeremy C. Stein. Agency, information and corporate investment. *Handbook of the Economics of Finance*, pages 110–163, 2003.

John Y. Zhu. Optimal contracts with shirking. *Review of Economic Studies*, 0:1–28, 2012.