Incentives, Project Choice and Dynamic Multitasking[*]
(Job Market Paper)

Martin Szydlowski[†]

January 14, 2013

Abstract

I study the optimal choice of investment projects in a continuous-time moral hazard model with multitasking. While in the first best, projects are invariably chosen by the net present value (NPV) criterion, moral hazard introduces a cutoff for project selection which depends on both a project's NPV as well as its risk-return ratio. The cutoff shifts dynamically depending on the past history of shocks, the current firm size, and the agent's continuation value. When the ratio of continuation value to firm size is large, investment projects are chosen more efficiently, and project choice depends more on the NPV and less on the risk-return ratio.

The optimal contract can be implemented with an equity stake, bonus payments, as well as a personal account. Interestingly, when the contract features equity only, the project selection criterion resembles a hurdle rate.

1   Introduction

In the neoclassical investment framework, firms operate a single technology and continue the same activity on a different scale as they grow. Within this standard paradigm, the literature has studied how agency problems affect the optimal level of investment. However, the paradigm not take into account that firms are crucially dependent on choosing the right projects. With agency, executing different projects makes it necessary to dynamically change managerial incentives. The question is then how to incentivize the manager to select the optimal portfolio of projects and how the cost of providing incentives distorts the project selection over time.

In this paper, I characterize a firm's optimal project choice in a continuous-time moral hazard model. I show that even though firm and manager are risk-neutral, project choice is not determined by the NPV criterion alone, but instead by a project-specific markup over NPV, which changes dynamically depending on the firm's past payoffs. This markup is driven entirely by the cost of incentives.

In my model, the firm hires a manager and has access to a fixed portfolio of projects. Each project yields a risky payoff stream, which is characterized by the project's risk-return profile. The manager's effort is required for projects to be profitable, but may be unobserved, which is the source of the agency friction. When the manager's effort allocation is observed, the firm's project selection follows the NPV criterion, since the firm is risk-neutral. That is, whenever a project's average payoff is higher than the cost of effort, the project is chosen irrespective of its risk. This criterion is static, and the portfolio of chosen projects never changes.

With moral hazard, there is both over- and underinvestment in projects relative to the NPV criterion.[1] While underinvestment is driven by the cost of incentives, overinvestment is caused by the principal's inability to punish the manager in the presence of a limited liability constraint.[2] Intuitively, since the principal cannot demand money from the manager, the only option after sufficiently bad performance is to fire him and liquidate the firm.[3] Because the manager may shirk, the principal must incentivize each chosen project by making the contract dependent on the project's output. This increases the likelihood that a path of bad outcomes leads to liquidation, and is the source of the cost of incentives. Underinvestment occurs precisely because a project's positive NPV may not compensate for the increase in liquidation risk. Overinvestment occurs because the principal seeks a less inefficient punishment scheme. Assigning projects to the manager increases his effort cost and serves as an alternative to firing, even when the NPV of the assigned projects is negative.

Since a project's markup reflects the cost of incentives, it is higher for projects with a high risk-return ratio. Intuitively, a high ratio means that it is difficult to detect shirking, and therefore the project is difficult to incentivize. The dynamics of the markup are determined by how the cost of incentives changes with the firms past payoffs, and how far the contract is from the liquidation boundary. When the past performance is sufficiently good, the cost of incentives decreases and the firm chooses projects which are more difficult to incentivize. If the projects' risk-return ratio is superlinear, this translates into taking more risky and more profitable projects. When liquidation becomes sufficiently unlikely, the firm's project selection approaches the first best. All negative

---

[1]Throughout the paper, I use the term underinvestment whenever a project with positive NPV is not taken, and overinvestment whenever one with negative NPV is taken. Since only positive-NPV projects are taken in the first best, underinvestment means equivalently that a project which is taken in the first best is not taken under moral hazard. The analog holds for overinvestment.

[2]Given risk neutrality, this is necessary to ensure that the optimal contract is not trivial. See Section 3 for details.

[3]It is possible to study a variant of the model in which the manager is replaced after bad performance, and the firm is subject to hiring costs. The intuition I outlined and the qualitative results in my model would be unchanged in this case.

NPV proejcts are phased out, and all positive NPV projects are taken regardless of their risk.

My result that firms close to liquidation forgo risky projects is supported by Rauh (2009), who shows empirically that pension funds with weak credit ratings, which may be interpreted as a proxy for the fund's bankruptcy probability, choose safer investments, while financially sound ones do the opposite. My finding also differs from the seminal risk shifting result in Jensen and Meckling (1976). In that paper, the possibility of liquidation leads firms to take on excessive risk, while in my work, it deters the firm from taking risky projects.

In addition to choosing projects, the firm in my setting has a capital stock and operates a neoclassical investment technology. This allows me to make predictions on how the optimal project choice changes with the firm's capital stock. Holding past performance constant, smaller firms choose their projects more efficiently, while larger ones forgo positive NPV projects which are sufficiently risky, and may overinvest in projects. By incorporating capital, my model nests DeMarzo et al. (2012), who study capital investment in a continuous-time moral hazard framework, and find that the agency friction leads to too little investment. The distinction between projects and capital investment enables me to address a set of qualitatively different questions. While DeMarzo et al. (2012) are study the *level* of investment at any given point in time, I make predictions on *which kinds* of projects are chosen.

In addition, my framework also yields insights into firms' use of hurdle rates, the exercise of real options under agency, and managerial incentive contracts.

When the principal only observes the sum of all project outputs, project choice resembles a hurdle rate. In particular, the NPV of each chosen project is above the same threshold, which depends on the project with the lowest NPV currently chosen. This hurdle rate allocation is not efficient and the contract carries excessive risk. Consequently, my model suggests that hurdle rates, which are widely observed in practice, arise when the firm is unable to condition the contract on individual projects, or unable to find incentive schemes which condition on this information.[4]

Commonly, deviations from the NPV criterion are explained by real options.[5] In the real options framework, positive NPV projects are not taken because they carry an option value of waiting, and negative NPV projects are taken if they entail the option to start additional projects in the future. While real options rely on either irreversibilities or fixed costs, my model can generate deviations from the NPV criterion in the absence of both, and solely as a consequence of moral hazard. Since fixed costs of starting or stopping projects are relevant in practice, I introduce them into the contract in Section 6. I show that my model serves as an approximation to a real options framework under agency with small fixed costs. The approximation result has a precise interpretation. We can compute each project's marginal benefit according to the criterion in my model, and choose projects accordingly. The loss in value from this scheme is negligible, and wrong

---

[4]I will explicitly study the relation between hurdle rates and managerial compensation structures observed in practice in Section 5.

[5]See Dixit et al. (1994) for further references to this literature.

projects are almost never chosen, provided that the fixed costs are small. Thus, my model can serve as a guideline for choosing projects in the presence of fixed costs.

In Section 5, I derive an implementation of the optimal contract which features an equity stake, a fixed wage, and bonus payments. This implementation rationalizes the majority of managerial contracts found in reality, and is due to the introduction of multiple projects. As Murphy (1999) documents, most CEO's contracts consist of equity, wage and a bonus, and the latter is a linear weighted function of the CEO's performance across different categories. This is exactly the case in my model. In contrast, an implementation which consists of equity only can implement the hurdle rate allocation, but fails to implement the second best. Thus, hurdle rates arise in firms whose managerial contracts put too much emphasis on equity.

My implementation also rationalizes the manager buying and selling equity at ex-ante determined transfer prices. The optimal equity share in my model is not static, as in DeMarzo and Sannikov (2006) or DeMarzo et al. (2012), and has to be adjusted when the project selection changes. These equity transfers may distort incentives, since if the manager expects to be stripped of shares in the future, he may be less likely to put in effort. The transfer prices are designed to exactly offset this incentive effect.

The manager's continuation value in the implementation equals the sum of the firm's cash balances and the value of a personal account.[6] Keeping the account value constant implies that firms with higher cash buffer, relative to firm size, choose their projects more efficiently.

The paper proceeds as follows. Section 2 provides an overview of 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 both under output- and project-based incentives. It also considers extensions of the original setting in which the agent can steal instead of exerting effort, and the firm can allocate funds between projects. The implementation outlined in the paragraphs above is derived in Section 5. Finally, Section 6 provides a discussion how my setup relates to the real options framework while Section 7 concludes.

2 Related Literature

The present model is related to three strands of literature. The techniques employed to characterize the dynamic contract stem from the literature on continuous time contracting, most notably Schattler and Sung (1993) and Sannikov (2008). Recent contributions in this literature which share certain features with my setup include Biais et al. (2010), who study investment and downsizing of firm size as a way to incentivize accident prevention, He (2009), in whose model the agent's effort

---

[6]The use of cash balances is similar to DeMarzo et al. (2012), while the existence of a personal account is novel to my model, and due to the changing nature of the manager's optimal pay-performance sensitivity.

directly affects the evolution of firm size, and Fong (2007) who studies a binary effort decision with two agents.

The closest paper to mine is DeMarzo et al. (2012), which is the first to study firm investment in a continuous-time agency framework. My model features multiple projects with varying risk-return profiles in addition to a continuous investment variable, and nests DeMarzo et al. (2012). My main focus however is on the optimal choice of projects, instead of the level of capital investment, and my paper is aimed to be complementary to the analysis of DeMarzo et al. The two models also differ in terms of implementation. Due to the multitask structure, my implementation uses transfer prices for equity, bonus payments and a personal account in addition to cash, while the implementation in DeMarzo et al. (2012) relies on cash and equity only. Also, the optimal equity share in my setting changes over time, and an implementation with constant equity share cannot be optimal.

The problem of multitasking has received significant attention since the seminal article of Holmstrom and Milgrom (1991).[7] Due to the complex nature of the problem, dynamic studies of multitasking are rare. The most recent ones include Manso (2006), who studies the trade off between two tasks interpreted as exploration and exploitation, and Miquel-Florensa (2007), who answers under whether two tasks should be executed sequentially or in parallel, depending on the strength of the externalities between them.

In a continuous time setup, Hartman-Glaser et al. (2010) consider a multitasking model where an underwriter issues a mortgage backed security, and may shirk in selecting the mortgages, which will default with different rates. They find that bundling the mortgages is optimal, which is reminiscent of a similar static result by Laux (2001), and the underwriter will either exert effort in all mortgages or none.

Finally, my model is related to the literature on optimal investment. The real options literature[8] offers a complementary view on the issue of project choice, in which both fixed costs and an option value of waiting drive deviations from the NPV criterion. Although the real options framework has been extended to incorporate agency frictions, see p.e. Grenadier and Wang (2005), Grenadier and Malenko (2010) and Morellec and Schürhoff (2010), studies are mostly limited to the choice of a single project. This is because taking one project will affect the value of other projects, via irreversibilities or fixed costs, which makes it difficult to characterize the the optimal choice of multiple projects. In my model, the externality between projects is well behaved, which allows for the characterization of an entire project portfolio.

The capital budgeting literature, see Harris and Raviv (1996) and Harris and Raviv (1998), studies the choice of projects when a division agent has private information about project quality and has an incentive to misreport. In Harris and Raviv (1996) both over-and under-investment relative to the NPV criterion can occur, depending on whether the project is of low or high quality, and the

---

[7]For a recent contribution and further references, see Bond and Gomes (2009).

[8]See e.g. Dixit et al. (1994) for a comprehensive overview.

optimal contract can be implemented by allocating a fixed budget to the agent. In a similar setup, Berkovitch and Israel (2004) derive an implementation which takes the form of an internal rate of return, which is similar to my result on the hurdle rate. Finally, Malenko (2011) considers a dynamic version of the problem, and derives the capital budgeting mechanism in continuous time.

