

Incentives, project choice, and dynamic multitasking*

Martin Szydlowski[†]

January 10, 2019

Abstract

I study the optimal choice of projects in a continuous-time moral hazard model with multitasking. I characterize the distortions caused by moral hazard and the dynamics of the firm's project choice. Both overinvestment and underinvestment relative to an NPV criterion can occur on the path of the contract. As past performance increases, the firm chooses projects which require higher pay-performance sensitivity. When the continuation value is large, investment projects are chosen more efficiently, and project choice depends more on the NPV and less on the incentive costs.

I implement the optimal contract with an equity stake, bonus payments, and a personal account.

Keywords: Continuous-time contracting, project choice, multitasking, bonus payments.

JEL Classification: D86, G11, G31, G32, M12, M52.

*I am indebted to Jeffrey Ely, Michael Fishman, and Bruno Strulovici for their guidance. I would like to thank Bruno Biais, Peter DeMarzo, Janice Eberly, Arvind Krishnamurthy, Alessandro Pavan, Mark Satterthwaite, and Yuliy Sannikov for helpful discussions, Zhiguo He for sharing MATLAB code, seminar participants at LSE, MIT, Penn State, UBC, UNC, the University of Montreal, and UW-Madison, and conference participants at the Canadian Economic Theory Conference, the North American Summer Meeting of the Econometric Society, the International Conference on Game Theory at Stony Brook, the Society for Economic Dynamics, the Transatlantic Doctoral Conference, and the Western Finance Association Conference. I am grateful to the editor and two anonymous referees for their helpful comments and suggestions.

[†]Department of Finance, Carlson School of Management, University of Minnesota, 321 19th Ave S, Minneapolis, MN 55401. Email: szydl002@umn.edu.

1 Introduction

Managers oversee different projects. They select markets to enter, direct product development, decide which employees to promote or fire, and allocate capital across divisions. Incentive contracts must guarantee that the manager allocates his effort to the most profitable projects, not just that he works hard. But when effort is unobservable, which project is profitable depends on how the cost of incentives interacts with its payoff characteristics. With forward looking managers, as past performance changes so must the cost of incentives. Project choice is therefore endogenous and changes over time.

Despite the prevalence of long-term contracts in the real world, almost all of the literature on moral hazard with multiple tasks is static. The following important questions are so far unanswered. (1) Which projects should be chosen at any given time? (2) How does the project choice change with the firm's past performance? (3) Is there a selection criterion the firm can use which takes the agency costs into account? (4) Can a simple scheme, based on real-world instruments, implement the optimal contract?

To answer these questions, I build a continuous-time moral hazard model with multiple projects. I characterize how moral hazard distorts project choice, both across projects and across time, and I examine the implications for the optimal provision of incentives.

In my model, a firm has access to a fixed number of projects and a manager decides whether to work or shirk in each. When he works, the project generates output which follows an arithmetic Brownian Motion, with different drift and volatility for each project.¹ Both principal and agent are risk-neutral. The agent is protected by limited liability and the principal conditions payments to the agent on the observed history of output. Incentivizing effort is costly since the agent has to be compensated for the effort cost. Also, whenever the agent exerts effort in a project, the contract must be sufficiently sensitive with respect to that project's output. Because of limited liability, sufficiently bad performance can lead to termination, which again is costly for the principal. Thus, there is a tradeoff between the project's flow payoff and the induced termination risk, which can depend on the history of the contract.

I characterize this tradeoff and show that it can be expressed as a function of the project's NPV and its risk-return ratio. The former captures the social value of the project while the

¹Throughout the paper, I refer to drift as the return and to volatility as the risk of the project.

latter affects the implicit cost of incentives. The distortions of moral hazard can thus be interpreted as deviations from the NPV criterion.

These deviations from NPV can take the form of *overinvestment* or *underinvestment*. That is, the principal may allow the agent to shirk on a positive NPV project but she may also direct him to work on a project with negative NPV on the path of the contract. The intuition for underinvestment is standard: because of the cost of incentives, the principal foregoes projects which are socially valuable. The overinvestment result is new and requires a dynamic setting with multiple projects. It is caused by the principal's inability to punish the manager with lower payments in the presence of limited liability.² With a single project, the principal cannot demand money from the manager after bad performance, so the only option is to fire him, which is inefficient.³ With multiple projects, the principal can instead incentivize effort in a project she would otherwise not choose. This lowers the manager's realized flow utility because of the effort cost, so it serves as an alternative punishment device. Because it substitutes for costly firing, this punishment can be valuable to the principal even when the NPV of the project is negative. Importantly, overinvestment is only temporary on the path of the contract. As performance improves, the principal stops incentivizing any negative NPV projects and instead chooses ones with positive NPV.⁴

I characterize the dynamics of the project choice and derive the set of chosen projects at any point in time. This delivers several testable implications. Firms with sufficiently high past performance choose projects which require higher pay-performance sensitivity. Those are projects whose risk-return ratios are relatively low.⁵ The firm's choice is also more in line with the NPV criterion. If across projects, higher return implies the project is more costly to incentivize, firms choose projects with higher risk and higher return when their past performance is better. Conversely, they choose safer projects with lower return when their performance is worse.⁶

²Managers have limited liability with respect to the firm's profits. In practice, the scope of clawback provisions is often limited, see [Babenko et al. \(2017\)](#).

³[Zhu \(2013\)](#) has studied shirking in a dynamic moral hazard model with linear utilities and limited liability. Throughout this paper, I focus on the case when shirking would never be optimal to avoid complications stemming from a degenerate continuation value process which arise in Zhu's paper.

⁴In other words, a negative NPV project is only valuable because there are other positive NPV projects which can be chosen at other points in time.

⁵The risk-return ratio is the average project return conditional on the agent exerting effort divided by the volatility of the project output. I define this ratio formally in Section [3.3](#).

⁶This finding also differs from the seminal risk-shifting result in [Jensen and Meckling \(1976\)](#). In that

We can alternatively understand the model as a theory of managerial task allocation. High performing managers are allocated more profitable but more difficult to incentivize tasks, while low performing ones work on tasks which are easy to incentivize (i.e. require a low pay-performance sensitivity) but not very valuable. When a manager’s performance decreases, he may even be allocated “busywork”, which is not valuable but easy to incentivize.

From an applied perspective, the model provides a novel explanation for overinvestment.⁷ The model also provides a rationalization for the use of bonus contracts. As [Murphy \(1999\)](#) documents, most CEO contracts consist of equity and a bonus, and the latter is a linearly weighted function of the CEO’s performance across different categories. The optimal contract in my paper can be implemented using an equity stake and bonus payments. Bonus payments are used since they help fine-tune the manager’s incentives when there are multiple projects, by conditioning the manager’s payoffs on the output of individual projects. This cannot be achieved with an equity stake alone.

The manager’s pay-performance sensitivity can be decomposed into the sensitivity of the equity stake and the one of his bonus payments, which are both determined by the choice of projects and therefore the manager’s past performance. Again, this yields testable implications since we can back out the pay-performance sensitivities from the risk-return characteristics of the currently chosen projects.

The paper makes several technical contributions. First, there might be interactions between projects, i.e. incentivizing one project might alter the cost of incentives for others. Then, the optimal project choice would have to be determined via a discrete optimization problem, which is difficult to solve. In my setting, incentivizing effort in one project does not affect the *current* cost of incentives for other projects,⁸ although it does alter the dynamics of the contract and therefore *future* cost of incentives. Therefore, at any point in time, whether the agent should exert effort can be determined independently for each project, using a criterion which links the project’s characteristics to the current cost of incentives.

paper, the possibility of liquidation leads firms to take on excessive risk, while in my work, it deters the firm from taking risky projects.

⁷Overinvestment is well documented in the empirical literature. For example, [Harford \(1999\)](#) finds that acquisitions of cash-rich firms decrease value. [Richardson \(2006\)](#) uses accounting measures to document investment in negative NPV projects.

⁸This is because I am assuming that the effort cost is linearly separable. Doing so allows me to cleanly separate the dynamic effects of moral hazard. Without moral hazard, the choice of projects is independent over time. With moral hazard, past profits in one project affect the future value of another project.

Second, incentivizing effort in a project implies the agent’s continuation value must feature a volatility term with respect to the noise of that project’s output. Since effort is not continuous, changing the set of projects the agent exerts effort in implies a jump in the volatility term of the continuation value, so existence of the contract has to be carefully considered. I show that the optimal contract can nonetheless be characterized via a Hamilton-Jacobi-Bellman equation.

The paper proceeds as follows. Section 2 provides an overview of the related literature. Section 3 introduces the model, and illustrates basic results on the incentive scheme and the principal’s value function. Section 4 is the core of the paper and discusses the optimal project selection scheme under both aggregate and project-based incentives. The implementation outlined in the paragraphs above is derived in Section 5. Section 7 concludes.

2 Related Literature

My paper contributes to the continuous-time contracting literature and uses the martingale representation approach (Schattler and Sung (1993) and Sannikov (2008)).

Underinvestment is a common result in this literature. He (2009) studies a model where the payoffs follow a geometric Brownian Motion. The agent controls the evolution of firm size by choosing to either work or shirk. For low continuation values, the contract may feature shirking. We can interpret this as underinvestment, since the firm grows more slowly.

Biais et al. (2010) study investment and downsizing as a way to incentivize accident prevention. In their model, effort lowers the arrival rate of a Poisson shock that lowers firm value. When the agent’s continuation value is low, reducing promised pay cannot incentivize effort, because the agent’s limited liability constraint would be violated. Then, the principal provides punishments by downsizing the project. In this sense, Biais et al. (2010) rationalize underinvestment.

DeMarzo et al. (2012) is the first paper to consider neoclassical investment in a continuous-time agency framework. Their framework builds on DeMarzo and Sannikov (2006), i.e. there is a single task, payoffs follow an arithmetic Brownian Motion, the agent either works or shirks, and everyone is risk neutral. In addition, the principal may invest in capital, which increases average profits. Inefficient termination reduces the value of capital, so there is underinvestment.

Zhu (2013) also builds on DeMarzo and Sannikov (2006), but he considers contracts which allow shirking. Shirking may be optimal for low continuation values, where it is used as a punishment, or for high continuation values, where it is used as a reward. Fong (2007) considers incentive contracts for surgeons, who have privately observed quality and who can decide whether to avoid risky patients. This “risk selection” increases the surgeon’s average performance but it may reduce patient welfare. The optimal contracts for both good and bad surgeons are distorted. The bad surgeon always engages in selection behavior and the good surgeon engages in it when his continuation value is low. In this sense, the optimal contract features shirking.

My underinvestment result is similar to the above papers. For low continuation values, the cost of incentives is high, because termination is likely. This lowers the value of effort (Fong (2007), He (2009), and Zhu (2013)), capital investments (Biais et al. (2010) and DeMarzo et al. (2012)), and, in my case, risky projects.

My overinvestment result is new compared to this literature. It relies on the following conceptual insight. Incentivizing effort in a negative NPV project can effectively relax the agent’s limited liability constraint and help avoid inefficient termination. This is because negative NPV projects can act as punishment devices. To obtain such a result, the principal needs to have different projects, with different NPVs and risk/return ratios. My implementation, which features equity and project-specific bonus payments is, to my knowledge, also new and has not been obtained by other papers following Sannikov (2008) or DeMarzo and Sannikov (2006). This is because multiple projects are necessary to capture the key stylized facts of bonus schemes.⁹

The problem of multitasking has received significant attention since the seminal article of Holmström and Milgrom (1991). Most of this literature assumes correlation between task outputs (see e.g. Holmström and Milgrom (1991)) or complementarity in the profit or total effort cost (see e.g. Dewatripont et al. (2000) and the discussion therein or Bond and Gomes (2009) for a recent contribution). In my paper, there is no such dependence. That is, project outputs are independent and the effort cost is linear. This allows me to establish a clean

⁹As Murphy (1999) describes, CEOs are evaluated according to “customer satisfaction, operational and/or strategic objectives (such as increasing plant capacity, bringing a new computer system on line by a particular date, reducing time-to-market, etc.) and measures of plant safety” in addition to profits or earnings. The total bonus is additive in the individual performance categories. With a single project, it is impossible to rationalize such a structure.

benchmark. Without moral hazard, projects would be chosen independently of each other. With moral hazard, certain projects are always chosen at the same time on the equilibrium path. Thus, any dependence between projects arises endogenously.