Since in the capital budgeting literature, projects only have an unidimensional quality associated with them instead of risk and return, it is difficult to compare my results. If the average project payoff in my framework is interpreted as quality, and the relation between payoff and the SN ratio is positive and sufficiently large, then my model will imply that there are too many low quality projects and too few high quality projects in the firm's portfolio, in line with the above.

Another related area is delegated portfolio management as found in Cadenillas et al. (2007), He and Xiong (2008), Ou-Yang (2003) and Makarov and Plantin (2010). The key difference between my model and the portfolio choice framework, is that, very similar to the real options literature, project choice is a binary decision. This allows me to characterize selection criteria as well as the delay in project implementation stemming from the agency friction.

3 Model Setup

3.1 Projects and Investment Technology

I study a firm which has access to a fixed number of projects and hires an agent. The shareholders of the firm act as the principal. Time is continuous, indexed by $t \in \mathbb{R}_+$, 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 positive only if the agent works, and otherwise it is zero. Formally, for any project $i \in \{1, \ldots, N\}$, the cumulative output x_{it} is given by

$$dx_{it} = \mu_i a_{it} dt + \sigma_i dB_{it}. \tag{1}$$

The Brownian Motion B_{it} to captures the idiosyncratic noise in the project's payoff, while the binary variable $a_{it} \in \{0, 1\}$ denotes the agent's decision to work or shirk. μ_i is the project's average profitability when the agent exerts effort, 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.[9]

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 $\mathcal{A} = \{0, 1\}^N$. The event $a_{it} = 1$ shall be interpreted as project i being *assigned* to the agent, or alternatively project i being *implemented* at time t.

---

[9]Formally, $B_{it} \perp B_{js}$ for all times $t, s \geq 0$. This helps 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 firm has a capital stock and operates a neoclassical investment technology.[10] A higher capital stock increases the payoff of each project linearly. Given managerial effort allocation a_t and capital stock π_t, the incremental payoff of the principal is[11]

$$\pi_t \sum_i a_{it} dx_{it}.$$

Thus, total payoffs depend on both the capital stock, which I also interpret as firm size, and the number of implemented projects.

I denote investment as a fraction of capital as I_t. The capital stock increases with investment, depreciates over time at rate $\delta > 0$, and follows the law of motion

$$d\pi_t = (I_t - \delta) \pi_t dt. \tag{2}$$

Investment is subject to an adjustment cost, for which I assume the functional form $\pi_t \kappa (I_t)$. The function $\kappa (.)$ is increasing, convex, and satisfies $\kappa (0) = 0$. Unlike project choice, investment in capital is not subject to agency.

3.2 Utility Functions and the Contract Space

The vector of project-specific Brownian Motions $B_t := (B_{1t}, ..., B_{Nt})$ is defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $\mathcal{F}_t$, which satisfies the usual conditions.[12] 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[13] $a = \{a_t \in \mathcal{A} : t \geq 0\}$, and a firing time τ. Effort and consumption are progressively measurable with respect to $\mathcal{F}_t$, and τ is a stopping time.

Both principal and agent are risk-neutral. The agent's effort cost is linear, symmetric in effort for each project, and given by $h \sum_i a_{it}$. This specification, together with the mutual independence of the noise in project outputs, ensures that any dependence between projects along the path of the optimal contract is driven by the agency friction, instead of assumptions on technology or

---

[10]This allows my paper to nest the model of DeMarzo et al. (2012). All assumptions in the current paragraph are analogous to that paper. All qualitative results in my paper continue to hold when investment is omitted. In this case, the scaled HJB Equation (8) is the firm's actual HJB equation, and all terms involving I_t disappear.

[11]Throughout the paper, I abbreviate the sum over projects $\sum_{i=1}^N$ as $\sum_i$ where no confusion can arise.

[12]See Karatzas and Shreve (1991), p. 10.

[13]The discreteness of the effort process poses a potential problem. If the agent's effort a_{it} is not of bounded variation on an interval of time, the agent's continuation value process may not be sufficiently well behaved to guarantee a unique strong solution. The problem can be solved either by assuming a small positive switching cost which is incurred by the principal whenever the effort changes, so that it will never be optimal to change the project allocation more than once on a sufficiently small interval of time, or by considering an ε-optimal strategy which leaves the project allocation constant on such interval. The model with switching costs is studied in Section 6, and the existence of ε-optimal strategies is proven in Proposition (17) in the Appendix.

preferences. The agent's discounted lifetime utility W_0 is given by

$$W_0 = E\left[\int_0^\tau e^{-\gamma t}\left(dc_t - \pi_t h \sum_i a_{it} dt\right)|\mathcal{F}_0\right]. \tag{3}$$

The agent is protected by a limited liability constraint. That is, for any time t, the increment in his consumption payments dc_t may not be negative.[14]

The principal receives the payoffs from each project, pays the agent's compensation and bears the adjustment costs from investment in capital. Her expected discounted payoff is given by

$$J_0 = E\left[\int_0^\tau e^{-rt}\left(\left(\pi_t \sum_i \mu_i a_{it} - \pi_t \kappa(I_t)\right) dt - dc_t\right)|W_0, \pi_0\right]. \tag{4}$$

I assume that principal and agent have different discount factors, and that the agent is less patient, i.e. $r < \gamma$. As noted in DeMarzo and Sannikov (2006), this prevents the principal from postponing the agent's consumption forever. I also impose an upper bound $\bar{I}$ on relative investment, so that $I_t \leq \bar{I} < r + \delta$, to ensure that the firm's value function is bounded.[15] Finally, when the firm is shut down, the principal can recover a fraction $l > 0$ of current capital, and her payoff is $l\pi_t$. The agent receives an outside payoff of zero in this case.

3.3 Incentive Compatibility

To make the dynamic contracting problem tractable, I show that at any given point in time, the entire history of the contract can be summarized by a two state variables, the capital stock and the agent's expected continuation utility.[16] I focus on the continuation value in this section. In the next section, I establish that the state variables can be transformed into a single one.

For any incentive compatible contract (a, c, τ), this continuation utility is given by

$$W_t = E\left[\int_t^\tau e^{-\gamma(s-t)}\left(dc_s - \pi_t h \sum_i a_{it} ds\right)|\{a_s, c_s\}_{s \geq t}, \mathcal{F}_t\right].$$

Lemma 1 below uses the martingale representation theorem[17] to derive a law of motion for the continuation value W_t, which follows a diffusion process with respect to the multidimensional

---

[14]Precisely, we have $dc_t = c_t - \lim_{t' \uparrow t} c_{t'} \geq 0$ almost surely for all $t \in \mathbb{R}_+$. 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 $\underline{c} > 0$.

[15]If $I = r + \delta$ the principal's value (or equivalently shareholders value) of the firm might be infinite, since the firm could grow at a fast enough rate to negate any discounting.

[16]Using the agent's continuation value as a state variable is a common technique in dynamic contracts. See e.g. Spear and Srivastava (1987) for an illustration.

[17]See Karatzas and Shreve (1991), Theorem 4.15, p. 182 for the statement and Sannikov (2008) for its application to contracts in continuous time.

Brownian Motion B_t. Intuitively, given any project selection rule a and consumption schedule c, the only source of uncertainty in the model is the vector of Brownian noise terms B_t, and therefore at each point in time the agent's continuation value must be a function the past realizations of noise. The lemma also states an incentive compatibility condition which ensures that the agent exerts effort.

Lemma 1. *For any progressively measurable effort process a and consumption process c, there exists a collection of progressively measurable and square integrable stochastic processes $\left\{(\psi_{it})_{i=1}^N : 0 \leq t \leq \tau\right\}$, such that*

$$dW_t = \left(\gamma W_t + \pi_t h \sum_i a_{it}\right) dt - dc_t + \pi_t \sum_i \psi_{it} dB_{it}. \tag{5}$$

The contract is incentive compatible (IC) if and only if

$$\psi_{it} \geq \frac{\sigma_i}{\mu_i} h \tag{6}$$

whenever $a_{it} = 1$.

The parameter ψ_{it} measures the sensitivity of the agent's continuation value with respect to the noise in project i's output. Since the principal can control both consumption and effort, she is able to determine how much the continuation value responds to project outputs, and we can think of ψ_{it} being chosen directly in the optimal contract. When the output of project i features an unexpected jump by dB_{it}, the agent's continuation value changes by $\psi_{it} dB_{it}$. To see how this impacts his decision to exert effort, consider a deviation for a short period of time dt, during which the agent is shirking in project i.

Without exerting effort, his utility rises by $\pi_t h dt$. Because the principal would not know that the agent is shirking, she expects that $dB_{it} = \frac{\pi_t}{\sigma_i} (dx_{it} - \mu_i dt)$, while the true process is $dx_{it} = \sigma_i dB_{it}$. Hence, the principal's expectation about the realization of noise falls short by $-\pi_t \frac{\mu_i}{\sigma_i} dt$, and by the representation (5), the agent loses $\psi_i \pi_t \frac{\mu_i}{\sigma_i} dt$ in continuation utility. To induce effort, this loss must be larger than $\pi_t h$, which leads to Equation (6).

Lemma 1 also illustrates why the signal to noise ratio $SN_i = \frac{\mu_i}{\sigma_i}$ is important for providing incentives. When the agent shirks, he affects the principal's beliefs about the realization of dB_{it}. When the project is relatively safe, and the ratio is large, observing a shortfall in output by $\mu_i dt$ while the agent is working is a very unlikely event, and corresponds to a large negative realization of the Brownian noise. Thus, it is easy to detect shirking and the agent's continuation value does not have to react much to output to provide incentives .[18] Since $\frac{\sigma_i}{\mu_i}$ is the risk-return ratio of the project, an alternative interpretation is that shirking is difficult to detect for projects with high risk-return ratio, and easy to detect for projects with low risk-return ratio.

Analogously to the discrete time contracting literature, Equation (5) should be interpreted as a

---

[18]Formally, in Equation (5), the shortfall in output is equivalent to a very large negative realization of dB_{it}, which for given ψ_{it} implies that W_t falls by a relatively large amount, while the opposite is true for when σ_i is large.

promise keeping constraint. Given a continuation value W_t, higher consumption dc_t implies that ceteris paribus, the agent's promised value at the end of a small interval of time W_{t+dt} will be smaller, while demanding more effort implies that the principal has to promise more utility to the agent in the future.

3.4 The Optimal Contract

With the result of Lemma 1, the optimal contract can be expressed as a choice of processes $\left\{(\psi_{it})_{i=1}^N, c_t, I_t : 0 \leq t \leq \tau\right\}$, and a firing time τ. The principal seeks to maximize the firm value (4), subject to the promise keeping constraint (5), the law of motion for firm size (2), and the incentive compatibility condition (6). Formally, we have

$$J(W_0, \pi_0) = \max_{\{\psi_{it}, c_t, \tau, I_t\}} E\left[\int_0^\tau e^{-rt}\left(\pi_t \sum_i \mu_i a_{it} dt - dc_t - \pi_t \kappa\left(I_t\right) dt\right) + e^{-r\tau}\gamma\pi_t \,|\mathcal{F}_0\right] \quad (7)$$

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

$$d\pi_t = \pi_t\left(I_t - \delta\right) dt$$

$$\psi_{it} \geq \frac{\sigma_i}{\mu_i} h \ \text{ if } \ a_i = 1.$$

The principal's problem depends on two variables, the initial continuation value W_0 and the initial capital stock π_0. To simplify the analysis, I show that the principal's value function can be expressed as a function of a single state variable, which is the agent's scaled continuation value $w_t = \frac{W_t}{\pi_t}$. Intuitively, all components on the right hand side of the principal's value function $J\left(W, \pi\right)$ are multiples of π_t, except for the payout to the agent dc_t. Since this payout is a choice variable, the principal can simply maximize over the payout relative to capital $\frac{dc_t}{\pi_t}$.[19] Similarly, the laws of motion for the firm size and the agent's continuation value are also linear in π_t. In particular, dividing Equation (5) by π_t implies

$$\frac{dW_t}{\pi_t} = \left(\gamma \frac{W_t}{\pi_t} + h \sum_i a_{it}\right) dt - dc_t + \sum_i \psi_{it} dB_{it},$$

which determines the law of motion for the scaled value w_t. This suggests that the principal's value function is linear in π_0, and has the functional form $J\left(W_0, \pi_0\right) = \pi_0 j\left(w_0\right)$ for some function $j\left(.\right)$ to be determined. I verify this intuition in the proof of Proposition 2 in Section A.1 of the Appendix.