In my model, negative NPV projects (i.e. overinvestment) provide value because they effectively relax the agent's limited liability constraint. [Bohren and Kravitz \(2016\)](#) exploit a similar mechanic, albeit without any negative NPV projects.¹⁰ In a static model with multiple tasks and multiple agents, they show that bundling tasks into a single contract and assigning multiple agents to each task also relaxes the limited liability constraint.¹¹

Because of the complexity of multitask models, dynamic studies of multitasking are rare. [Manso \(2011\)](#) studies the trade-off between two tasks interpreted as exploration and exploitation and [Miquel-Florensa \(2007\)](#) characterizes whether two tasks should be executed sequentially or in parallel, depending on the strength of the externalities between them.

My model is also related to the literature on capital budgeting. For example, [Harris and Raviv \(1996\)](#) and [Harris and Raviv \(1998\)](#), study the choice of projects when a division manager has private information about project quality and has an incentive to misreport. Models in capital budgeting can explain underinvestment (e.g. [Bernardo et al. \(2006\)](#)) and overinvestment (e.g. [Harris and Raviv \(1996\)](#) and [Bernardo et al. \(2009\)](#)) in static settings, but they provide no guidance on how the firm's project choice changes over time or how it depends on past performance.¹² An exception is [Malenko \(2018\)](#), who considers a dynamic version [Harris and Raviv \(1996\)](#). He derives the capital budgeting mechanism in continuous time. Projects arrive at a random times and they have uncertain quality, which is privately observed by an agent. The optimal contract specifies how much capital the principal contributes to each project and which projects are audited to verify the manager's reports. The contract can be implemented with a discretionary account for the agent and a quality threshold above which projects are audited. In the papers on capital budgeting, projects have a one-dimensional quality. By contrast, my results on project choice dynamics (Section 4.2) are driven by a tradeoff between risk and return.

¹⁰See their Assumption 1.

¹¹I am grateful to a referee for pointing out this connection.

¹²In [Harris and Raviv \(1996\)](#) an agent has private information about the value of a project and prefers to overstate this value to receive a private benefit. The optimal contract then may optimally allow for some overinvestment.

3 Model Setup

3.1 Projects

I study a firm that has access to a fixed number of projects and hires a manager. The shareholders of the firm act as the principal and the manager as the agent. Time is continuous, indexed by t , and the horizon is infinite. Each project yields risky payoffs, and the agent decides in which projects to exert effort at any given time. A project's average payoff is μ_i if the agent works and zero otherwise. Formally, for any project $i \in \{1, \dots, N\}$, the cumulative output x_{it} is given by

$$dx_{it} = \mu_i a_{it} dt + \sigma_i dB_{it} \quad (1)$$

with initial condition $x_{i0} = 0$ for all i . The Brownian Motion B_{it} captures the noise in the project's payoff, while $a_{it} \in \{0, 1\}$ denotes the agent's decision to work or shirk. The project's average payoff given effort is μ_i and σ_i measures the amount of noise. For two projects i and j , the Brownian Motions B_{it} and B_{jt} are mutually independent, so that the path of each project's payoffs is determined by the agent's effort and the noise in that project alone.¹³

I denote the agent's allocation of effort among projects with the vector $a_t = (a_{it})_{i=1}^N$ and the set of all possible allocations with $\mathcal{A} = \{0, 1\}^N$. Whenever $a_{it} = 1$ on the path of the contract, I say project i is being *chosen* or *selected* at time t . When a project is chosen on an interval left of t' , but not chosen at t' , I say that the project is *stopped* at t' .¹⁴ The project is *started* when the opposite holds. I also call a_t the *project choice* or *project allocation* at time t . Additionally, the principal can decide to temporarily shut down a project so it does not produce any output.¹⁵ I denote these shutdown decisions as $k_{it} \in \{0, 1\}$ and I let $k_t = (k_{it})_{i=1}^N$.

Throughout this paper, I focus on the case in which letting the agent shirk in all projects is never optimal.¹⁶ I establish sufficient conditions in Appendix A.3.

¹³Formally, $B_{it} \perp B_{js}$ for all times $t, s \geq 0$. This helps to ensure that in the optimal contract, any dependence between projects is driven by the agency friction, and not by assumptions on the firm's technology. The framework can be extended to allow for direct complementarities between projects.

¹⁴Formally, project i is stopped at time t' if for some $\varepsilon > 0$, $a_{it} = 1$ for $t \in (t' - \varepsilon, t')$ and $a_{it'} = 0$.

¹⁵See also the discussion in the Section 3.2.

¹⁶This guarantees that the continuation value process for the agent has non-degenerate volatility, which

3.2 Utility Functions and the Contract Space

The vector of project-specific Brownian Motions $B_t := (B_{1t}, \dots, B_{Nt})$ is defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration \mathcal{F}_t , which satisfies the usual conditions.¹⁷ While each project's output can be fully observed by the principal and contracted upon, effort is unobservable. The principal commits to a contract, which consists of a cumulative consumption process $c = \{c_t \in \mathbb{R}_+ : t \geq 0\}$, a prescribed effort process $a = \{a_t \in \mathcal{A} : t \geq 0\}$, project shutdown decisions $k = \{k_t \in \{0, 1\}^N, t \geq 0\}$, and a firing time τ . Effort, consumption, and shutdowns are progressively measurable with respect to \mathcal{F}_t , and τ is an \mathcal{F}_t -stopping time. I characterize incentive-compatible contracts, in which the prescribed effort process equals the agent's actual effort process \hat{a}_t .

Both principal and agent are risk-neutral. They have different discount factors, r and γ respectively, and the agent is less patient, i.e. $r < \gamma$.¹⁸ The agent's total effort cost is linear and symmetric in effort for each project, and given by $h \sum_i a_{it}$. This specification, together with the mutual independence of the Brownian Motions in the project outputs, ensures that any potential dependence between projects along the path of the optimal contract is driven by the agency friction instead of assumptions on technology or preferences. The agent's value W_0 is given by

$$W_0 = E \left[\int_0^\tau e^{-\gamma t} \left(dc_t - h \sum_i a_{it} dt \right) \middle| \mathcal{F}_0 \right]. \quad (2)$$

The agent is protected by a limited liability constraint, so that the cumulative consumption process c_t is nondecreasing.¹⁹

The principal receives the payoffs from each project and pays the agent's compensation.

avoids the technical difficulties of having to deal with sticky Brownian Motion, as in [Zhu \(2013\)](#), and considerably simplifies the analysis.

¹⁷See [Karatzas and Shreve \(1991\)](#), p. 10.

¹⁸As noted in [DeMarzo and Sannikov \(2006\)](#), this prevents the principal from postponing the agent's consumption forever.

¹⁹Intuitively, the increment in the agent's consumption payments dc_t may not be negative. This assumption rules out trivial contracts in which it is costless to incentivize effort by demanding arbitrarily high payments from the agent when a low path of output realizes for one of the projects. Qualitatively, the results in my paper remain unchanged if I assume $dc_t \geq -\underline{c}dt$ for some $\underline{c} > 0$.

Her expected value is

$$J_0 = E \left[\int_0^\tau e^{-rt} (k_{it} dx_{it} - dc_t) + e^{-r\tau} l | \mathcal{F}_0 \right]. \quad (3)$$

When the firm is shut down, the principal can recover a salvage value $l > 0$. The agent receives an outside payoff of zero in this case.

If a project is shut down, i.e. $k_{it} = 0$, it generates no output for the principal, because in her flow value k_{it} multiplies dx_{it} . I adopt this specification to simplify the interpretation of “choosing” projects.

On the path of the optimal contract, we have $k_{it} = a_{it}$. The principal would never incentivize effort in a project, but then have it generate no output, so we must have $k_{it} = 1$ whenever $a_{it} = 1$. Without loss of generality, we also have $k_{it} = 0$ whenever $a_{it} = 0$. If the agent does not exert effort in a project, we have $E [dx_{it} | \mathcal{F}_0] = E [dB_{it} | \mathcal{F}_0] = 0$ for any k_{it} , so $k_{it} = 0$ is optimal.

Using the law of iterated expectations, the principal’s value in Equation (3) then becomes

$$J_0 = E \left[\int_0^\tau e^{-rt} (a_{it} \mu_{it} dt - dc_t) + e^{-r\tau} l | \mathcal{F}_0 \right].$$

To save notation, I suppress the distinction between a_{it} and k_{it} . In the following, I focus on contracts which specify (a, c, τ) .

3.3 Incentive Compatibility and Principal’s Problem

At any point in time, the history of the contract can be summarized by the agent’s continuation utility, which acts as a state variable.²⁰ For any incentive-compatible contract (a, c, τ) , this continuation utility is given by

$$W_t = E \left[\int_t^\tau e^{-\gamma(s-t)} \left(dc_s - h \sum_i a_{is} ds \right) | \mathcal{F}_t \right].$$

²⁰Using the agent’s continuation value as a state variable is a common technique in dynamic contracts. See [Spear and Srivastava \(1987\)](#) for an illustration.

Lemma 1 below uses the martingale representation theorem²¹ to derive a law of motion for the continuation value W_t . The lemma also states an incentive compatibility condition that ensures that the agent exerts effort.

Lemma 1. *For any effort process a and consumption process c , there exists a collection of progressively measurable and square integrable stochastic processes $\{(\psi_{it})_{i=1}^N : 0 \leq t \leq \tau\}$, such that the agent's continuation value satisfies the law of motion*

$$dW_t = \left(\gamma W_t + h \sum_i a_{it} \right) dt - dc_t + \sum_i \psi_{it} dB_{it}. \quad (4)$$

Let $\psi_i = h\sigma_i/\mu_i$. The contract is incentive compatible (IC) if and only if

$$\psi_{it} \geq \psi_i, \quad (5)$$

whenever $a_{it} = 1$ and $\psi_{it} \leq \psi_i$ whenever $a_{it} = 0$.²²

The parameter ψ_{it} measures the pay-performance sensitivity (PPS) of the agent's continuation value with respect to the noise in project i 's output. When the agent shirks on a small interval of time dt , his utility rises by $h dt$, since he does not incur the effort cost. However, the output dx_{it} decreases by $\mu_i dt$ on average. From the principal's perspective, this is the same as dB_{it} decreasing by $\mu_i/\sigma_i dt$. The agent then loses $\psi_{it}\mu_i/\sigma_i dt$ in continuation utility. To induce effort, this loss must be larger than $h dt$.

Lemma 1 also illustrates why the risk-return ratio $RR_i = \mu_i/\sigma_i$ is important for providing incentives. When the agent shirks, he affects the principal's beliefs about the realization of dB_{it} . When the risk-return ratio is large, observing a shortfall in output by $\mu_i dt$ while the agent is working is a relatively unlikely event and corresponds to a large negative realization of the Brownian noise. In this sense, it is easy to detect shirking and the principal does not need a large PPS to provide incentives. The opposite holds for a large volatility σ_i .

With the result of Lemma 1, the optimal contract can be expressed as a choice of consumption payments dc_t , prescribed effort a_t , pay-performance sensitivities for each project

²¹See Karatzas and Shreve (1991), Theorem 4.15, p. 182.

²²In Proposition 1, I show that the principal's value function is strictly concave. On the path of the optimal contract, we have $\psi_{it} = 0$ whenever $a_{it} = 0$ and $\psi_{it} = \psi_i$ whenever $a_{it} = 1$.

ψ_{it} , and a firing time τ . The principal maximizes the firm value (3), subject to the promise-keeping constraint (4) and the incentive compatibility condition (5). The principal's problem is

$$\begin{aligned}
J(W_0) &= \max_{\{\psi_{it}, a_{it}, c_t, \tau\}} E \left[\int_0^\tau e^{-rt} \left(\sum_i \mu_i a_{it} dt - dc_t \right) + e^{-r\tau} l \mid \mathcal{F}_0 \right] \\
s.t. \quad dW_t &= \left(\gamma W_t + h \sum_i a_{it} \right) dt - dc_t + \sum_i \psi_{it} dB_{it} \\
\psi_{it} &\geq \frac{\sigma_i}{\mu_i} h \text{ if } a_{it} = 1.
\end{aligned} \tag{6}$$

I characterize the solution in the following Section.

3.4 Shape of the Value Function

The principal's value function is the solution to a Hamilton-Jacobi-Bellman (HJB) equation, with the continuation value W_t as the only state variable. This solution is used in the following sections to characterize the choice of projects. The proof of the proposition below, which includes a verification argument, is in Appendix A.2.