---

[19]Formally, the principal optimizes over $d\tilde{c}_t = \frac{dc_t}{\pi_t}$, and the term $-dc_t$ is replaced by $-\pi_t d\tilde{c}_t$. With a slight abuse of notation, I keep denoting the scaled payouts as dc_t.

4 Properties of the Optimal Contract

4.1 Shape of the Value Function

In this section, I show that the principal's value function is the solution to a version of the Hamilton-Jacobi-Bellman (HJB) equation, with the scaled continuation value w_t as the only state variable. This solution is used in the following sections to characterize the choice of projects, investment in capital, and the payout and firing policies of the firm.[20]

Proposition 2. *Let $n_t = \sum_i a_{it}$ denote the number of projects taken at time t. The HJB equation*

$$rj(w) = \sup_{a,I} \sum_i \mu_i a_i - \kappa(I) + j'(w)((\gamma - I + \delta)w + hn) \tag{8}$$
$$+ j''(w)\frac{1}{2}\sum_i \psi_i^2 a_i + (I - \delta)j(w)$$

with the boundary conditions

$$j(0) = l$$
$$j'(\bar{w}) = -1$$
$$j''(\bar{w}) = 0$$

has a unique twice continuously differentiable solution on the interval $[0, \bar{w}]$, and equals the principal's optimal value function. The region $(0, \bar{w})$ is partitioned into regions on which a particular project selection a is optimal. 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. The third derivative $j'''(w)$ exhibits a jump whenever project selection changes.

Figure 1 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 $l\pi_t$. This is reflected in the boundary condition $j(0) = l$. Intuitively, because the agent is protected by limited liability, the worst future path of consumption the principal can pay to him involves zero consumption forever.[21] Exerting effort in the future would imply a negative continuation value, as can be seen from Equation (5). By shirking forever the agent can guarantee himself a continuation payoff of at least zero. Thus, once $w_t = 0$, any incentive compatible contract involves shirking forever and no consumption payments. The principal's expected value in this case is zero, and it is optimal for her to fire the agent in order to receive the scrap value of $l\pi_t$.

---

[20] The shape of the value function and the termination and payout policies are analogous to DeMarzo and Sannikov (2006) and DeMarzo et al. (2012), and are explained intuitively below for the convenience of the reader. Proposition 2 confirms that the characterization is robust to allowing for multiple projects, which introduces additional challenges in establishing the results. See Appendix A.2 for details.

[21] Formally, $dc_s = 0$ for all $s \geq t$.

The existence of the firing boundary also explains why the principal's value function is concave, and why it may be increasing in the agent's continuation value. Introducing more risk in the contract, exemplified by higher ψ_{it}, makes the agent's value react more strongly to the noise in output, and increases the likelihood of hitting the liquidation boundary due to a sequence of bad outcomes along the path of the contract. This termination is inefficient, and therefore an increase in risk must lower the principal's utility. A higher continuation value for the agent implies either less effort or higher expected consumption payments in the future, both of which lower the principal's payoff. At the same time, a higher w_t implies a lower likelihood of firing, which is beneficial for the principal. When w_t is high, the first effect dominates and the principal's value is decreasing in w_t. When the agent's value is low, and termination is sufficiently likely, the reduction in firing probability dominates, and the principal's value is increasing in w_t.[22]

When $w_t = \bar{w}$, the agent is paid a discrete amount, and $dc_t > 0$. Whenever the principal pays the agent, her value changes by $-\left(1 - j'(w_t)\right) dc_t$.[23] Therefore, $dc_t > 0$ whenever $j'(w_t) \leq -1$, and $dc_t = 0$ otherwise. Since the principal's value is concave in w, and the agent's value jumps downwards whenever $dc_t > 0$ by Equation (5), the point at which the agent is paid is unique, which leads to the boundary condition $j'(\bar{w}) = -1$. The second condition $j''(\bar{w}) = 0$ is the super contact condition, and guarantees that the payment threshold $\bar{w}$ is chosen optimally.[24]

4.2 Project Choice

Without moral hazard, a project is chosen at all points in time if the average payoff is higher than the effort cost, i.e. $\mu_i \geq h$, and never chosen otherwise. Hence, project choice follows the NPV criterion, and is independent of the agent's continuation value W_t or firm size π_t.

With moral hazard, the choice of projects is determined from the HJB Equation (8). The equation implies that at any point in time, the principal's problem is separable in each project once the scaled continuation value is taken into account. Therefore, the marginal benefit of each project can be determined separately, and is given by

$$b_i(w_t) = \mu_i + j'(w_t) h + j''(w_t) \frac{1}{2} \psi_i^2. \tag{9}$$

A project is executed whenever the marginal benefit is positive. Whether this is the case depends on the project's average payoff μ_i, the cost of compensating the agent for effort $j'(w_t) h$, and the

---

[22]Note 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 this would be mutually beneficial. The overall value for the principal is lower compared to 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.

[23]The first term is the direct loss in cash paid to the agent, while the second term measures how the change in the agent's continuation value affects the principal.

[24]Since $j'(\bar{w}) = -1$, the principal is indifferent between paying and not paying the agent at $\bar{w}$. If for example $j''(\bar{w}) < 0$, it would be optimal for the principal wait until w_t reaches some $w' > \bar{w}$, and then pay the agent, since at this point $-\left(1 - j'(w')\right) dc_t > 0$. The optimal payment threshold $\bar{w}$ is the one at which this is not profitable.
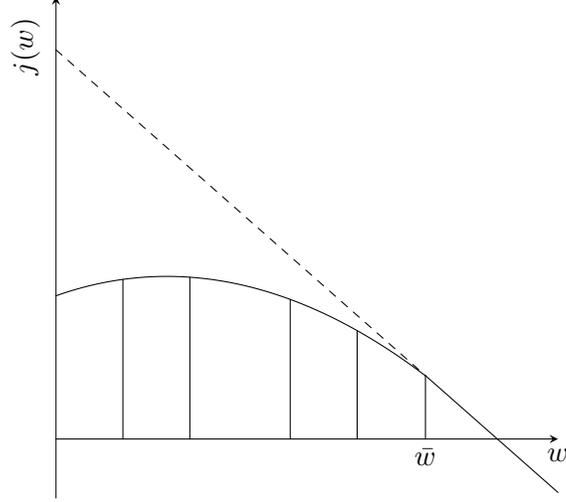
Figure 1: Shape of j(w) and continuation regions

cost of providing incentives $j''(w_t) \frac{1}{2} \psi_i^2$. While the marginal benefit of project i may depend on both past and current choices regarding other projects, any potential dependence is summarized by the dynamics of w_t and the value function $j(w)$. To see how project choice and NPV relate, we can rewrite Equation (9) as

$$b_i(w) = r\text{NPV}_i + \left(j'(w) + 1\right) h + \frac{1}{2} j''(w) \frac{h^2}{\text{SN}_i^2}. \tag{10}$$

The marginal benefit of implementing a project depends positively on both the net present value $\text{NPV}_i = \frac{\mu_i - h}{r}$ and the project's signal to noise ratio $\text{SN}_i = \frac{\mu_i}{\sigma_i}$.[25] While a higher NPV implies higher expected payoffs from the project, the signal to noise ratio works though 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 (5), this is equivalent to a lower volatility of w_t, which reduces the likelihood of hitting the boundary at which the agent is fired. The term $\left(j'(w) + 1\right) h$ measures the increase in social value from moving the continuation value away from the liquidation boundary, and is always positive.

Setting $b_i(w) = 0$, we can derive the minimal NPV which the firm requires to implement a project,

$$\underline{\text{NPV}}(\text{SN}, w) = -\frac{1}{r} \left(j'(w) + 1\right) h - \frac{1}{2r} j''(w) \frac{h^2}{\text{SN}^2}.$$

Consequently, all projects with higher than the minimal NPV are implemented, while all others are not. The threshold is a function of the current scaled continuation value, as well as the project's signal to noise ratio. Figure 2 illustrates the non-linear relationship between $\underline{\text{NPV}}$ and SN, and

---

[25]For an alternative interpretation, note that $\psi_i = \frac{\sigma_i}{\mu_i} h$, where $\frac{\sigma_i}{\mu_i}$ can be understood as the project's risk-return ratio.
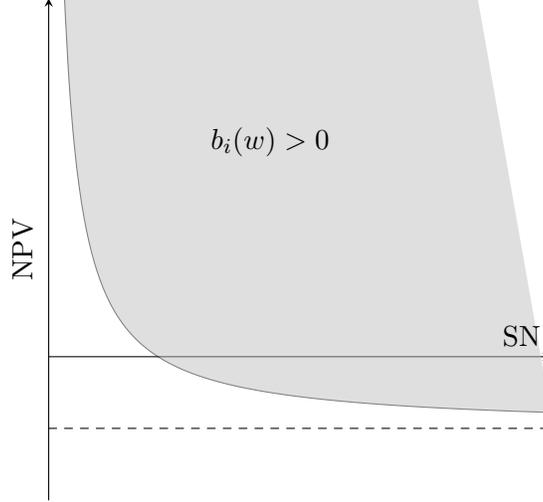
Figure 2: NPV vs. SN boundary

outlines the set of projects which are chosen when $w_t = w$.

In the optimal contract, both over- and underinvestment can occur. Underinvestment is due to the cost of incentives, and occurs whenever a project's NPV is not sufficient to compensate for the increase in termination probability, while overinvestment is caused by the agent's limited liability constraint.[26] The principal cannot demand monetary compensation from the agent after bad performance. Instead, her only means of punishment is to terminate all projects once w_t hits zero, which also precludes her from getting any payments in the future besides the scrap value $l\pi_t$. To avert termination, the principal can allocate more projects to the agent, which increases his disutility of effort, and serves as a less inefficient punishment device. By Equation (5), this increases the average growth rate of the continuation value, and therefore lowers the probability of hitting the firing boundary in the future. This effect may compensate for the negative NPV of a project, and the firm benefits from lowering the future expected termination probability, in exchange for current losses. However, the projects assigned as a punishment cannot be too difficult to incentivize. Otherwise the principal would have to increase the volatility of the agent's continuation value, which makes firing more likely and undermines the intended effect. In accordance with Figure 2, we therefore observe underinvestment in projects which are relatively difficult to incentivize, and have a low signal to noise ratio, and overinvestment in projects where shirking is easy to detect.[27]

---

[26]Standard explanations for overinvestment include 'empire building' as in Jensen (1986), and private benefits to the agent. In my model the effects are entirely driven by the agency friction.