Proposition 1. *Let $n(a_t) = \sum_i a_{it}$ denote the number of projects chosen at time t . The HJB equation*

$$rJ(W) = \max_{a \in \mathcal{A}} \sum_i \mu_i a_i + J'(W) (\gamma W + hn(a)) + J''(W) \frac{1}{2} \sum_i \psi_i^2 a_i \tag{7}$$

with boundary conditions

$$\begin{aligned}
J(0) &= l \\
J'(\bar{W}) &= -1 \\
J''(\bar{W}) &= 0
\end{aligned}$$

has a unique twice continuously differentiable solution on the interval $[0, \bar{W}]$ which equals

the principal's optimal value function. The optimal PPS for each project is given by

$$\psi_{it} = \psi_i$$

whenever $a_{it} = 1$ and $\psi_{it} = 0$ whenever $a_{it} = 0$. The value function is strictly concave on $(0, \bar{W})$ and three times continuously differentiable on any subset of $(0, \bar{W})$ with nonempty interior on which project choice is constant. If $W_0 > \bar{W}$, then the principal pays the agent a discrete lump sum so that the continuation value reaches \bar{W} .

Figure 1 qualitatively illustrates the shape of the value function, which has the following features. When W_t hits zero, the agent is fired, and the principal receives the scrap value of l . The principal's value is concave, because introducing more risk in the contract, via higher ψ_{it} , increases the likelihood that W_t hits zero and the contract is terminated. Increasing W_t has two effects: a reduction in the likelihood of termination, but also an increase in expected payments to the agent. When W_t is small, the first effect dominates and the principal's value may be increasing in W_t .²³ When $W_t = \bar{W}$, the agent is paid, i.e. $dc_t > 0$.²⁴ The smooth pasting and super contact conditions, $J'(\bar{W}) = -1$ and $J''(\bar{W}) = 0$, ensure that the threshold \bar{W} is optimal.

Characterizing the value function faces several technical hurdles. First, the components of the volatility in Equation (4) have discrete jumps. This is because the volatility of W_t with respect to B_{it} is ψ_i when the agent works on project i and zero when he shirks. Therefore, one has to take care in showing that the stochastic differential equation for W_t has a unique solution. Specifically, I prove existence of ε -optimal controls for the principal which ensure that W_t has a unique weak solution and which approximate the optimal value arbitrarily closely.²⁵ In all numerical solutions I have encountered, including those I have not reported in the paper, there is a finite number of cutoffs at which project choice changes. In these cases, W_t admits a unique weak solution even without resorting to ε -optimal controls.²⁶ Many

²³Note that this regime is not renegotiation proof as the principal's value function is increasing in the agent's value. If renegotiation is allowed, the principal may agree to promise the agent a higher value, since doing so would be mutually beneficial. The overall value for the principal is lower compared with the case with commitment, since the agent's incentives are diminished if he anticipates the contract to be renegotiated. Thus, the principal will always commit if she has the ability. See also [DeMarzo and Sannikov \(2006\)](#) for a discussion of renegotiation in a related setting.

²⁴See e.g. [DeMarzo and Sannikov \(2006\)](#). This corresponds to the linear segment in Figure 1.

²⁵I do this as part of the verification argument in Appendix A.2, in the "Proof of Proposition 8."

²⁶See Figure 4 for an example.

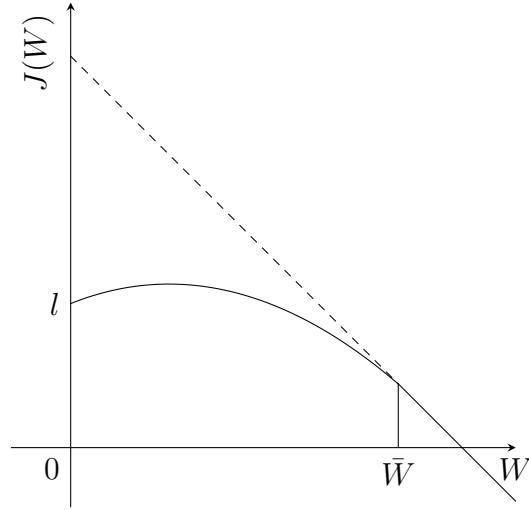


Figure 1: Shape of the principal's value function.

standard sources rely on strong solutions and require Lipschitz continuity of the volatility with respect to the state under the optimal control, which is clearly violated in my case.²⁷ Second, the optimization problem on the right hand side of the HJB equation is discrete. As is common in optimal control, a verification argument is needed to establish that the solution to the HJB equation equals the principal's value function. For this argument to apply, the HJB equation must admit a twice continuously differentiable solution, despite this discreteness. The argument to establish concavity also requires twice continuous differentiability.²⁸

4 Results

4.1 Project Choice

I now characterize the optimal project choice. I show that moral hazard can cause both underinvestment and overinvestment in projects relative to an NPV criterion.

²⁷That is, $\hat{\sigma}(W) = \sigma(W_t, a(W_t))$ must be (at least locally) Lipschitz continuous. See e.g. [Fleming and Soner \(2006\)](#), p. 159. [Yong and Zhou \(1999\)](#), p. 68, establish existence results for an equation similar to mine, but require a convex control space. For unique weak solutions, standard sources require that the controlled volatility is continuous in the state, see e.g. [Karatzas and Shreve \(1991\)](#), p. 327 or [Yong and Zhou \(1999\)](#), p. 45, which is also violated.

²⁸I show this in [Appendix A.2](#).

The project choice is determined via the HJB Equation (7). Conditional on the continuation value, the optimization problem is separable in each project. Therefore, the instantaneous marginal benefit of each project can be determined separately. It is given by²⁹

$$b_i(W_t) = \mu_i + J'(W_t)h + J''(W_t)\frac{1}{2}\psi_i^2. \quad (8)$$

A project is chosen whenever the marginal benefit is positive, which depends on the project's average payoff μ_i , the cost of compensating the agent for effort $J'(W_t)h$, and the cost of providing incentives $J''(W_t)\psi_i^2/2$. While the marginal benefit of project i implicitly depends on both past and current choices of other projects, any dependence is summarized by the value function $J(W)$.

To relate project choice to the NPV criterion, we can rewrite Equation (8) as

$$b_i(W) = r\text{NPV}_i + (J'(W) + 1)h + J''(W)\frac{1}{2}\frac{h^2}{\text{RR}_i^2}. \quad (9)$$

Here, I define a project's NPV as $\text{NPV}_i = (\mu_i - h)/r$. The marginal benefit of choosing a project depends positively on both the NPV and the project's risk-return ratio $\text{RR}_i = \mu_i/\sigma_i$. While a higher NPV implies higher expected payoffs from the project, the risk-return ratio works through the agent's incentives. The higher the ratio, the easier it is to detect shirking, and the agent's incentives can be weaker without ceasing to motivate effort. By the representation in Equation (4), this is equivalent to a lower volatility of W_t . The term $(J'(W) + 1)h$ measures the increase in social value from moving the continuation value away from the termination boundary, and it is always positive. Setting $b_i(W) = 0$, I can derive a minimum NPV criterion which determines project choice.³⁰

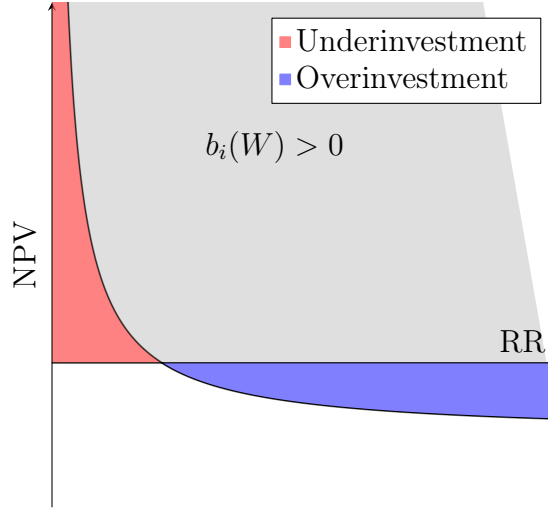
Proposition 2. *Conditional on W , a project with a given risk-return ratio is chosen when-*

²⁹Recall that $\psi_i = \sigma_i h/\mu_i$ (Lemma 1). We can rewrite the HJB equation (7) as

$$rJ(W) = \sum_i \max\{0, \mu_i + J'(W)h + 0.5J''(W)\psi_i^2\} + \gamma W J'(W).$$

The terms inside the sum are the principal's optimal project choice.

³⁰The derivation for this result is in the text above, so the proof is omitted.



The solid black line qualitatively illustrates the tradeoff between NPV and the risk-return ratio. At the line, $b_i(W)$ equals zero. The area above the line, which consists of the gray and blue portions, represents combinations of NPV and risk-return for which $b_i(W)$ is positive.

Figure 2: NPV vs. risk-return boundary

ever its NPV exceeds

$$\underline{NPV}(RR, W) = -\frac{1}{r} (J'(W) + 1) h - J''(W) \frac{1}{2r} \frac{h^2}{RR^2}. \quad (10)$$

This threshold is a function of the current continuation value and the project's risk-return ratio. Figure 2 qualitatively illustrates the non-linear relationship between \underline{NPV} and RR, and outlines the set of projects that are chosen at continuation value W .

Depending on W and the risk-return ratio, the threshold may be positive or negative. If it is positive, a project with positive NPV might not be chosen, leading to underinvestment. If it is negative, which happens whenever the risk-return ratio is sufficiently high, a project with negative NPV may be chosen, leading to overinvestment.

Thus, moral hazard leads to both over- and underinvestment in my setting. Underinvestment is due to the cost of incentives. Overinvestment occurs since assigning projects can be used as a punishment device. The optimal contract provides incentives by inducing a low continuation value for the agent utility after bad performance.³¹ This can be done

³¹This can be seen from equation (4). The pay-performance sensitivity ψ_{it} is positive whenever the agent

either via a high likelihood of firing or by a low flow utility. Since firing is inefficient, the principal would prefer to lower the agent’s flow utility, but she is constrained by limited liability. Incentivizing the agent to work on a project allows the principal to circumvent the limited liability constraint, because it lowers his realized flow utility via the effort cost. This is valuable, since it allows her to provide incentives while relying less on costly termination.

Formally, this effect is captured in $b_i(W)$ in equation (9). Choosing an additional project increases the drift of the continuation value in equation (4). Equivalently, it lowers the realized flow utility of the agent in equation (2). Since the drift of W_t is higher, the likelihood of hitting the termination boundary at $W = 0$ is lower. The value of this for the principal is captured by $J'(W)$, which can be positive when W is low. However, to incentivize a project, the principal must increase the volatility of W_t , which in turn increases the likelihood of termination. In equation (9), this is captured by the last term, $J''(W)h^2/RR_i^2$, which is always negative. When the risk-return ratio is sufficiently large, the overall expression can still be positive. Then, choosing a negative NPV project can be optimal.

An important qualification for this result is that (1) the continuation value must be sufficiently low, i.e. termination must be sufficiently likely, (2) the NPV of the project cannot be *too* negative, and (3) the risk-return ratio must be relatively high, i.e. the project must be relatively easy to incentivize.³² To understand (1) more clearly, notice that increasing the drift of W_t has two effects. On average, the payment boundary \bar{W} is hit earlier, and therefore expected payments to the agent are higher. This decreases the principal’s value. However, as just outlined, it also decreases the likelihood of termination, which increases the principal’s value. $J'(W)$ is positive whenever the second effect outweighs the first. This happens when W is sufficiently low and termination is sufficiently likely.

The overinvestment and underinvestment effects are shown in Figure 2. For a given continuation value W , the gray area is the region where a project’s marginal benefit is positive for a given combination of NPV and risk-return ratio. The lower bound of this area is equation (10).

The intuition that overinvestment can substitute for termination allows me to derive a simple sufficient condition for when it occurs on the path of the contract. Whenever the

exerts effort in a project, meaning W_t must decrease whenever dB_{it} is negative.

³²In other words, the pay-performance sensitivity ψ_i must be relatively low. Of course, if the principal had access to a positive or zero-NPV project with small required pay-performance sensitivity, she would prefer to use that project to act as a punishment device rather than one with negative NPV.

scrap value l is sufficiently small, i.e. termination is sufficiently costly, overinvestment will occur as long as the project’s payoff is not too negative and the required pay-performance sensitivity is not too large. In the following Lemma, I prove an even stronger result: when l is small, projects with negative *payoff* μ_i , not just NPV, may be chosen.

Proposition 3. *Suppose there exists a project i such that $\psi_i = \varepsilon$ and $\mu_i = -\varepsilon$ with ε sufficiently small. Then, as $l \rightarrow 0$, there exists a region of W such that $b_i(W) > 0$.*

The proof is straightforward and only sketched here. As $l \rightarrow 0$, I can show that $J'(0) > 0$, otherwise concavity of the value function would imply that the principal’s value is negative close to $W = 0$. By continuity of $J'(W)$, the first derivative remains positive on some region right of 0. On this region,

$$b_i(W) = -\varepsilon + J'(W)h + J''(W)\frac{1}{2}\varepsilon^2,$$

which is positive for ε sufficiently small.

Discussion Before moving to dynamics, I would like to address how the results relate to other work.

Zhu (2013) studies an extension of DeMarzo and Sannikov (2006) in which there is a single project and the principal can let the agent shirk temporarily. In both my paper and his, shirking is optimal whenever the decrease in the risk of termination, which comes from the principal having to provide sufficient incentives to motivate effort, outweighs the loss of payoff from shirking. This happens if the payoff loss is sufficiently small.³³ This is analogous to underinvestment in my paper, since not choosing a positive NPV projects decreases the principal’s flow payoff, but also decreases the cost of incentives.

The key difference is that in in my paper, it may be optimal to incentivize effort in a project which has negative NPV. Since Zhu’s model only has a single task, and that task has positive “NPV”,³⁴ this result cannot be obtained in his setting. If the value of incentivizing

³³See Zhu’s Lemma 3.3 and Section 6.2.1.

³⁴In my model, NPV is defined as $(\mu_i - h)/r$. In Zhu’s model, when the agent works the principal gets μ , while when he shirks she gets $\mu - A$. When working, the agent receives no private benefit, while when shirking, he receives $\lambda A > 0$ where $0 < \lambda < 1$. The analog definition of NPV is hence $\mu - ((\mu - A) + \lambda A) = (1 - \lambda)A > 0$. Thus, NPV is positive in Zhu’s model. In the first-best, the principal would choose any project with positive NPV in my model. The analog holds in Zhu’s paper. There, the principal would implement effort forever if and only if $(1 - \lambda)A > 0$ in the first best.

effort were negative, Zhu’s optimal contract would simply never implement effort.

Importantly, choosing a negative NPV project does not follow the simple intuition of trading flow payoff for lower instantaneous incentive costs. In Zhu’s model, shirking decreases the principal’s flow payoff, decreases the drift of the continuation value, and decreases the pay-performance sensitivity. In my model, it may be optimal to choose a negative NPV project *in addition* to a positive NPV project, purely to punish the agent. This *increases* the principal’s flow payoff (even though NPV, which represents social value, is negative), *increases* the drift of the continuation value, and *increases* total pay-performance sensitivity. In sum, the overinvestment effect identified in my paper cannot be obtained in Zhu’s setting and it is qualitatively different from shirking his model.³⁵

In DeMarzo et al. (2012), there is a single task and the principal can invest to increase firm size subject to a cost. Due to the cost of incentives, investment is below the first-best level. Thus, in DeMarzo et al. (2012), moral hazard can only create underinvestment, whereas my setting generates both underinvestment and overinvestment depending on the project characteristics.

In Section 5, I show the contract can be implemented with equity and bonus payments. These bonuses are a linearly weighted function of individual project outputs. Such implementation is new in the literature. It cannot be obtained in single-project settings such as DeMarzo and Sannikov (2006), Zhu (2013), or DeMarzo et al. (2012).

My overinvestment result is also different from the “cross pledging” result in static moral hazard models.³⁶ There, the principal can reduce the agent’s information rents by incentivizing effort in a positive and a negative NPV project together, provided she can condition

³⁵To see this explicitly, here is Zhu’s HJB equation (eq. (5) on p. 5),

$$rB(W) = \max \left\{ \mu + \gamma W B'(W) + \frac{\lambda^2}{2} B''(W), \mu - A + (\gamma W - \lambda A) B'(W) \right\},$$

where the first term is the HJB under working and the second term under shirking. Thus, shirking reduces the flow payoff to the principal, to $\mu - A$, reduces the drift of the continuation value to the agent, i.e. $E(dW_t) = (\gamma W_t - \lambda A) dt$, and reduces the pay-performance sensitivity, so the last term $B''(W) \lambda^2/2$ disappears under shirking. This is analogous to what happens with underinvestment in my paper. Not exerting effort in a project reduces the principal’s flow payoff by μ_i , reduces the drift of the continuation value by h and reduces the total pay-performance sensitivity by ψ_i^2 , but not necessarily analogous to overinvestment, since that may increase flow payoff to the principal, the drift of the continuation value, and the total pay-performance sensitivity.

³⁶See e.g. Diamond (1984). I am grateful to Bruno Biais for pointing out this connection.

payments on the success of both. This reduces the agent's value, because it reduces information rents. Thus, cross pledging is only possible if the agent's participation constraint is not binding. In my model, the negative NPV project is valuable only because it serves as a temporary punishment device. This is different than the rent reduction effect. We could start out with a contract where the agent's ex-ante participation constraint is binding, but the principal would still choose negative NPV projects on the optimal path.

4.2 Project Choice Dynamics

I now characterize how the project choice changes over time. The main results are as follows. First, as W approaches \bar{W} , the distortions from moral hazard disappear, and project choice approaches the first-best. Second, as W increases, the firm chooses projects with higher costs of incentives.³⁷ These are the projects which require a higher pay-performance sensitivity to motivate effort and have lower risk-return ratios. Third, if projects with higher return have higher costs of incentives, the firm chooses higher return projects as W increases.

Exploiting the fact that high continuation value means good past performance, I can convert these results into empirical predictions. Better performing firms have a project choice which is closer to NPV, while worse performing firms are more likely to either over- or underinvest. As past performance increases, the firm chooses projects which are more profitable, more risky, and which require a higher pay-performance sensitivity to motivate effort.

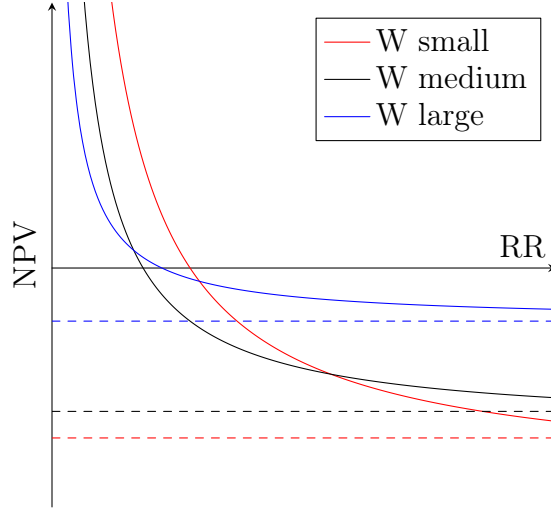
The intuition for these results is that when the continuation value increases, the likelihood of termination declines, so that the firm's project choice shifts toward high-NPV projects with a lower risk-return ratio and high cost of incentives. To get a clean characterization, I focus here on the case where the agent is sufficiently patient, i.e. γ is sufficiently close to r for the remainder of this section. Then, the cost of incentives is decreasing as W increases.³⁸ I defer all formal proofs to Appendix A.4.

The first result, that distortions disappear as $W \rightarrow \bar{W}$, follows from the characterization in Proposition 1.

Corollary 1. *As W approaches \bar{W} , the optimal project choice converges to the project choice*

³⁷The costs of incentives for a project i are $J''(W_t) \psi_i^2/2$ or, equivalently, $J''(W) h^2/(2RR_i^2)$.

³⁸Formally, $J'''(W) > 0$ for all $W \in [0, \bar{W}]$.



This figure shows, qualitatively, the combinations of NPV and the risk-return ratio for which $b_i(W)$ is zero for different values of W .

Figure 3: Project cutoffs as a function of W

under the NPV criterion.

As $W \rightarrow \bar{W}$, $J'(W) \rightarrow -1$ and $J''(W) \rightarrow 0$, which implies

$$b_i(W) \rightarrow rNPV_i.$$

This result is illustrated qualitatively in Figure 3. The red function shows the project boundary for small W and the black and blue functions represent the boundaries for successively larger W . As $W \rightarrow \bar{W}$, the break-even NPV line in the figure approaches the x-axis.

Firms start projects with higher costs of incentives if the marginal benefit function of those projects is increasing in the continuation value. This is because if a project is started at some value \hat{W} , it must be that $b_i(W)$ crosses zero from below at \hat{W} . However, the marginal benefit $b_i(W)$ is not necessarily increasing in the continuation value W for every project. As W increases the cost of incentives, which is proportional to $J''(W)$, decreases, but $J'(W)$, which measures the cost of compensating the agent, increases because $J(W)$ is concave. However, I can show that if a project requires a higher pay-performance sensitivity than

those already chosen, $b_i(W)$ is increasing. Formally, we have the following Proposition.³⁹

Proposition 4. *The marginal benefit $b_i(W)$ is increasing in W if*

$$\psi_i^2 \geq \frac{1}{n(a(W))} \sum_j \psi_j^2 a_j(W), \quad (11)$$

where $a(W) = (a_1(W), \dots, a_N(W))$ is the optimal project choice at continuation value W .

The proposition implies that if a project has sufficiently high cost of incentives, its marginal benefit is increasing. If a project is started as W increases, it must be among those projects. Conversely, if it is stopped it must be among those with relatively low cost of incentives. In this sense, firms choose projects which are more difficult to incentivize.

An important question is how the average risk or return of the chosen projects changes with W . I can obtain such results under additional assumptions on the relation between risk and return. The firm will take projects with higher return as W increases if higher return implies the project requires a higher pay-performance sensitivity (or, equivalently, has a lower risk-return ratio). Specifically, suppose all projects are ordered: any project with higher return also requires a higher pay-performance sensitivity. Then, the marginal benefits of any two projects i and j satisfy a single-crossing condition in W . The firm chooses projects which have higher returns, higher risk, but are more difficult to incentivize as W increases.⁴⁰

Proposition 5. *Suppose for any two projects i and j , $\mu_i > \mu_j$ implies $\psi_i > \psi_j$. Either $b_i(W) \geq b_j(W)$ for all W or $b_i(W)$ crosses $b_j(W)$ from below at a unique cutoff W_{ij} .*

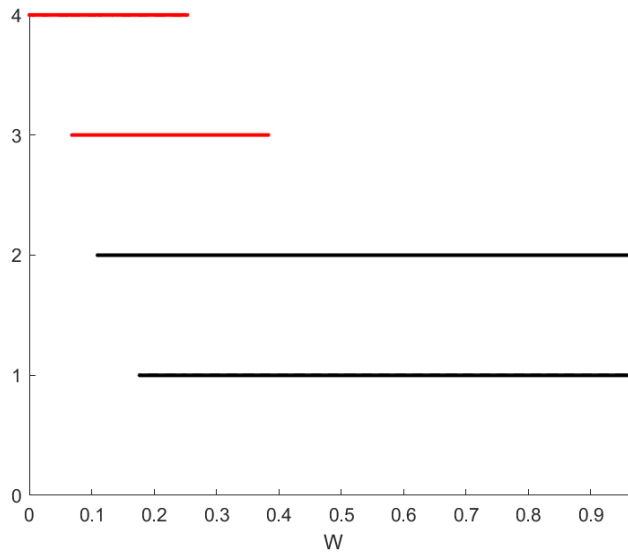
³⁹The intuition is as follows. We have $b_i(W) = \mu_i + J'(W)h + J''(W)\psi_i^2/2$. The last term is increasing in W , but the middle term is decreasing by concavity of J . If $b'_i(W) < 0$ for some projects, these projects must have low ψ_i^2 . Conversely, for projects with high ψ_i^2 , $b_i(W)$ must be increasing since the weight on $J''(W)$ is higher. These are exactly the projects for which the incentive cost is relatively high. To derive the characterization in equation (11), I show in Appendix A.4 that if $b_i(W)$ is decreasing for a project with higher than average cost of incentives, then $J'(W)$ would have to be so large that the principal's value exceeds the first-best value, which is impossible.