[27]A secondary effect in favor of overinvestment is that the agent is paid only when w_t reaches $\bar{w}$. When the continuation value is low, the probability that w_t hits zero before reaching $\bar{w}$ is high, and thus in expectation, the principal only has to compensate the agent for a fraction of the incurred effort cost. This explains why $j'(w) > -1$ for $w < \bar{w}$.
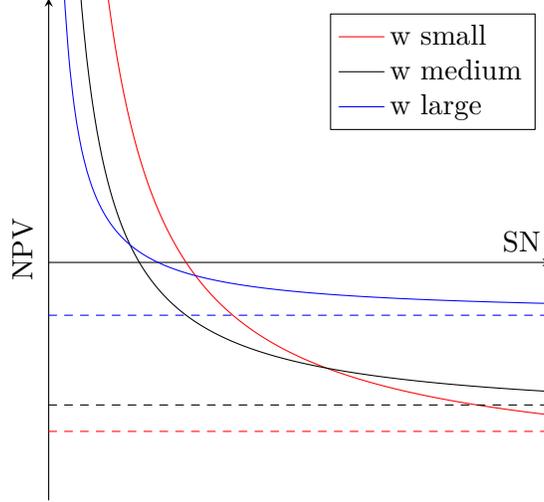
Figure 3: Project Cutoffs as a function of w

4.3 Project Portfolio Dynamics

The choice of projects evolves over time as w_t changes, and each project's marginal benefit function is non-monotonic in w. A higher continuation value decreases the likelihood of firing. This increases the probability that the agent collects payments, and thus the cost of compensating him for effort $j'(w)h$. It also lowers the cost of incentivizing an additional project. For w_t sufficiently high, this implies that the cost of incentives is decreasing in the agent's value. It may be increasing when w_t is low however. In this case, the principal's value is increasing in w_t, and for higher w she stands to lose more if a sequence of bad realizations of output drives the contract into termination, which outweighs the decrease in termination probability.

When the continuation value approaches the payout boundary $\bar{w}$, the cost of incentives becomes negligible and the relative importance of a project's SN ratio vanishes, while the instantaneous cost of compensating for effort is exactly h.[28] Thus, the marginal benefit function at $\bar{w}$ is

$$b_i(\bar{w}) = \mu_i - h,$$

and the project allocation converges to the NPV criterion. Any negative NPV projects with low ψ_i, which have been taken at a lower continuation value are phased out, while any projects with positive NPV and high ψ_i are taken. This result is illustrated in Figure 3. The red function shows the project boundary for small w, while the black and blue functions represent the boundaries for successively larger w. As $w \to \bar{w}$, the break-even NPV line in the figure approaches the x-axis.

When the continuation value grows, and the cost of incentives declines, it is intuitive to think that the firm's optimal portfolio shifts towards high-NPV, low-SN projects, while projects with negative

---

[28]Formally, this follows from the boundary conditions in Proposition 2.

NPV are sorted out. Proposition 3 details in which sense this intuition holds. Whenever the cost of incentives is decreasing in w, the marginal benefit of any project which is sufficiently difficult to incentivize is increasing. Thus, if at a certain cutoff a new project is taken, its risk-return ratio must be relatively high compared to other projects. The opposite holds for projects which are relatively easy to incentivize, which see their marginal benefit decrease with w. If a project is removed, it must be among these.

Proposition 3. *If $j'''(w) \leq 0$, $b_i'(w) < 0$ 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*

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

then $b_i'(w) > 0$.

If the relationship between risk and return is linear for every project, the result in Proposition 3 can be sharpened, and for w sufficiently large, an increase implies that the firm adds projects which have both higher risk and higher return.

Corollary 4. *Suppose the $\sigma_i = K\mu_i$ for all i and $K > 0$. Then for all w, the projects with the highest return μ_i are chosen, and $b_i'(w) > 0$ for all i whenever $j'(w) \leq 0$.*

Finally, I characterize the externalities between projects induced by the agency problem. In the first best, project choice is static and projects are chosen independently of each other. Under moral hazard, taking w as given, $b_i(w)$ is still independent of $b_j(w)$ for $j \neq i$. In this sense, the choice of projects is independent conditional on the scaled continuation value. This is surprising in the light of the literature on static multitasking under moral hazard. For instance, Laux (2001) shows that in a setting with a risk neutral principal and agent, and limited liability, bundling projects increases the principal's payoff, since it allows to extract more of the agent's rents by loosening the limited liability constraint. In my setting, there is no such first order effect of project choice on payoff, since the value function is twice continuously differentiable.[29] Intuitively, there is no hysteresis effect as in the real options literature and the firm can freely switch between projects, which rules out first-order externalities.

There is, however, a second order effect. Choosing a project generates an externality not on the current payoff of other projects, but on the rate at which their value changes with w.

Proposition 5. *At any threshold $\hat{w}$ where a project is added or removed, $j_+'''(\hat{w}) < j_-'''(\hat{w})$.*

---

[29]If I were to introduce a fixed cost with project implementation, the value function would not be C^2 at the cutoffs and hence there would be a first order externality between projects. Implementing another projects causes a discrete jump in j'', and hence in b_i for all projects. See Section 6.

4.4 Project Choice and Investment

My setup nests the framework of DeMarzo et al. (2012), which deals with optimal investment in capital. In this section, I compare how moral hazard distorts the choice of capital and projects.

In the optimal contract, optimal investment $I(w)$ is strictly below the first best level of investment, which is consistent with DeMarzo et al. (2012). The first order condition for investment in Equation (8) implies that

$$\kappa'(I(w)) = j(w) - j'(w)w.$$

Since κ is convex and the right hand side is increasing by the concavity of $j(w)$, $I(w)$ is increasing. At $\bar{w}$, investment is still below the first best level, which can be seen from plugging the boundary conditions into Equation (8), and comparing the resulting expression to the first best payoff

$$j_{fb}(w) = \max_{I} \frac{\sum_i (\mu_i - h)^+}{r - I + \delta} - w,$$

where $(\mu_i - h)^+ = \max\{\mu_i - h, 0\}$. This explains why for all w, $I(w)$ is lower than the optimal investment without agency.

In my setting, a low scaled continuation value implies both less efficient investment and portfolio choice. However, it does not imply underinvestment in both projects and $I(w)$, and it is possible that low I is coupled with overinvestment in projects.

4.5 Output- vs. Project-Based Incentives

In Section 4.2, I have shown that in the optimal contract the principal adjusts the pay-performance sensitivity ψ_{it} separately for each project. In this section, I show that the project selection criterion resembles a hurdle rate when the principal is forced to condition the contract on the total output, $dx_t = \sum_i dx_{it}$, and thus has to bundle all projects. In this case, all projects chosen at a particular time have an NPV which lies above a reference level, which is determined by the cost of incentives for the entire portfolio of chosen projects.

Repeating the argument from Lemma 1 shows that the agent's continuation value satisfies

$$dW_t = \left(\gamma W_t + h \sum_i a_{it}\right) dt - dc_t + \bar{\psi}_t \pi_t \sum_i \sigma_i dB_{it}. \tag{11}$$

Unlike in Equation (5), the vector Brownian Motion B_t enters with a single factor $\bar{\psi}_t$, because the principal is forced to condition on the total output. The agent's IC constraint becomes

$$a_{it} = 1 \Rightarrow \bar{\psi}_t \geq \frac{h}{\mu_i}. \tag{12}$$

At any point in time, the firm chooses a certain portfolio of projects, and by Equation (12), the

project with the lowest NPV determines $\bar{\psi}$, the risk exposure required to incentivize all projects in the portfolio. To see why this is the case, consider a variant of the intuition outlined in Section 3.3. If the agent shirks on project i for a small period of time dt, he saves hdt in effort cost, but at the same time forgoes $\mu_i dt$ in payoff, and by the representation (11), suffers a reduction in continuation value by $-\mu_i \bar{\psi}$. The deviation is not profitable when $\bar{\psi} \geq \frac{h}{\mu_i}$, and since the principal can choose only one $\bar{\psi}$, it must be high enough to incentivize effort on all implemented projects. Since the principal's value function is concave, we have

$$\bar{\psi}_t = \max_{i:a_{it}=1} \frac{h}{\mu_i}.$$

Therefore, given a portfolio and a value for $\bar{\psi}$, each project in the portfolio must satisfy

$$\text{NPV}_i \geq h\frac{\bar{\psi}-1}{\bar{\psi}},$$

otherwise it would not be profitable. Therefore, the portfolio appears to have been chosen using a hurdle rate or minimum NPV criterion. Since the hurdle depends on both the scaled continuation value w as well as the current project selection, it shifts non-monotonically as w changes, but it will converge to the NPV criterion when w approaches $\bar{w}$.

Under project based incentives, the total risk in the contract is given by $\sum_i \psi_i^2 = \sum_i \sigma_i^2 \left(\frac{h}{\mu_i}\right)^2$, while output based incentives raise it to $\sum_i \sigma_i^2 \cdot \max_{i:a_i=1} \left(\frac{h}{\mu_i}\right)$. Hence, for any effort profile a with at least two projects implemented, the agent's risk exposure is strictly higher, as long as $\frac{h}{\mu_i} \neq \frac{h}{\mu_i}$ for some implemented projects i and j. Therefore, conditioning the incentive contract on total output alone cannot be efficient, since it introduces unnecessary risk in the contract.[30]

4.6 Allocation of Funds in Projects

Suppose that each period, the principal can distribute $k\pi$ resources among the projects to increase the effectiveness of managerial effort, where $k \in (0, 1)$. The resource allocation satisfies

$$\sum_i \pi_i \leq k\pi, \tag{13}$$

and is complementary to the agent's effort, so that

$$dx_{it} = \pi_{it}\mu_i a_{it}dt + \sigma_i \pi_t dB_{it}.$$

---

[30]This intuition can be verified by writing out the principal's HJB equation with output based incentives, and comparing it to Equation (8).

Total instantaneous cash flow hence follows

$$dx_{it} = \pi_t \left(\sum_i \tilde{\pi}_{it} \mu_i a_{it} dt + \sigma_i dB_{it} \right),$$

where $\tilde{\pi}_{it} = \frac{\pi_{it}}{\pi_t}$ is the fraction of resources allocated to project i.

In the first best, the firm engages in an extreme form of winner picking, since only the project with the highest NPV receives all the funds. With agency however, project funding not only acts to increase the cash flow, but also serves to change incentives. Given a funding allocation $\tilde{\pi}_i$, the risk exposure required to motivate effort is given by

$$\psi_i \geq h \frac{\sigma_i}{\mu_i} \frac{1}{\tilde{\pi}_i},$$

and project funding serves to lower the required pay-performance sensitivity, because it improves the signal to noise ratio of the project output dx_{it}, which makes shirking easier to detect. In this sense, funding has an added benefit next to improving the efficiency of the agent's effort. The principal's scaled HJB Equation (8) now changes to

$$
\begin{aligned}
rj(w) \;=\; & \sup_{a,I,\tilde{\pi}_i} \sum_i \tilde{\pi}_i \mu_i a_i - \kappa(I) + j'(w) \left((\gamma - I + \delta) w + hn \right) \\
& + j''(w) \frac{1}{2} \sum_i \left(h \frac{\sigma_i}{\mu_i} \frac{1}{\tilde{\pi}_i} \right)^2 a_i + (I - \delta) j(w) - \lambda \left(\sum_i \tilde{\pi}_i - k \right),
\end{aligned}
$$

where λ is the Lagrange multiplier associated with resource constraint (13). Given project i is implemented, its capital allocation solves the FOC[31]

$$\mu_i - \lambda - j''(w) \frac{h^2}{\text{SN}_i^2} \frac{1}{\tilde{\pi}_i^3} = 0,$$

which implies that

$$\tilde{\pi}_i = \left(\frac{-j''(w) h^2}{\text{SN}_i^2 (\lambda - \mu_i)} \right)^{\frac{1}{3}}.$$

Hence, project funding is decreasing in the project's SN ratio, and low risk projects receive lower funding compared to high risk projects, since for high risk projects, the marginal value of lowering the cost of incentives is higher. The link between return and funding remains positive, and higher payoff projects receive relatively more funds.

As the following Lemma shows, project funding increases in w only for projects with sufficiently high NPV. This is intuitive, since as w rises, the costs of exposing the agent to risk decline, and therefore the motive to distort funds away from high payoff and towards high risk projects diminishes as well.