⁴⁰Proposition 5 resembles the result in Radner and Shepp (1996), who study project choice and payout decisions in a model with exogenous bankruptcy, but without moral hazard. Under a similar condition as in Proposition 5, they find that firms further from bankruptcy (when the cash stock is high) choose more profitable projects. In their model, projects have no costs and strictly positive average payoffs. Thus, their model only induces underinvestment relative to the NPV criterion. I am grateful to an anonymous referee for pointing out this connection.

That is, the principal derives a higher (marginal) value from project i than from project j if and only if W is above a cutoff, unless the value of project i is always larger. Figure 4 illustrates this result, which can be seen by constructing the difference between $b_i(W)$ and $b_j(W)$, which satisfies

$$\Delta_{ij}(W) = \mu_i - \mu_j + J''(W) \frac{1}{2} (\psi_i^2 - \psi_j^2).$$

This expression is increasing because $J''(W)$ is increasing.⁴¹ This immediately implies single crossing in W .



The solid lines indicate the regions in which different projects are chosen. Projects 1 and 2 (black) have positive NPV and projects 3 and 4 (red) have negative NPV. I set $\gamma = 0.046$, $r = 0.05$, $l = 0.97$, and $h = 0.1$. The projects are characterized by $\mu_1 = 0.15$, $\sigma_1 = 0.3$ and $\psi_1 = 2$, $\mu_2 = 0.11$, $\sigma_2 = 0.21$ and $\psi_2 = 1.909$, $\mu_3 = 0.04$, $\sigma_3 = 0.05$ and $\psi_3 = 1.25$, and $\mu_4 = 0.02$, $\sigma_4 = 0.02$ and $\psi_4 = 1$.

Figure 4: Project Choice with 4 Projects

⁴¹This follows from the assumption that γ is sufficiently close to r . See Proposition 12 in Appendix A.4.

5 Implementation with Bonus Contracts

The optimal contract can be implemented with equity and project-specific bonus payments. Here, the equity share is used to provide a baseline level of incentives among all chosen projects, while bonus payments provide additional incentives beyond the equity share. By adjusting the sensitivity of the bonus payments to output, the principal can incentivize effort. [Murphy \(1999\)](#) finds that the majority of incentive contracts feature a mix of equity and bonus payments, with the latter being a linearly weighted function of the agent’s performance across a range of categories, just as in my implementation.⁴²

Implementing the contract requires adjusting the equity share as the choice of projects changes. To ensure that these changes in shares are incentive neutral, the agent must be compensated whenever he loses shares. Effectively, this means the agent trades the shares at transfer prices chosen by the principal. These prices depend on past performance. An explicit derivation is in [Appendix A.5](#).

Proposition 6. *The optimal contract can be implemented with equity and project specific bonus payments. The agent trades equity at pre-determined transfer prices and the proceeds are held in a bonus account.*

Intuitively, the equity share depends on the payoffs of all projects and the volatility of the agent’s continuation value increases in the equity share. The optimal share must hence be low enough to prevent unnecessary risk in the contract. It equals the equity share that incentivizes effort in the project with the lowest return. The rest of the incentives is provided by adjusting the sensitivity of the bonus payments, which map each project’s payoffs to payments into the agent’s bonus account.

6 Discussion of Applied Results

Introducing multiple projects significantly alters the firm’s behavior and the optimal compensation contract compared to models where effort is static (e.g. [DeMarzo and Sannikov \(2006\)](#) and [DeMarzo et al. \(2012\)](#)). There, the firm is restricted to always taking the same

⁴²For this implementation to be valid, it is necessary to assume that the principal can “shut off” projects in which the agent does not exert effort.

project. As a result, while promised compensation varies with performance, the manager's pay-performance sensitivity can never change. Specifically, he receives a constant equity share which does not depend on past performance. Therefore, such models are unable to obtain any relation between the firm's performance and the optimal compensation contract.

In [Sannikov \(2008\)](#), effort is variable, and the model delivers a link between past performance and instantaneous pay-performance sensitivity. However, the optimal contract does not have an implementation which resembles real-world instruments. Allowing the firm to allocate different projects to the manager allows me to resolve this problem. As I have shown in [Section 5](#), I can explicitly derive how the manager's optimal equity share and the sensitivity of the bonus payments change over time and how they depend on the projects currently chosen. Also, I can derive predictions on how the firm's cash flow volatility changes with past performance, since taking different projects with different risk will yield different volatilities for the total cash flow.

The model yields several testable predictions about the firm's project choice and the optimal compensation contract. Project choice depends on past performance and firms with better performance choose riskier projects. Since the firm's total cash flow is the sum of the cash flows of all individual projects, the project choice determines the realized volatility of the firm's total cash flow. Thus, firms with better past performance also have higher cash flow volatility.

In the model, past performance determines the distance to the termination boundary. If we interpret this as default, firms who are closer to default will choose safer projects and those who are far away will choose riskier ones. This prediction is broadly consistent with results found in [Rauh \(2009\)](#). In the implementation in [Section 5](#), I have shown that the agent's continuation value can be interpreted as the sum of the firm's free cash and a personal incentive account. Thus, firms with higher free cash will choose riskier projects, since that corresponds to a higher continuation value. By the same logic, larger firms will choose less risky projects, since the continuation value is lower.

Recently, [Cassell et al. \(2012\)](#) have found that managers with higher inside debt holdings, i.e. deferred compensation and pension benefits, choose less risky investments, because they are more likely to lose those if the firm goes bankrupt. This result is consistent with the mechanics in my model. Here, if the manager's continuation value is low, he is more likely to be terminated than to receive any of his promised payouts. The firm then optimally chooses

safer projects because the incentive cost is high.

An alternative interpretation for the model is to understand the projects as individual tasks allocated to the manager instead of actual investment projects. In this case, more risky and difficult to incentivize projects are allocated to relatively successful managers, while unsuccessful managers work only on relatively safe ones. If higher risk implies higher return, this means that the most lucrative activities are allocated to the managers with the best track record. Since the continuation value in Equation (4) has positive drift, on average, more experienced managers, i.e. those with longer tenure, get more risky tasks and receive a higher pay-performance sensitivity. This is consistent with the findings in [Cremers and Palia \(2011\)](#).

In addition, my model sheds light on the connection between managerial incentives and the firm's investment choices, which has received considerable attention in finance. In my model, total pay performance sensitivity is higher whenever the firm chooses riskier projects. Thus, firms with higher pay-performance sensitivity also have higher cash flow volatility. This is consistent with results from [Massa and Patgiri \(2009\)](#), who study mutual funds and find that those with higher pay-performance sensitivity also have more risky and more profitable investments.⁴³

7 Conclusion

I analyze a continuous-time moral hazard problem in which the agent's effort is distributed among different projects. Unlike past studies, project choice is simultaneous, and the possible feedback effect between projects is explicitly considered. The model sheds light on the optimal choice of projects under moral hazard and the prevalence of bonus contracts in CEO compensation. Further, it explains the use of different criteria to evaluate projects aside from NPV, which is broad practice in companies, as shown by [Graham and Harvey \(2001\)](#).

In the optimal contract, a project is chosen whenever its NPV is above a cutoff that

⁴³Most of the literature, including [Massa and Patgiri \(2009\)](#), focuses on the convexity of the manager's payoffs, as measured by the vega of the manager's option grants (see also [Guay \(1999\)](#) and [Coles et al. \(2006\)](#)) as an explanatory variable, although some results on concerning the slope are available. [Coles et al. \(2006\)](#) show in their Table 3 that a higher option delta in the manager's compensation contract yields higher R&D expenses, which can be interpreted as the firm choosing riskier projects. [Massa and Patgiri \(2009\)](#) also provide explicit results on the slope.

depends on the firm's current cash stock and the project's risk-return ratio. This cutoff changes stochastically over time and depends on the agent's cumulative past performance. Firms with a large cash stock feature a relatively efficient investment portfolio, comprising only projects with positive NPV, whereas firms with a low cash stock suffer from an inefficient choice of investment projects, passing up positive NPV projects when their risk is too high. The absolute benefit of projects with above-average risk increases with the firm's cash stock when the cost of incentives is decreasing, whereas the benefit of projects which are relatively safe decreases. The project choice follows the NPV criterion whenever the cash stock is large enough. The agent is assigned riskier projects after a history of sufficiently good performance, whereas a poorly performing agent either will see the number of projects assigned to him diminished or will be given relatively safe projects as a punishment.

My model can be applied to investment situations whenever the choice of projects is discrete. This allows studying issues such as R&D efforts, the opening of new manufacturing plants, natural resource exploration, and diversification into different markets, to name a few. The empirical literature on firm investment has overwhelmingly focused on a firm's aggregate investment, which is treated as a continuous variable. My model constitutes a theoretical benchmark that makes predictions on a firm's entire project portfolio and may be used to test against data, once estimates of the individual projects' average payoff and volatility have been obtained, instead of providing insights into the choice of one investment project in isolation.

A Proofs

A.1 Proof of Lemma 1

That W_t satisfies Equation (4) follows from a simple modification of [DeMarzo and Sannikov \(2006\)](#), Lemma 2, p. 2689. I therefore omit his proof. To derive the incentive compatibility condition in Equation (5), suppose the agent's actual effort process \hat{a}_t differs from the prescribed effort process a_t . At any such time, the agent's effort cost changes by $h(\hat{a}_{it} - a_{it})$ and his expected payoffs change by $\psi_{it}(\hat{a}_{it} - a_{it})\mu_i/\sigma_i$ for any project i . This follows from applying Girsanov's Theorem to B_{it} . The total payoff from effort process \hat{a}_t is thus

$$W_0 + E \left[\int_0^\tau e^{-\gamma t} \sum_i (\psi_{it} - h) (\hat{a}_{it} - a_{it}) dt \right].$$

Thus, $\psi_{it} \geq \psi_i$ ensures that $\hat{a}_{it} = 1$ is optimal. If $\psi_{it} < \psi_i$ agent chooses $\hat{a}_{it} = 0$. This establishes the result.

A.2 Existence and Uniqueness of the Value Function and Verification

In this section, I establish the existence of a twice continuously differentiable solution to the HJB equation⁴⁴

$$\begin{aligned} rJ(W) = & \max_a \sum_i \mu_i a_i + J'(W) (\gamma W + hn(a)) \\ & + J''(W) \frac{1}{2} \sum_i \psi_i^2 a_i \end{aligned} \tag{12}$$

⁴⁴The arguments in many standard sources cannot be used. [Krylov \(1980\)](#), Theorem 5, p. 24, provides sufficient conditions for a twice continuously differentiable solution to the HJB equation, but he requires a convex control space. By contrast, the project choice in my model changes discretely. [Gilbarg and Trudinger \(2001\)](#) establish sufficient condition for PDEs, of which HJB equations are a special case. They require that the coefficients are Holder continuous in the state, which is violated in my model because the volatility of W_t has discrete jumps. [Fleming and Soner \(2006\)](#), p. 161, provide sufficient conditions for the case where coefficients may depend on time, but they require the state space to be bounded a priori (in my model \bar{W} is endogenous). [Yong and Zhou \(1999\)](#), Ch. 4, only provide a verification theorem.

with boundary conditions $J(0) = 0$, $J'(\bar{W}) = -1$ and $J''(\bar{W}) = 0$, and verify that this solution equals the value function under the optimal contract. Throughout, I focus on the case where there exists at least one project with positive NPV, i.e. $\mu_i - h > 0$, and where for all W the principal implements effort in at least one project, so that the volatility of the continuation value process is never degenerate. I find parameter values such that this is indeed optimal in Proposition 10 in Section A.3 below.

For simplicity, I assume that for all i , $\mu_i - h$ is either strictly greater or smaller than zero, which implies that at the first best, the principal cannot be indifferent between taking a project or not. At \bar{W} , the boundary conditions imply that

$$rJ(\bar{W}) = \sum_i (\mu_i - h)^+ - \gamma\bar{W},$$

where $(\mu_i - h)^+ = \max\{\mu_i - h, 0\}$. Therefore, they are equivalent to $J(\bar{W}) = J_*(\bar{W})$ and $J'(\bar{W}) = -1$, where $J_*(W)$ is given by

$$J_*(W) = \frac{\sum_i (\mu_i - h)^+ - \gamma W}{r}.$$

The conditions including $J_*(W)$ are easier to work with, and I will use them for the remainder of the argument, which relies on a variant of the shooting method. I fix a sufficiently large but finite domain $[0, W_{max}]$ and define the function $H : [0, W_{max}] \times \mathbb{R}^2 \rightarrow \mathbb{R}$ by

$$H(W, u, p) = - \min_a \frac{ru - \sum_i \mu_i a - p(\gamma W + hn(a))}{\frac{1}{2} \sum_i \psi_i^2 a_i}. \quad (13)$$

The HJB equation is equivalent to

$$J''(W) + H(W, J(W), J'(W)) = 0. \quad (14)$$

By Berge's Maximum Theorem,⁴⁵ $H(W, u, p)$ is jointly continuous in its parameters. This implies that for any starting slope s , the initial value problem (IVP) satisfying (14) with boundary conditions $J(0) = l$ and $J'(0) = s$ has a twice continuously differentiable solution

⁴⁵See Aliprantis and Border (2006), Theorem 17.31, p. 570.

on the domain $[0, W_{max}]$, and is uniformly continuous with respect to s .⁴⁶ I denote a solution with starting slope s by $J_s(W)$.

For a large negative number b , choose W_b such that $J_*(W_b) = b$, and define the boundary $\mathcal{B} \subset \mathbb{R}^2$ as⁴⁷

$$\mathcal{B} = [(0, b), (J_*(W_b), b)] \cup \{(y, W) \in [b, J_*(0)] \times [0, W_b] : y = J_*(W)\}.$$

Finally, let $\bar{W}(s) = \inf \{W : J_s(W) = J_*(W)\}$ be the first point at which $J_s(W)$ hits $J_*(W)$.⁴⁸ The following proposition is crucial for establishing uniqueness of the solution, and establishes concavity.

Proposition 7. *Any solution to the IVP (14) for which $0 > J'_s(\bar{W}(s)) \geq -1$ holds is strictly concave on $(0, \bar{W}(s))$.*

Proof. Since s is constant throughout the proof, I omit it for the sake of notation. Note that $J'(\bar{W}) \geq -1$ implies that $J''(\bar{W}) \leq 0$, otherwise $J(\bar{W}) > J_*(\bar{W})$. By the boundary conditions we have $a_i = 1$ whenever $\mu_i \geq h$ at \bar{W} . By the envelope theorem, $J'''(W)$ exists on a neighborhood left of \bar{W} ,⁴⁹ and is given by

$$J'''(W) = \frac{-(\gamma - r)J'(W) - J''(W)(\gamma W + hn(a))}{\frac{1}{2} \sum_i \psi_i^2 a_i} > 0,$$

which follows from continuity of $J'(W)$ and $J''(W)$. Therefore, $J''(W) < 0$ for W sufficiently close to \bar{W} . If $J''(W) \geq 0$ for some $W < \bar{W}$, the point $\hat{W} = \sup \{W < \bar{W} : J''(\hat{W}) \geq 0\}$ exists. If $J'(\hat{W}) \geq 0$, we have

$$rJ(\hat{W}) \geq \max_a \sum_i \mu_i a_i$$

⁴⁶See for example [Hartman \(2002\)](#), Chapters 2 and 5.

⁴⁷Because the optimal value function satisfies $J(W) \geq l - W$ for all W , and $J'_*(W) < -1$, a pair (b, W_b) can always be found, and restricting the solution of the HJB equation to lie in \mathcal{B} is without loss of generality.

⁴⁸ $\bar{W}(s)$ may not exist for all s , for example when s is a large negative number. However, $\bar{W}(s)$ is only used in the argument in cases where J_s actually hits J_* , so this is not an issue.

⁴⁹This follows because the boundary conditions imply $b_i(\bar{W}) = \mu_i - h$, which is either greater or smaller than zero, and $b_i(W)$ is continuous.

and therefore $a_i = 1$ and

$$J(\hat{W}) \geq \frac{\sum_i \mu_i}{r}.$$

The right hand side is the value obtained in the first best with zero effort cost. $J(W)$, which is the value under moral hazard and strictly positive effort cost, must always lie strictly below it. Thus, the above inequality implies a contradiction.

Therefore, we need $J'(\hat{W}) < 0$. If $\hat{W} \in \text{int}\mathcal{C}_a$ for some continuation region \mathcal{C}_a , then $J'''(\hat{W})$ exists and is given by

$$J'''(\hat{W}) = \frac{-(\gamma - r) J'(\hat{W})}{\frac{1}{2} \sum_i \psi_i^2 a_i} > 0$$

which makes it impossible for J'' to cross zero from above, as required by the definition of \hat{W} . If \hat{W} does not lie on the interior of any continuation region, there exists a project i such that $b_i(\hat{W}) = \mu_i + J'(\hat{W})h = 0$. Take some $W > \hat{W}$. Because $J'' < 0$ on (\hat{W}, W) , J' is decreasing on this region, and $b_i(W) < b_i(\hat{W}) = 0$. Thus, the project cannot be taken again on (\hat{W}, W) , and the project choice must stay constant on some neighborhood right of \hat{W} . This implies that the right derivative $J''_+(\hat{W})$ exists, and it is positive since J' and J'' are negative. But then again $J''(W)$ cannot cross zero from above at \hat{W} . Therefore, $J''(W) < 0$ for all $W < \bar{W}$. \square

Lemma 2. *There exists at most one s^* such that $J'_{s^*}(\bar{W}(s^*)) = -1$.*

Proof. I first establish two auxiliary results. First, for two initial slopes s and s' with $s' > s$, we have $J_{s'}(W) > J_s(W)$ on $(0, W_{max})$. To see that this is the case, let $\hat{W} = \inf\{W : J_{s'}(W) \geq J_s(W)\}$. By construction, $J_{s'}(\hat{W}) > J_s(\hat{W})$. Since $H(W, u, p)$ is decreasing in its second argument, we have

$$H(\hat{W}, J_s(\hat{W}), J'_s(\hat{W})) = H(\hat{W}, J_s(\hat{W}), J'_{s'}(\hat{W})) > H(\hat{W}, J_{s'}(W), J'_{s'}(\hat{W})),$$

which implies that $J''_{s'}(\hat{W}) > J''_s(\hat{W})$. Therefore, $J'_{s'}(W)$ cannot cross $J'_s(W)$ from above at \hat{W} , which is a contradiction. Since $J_*(W)$ is a strictly decreasing function, this result also implies that $\bar{W}(s') < \bar{W}(s)$ whenever $s' > s$.

Now, suppose that for $s' > s$, $J'_{s'}(\bar{W}(s')) = J'_s(\bar{W}(s)) = -1$. By the preceding argument,

$$-1 = J'_{s'}(\bar{W}(s')) > J_s(\bar{W}(s')),$$

and by Proposition 7, J_s is strictly concave, and therefore $J_s(\bar{W}(s)) < J_s(\bar{W}(s')) < -1$. \square

To conclude the proof, I define a mapping $\mathcal{S}(s) = J'_s(\bar{W}(s))$, which is continuous since the solutions to the IVP are continuous with respect to s . If there exists a pair $\{\underline{s}, \bar{s}\}$ with $\bar{s} > \underline{s}$ such that $\mathcal{S}(\bar{s}) > -1$ and $\mathcal{S}(\underline{s}) < -1$, there exists an s^* for which $\mathcal{S}(s^*) = -1$, which is a consequence of the continuous mapping theorem. The Lemma above then guarantees uniqueness.

Lemma 3. *There exist two values $\bar{s} > \underline{s}$ such that all $s \geq \bar{s}$, $\mathcal{S}(s) \geq 0$ and for all $s \leq \underline{s}$, $\mathcal{S}(s) \leq -1$.*

Proof. First, consider the map $\mathcal{T}(s) = \{(y, W) : W = \inf \{u : (J_s(u), u) \in \mathcal{B}\}\}$, which associates to each s the first point where J_s hits \mathcal{B} . For $s \leq \underline{s}$, where \underline{s} is chosen sufficiently small, J_s hits b , and the first hitting point of b can be made arbitrarily close to zero by the choice of \underline{s} . Similarly, choosing $s \geq \bar{s}$, where \bar{s} is large and positive implies that J_s hits J_* at some W close to zero, and that $\mathcal{S}(\bar{s})$ is positive. Define

$$\mathcal{B}_\varepsilon = [(\varepsilon, b), (J_*(W_b), b)] \cup \{(y, W) \in [b, J_*(0)] \times [\varepsilon, W_b] : y = J_*(W)\}$$

for some positive but small ε . Since the solution to the IVP (14) is uniformly continuous in s on $[0, W_{max}]$, $\mathcal{T}(s)$ is continuous. Because it maps an appropriately chosen interval $[\underline{s}, \bar{s}]$ into the compact set \mathcal{B}_ε , the mapping is onto, by the continuous mapping theorem. Therefore, there exists a subinterval of $[\underline{s}, \bar{s}]$ for which J_s hits J_* .

Now, take \hat{W} sufficiently close to W_b , so that $J_*(\hat{W})$ is negative and large. By the preceding argument, there exists a slope s such that J_s hits J_* at \hat{W} . Suppose that $0 > J'_s(\hat{W}) > -1$. By Proposition 7, J_s is strictly concave on $(0, \hat{W})$, and $J'_s(W) > -1$ for $W < \hat{W}$. But then $J_s(\hat{W}) > J_s(0) - \hat{W} > J_*(\hat{W})$, since $J'_*(W) < -1$, and W_b was chosen sufficiently large. This is a contradiction.

If $J'_s(\hat{W}) > 0$, then necessarily $J''_s(\hat{W}) < 0$, otherwise $J_s(\hat{W}) > J_*(W)$. There must exist a region left of \hat{W} on which $J''_s(W) > 0$, otherwise concavity would imply that

$J'_s(W) > 0$ for all $W > \hat{W}$, and $J_s(\hat{W}) = b$ could not hold. In particular, there must exist some $W < \hat{W}$ for which $J''_s(W) > 0$ and $J'_s(W) > 0$. Then, a similar argument as in Proposition 7 establishes that $J_s(W) > J_{fb}(W)$, which is a contradiction. This shows that there exists a number \underline{s} such that for all $s \leq \underline{s}$, $\mathcal{S}(s) \leq -1$. \square

I now verify that the unique solution to the HJB equation is indeed the optimal contract.

Proposition 8. *Let G_0 be the payoff from an incentive compatible contract (a, c, τ) . Then $J_0 \geq G_0$.*

Proof. Take any incentive compatible contract (a, c, τ) . By the martingale representation result in Lemma 1, W_t follows Equation (4). Define the realized payoff from using the contract until time $t \leq \tau$ as

$$G_t = \int_0^t e^{-rs} \left(\sum_i dX_{is} - dc_s \right) + e^{-rt} J(W_t),$$

where dX_{is} is the output process under the effort implemented in the contract. By Itô's Lemma,

$$\begin{aligned} dG_t &= e^{-rt} \left(\sum_i \mu_i a_{it} dt + \sum_i \sigma_i dB_{it} - dc_t \right) \\ &+ e^{-rt} \left(J'(W_t) \left(\gamma W_t dt + hn(a_t) dt - dc_t + \sum_i \psi_{it} dB_{it} \right) + J''(W_t) \frac{1}{2} \sum_i \psi_{it}^2 dt \right) \\ &- r e^{-rt} J(W_t) dt. \end{aligned}$$

By the construction of the HJB Equation (12), for any incentive compatible contract

$$-rJ(W) + \sum_i \mu_i a_i + J'(W) (\gamma W + hn(a)) + J''(W) \frac{1}{2} \sum_i \psi_i^2 \leq 0, \quad (15)$$

and for any consumption payout policy, $-(1 + J'(W)) dc_t \leq 0$ on $(0, \bar{W})$. Since $J'(W)$ is bounded, this term is square integrable, and G_t is a supermartingale.