Lemma 6. $\tilde{\pi}_i'(w)$ *is positive whenever* $\mu_i - \lambda > \frac{-\lambda'(w)}{j'''(w)} j''(w) h$ *and negative otherwise. Moreover,*

---

[31]Note that $\lambda > \mu_i$ for all implemented projects, since otherwise $\tilde{\pi}_i \to \infty$, which violates (13).

$$\lambda'(w) \propto -j'''(w).$$

Proof. We have

$$\frac{\partial \tilde{\pi}_i}{\partial w} = \frac{1}{3} \left(\frac{-j''(w) h^2}{\mathrm{SN}_i^2 (\lambda(w) - \mu_i)} \right)^{-\frac{2}{3}} h^2 \mathrm{SN}_i^{-2} \frac{j'''(w) (\mu_i - \lambda(w)) + \lambda'(w) \cdot j''(w)}{(\lambda(w) - \mu_i)^2}$$

which is positive whenever the condition holds. The result on $\lambda'(w)$ can be obtained by plugging the above expression into (13), and solving for $\lambda'(w)$. $\square$

4.7 Project Dynamics with Stealing

In this section, I assume the agent can steal the project output instead of shirking, and there is no investment in capital. The principal chooses between either not executing a project, in which case there is no output to steal, or executing the project and deterring the agent from stealing.[32] This setup simplifies the dynamics of the project selection, because in equilibrium the principal does not have to compensate the agent for effort. Instead, she trades off the project's average payoff against the costs of providing incentives.

The agent receives a private benefit of $\phi \mu_i$ per unit of time when he steals, where $\phi \in (0,1)$. Stealing entails a loss of social value given by $\mu_i (1 - \phi) > 0$. The HJB Equation (8) becomes

$$r j(w) = \max_{a \in \mathcal{A}} \sum_i \mu_i a_i + j'(w) \gamma w + j''(w) \frac{1}{2} \sum_i \psi_i^2 a_i, \tag{14}$$

and each project's marginal benefit function simplifies to

$$b_i(w) = \mu_i + j''(w) \frac{1}{2} \psi_i^2.$$

The dynamics of the project selection criterion are driven by the cost of incentives alone, and are described in the proposition below.

Proposition 7. *When the agent can steal the project output and the principal cannot invest in capital, then there exists a unique threshold $w_3 \geq 0$ such that $j'''(w) < 0$ for all $w < w_3$ and $j'''(w) > 0$ for all $w > w_3$. Left of w_3, no projects are added as w increases, and right of w_3, no projects are removed as w increases.*

The Proposition implies $|j''(w)|$, and with it the cost of providing incentives, is highest for intermediate values of w. The principal's willingness to tolerate additional risk in the contract depends on the likelihood of termination, and the potential gains and losses from high or low output. When

---

[32]Thus, I rule out the case where the principal executes the project, but allows the agent to steal. This may be justified with either legal or reputational concerns of the firm. Moral hazard frameworks with stealing include DeMarzo and Sannikov (2006) and DeMarzo et al. (2012).

w is low, the principal's expected payoff is close to the termination payoff. The principal loses little if a bad outcome occurs, and gains much if a good outcome propels the contract into a region where termination is unlikely. Therefore, while she is still risk averse, her tolerance for additional volatility is relatively high. As w increases, termination becomes less likely, but at the same time the principal's value, and with it the loss from a bad outcome, increases. For $w < w_3$, this effect dominates, and $j''(w)$ is decreasing. If w is sufficiently high, the first effect dominates, and $j''(w)$ increases towards zero. In line with this intuition, when w is low, the principal takes on more projects than when w is at intermediate values, and thus gamble by increasing the volatility in the contract.

The case with stealing and no investment also simplifies the dynamics of the hurdle rate, and the associated project selection when the contract conditions on total output, as in Section 4.5.

Proposition 8. *In the case of output based incentives, let $\underline{\mu}(w)$ be the hurdle rate when the agent's continuation value is w, and $\bar{\psi}(w) = \max_{i:a_i=1} \frac{h}{\mu_i}$ be the pay performance sensitivity associated with the optimal project portfolio at w. When the agent can steal and the principal cannot invest, the result in Proposition 7 holds. For $w < w_3$, $\underline{\mu}(w)$ is increasing on any interval where project choice stays constant, and jumps up when project choice changes, while $\bar{\psi}(w)$ jumps down when project choice changes. The opposite holds for $w > w_3$.*

For $w < w_3$, the cost of incentivizing the project portfolio is increasing, and so is the hurdle rate $\underline{\mu}(w)$, as long as the project selection remains constant. When the selection changes, it must be that a project is removed, and $\bar{\psi}(w)$ jumps downwards, which implies a lower cost of incentives for the new portfolio, and therefore a lower hurdle rate.

5 Implementation

In this section, I discuss how the optimal contract can be implemented. I consider two setups, depending on whether the firm can or cannot issue equity on individual projects. In the first case, the firm holds a cash balance and assigns an equity share in every active project to the agent. The shares are vested, meaning that the agent may lose shares in current projects, or gain shares in new ones depending on his performance.

When shares are issued only on the firm, and not the individual projects, equity is not sufficient to implement the optimal contract. The intuition for this is analogous to Section 4.5, and I show that the hurdle rate allocation from that section is implemented. To achieve the second best, the implementation must feature a measure of the agent's performance in the individual projects, which in my model shares many features of bonus contracts observed in practice.

5.1 Project-Specific Equity

As a benchmark, consider a firm which can issue equity for each individual project, and holds a cash balance to finance its operations. By choosing the appropriate share of equity in each project, the firm can implement the optimal contract.

Specifically, let M_t denote the total stock of cash, which can be allocated among the projects so that $M_t = \sum_i M_{it}$, where M_{it} is the cash stock associated with project i. M_{it} earns a total interest of rM_{it}, where r is the interest rate, has inflows from the project's output dX_{it}, and outflows from the dividends paid on the equity, the share of the cost of investment $\kappa(I_t)$, and the payout to the agent.[33] Thus, M_{it} evolves according to

$$dM_{it} = rM_{it}dt + dX_{it} - d\text{Div}_{it} - dc_{it} - \alpha_i \kappa(I_t)\, dt. \tag{15}$$

I assume that equity holders require a minimal dividend payoff, which satisfies

$$d\text{Div}_{it} = (r - \gamma) M_{it}dt - \alpha_i \kappa(I_t)\, dt. \tag{16}$$

The agent is endowed with a personal account[34] with balance A_t, which pays interest at rate γ, and is used in part to pay the agent. At any point in time, the agent receives an equity share Ψ_{it} in any active project. Whenever a new project is executed, he buys equity in that project at a pre-determined price, and when a project is halted, he sells off the equity. Proceeds from these sales and purchases are deposited in the personal account.[35] Finally, the agent may not access funds inside the account, except for when a dividend dc_t is paid. Formally, the account balance satisfies

$$dA_t = \gamma A_t dt + p_{it} d\Psi_{it} - dc_t^A, \tag{17}$$

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

Proposition 9 below shows that the firm can implement the optimal contract by setting the correct transfer prices and equity shares, and firing the agent when the sum of his stake in the firm and his personal account is sufficiently low.

Proposition 9. *Suppose the firm holds a cash balance M_t, which satisfies $\sum_i M_{it} = M_t$ and Equations (15), (16), and (17) hold. Further, suppose the equity share in each active project is $\Psi_{it} = \frac{\psi_{it}}{\sigma_i}$, the agent is fired whenever $\sum_i \Psi_{it} M_{it} + A_t = 0$, and for each i and t, the transfer price*

---

[33]The terms dc_{it}, M_{it} and α_i are for accounting purposes only. Since the agent must receive a payout dc_t when w hits $\bar{w}$, any assignment of payouts to the projects such that $\sum_i dc_{it} = dc_t$ yields the same result. The same holds for the assignment of investment costs towards projects, which are split according to share α_i, with $\sum_{i:a_i=1} \alpha_i = 1$.

[34]The 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.

[35]I assume that it is possible for the account to have a negative balance.

equals $p_{it} = M_{it}$. *Then the contract from Section [4.5] is implemented and the agent's continuation value satisfies* $W_t = \sum_i \Psi_{it} M_{it} + A_t$.

In the optimal contract, despite the fact that project choice is discrete, the agent's continuation value W_t is a continuous function of time. An intuitive, but wrong, implementation of the contract is to simply strip the agent of his equity share whenever a project is halted without further compensation. This would imply a jump in the continuation value function, since by losing the equity share, the agent loses the claim to future dividends, and distort incentives right before the jump. The transfer prices p_{it} are chosen to exactly offset this change in the agent's value, which is given by M_{it}.[36]

The implementation depends not only on the cash holdings of the firm, as in DeMarzo et al. (2012), but also on the sum of the value of the personal account A_t and the naively calculated value of the managerial share in the firm's cash stock $SC_t = \sum_i \Psi_{it} M_{it}$. Whenever the sum reaches an upper bound, the firm pays dividends, while when the sum reaches zero, the agent is terminated. When dividend payments dc_t are made, it is easy to see that it does not matter whether the money is awarded to the agent from the equity stake, or an equivalent payout from the personal account. Hence the personal account may serve as another way to reward the agent without paying special dividends on equity.

5.2 Firm Level Equity

In reality, firms issue stock based on the entire firm's performance, instead of individual projects. In the following, I describe how to implement the contract when the equity stake can only be conditioned on the total cash holdings M_t, and the agent is restricted to a single equity share $\bar{\Psi}_t$.

The equity stake grants the agent a fraction of ownership in the firm's total cash holdings $M_t = \sum_i M_{it}$, as well as any dividend payments. There is a single dividend process, which satisfies simply

$$d\text{Div}_t = (r - \gamma) M_t dt - \kappa (I_t) dt, \tag{18}$$

while the cash holdings process follows

$$dM_t = rM_t dt + dX_t - dc_t - d\text{Div}_t - \kappa (I_t) dt, \tag{19}$$

and again the personal account is used to escrow proceeds from the agent's equity transactions. Its balance evolves as

$$dA_t = \gamma A_t dt + p_t d\Psi_t - dc_t^A. \tag{20}$$

Inserting the dividend process into Equation (19) shows that holding a single equity stake $\bar{\Psi}_t$

---

[36]In DeMarzo et al. (2012), DeMarzo and Sannikov (2006) and many other works, the agent's effort level and therefore the optimal risk exposure is constant. In these settings, the optimal contract can be implemented via an equity share which is constant over time, and changes in his equity share are not an issue.

conditions the agent's future payouts on the total output of all active projects, $dX_t = \sum_{it} dX_{it}$. Therefore, this implementation will not achieve the second best, but instead implement the hurdle rate allocation from Section (4.5).

Proposition 10. *Suppose that cash balance, dividend payouts and personal account balance follow* (19), (18) *and* (20) *respectively. Further, suppose the equity share is*

$$\Psi_t = \max_{i:a_i=1} \frac{h}{\mu_i},$$

the agent if fired whenever $\Psi_t M_t + A_t$ reaches zero, and the transfer price is $p_t = M_t$. Then the contract from Section 4 is implemented and the agent's continuation value satisfies $W_t = \Psi_t M_t + A_t$.

To see intuitively why only the hurdle rate allocation can be implemented, note that the growth of the firm's cash stock declines by $\pi_t \mu_i dt$ on average if the agent shirks. Given equity share Ψ_t, the growth in the value of the agent's holdings declines by $\Psi_t \pi_t \mu_i dt$, while he saves $\pi_t h dt$ by not exerting effort. Thus, shirking is not optimal if for all active projects

$$\Psi_t \geq \max_{i:a_i=1} \frac{h}{\mu_i},$$

which is the same expression as for the optimal risk exposure in Section 4.5. Thus, the implementation provides the same incentives.