For all finite t , the principal's profit is

$$E[G_\tau] = E[G_{t \wedge \tau}] + E \left[\mathbb{1}_{\{t < \tau\}} e^{-rt} E \left[\int_t^\tau e^{-rs} \left(\sum_i dX_{is} - dc_s \right) + e^{-r\tau} l - J(W_t) \right] \right].$$

The second term is bounded from above by, $J_{fb}(0) - W_t - J(W_t)$. Since $J(W_t) = J(W_t) \geq (l - W_t)$, and $J'(W_t) \geq -1$, we have $(W_t + J(W_t)) \geq l$, so that

$$J_{fb}(0) - W_t - J(W_t) \leq (J_{fb}(0) - l).$$

By the supermartingale property of G_τ , $E[G_{t \wedge \tau}] \leq G_0 = J(W_0)$, and the profit satisfies the following bound

$$E[G_\tau] \leq G_0 + e^{-rt} (J_{fb}(0) - l). \quad (16)$$

□

Proposition 9. *There exists an ε -optimal strategy such that the agent's continuation value in Equation (4) admits a unique weak solution.*

Proof. Existence is guaranteed since both drift and volatility are bounded and measurable (see Krylov (1980), p. 87). From Krylov (2004), we know that weak uniqueness holds if the coefficients are discrete on a finite set. I now construct ε -optimal strategies which partition $[0, \bar{W}]$ into intervals of length ε , so that the project choice is constant on each interval.

The loss in payoff in the HJB equation between the optimal project selection a and an arbitrary a' is

$$L(W, a, a') = \sum_i \left(\mu_i + J'(W) h + \frac{1}{2} J''(W) \psi_i^2 \right) (a_i - a'_i)$$

which is bounded on $(0, \bar{W})$ by a constant \bar{L} , because $J'(W)$ and $J''(W)$ are continuous. Let the grid $\mathcal{W} = \{W_1, W_2, \dots, W_N\}$ for some finite N such that $W_n < W_{n+1}$ for $n = 1, \dots, N - 1$ and $\max_n W_{n+1} - W_n \leq \varepsilon$. For any W , define the Markov control $\hat{a}(W) \in \mathcal{A}$ as follows: $\hat{a}(W) = a(W_n)$ if $W \in [W_n, W_{n+1}]$. I only need to show I can construct a grid \mathcal{W} such that as $\varepsilon \rightarrow 0$, the expected loss vanishes.

If a given project choice a is optimal on an interval of W (i.e. the optimal policy $a(W)$ is constant on that interval), then the loss $L(W, a(W), \hat{a}(W))$ becomes zero for a positive but small ε . If W is on the boundary of two nonempty intervals in which two different project choices a and a' are optimal, take W without loss of generality to be the midpoint of the grid $[W - \varepsilon/2, W + \varepsilon/2]$. The expected loss in this interval starting from W is bounded below

$$\lambda(W) = \bar{L}E \left[\int_0^{\tau_\varepsilon} e^{-rt} dt \right]. \quad (17)$$

Here, the expectation operator is taken under the action $\hat{a}(W)$, and τ_ε denotes the first exit time of W_t from the interval. As $\varepsilon \rightarrow 0$, this expression converges to zero.⁵⁰

If W is an accumulation point of regions in which two actions a and a' are optimal, then $L(W, a, a') = 0$. If W is the midpoint of a grid with size ε , and $\hat{a}(W)$ is fixed at a' , then by continuity of L in W , $L(W, a(W), \hat{a}(W))$ converges to zero as $\varepsilon \rightarrow 0$.

Then, repeating the previous verification argument with $\hat{a}(W)$ implies that for ε sufficiently small, the loss relative to the optimal project selection from the HJB Equation (12) is bounded by $\bar{L}\varepsilon/r$. \square

A.3 Optimality of No-Shirking

I now show that it is optimal to never let the agent shirk whenever the payoff from termination for the principal is sufficiently large, even if there are projects which have negative NPV.

Proposition 10. *For l sufficiently large, the optimal contract implements effort in at least one project for all $t \geq 0$.*

Proof. It is optimal to always implement at least one project if for all $W \in [0, \bar{W}]$, there exists some project such that $b_i(W) > 0$. By the HJB equation (7), this is equivalent to

$$rJ(W) - \gamma W J'(W) = \sum_i b_i(W) a_i^* > 0. \quad (18)$$

⁵⁰Precisely, since $\hat{a}(W)$ is constant, W_t satisfies $dW_t = (\gamma W_t + h\hat{n}) dt + \sum_i \psi_i \hat{a}_i dB_{it}$, and $\lambda(W)$ solves the differential equation $r\lambda(W) = \bar{L} + \lambda'(W)(\gamma W + h\hat{n}) + \lambda''(W) \hat{\psi}^2/2$ subject to the boundary conditions $\lambda(W - \varepsilon/2) = \lambda(W + \varepsilon/2) = 0$, where $\hat{\psi} = \sum_i \psi_i \hat{a}_i$. The boundary value problem is linear in W , and the existence of a C^2 solution is standard, see e.g. Friedman (1975), p. 134, Theorem 2.4. Then, an estimate from Hartman (2002), p. 428, allows to find a bound on $\lambda'(W)$ which is uniform in ε . This implies that $|\lambda(W)| \leq M\varepsilon$ on $[W - \varepsilon/2, W + \varepsilon/2]$ for some $M > 0$ and all $\varepsilon > 0$, which is the desired result.

Here, a_i^* denotes the optimal choice of projects in the HJB equation. Consider the IVP with initial condition $J(0) = l$ and starting slope s , as in Section A.2. I first show that for all initial slopes s and initial values $l > l'$, the solutions to the IVP, $J_{s,l}(W)$ and $J_{s,l'}(W)$ satisfy $J'_{s,l}(W) > J'_{s,l'}(W)$ for all $W > 0$. To see this, note that

$$J''_{s,l}(0) = -H(0, l, s) > -H(0, l', s) = J''_{s,l'}(0),$$

where $H(W, p, q)$ is given by (13), implies $J'_{s,l}(W) > J'_{s,l'}(W)$ on some neighborhood right of 0. Now, suppose $\hat{W} = \inf \{W > 0 : J'_{s,l}(W) < J'_{s,l'}(W)\}$ exists and note that we must have $\hat{W} > 0$. Since $J'_{s,l}$ and $J'_{s,l'}$ are continuous, we have $J'_{s,l}(\hat{W}) = J'_{s,l'}(\hat{W})$. For $W \in (0, \hat{W})$, we have $J'_{s,l}(W) > J'_{s,l'}(W)$ by construction of \hat{W} , and therefore $J_{s,l}(\hat{W}) > J_{s,l'}(\hat{W})$. This implies

$$\begin{aligned} J''_{s,l}(\hat{W}) &= -H(\hat{W}, J_{s,l}(\hat{W}), J'_{s,l}(\hat{W})) \\ &> -H(\hat{W}, J_{s,l'}(\hat{W}), J'_{s,l}(\hat{W})) = -H(\hat{W}, J_{s,l'}(\hat{W}), J'_{s,l'}(\hat{W})) = J''_{s,l'}(\hat{W}). \end{aligned}$$

But then $J'_{s,l}(W)$ cannot cross $J'_{s,l'}(W)$ from above, which implies a contradiction. Hence, $J'_{s,l}(W) > J'_{s,l'}(W) \forall W > 0$.

Now, pick $l = 0$ and some fixed $s > 0$ and let $W_{s0} = \inf \{W > 0 : J'_{s,0}(W) = 0\}$. By construction, $J_{s,0}(W_{s0}) > 0$ and by the previous result, for any $l > 0$,

$$J_{s,l}(W_{s0}) = \int_0^{W_{s0}} J'_{s,l}(W) dW + l \geq \int_0^{W_{s0}} J'_{s,0}(W) dW + l = J_{s,0}(W_{s0}) + l.$$

Since $J_*(0) = \sum_i (\mu_i - h)^+ / r$ and $J_*(W)$ is strictly decreasing, this implies

$$J_{s,l}(W_{s0}) > J_*(W_{s0})$$

for $s > 0$ fixed and l sufficiently large. Thus, that solution to the IVP hits $J_*(W)$ with a positive slope and cannot be a solution to the BVP (12) since it violates the boundary conditions $J_*(\bar{W}) = J_{s,l}(\bar{W})$ and $J'_{s,l}(\bar{W}) = -1$.

Let s^* be the starting slope that solves the BVP (12), $J_l(W)$ the solution to the BVP, and $\bar{W}(l)$ be the value at which $J'_l(W) = 0$. The previous result implies that $s^* \rightarrow 0$ for

l sufficiently large. Also, since $J_l(W) \geq J_l'(W)$ for all $W \geq 0$, we have $\bar{W}(l) \leq \bar{W}(l')$ and thus for all $l > 0$, $\bar{W}(l) \leq \bar{W}(0)$. Since for all l , $J_l(W)$ is strictly concave, we have $J_l'(W) \leq s^*$. Plugging this into condition (18) yields

$$\begin{aligned} rJ_l(W) - \gamma W J_l'(W) &\geq \\ rJ(W) - \gamma W s^*. \end{aligned}$$

For s^* sufficiently small (or equivalently l sufficiently large), we have

$$rJ(W) - \gamma s^* > 0$$

for all $W \leq \bar{W}(l)$. Thus,

$$\sum_i b_i(W) a_i^* > 0$$

for all $W < \bar{W}(l)$, which establishes the result. \square

A.4 Proofs on Project Dynamics

The following proposition offers a complete characterization of when firms choose projects with higher pay-performance sensitivity. Notice throughout Section (4.2) I have assumed that γ is sufficiently close to r . In the proofs below, I am making the assumption explicit.

Proposition 11. *If $J'''(W) \leq 0$, then $b_i'(W) < 0$ for all projects i . If $J'''(W) > 0$, there exists a cutoff $\bar{\psi}(W)$ such that $b_i'(W) > 0$ if and only if $\psi_i > \bar{\psi}(W)$. If*

$$\psi_i^2 > \frac{1}{n(a)} \sum_j \psi_j^2 a_j,$$

*then $b_i'(W) > 0$. There exists a $W_0 \in [0, \bar{W})$ such that $J'''(W) > 0$ for all $W > W_0$.*⁵¹

To establish the result, we need the following Lemma.

Lemma 4. *Whenever $J'(W) < 0$ and $J'''(W)$ exists, we have $J'''(W) > 0$.*

⁵¹Precisely, the W_0 here satisfies $J'(W_0) = 0$.

Proof. Whenever J''' exists it is given by

$$J'''(W) = \frac{-(\gamma - r) J'(W) - J''(W) (\gamma W + hn(a))}{\frac{1}{2} \sum_i \psi_i^2 a_i}, \quad (19)$$

which follows from differentiating the HJB equation (7) for a fixed project choice. This expression is positive because J is concave. \square

Lemma 5. *For any W , and all projects k , if $b'_k(W)$ exists, it is strictly positive if both $J'(W) < 0$ and*

$$\psi_j^2 \geq \frac{1}{n(a)} \sum_i \psi_i^2 a_i.$$

Whenever $J'''(W) < 0$, $b'_k(W) < 0$ for all k and if $J'''(W) > 0$, then $b'_k(W) > 0$ if $\psi_k^2 > -\frac{2J''(W)}{J'''(W)}$.

Proof. We have $b'_k(W) = J''(W) h + J'''(W) \psi_k^2/2$, which follows from differentiating equation (8). Using equation (19) in the proof of the previous Lemma, we have

$$b'_k(W) = J''(W) h - \frac{1}{2} \psi_k^2 \frac{(\gamma - r) J'(W) + J''(W) (\gamma W + hn(a))}{\frac{1}{2} \sum_i \psi_i^2 a_i},$$

which is positive whenever

$$h \frac{1}{2} \sum_i \psi_i^2 a_i - \frac{1}{2} \psi_k^2 (\gamma W + hn(a)) < \frac{(\gamma - r) J'(W)}{J''(W)} \frac{1}{2} \psi_k^2.$$

Since the right hand side is positive whenever $J'(W) < 0$, a sufficient condition is

$$\psi_k^2 \geq \frac{1}{n(a)} \sum_i \psi_i^2 a_i.$$

Note that a sharp condition is $\psi_k^2 > \frac{-2J''(W)}{J'''(W)}$, which is equivalent to

$$\psi_k^2 > \frac{2}{(\gamma W + hn(a))} \left(h \frac{1}{2} \sum_i \psi_i^2 a_i - \frac{(\gamma - r) J'(W)}{J''(W)} \right).$$

\square

The result that for γ close to r the cost of incentives is decreasing was used in the main text to characterize the dynamics. The following proposition formalizes this result.