Proposition 10 has an interesting interpretation. In Section 4 I have shown that hurdle rates can arise as a suboptimal outcome. Thus, when the agent's contract consists predominantly of an equity share, the inefficient hurdle rate contract is the only one that is implementable. This suggests that the widespread use of hurdle rates may not be optimal, as for example Berkovitch and Israel (2004) suggest, but instead the result of flawed incentive contracts, which put too much emphasis of firm-level equity to measure performance.

5.3 Firm Level Equity and Bonus Contracts

In order to implement the optimal contract, it is necessary introduce a project-dependent component. In a survey on managerial compensation, 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.[37] Below, I derive an implementation which rationalizes these features.

Formally, the laws of motion for cash stock and dividends are the same as in Equations (19) and (18). The agent receives a constant equity share Ψ_t, and bonus payments P_t, which are linear in

---

[37]Usually, the performance categories and the weights are set by the board of directors. I abstract from agency issues in the relationship between the board and the firm's shareholders. This allows me to interpret the bonus scheme as being set by the principal.

individual project outputs

$$dP_t = \sum_i \beta_{it} dX_{it}, \tag{21}$$

with positive weights β_{it}. Since in the optimal contract, the agent is not paid unless w_t hits an upper bound, the bonus payments flow into the personal account, whose balance evolves as

$$dA_t = \gamma A_t dt + p_t d\Psi_t + dP_t - dc_t^A. \tag{22}$$

Proposition 11. *Suppose that the cash balances, dividend process and personal account balance are given by Equations* (19), (18) *and* (22). *Further, suppose the managerial equity share satisfies*

$$\Psi_t = \min_{i:a_i=1} \Psi_{it},$$

the transfer prices are $p_t = M_t$, *and the weights in the bonus payment process* (21) *are*

$$\beta_{it} = \Psi_{it} - \Psi_t$$

whenever $a_{it} = 1$. *Then the optimal contract is implemented.*

Since the shareholders prefer to fine-tune the agent's risk exposure, the 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. Although the agent does not receive the bonus payment immediately, it raises the balance on his account, and thus brings him closer to the payout boundary, raising his expected continuation value as a response to past performance.

6 Relations to Real Options

6.1 Overview

In the real options framework, an investor chooses when to undertake a project with a fixed cost whose value changes stochastically over time. The project carries an option value of waiting for a higher payoff in the future and is started at a strictly positive NPV in order to compensate for the loss of the option. If starting the project gives access to additional options, such as follow-up projects, these can compensate for the loss of the initial option value, and starting a project with negative NPV may be optimal.[38] Thus, the real options framework can rationalize both over- and underinvestment in projects relative to the NPV criterion.

My model generates both results without relying on irreversibilities or fixed costs, which are crucial

---

[38]This was shown by Baldwin (1982) for product market competition, and Roberts and Weitzman (1981) in the context of sequential investment.

for the real options literature, and without which the option values would necessarily be zero. Instead, over- and underinvestment arise solely because of the distortions from moral hazard. A project with more volatile payoffs needs a higher NPV to be chosen at a particular date not because its option value is higher, but because the principal faces an increased risk of termination from providing the proper incentives. Similarly, a relatively safe project with negative NPV may be chosen not because it grants access to more projects, but serves as a more efficient means to punish the agent.

An inherent difficulty in real options models is characterizing the choice of multiple projects simultaneously. Since each project is associated with fixed costs of starting, and potentially stopping, choosing a particular one affects the marginal benefit of all other projects. I confirm this intuition in Proposition 12, where I show that introducing fixed costs into my framework induces jumps in the second derivative of the principal's value function, and thus in the marginal benefit function (9) whenever a project is started or stopped. When the fixed costs are small, the model in Section 3 serves as an approximation to one which is closer to the real options literature, which is detailed in Proposition 13.

6.2   Project Choice under Fixed Costs

Whenever the principal changes the project allocation from a to a', she incurs an instantaneous cost of $k(a, a')$. The cost function satisfies $k(a, a) = 0$, $k(a, a') > 0$ whenever $a' \neq a$, and

$$k\left(a, a''\right) < k\left(a, a'\right) + k\left(a', a''\right)$$

for all $a, a', a'' \in \mathcal{A}$. The last inequality implies that is is never optimal to switch from a to a'' via some intermediate allocation a', compared to switching directly.[39]

Since the switching cost is strictly positive, it is never optimal to switch an infinite amount of times on a finite interval of time, since the incurred cost would be infinite. Therefore, it is possible to describe the times at which a change in the project allocation occurs as a sequence of stopping times $\{\tau_s\}_{s=1}^{\infty}$, with $\tau_s < \tau$ for all s, and the principal's value function can be written as

$$
\begin{aligned}
J(W, \pi) \;=\; E\Bigg[&\int_0^{\tau} e^{-rt}\left(\pi_t \sum_i \mu_i a_{it} dt - dc_t - \pi_t \kappa\left(I_t\right) dt\right) + \\
&+ e^{-r\tau}\gamma\pi_t - \sum_{s=1}^{\infty} e^{-r\tau_s} k\left(a_{\tau_s-}, a_{\tau_s}\right) \Bigg| \mathcal{F}_0 \Bigg].
\end{aligned}
\tag{23}
$$

Here, $a_{\tau_s-} = \lim_{t\uparrow\tau_s} a_t$ is the project choice right before the switch, and a_{τ_s} is the one right after. The principal faces the same constraints as in problem (7), namely the incentive compatibility

---

[39]This prevents the principal from choosing a project allocation a' for an infinitesimal amount of time. The previous assumptions are satisfied if each project carries strictly positive fixed costs of starting and stopping for example.

conditions and the laws for W_t and π_t.

Introducing fixed costs affects the project selection policy in two important ways. First, project selection is not linearly independent anymore and choosing a project has a first-order effect on the marginal benefit of all projects. Second, the simple HJB equation approach is no longer valid for characterizing the value function, because it is not twice differentiable.

Proposition 12. *Let $\mathcal{L}_{a,I}$ denote the second order differential operator when the investment is I and project portfolio a is chosen, i.e. for any function $\phi \in C^2$*

$$\mathcal{L}_{a,I}\phi\left(w\right) = \left(\left(\gamma - I + \delta\right)w + h\sum_i a_i\right)\phi'\left(w\right) + \phi''\left(w\right)\frac{1}{2}\sum_i \psi_i^2. \tag{24}$$

The solution to problem (23) *is determined by the following system of quasi variational inequalities for all a and w*

$$\min\left\{\min_{I\in\left[0,\bar{I}\right]} rj_a\left(w\right) - \mathcal{L}_{a,I}j_a\left(w\right) - \sum_i \mu_i + \kappa\left(I\right) + \left(I - \delta\right)j_a\left(w\right), \right. \tag{25}$$

$$\left. j_a\left(w\right) - \max_{a'\neq a} j_{a'}\left(w\right) - k\left(a, a'\right)\right\} = 0.$$

For any $a \in \mathcal{A}$, j_a is continuously differentiable for all w, twice continuously differentiable except for a finite number of points, and satisfies the above equation in a viscosity sense.

Further, let $w\left(a, a'\right)$ denote the threshold at which project choice switches from a to a'. Then for any $a \neq a'$ and $w\left(a, a'\right)$ the following conditions hold

$$j_a\left(w\left(a, a'\right)\right) = j_{a'}\left(w\left(a, a'\right)\right) \tag{26}$$
$$j'_a\left(w\left(a, a'\right)\right) = j'_{a'}\left(w\left(a, a'\right)\right),$$

and

$$j''_a\left(w\left(a, a'\right)\right) \geq j''_{a'}\left(w\left(a, a'\right)\right). \tag{27}$$

Due to the fixed costs of setting up and scrapping projects, the value function depends on the current project portfolio a, which is expressed by the notation $j_a(w)$. Equation (25) encodes the optimal choice of a as a function of w. When a particular project selection a is optimal, we have

$$j_a\left(w\right) > \max_{a'\neq a} j_{a'}\left(w\right) - k\left(a, a'\right). \tag{28}$$

By Equation (25), the HJB Equation (8) holds on this region, albeit with project choice fixed at a, and under a different set of boundary conditions, which are given by (26). In general, the inequality (28) does not imply

$$j_a\left(w\right) > \max_{a'\neq a} j_{a'}\left(w\right),$$

so that the firm may not alter a locally suboptimal project choice when the benefit of changing the portfolio does not outweigh the fixed costs.

The principal's value function is in general not twice differentiable at the thresholds $w(a, a')$. By Equation (27), $j_a''(w)$ jumps downward after a change in projects occurs, and the firm becomes more risk averse. Therefore, the marginal benefit function for every project is experiences a downward jump at the threshold, even if the project itself is taken, or not taken, at both sides. Because of these direct spillover effects, the optimal choice of projects has to be determined jointly, instead of using a simple marginal benefit criterion. In particular, if the new portfolio has strictly more projects than the old one, the optimality conditions around the threshold $w(a, a')$ imply

$$\sum_{i:\text{ added}} \left(\mu_i + j_{a'}'(w) h + j_{a'}''(w) \frac{1}{2} \psi_i^2 \right) = \left(j_a''(w) - j_{a'}''(w) \right) \frac{1}{2} \sum_{i:a_i=a_{i'}=1} \psi_i^2$$
$$+ k(a, a'),$$

and the marginal benefit of adding projects must not only exceed the fixed costs, but also compensate for the increase in risk aversion.

It is possible to recover the previous analysis as a limiting case when fixed costs are relatively small.

Proposition 13. *Let $\hat{w}(a, a')$ and $w(a, a')$ denote thresholds at which the optimal project choice changes from a to a' in the case without[40] and with fixed costs, and let j and j_a denote the respective value functions. Then as $\max_{a,a' \in \mathcal{A}} k(a, a') \to 0$ we have for all $a, a' \in \mathcal{A}$, $\hat{w}(a, a') \to w(a, a')$, and for all w, $j_a(w) \to j(w)$ and $\left| j_{a+}''(w) - j_{a-}''(w) \right| \to 0$.*

The proposition specifies in which sense we may take the model in Section 3 as an approximation to a real options model with fixed costs. When these costs are small, the value functions converge towards j, the value function without costs, and the marginal benefit function b_i is approximately continuous. In the limit, the same criterion can be used for determining the project selection portfolio as in Section 4 and that the cutoffs for optimal project choice coincide.

When the fixed costs are small, we can use the marginal benefit functions of Section 4 to select projects. While for some w, the wrong projects are chosen by this procedure, the set of w for which this is the case is small. Equivalently, the amount of time the wrong projects are selected is negligible.

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

---

[40]Note that without fixed costs we have the symmetry $\hat{w}(a, a') = \hat{w}(a', a)$.

projects under moral hazard, as well as the distribution of funds among projects and the persistence 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, projects are selected whenever their NPV is above a cutoff depending on the firm's current cash stock, as well as 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 relative to firm size feature a relatively efficient investment portfolio, comprised only of 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, while the benefit of projects which are relatively safe decreases. The first best project selection schedule is attained whenever the cash stock is large enough. The agent is given riskier projects after a history of sufficiently good performance, while a poorly performing agent will see either the number of projects assigned to him diminished, or be given relatively safe projects as a punishment.

There is a negative externality between projects, which, unlike in the static multitasking literature is of second order only, and affects the rate at which a project's benefit changes with the state variable. If the firm can allocate funds between projects, fund allocation is distorted away from the most profitable to the most risky projects. This inefficiency diminishes as the cash stock becomes sufficiently large. Finally, the contract can be implemented using an equity share, as well as a bonus payment contingent on performance in the individual tasks.

As described in the introduction, the 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 which 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 Scaling of the Value Function

Given the combined stopping and control problem in (7), suppose that the principal's HJB equation satisfies the following HJB equation

$$rJ = \max_{(a,I)\in\mathcal{A}\times[0,\bar{I}]} \pi \sum_i \mu_i a_i - \pi\kappa(I) + J_w\left(\gamma W + \pi h \sum_i a_{it}\right) + \frac{1}{2}J_{ww}\pi^2 \sum_i \psi_i^2 + J_\pi \pi(I-\delta) \quad (29)$$

on some region $\mathcal{C}$ of $\mathbb{R}^2$, with the boundary conditions $J(0,\pi) = l\pi$, and $J_w(W,\pi) = -1$ and $J_{ww}(W,\pi) = 0$ on the boundary of $\mathcal{C}$ for which $W > 0$.

Taking a guess and verify approach, let $w = \frac{W}{\pi}$ and $\pi j(w) = J\left(\frac{W}{\pi}\right)$. Using $J_\pi = j(w) - w \cdot j'(w)$, $J_w = j'(w)$ and $J_{ww} = \frac{1}{\pi}j''(w)$, we can show that the HJB Equation (29) is equivalent to Equation (8), with boundary conditions $j(0) = l$, $j'(\bar{w}) = -1$ and $j''(\bar{w}) = 0$ for some $\bar{w} > 0$ to be determined. Since both laws of motion (2) and (5), as well as the termination value $l\pi$ obey the same scaling, this implies that the control problem in (7) is equivalent to the scaled control problem in w alone.

A.2 Existence and Uniqueness of the Value Function

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

$$rj(w) = \max_{a,I} \sum_i \mu_i a_i - \kappa(I) + j'(w)\left((\gamma - I + \delta)w + hn\right) \quad (30)$$
$$+ j''(w)\frac{1}{2}\sum_i \psi_i^2 a_i + (I - \delta)j(w)$$

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. 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

$$rj(\bar{w}) = \max_I \sum_i (\mu_i - h)^+ - \kappa(I) - (\gamma - I + \delta)\bar{w} + (I - \delta)j(\bar{w}),$$

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

$$j_*(w) = \max_I \frac{\sum_i (\mu_i - h)^+ - \kappa(I) - (\gamma - I + \delta)w}{r - I + \delta}.$$

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 $H : [0, w_{max}] \times \mathbb{R}^2 \to \mathbb{R}$ by

$$H(w, u, p) = -\min_{a, I} \frac{(r - I + \delta) u - \sum_i \mu_i a + \kappa(I) - p((\gamma - I + \delta) w + hn)}{\frac{1}{2} \sum_i \psi_i^2 a_i}.$$

The HJB equation is equivalent to

$$j''(w) + H(w, j(w), j'(w)) = 0. \tag{31}$$

By Berge's Maximum Theorem,[41] $H(w, u, p)$ is jointly continuous in its parameters. This implies that for any starting slope s, the initial value problem (IVP) satisfying (31) with boundary conditions $j(0) = 0$ and $j'(0) = s$ has a twice continuously differentiable solution on the domain $[0, w_{max}]$, and is uniformly continuous with respect to s.[42] 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 $\mathcal{B} \subset \mathbb{R}^2$ as[43]

$$\mathcal{B} = [(0, b), (j_*(w_b), b)] \cup \{(y, w) \in [b, j_*(0)] \times [0, w_b] : y = j_*(w)\}.$$

Finally, let $\bar{w}(s) = \inf\{w : j_s(w) = j_*(w)\}$ be the first point at which $j_s(w)$ hits $j_*(w)$.[44] The following proposition is crucial for establishing uniqueness of the solution, and establishes concavity.

Proposition 14. *Any solution to the IVP* (31) *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 $\bar{w}$,[45] and is given by

$$j'''(w) = \frac{-(\gamma - r) j'(w) - j''(w)((\gamma - I + \delta) w + hn)}{\frac{1}{2} \sum_i \psi_i^2 a_i} > 0,$$

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 $w < \bar{w}(s)$, the point $\hat{w} = \sup\{w < \bar{w} : j''(\hat{w}) \geq 0\}$ exists. If $j'(\hat{w}) \geq 0$,

---

[41] See Aliprantis and Border (2006), Theorem 17.31, p. 570.

[42] See for example Hartman (2002), Chapters 2 and 5.

[43] Because 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.

[44] $\bar{w}(s)$ 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.

[45] This 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.

we have

$$rj\left(\hat{w}\right) \geq \max_{a,I} \sum_i \mu_i a_i - \kappa\left(I\right) + \left(I - \delta\right) j\left(\hat{w}\right)$$

and therefore $a_i = 1$ for all I, and

$$j\left(\hat{w}\right) \geq \frac{\sum_i \mu_i - \kappa\left(I^*\right)}{r - I^* + \delta}$$

where I^* is given by the FOC $\kappa'\left(I^*\right) = j\left(\hat{w}\right)$. The first best value $j_{fb}\left(w\right)$ satisfies

$$j_{fb}\left(w\right) = \frac{\sum_i \left(\mu_i - h\right)^+ - \kappa\left(I_{fb}\right)}{r - I_{fb} + \delta} - w.$$

Since for all fixed I,

$$\frac{\sum_i \left(\mu_i - h\right)^+ - \kappa\left(I\right)}{r - I + \delta} < \frac{\sum_i \mu_i - \kappa\left(I\right)}{r - I + \delta}$$

we have

$$j_{fb}\left(\hat{w}\right) < \frac{\sum_i \left(\mu_i - h\right)^+ - \kappa\left(I_{fb}\right)}{r - I_{fb} + \delta} - w \leq \frac{\sum_i \mu_i - \kappa\left(I^*\right)}{r - I^* + \delta} - w \leq j\left(\hat{w}\right),$$

which implies a contradiction. Therefore, we need $j'\left(\hat{w}\right) < 0$. If $\hat{w} \in \text{int}\mathcal{C}_a$ for some continuation region $\mathcal{C}_a$, then $j'''\left(\hat{w}\right)$ exists and is given by

$$j'''\left(\hat{w}\right) = \frac{-\left(\gamma - r\right) j'\left(\hat{w}\right)}{\frac{1}{2}\sum_i \psi_i^2 a_i} > 0$$

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 $b_i\left(\hat{w}\right) = \mu_i + j'\left(\hat{w}\right) h = 0$. Take some $w > \hat{w}$. Because $j'' < 0$ on $\left(\hat{w}, w\right)$, j' is decreasing on this region, and $b_i\left(w\right) < b_i\left(\hat{w}\right) = 0$. Thus, the project cannot be taken again on $\left(\hat{w}, w\right)$, and the project choice must stay constant on this region. This implies that the right derivative $j'''_+\left(\hat{w}\right)$ exists, and it is positive since j' and j'' are negative. But then again $j''\left(w\right)$ cannot cross zero from above at $\hat{w}$. Therefore, $j''\left(w\right) < 0$ for all $w < \bar{w}$. $\qquad\square$

Lemma 15. *There exists at most one s^* such that $j'_{s^*}\left(\bar{w}\left(s^*\right)\right) = -1$.*

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

$$H\left(\hat{w}, j_s\left(\hat{w}\right), j'_s\left(\hat{w}\right)\right) = H\left(\hat{w}, j_s\left(\hat{w}\right), j'_{s'}\left(\hat{w}\right)\right) > H\left(\hat{w}, j_{s'}\left(w\right), j'_{s'}\left(\hat{w}\right)\right),$$

which implies that $j''_{s'}\left(\hat{w}\right) > j''_s\left(\hat{w}\right)$. Therefore, $j'_{s'}\left(w\right)$ cannot cross $j'_s\left(w\right)$ from above at $\hat{w}$, which is a contradiction. Since $j_*\left(w\right)$ is a strictly decreasing function, this result also implies that $\bar{w}\left(s'\right) < \bar{w}\left(s\right)$ 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 14, j_s is strictly concave, and therefore $j_s(\bar{w}(s)) < j_s(\bar{w}(s')) < -1$. □

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 16. *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 (31) 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 $\mathcal{B}_\varepsilon$, 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 14, 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 14 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 $s \leq \underline{s}$, $\mathcal{S}(s) \leq -1$. □

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

Proposition 17. *Let G_0 be the payoff from an arbitrary contract (a, c, τ) which is incentive compatible and for which the law for the agent's continuation value (5) has a unique strong solution. 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 (5). I restrict attention to contracts for which (5) has a unique strong solution. Define the realized payoff from using the contract until time $t \leq \tau$ as

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

where dX_{is} is the output process under the effort implemented in the contract. By Itô's Lemma and the transformation used in Section A.1,

$$
\begin{aligned}
dG_t &= e^{-rt} \pi_t \left(\sum_i \mu_{it} dt + \sum_i \sigma_i dB_{it} - \kappa (I_t) \, dt - \frac{dc_t}{\pi_t} \right) \\
&\quad + e^{-rt} \pi_t \left(j'(w_t) \left(\gamma w_t dt + h n_t dt - \frac{dc_t}{\pi_t} + \sum_i \psi_{it} dB_{it} \right) + j''(w_t) \frac{1}{2} \sum_i \psi_{it}^2 dt \right) \\
&\quad + e^{-rt} \pi_t (I_t - \delta) \left(j(w_t) - j'(w_t) w_t \right) dt - r \pi_t j(w_t) \, dt.
\end{aligned}
$$

By the construction of the HJB Equation (30), for any incentive compatible contract

$$-(r - I + \delta) j(w) + \sum_i \mu_i - \kappa (I) + j'(w) ((\gamma - I + \delta) w + hn) + j''(w) \frac{1}{2} \sum_i \psi_i^2 \leq 0, \quad (32)$$

and for any consumption payout policy, $-\pi_t (1 + j'(w)) dc_t \leq 0$ on $(0, \bar{w})$. Since $j'(w)$ is bounded, and π_t grows at a rate strictly less that r, this term is square integrable, and G_t is a supermartingale. For all finite t, the principal's profit is

$$
\begin{aligned}
E[G_\tau] &= E[G_{t \wedge \tau}] \\
&\quad + E\left[\mathbb{1}_{\{t < \tau\}} e^{-rt} E\left[\int_t^\tau e^{-rs} \left(\sum_i dX_{is} - \pi_s \kappa (I_s) \, dt - dc_s \right) + e^{-r\tau} \pi_\tau l - J(W_t, \pi_t) \right] \right].
\end{aligned}
$$

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

$$\pi_t j_{fb}(0) - \pi_t w_t - J(W_t, \pi_t) \leq \pi_t (j_{fb}(0) - l).$$

By the supermartingale property of G_τ, $E[G_{t \wedge \tau}] \leq G_0 = \pi_0 j(w_0)$, and the profit satisfies the following bound

$$E[G_\tau] \leq G_0 + e^{-rt} \pi_t (j_{fb}(0) - l). \quad (33)$$

Since for all contracts, $r > I_t - \delta$ holds uniformly in t, the transversality condition

$$\lim_{t \to \infty} e^{-rt} \pi_t = 0$$

holds.