Proposition 12. *There exists an $\varepsilon > 0$ such that for all $\gamma \in (r, r + \varepsilon)$, we have $J'''(W) > 0$ for any $W \in [0, \bar{W}]$.*

Proof. Suppose $J'''(W) \leq 0$ for some $W \in [0, \bar{W}]$. Then, using equation (19), we have

$$J'(W) \geq \frac{-J''(W)(\gamma W + hn^*)}{\gamma - r},$$

where n^* the optimal number of projects from the HJB equation (7). Plugging this expression into the HJB equation (7), we have

$$\begin{aligned} rJ(W) \geq & \max_a \sum_i \mu_i a_i \\ & - \frac{J''(W)}{\gamma - r} \left((\gamma W + hn^*)(\gamma W + hn(a)) - (\gamma - r) \frac{1}{2} \sum_i \psi_i^2 a_i \right). \end{aligned}$$

As $\gamma \downarrow r$, the term multiplying $J''(W)$ is strictly positive for all $W > 0$ so that

$$rJ(W) \geq \max_a \sum_i \mu_i a_i,$$

which implies a contradiction along the same lines and in the proof of Proposition 7. \square

I now restate Proposition 4 for the convenience of the reader and prove it.

Proposition 13. *There exists an $\varepsilon > 0$ such that for all $\gamma \in (r, r + \varepsilon)$, we have $b'_k(W) > 0$ for all k with*

$$\psi_k^2 > \frac{1}{n(a)} \sum_i \psi_i^2 a_i.$$

Proof. Choose k and W such that

$$r_k \equiv \psi_k^2 - \frac{1}{n(a)} \sum_i \psi_i^2 > 0$$

and $b'_k(W) \leq 0$. The last condition is equivalent to

$$-J'(W)(\gamma - r)\frac{1}{2}\psi_k^2 - J''(W)\left(\frac{1}{2}\psi_k^2\gamma W + hn^*r_i\right) \leq 0$$

where again n^* is the optimal number of projects and I have used the definition of $b_i(W)$ in equation (8). This condition can be rewritten as

$$J'(W) \geq \frac{-J''(W)}{(\gamma - r)\frac{1}{2}\psi_k^2} \left(\frac{1}{2}\psi_k^2\gamma W + hn^*r_i\right).$$

Plugging this inequality into the HJB equation (7) yields

$$\begin{aligned} rJ(W) &\geq \max_a \sum_i \mu_i a_i \\ &\quad - \frac{J''(W)}{(\gamma - r)\frac{1}{2}\psi_k^2} \left(\frac{1}{2}\psi_k^2(\gamma W + hn^*r_k)(\gamma W + hn(a)) - (\gamma - r)\frac{1}{2}\sum_i \psi_i^2 a_i \right). \end{aligned}$$

For $\gamma - r$ sufficiently small, the term multiplying $-J''(W)$ is positive for all W and r_k chosen as above and therefore

$$rJ(W) \geq \max_a \sum_i \mu_i a_i,$$

which again implies a contradiction. □

The result that $b_i(W) \rightarrow \text{NPV}_i$ as $W \rightarrow \bar{W}$ was sketched in the main text.

A.5 Proof of the Implementation in Section 5

I now derive the implementation in Proposition 6. I let M_t denote the cash stock. M_t earns a total interest of rM_t , where r is the interest rate, and it has inflows from the output of all projects $dX_t = \sum_i dX_{it}$ and outflows from the dividends paid on the equity and the payout to the agent. Thus, M_t evolves according to

$$dM_t = rM_t dt + dX_t - dc_t - d\text{Div}_t. \tag{20}$$

Here, Div_t is a dividend process required by shareholders which satisfies

$$dDiv_t = (\gamma - r) M_t dt. \quad (21)$$

The agent receives a constant equity share Ψ_t and bonus payments P_t , which are linear in individual project outputs

$$dP_t = \sum_i \beta_{it} dX_{it}, \quad (22)$$

with nonnegative weights β_{it} .

The agent is also endowed with a personal account⁵² with balance A_t , which pays interest at rate γ and is used in part to pay him. At any point in time, he receives an equity share Ψ_t in any currently chosen project. Whenever a new project is started, he buys equity in that project at a pre-determined price p_t , and when a project is stopped, he sells off the equity. Proceeds from these sales and purchases are deposited in the personal account.⁵³ Finally, the agent may not access funds inside the account, except for when a dividend dc_t is paid. Formally, the account balance satisfies

$$dA_t = \gamma A_t dt + p_t d\Psi_t + dP_t - dc_t^A. \quad (23)$$

Here, dc_t^A is the agent's consumption paid from the account, p_t is the transfer price on equity sales and purchases, and $d\Psi_t = \Psi_t - \lim_{t' \uparrow t} \Psi_{t'}$ denotes the amount of shares that are purchased if $d\Psi_t > 0$ or sold if $d\Psi_t < 0$.

The optimal equity share needs to be low enough to prevent unnecessary risk in the contract, which is achieved exactly by setting it to the minimal equity stake the agent would hold if project-specific shares could be issued. The rest of the incentives is provided by adjusting the sensitivity of the bonus payments.

I now show that by setting $p_t = M_t$, $\Psi_t = \min_{i:a_i=1} h/\mu_i$, and $\beta_{it} = (h/\mu_i - \Psi_t) a_{it}$, the principal can implement the optimal contract. First, I show that the agent's continuation utility satisfies $W_t = \Psi_t M_t + A_t$. Indeed, Equations (20), (21), (22), and (23) imply that the

⁵²The unit of account is irrelevant, and the balance on the agent's account can be interpreted in terms of cash or an incentive point scheme.

⁵³I assume that it is possible for the account to have a negative balance.

agent's HJB equation satisfies

$$\begin{aligned}
\gamma W_t &= \max_{a \in \mathcal{A}} \Psi_t \left(rM_t + \sum_i \mu_i a_{it} - \frac{dc_t}{dt} - dDiv_t \right) + \Psi_t \frac{dc_t}{dt} - h \sum_i a_i + dA_t + \frac{d\Psi_t}{dt} M_t \\
&= \max_{a \in \mathcal{A}} \Psi_t \left(\gamma M_t + \sum_i \mu_i a_i \right) - h \sum_i a_i + \gamma A_t + \sum_i e_{it} \left(\frac{h}{\mu_i} - \Psi_t \right) \mu_i a_i \\
&= \Psi_t \gamma M_t + \max_{a \in \mathcal{A}} \left(\sum_i (\Psi_{it} \mu_i a_i - h a_{it}) \right) + \gamma A_t \\
&= \Psi_t \gamma M_t + \gamma A_t
\end{aligned}$$

Given the law of motion for A_t and M_t we can verify that if $W_t = \Psi_t M_t + A_t$, then dW_t still satisfies Equation (4). Thus the optimal contract is implemented.

References

- Aliprantis, C. D. and K. Border (2006). *Infinite dimensional analysis: A hitchhiker's guide*. Springer Science & Business Media.
- Babenko, I., B. Bennett, J. M. Bizjak, and J. L. Coles (2017). Clawback provisions. *Working Paper*.
- Bernardo, A. E., H. Cai, and J. Luo (2009). Motivating entrepreneurial activity in a firm. *Review of Financial Studies* 22(3), 1089–1118.
- Bernardo, A. E., J. Luo, and J. J. Wang (2006). A theory of socialistic internal capital markets. *Journal of Financial Economics* 80(3), 485–509.
- Biais, B., T. Mariotti, J.-C. Rochet, and S. Villeneuve (2010). Large Risks, Limited Liability, and Dynamic Moral Hazard. *Econometrica* 78(1), 73–118.
- Bohren, J. A. and T. A. Kravitz (2016). Optimal contracting with costly state verification, with an application to crowdsourcing. *Working Paper*.
- Bond, P. and A. Gomes (2009). Multitask principal-agent problems: Optimal contracts, fragility, and effort misallocation. *Journal of Economic Theory* 144(1), 175–211.
- Cassell, C. A., S. X. Huang, J. M. Sanchez, and M. D. Stuart (2012). Seeking safety: The relation between CEO inside debt holdings and the riskiness of firm investment and financial policies. *Journal of Financial Economics* 103(3), 588–610.
- Coles, J. L., N. D. Daniel, and L. Naveen (2006). Managerial incentives and risk-taking. *Journal of Financial Economics* 79(2), 431–468.
- Cremers, M. and D. Palia (2011). Tenure and CEO pay. *Working Paper*.
- DeMarzo, P., M. Fishman, Z. He, and N. Wang (2012). Dynamic Agency and the q Theory of Investment. *Journal of Finance* 67(6), 2295–2340.
- DeMarzo, P. and Y. Sannikov (2006). Optimal security design and dynamic capital structure in a continuous-time agency model. *Journal of Finance* 61(6), 2681–2724.
- Dewatripont, M., I. Jewitt, and J. Tirole (2000). Multitask agency problems: Focus and task clustering. *European Economic Review* 44(4-6), 869–877.

- Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *Review of Economic Studies* 51(3), 393–414.
- Fleming, W. H. and H. M. Soner (2006). *Controlled Markov processes and viscosity solutions*. Springer Verlag.
- Fong, K. (2007). Evaluating Skilled Experts: Optimal Scoring Rules for Surgeons. *Working Paper*.
- Friedman, A. (1975). *Stochastic differential equations and applications*. New York: Academic Press.
- Gilbarg, D. and N. S. Trudinger (2001). *Elliptic partial differential equations of second order*. Springer Verlag.
- Graham, J. R. and C. R. Harvey (2001). The theory and practice of corporate finance: Evidence from the field. *Journal of Financial Economics* 60(2-3), 187–243.
- Guay, W. R. (1999). The sensitivity of CEO wealth to equity risk: an analysis of the magnitude and determinants. *Journal of Financial Economics* 53(1), 43–71.
- Harford, J. (1999). Corporate cash reserves and acquisitions. *Journal of Finance* 54(6), 1969–1997.
- Harris, M. and A. Raviv (1996). The capital budgeting process: Incentives and information. *Journal of Finance* 51(4), 1139–1174.
- Harris, M. and A. Raviv (1998). Capital budgeting and delegation. *Journal of Financial Economics* 50(3), 259–289.
- Hartman, P. (2002). *Ordinary Differential Equations*. Classics in Applied Mathematics. SIAM.
- He, Z. (2009). Optimal Executive Compensation when Firm Size Follows Geometric Brownian Motion. *Review of Financial Studies* 22(2), 859–892.
- Holmström, B. and P. Milgrom (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization* 7(special issue), 24–52.
- Jensen, M. C. and W. H. Meckling (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* 3(4), 305–360.

- Karatzas, I. and S. E. Shreve (1991). *Brownian motion and stochastic calculus*. Springer Verlag.
- Krylov, N. V. (1980). *Controlled diffusion processes*. Springer Verlag.
- Krylov, N. V. (2004). On weak uniqueness for some diffusions with discontinuous coefficients. *Stochastic Processes and their applications* 113(1), 37–64.
- Malenko, A. (2018). Optimal Dynamic Capital Budgeting. *Review of Economic Studies*, forthcoming.
- Manso, G. (2011). Motivating Innovation. *Journal of Finance* 66(5), 1823–1860.
- Massa, M. and R. Patgiri (2009). Incentives and mutual fund performance: Higher performance or just higher risk taking? *Review of Financial Studies* 22(5), 1777–1815.
- Miquel-Florensa, J. (2007). Optimal Incentives in Dynamic Multiple Project Contracts. *Working Paper*.
- Murphy, K. J. (1999). Executive compensation. *Handbook of Labor Economics* 3, 2485–2563.
- Radner, R. and L. Shepp (1996). Risk vs. profit potential: A model for corporate strategy. *Journal of Economic Dynamics and Control* 20(8), 1373–1393.
- Rauh, J. D. (2009). Risk shifting versus risk management: Investment policy in corporate pension plans. *Review of Financial Studies* 22(7), 2687–2733.
- Richardson, S. (2006). Over-investment of free cash flow. *Review of Accounting Studies* 11(2-3), 159.
- Sannikov, Y. (2008). A continuous-time version of the principal-agent problem. *Review of Economic Studies* 75(3), 957–984.
- Schattler, H. and J. Sung (1993). The first-order approach to the continuous-time principal-agent problem with exponential utility. *Journal of Economic Theory* 61(2), 331–371.
- Spear, S. E. and S. Srivastava (1987). On repeated moral hazard with discounting. *Review of Economic Studies* 54(4), 599–617.
- Yong, J. and X. Y. Zhou (1999). *Stochastic controls: Hamiltonian systems and HJB equations*, Volume 43. Springer Verlag.

Zhu, J. (2013). Optimal Contracts with Shirking. *Review of Economic Studies* 80(2), 812–839.