If the optimal contract generates a strong solution for the agent's continuation value process W_t, Equation (33) holds with equality as $t \to \infty$, because Equation (32) holds with equality. If the contract does not generate a strong solution, because the project choice and therefore the volatility of the continuation value does not have bounded variation, there always exists an ε-optimal strategy which does. The loss in payoff in the HJB equation between the optimal project selection a and an arbitrary a' is

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

which is bounded on $(0, \bar{w})$ by a constant $\bar{L}$, because $j'\left(w\right)$ and $j''\left(w\right)$ are continuous. Taking a grid for w of step size ε, define a Markov control $\hat{a}\left(w\right)$ which takes the value of the optimal project choice of the nearest point in the grid. For any w on the interior of a region where action a is optimal, the loss $L\left(w, a\left(w\right), \hat{a}\left(w\right)\right)$ eventually becomes zero for some $\varepsilon > 0$. If w is on the boundary of $\mathcal{C}_a$ and $\mathcal{C}_{a'}$ and both have a nonempty interior, take w without loss of generality to be the midpoint of the grid $\left[w - \frac{1}{2}\varepsilon, w + \frac{1}{2}\varepsilon\right]$. The expected loss in this interval starting from w is bounded below

$$\lambda\left(w\right) = \bar{L}E\left[\int_0^{\tau_\varepsilon} e^{-rt} dt\right]. \tag{34}$$

Here, the expectation operator is taken under the action $\hat{a}\left(w\right)$, and τ_ε denotes the first exit time of w_t from the interval. As $\varepsilon \to 0$, this expression converges to zero.[46]

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

Then, repeating the previous verification argument with $\hat{a}\left(w\right)$ implies that for ε sufficiently small, the loss relative to the optimal project selection from the HJB Equation (30) is bounded by $\bar{L}\frac{\varepsilon}{r}$. $\quad\square$

A.3 Properties of the Value function

A.3.1 Proof of Proposition 3

Lemma 18. *Whenever $j'\left(w\right) < 0$, $j'''\left(w\right) > 0$ if it exists.*

---

[46]Precisely, since $\hat{a}\left(w\right)$ is constant, w_t satisfies $dw_t = \left(\gamma w_t + h\hat{n}\right) dt + \sum_i \psi_i \hat{a}_i dB_{it}$, and $\lambda\left(w\right)$ solves the differential equation $r\lambda\left(w\right) = \bar{L} + \lambda'\left(w\right)\left(\gamma w + h\hat{n}\right) + \lambda''\left(w\right)\frac{1}{2}\hat{\psi}^2$ subject to the boundary conditions $\lambda\left(w - \frac{1}{2}\varepsilon\right) = \lambda\left(w + \frac{1}{2}\varepsilon\right) = 0$, 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 $\lambda'\left(w\right)$ which is uniform in ε. This implies that $|\lambda\left(w\right)| \leq M\varepsilon$ on $\left[w - \frac{1}{2}\varepsilon, w + \frac{1}{2}\varepsilon\right]$ for some $M > 0$ and all $\varepsilon > 0$, which is the desired result.

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

$$j'''(w) = \frac{-(\gamma - r) j'(w) - j''(w) ((\gamma - I + \delta) w + hn)}{\sum_i \psi_i^2 a_i},$$

and it is positive because j is concave. □

Proposition 19. *For any w, and all projects j if $b_j'(w)$ exists, it is strictly positive if $j'(w) < 0$ and*

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

Whenever $j'''(w) < 0$, $b_j'(w) < 0$ for all j and if $j'''(w) > 0$, then $b_j'(w) > 0$ if $\psi_i^2 > -\frac{2 j''(w)}{j'''(w)}$.

Proof. We have $b_j'(w) = j''(w) h + j'''(w) \frac{1}{2} \psi_j^2$. Using the expression of $j'''(w)$, we have

$$b_j'(w) = j''(w) h - \frac{1}{2} \psi_j^2 \frac{(\gamma - r) j'(w) + j''(w) ((\gamma - I + \delta) w + hn)}{\frac{1}{2} \sum_i \psi_i^2 a_i}$$

which is positive whenever

$$h \frac{1}{2} \sum_i \psi_i^2 a_i - \frac{1}{2} \psi_j^2 ((\gamma - I + \delta) w + hn) < \frac{(\gamma - r) j'(w)}{j''(w)}.$$

Since the right hand side is positive, a sufficient condition is $\psi_j^2 \geq \frac{1}{n} \sum_i \psi_i^2 a_i$. Note that a sharp condition is $\psi_j^2 > \frac{-2 j''(w)}{j'''(w)}$, which is equivalent to

$$\psi_j^2 > \frac{2}{((\gamma - I + \delta) w + hn)} \left(h \frac{1}{2} \sum_i \psi_i^2 a_i - \frac{(\gamma - r) j'(w)}{j''(w)} \right).$$

□

A.3.2 Proof of Proposition 5

Suppose that at $\hat{w}$, an additional project, indexed by $n + 1$, is added. Then, it must be the case that $b_{n+1}(w)$ crosses zero at $\hat{w}$ from below. The derivatives of $b_{n+1}(w)$ left and right of $\hat{w}$ are

$$
\begin{aligned}
j_-'''(\hat{w}) &= \frac{-(\gamma - r) j'(\hat{w}) - j''(\hat{w}) ((\gamma - I + \delta) \hat{w} + hn)}{\frac{1}{2} \sum_i \psi_i^2 a_i} \\
j_+'''(\hat{w}) &= \frac{-(\gamma - r) j'(\hat{w}) - j''(\hat{w}) ((\gamma - I + \delta) \hat{w} + h(n+1))}{\frac{1}{2} \sum_i \psi_i^2 a_i + \frac{1}{2} \psi_{n+1}^2},
\end{aligned}
$$

and

$$j_+'''(\hat{w}) - j_-'''(\hat{w}) = -\frac{1}{\frac{1}{2} \sum_i \psi_i^2 a_i} b_{n+1,+}'(\hat{w}).$$

36

Here $b'_{n+1,+}(\hat{w})$ denotes the right derivative of $b_{n+1}(w)$ at $\hat{w}$. Since $b_{n+1}(w)$ must cross zero, the derivative is positive, and thus, $j'''_+(\hat{w}) < j'''_-(\hat{w})$. An analogous calculation in the case where a project is removed yields

$$j'''_+(\hat{w}) - j'''_-(\hat{w}) = \frac{1}{\frac{1}{2}\sum_i \psi_i^2 a_i} b'_{n+1,+}(\hat{w}) < 0.$$

The jump in $j'''(w)$ can be equivalently expressed as

$$j'''_+(\hat{w}) - j'''_-(\hat{w}) = -\frac{1}{\frac{1}{2}\sum_i \psi_i^2 a_i + \frac{1}{2}\psi_{n+1}^2} b'_{n+1,-}(\hat{w})$$

in case a project is added, and

$$j'''_+(\hat{w}) - j'''_-(\hat{w}) = \frac{1}{\frac{1}{2}\sum_i \psi_i^2 a_i - \frac{1}{2}\psi_{n+1}^2} b'_{n+1,-}(\hat{w})$$

in case it is removed. Since the expressions must equal in the respective cases, we have

$$b'_{n+1,-}(\hat{w}) = \left(1 + \frac{\psi_{n+1}^2}{\sum_i \psi_i^2 a_i}\right) b'_{n+1,+}(\hat{w})$$

if a project is added, and

$$b'_{n+1,-}(\hat{w}) = \left(1 - \frac{\psi_{n+1}^2}{\sum_i \psi_i^2 a_i}\right) b'_{n+1,+}(\hat{w})$$

if it is removed.

A.3.3 Proof of Proposition 7

The principal's HJB equation is given by

$$rj(w) = \max_a \sum_i \mu_i + j'(w)\gamma w + j''(w)\frac{1}{2}\sum_i \psi_i^2,$$

with boundary conditions $j(0) = l$, $j'(\bar{w}) = -1$ and $j''(\bar{w}) = 0$. Since this is a special case of Equation (30), it can be verified that the analysis in Section A.2 carries over, so that $j(w)$ has the same qualitative features.

On any region where project choice is constant, we have

$$j'''(w) = \frac{-(\gamma - r)j'(w) - j''(w)\gamma w}{\frac{1}{2}\sum_i \psi_i^2 a_i}.$$

For w sufficiently close to zero, $j'''(w)$ is negative independently of the current project choice at any point where it exists, since $j'(0) > 0$. The marginal benefit function of each project is given

by $b_i(w) = \mu_i + j''(w)\frac{1}{2}\psi_i^2$, and its derivative satisfies

$$b_i'(w) = j'''(w)\frac{1}{2}\psi_i^2.$$

Therefore $b_i'(w) < 0$ whenever $j'''(w) < 0$. If $j'(w) < 0$, then $j'''(w) > 0$, independently of a and $b_i'(w) > 0$. Whenever another project is added, we have

$$j_+'''(w) = j_-'''(w)\frac{\sum_i \psi_i^2 a_i}{\sum_i \psi_i^2 a_i + \psi_{i'}^2} > 0,$$

so that $j'''(w)$ can never jump below zero in this case. To see that once $j'''(w)$ is positive, it cannot fall below zero on any region where the project choice is constant, note that if $j'''(w) = 0$, its derivative is given by

$$j^{(4)}(w) = \frac{-(2\gamma - r)j''(w)}{\frac{1}{2}\sum_i \psi_i^2 a_i} > 0,$$

and hence $j'''(w)$ cannot cross zero from above. Therefore, the cutoff w_3 is unique.

Since $b_i'(w)$ for all i and $w < w_3$, project choice can only change finitely many times, and each time a project is removed. Similarly, when $w > w_3$, $b_i'(w) > 0$ and whenever project choice changes, a project is added.

A.4 Proofs on Implementation

A.4.1 Proof of Proposition 9

First consider the process $SC_t \equiv \sum_i \Psi_{it} M_{it}$, which can be interpreted as the share of the firm's cash holdings the agent has at time t. M_{it} follows the process in Equation (15).

I define a process $\{A_t\}_{t \geq 0}$ such that the decomposition

$$W_t = SC_t + A_t \tag{35}$$

holds, and interpret A_t as the current balance in the agent's personal account. At the optimal contract, the agent's continuation value W_t is a diffusion and its path is a continuous function of time. By Equation (15), SC_t is continuous in t whenever no change is made in the project portfolio, and exhibits a jump of $\Psi_{it} M_{it}$ when project i is added, and $-\Psi_{it} M_{it}$ when project i is dropped.[47] Therefore, the process $\{A_t\}$ has to be chosen to compensate for jumps in SC_t, otherwise the representation (35) cannot hold. This is achieved by choosing $\{A_t\}$ such that $dA_t = \gamma A_t dt$ whenever SC_t is continuous, $dA_t = \gamma A_t dt - \Psi_{it} M_{it}$ whenever project i is added, and $dA_t = \gamma A_t dt + \Psi_{it} M_{it}$ whenever it is dropped. The interpretation of the simultaneous changes in SC_t and A_t is that the agent either sells or buys shares at the price of M_{it} per unit, which corresponds exactly to their

---

[47]Thus, the process SC_t is a semimartingale.

naive value in terms of the firm's cash holdings.

To verify that the agent's continuation value satisfies (35), by Itô's Lemma for semimartingales[48] we have

$$dW_t = \sum_i \left(\Psi_{it} dM_{it} + d\Psi_{it} M_{it} \right) + dA_t.$$

Further, since W_t is a diffusion at the optimal contract we can use the HJB equation approach to get

$$
\begin{aligned}
\gamma W_t &= \sup_{a_i} \sum_i \Psi_i \left(rM_{it} + \pi_t \mu_i a_i - \alpha_i \kappa\left(I_t\right) - (r - \gamma) M_{it} + \alpha_i \kappa\left(I_t\right) - \frac{dc}{dt} \right) \quad (36) \\
&= + \sum_i \Psi_i \frac{dc_{it}}{dt} - \pi_t h \sum_i a_i + \sum_i \frac{d\Psi_i}{dt} M_{it} + \frac{dA_t}{dt} \\
&= \sup_{a_i} \sum_i \Psi_i \left(rM_{it} + \pi_t \mu_i a_i - \alpha_i \kappa\left(I_t\right) - (r - \gamma) M_{it} + \alpha_i \kappa\left(I_t\right) - \frac{dc}{dt} \right) \\
&\quad + \sum_i \Psi_i \frac{dc_{it}}{dt} - \pi_t h \sum_i a_i + \gamma A_t
\end{aligned}
$$

From the above equation we see that if we let the optimal equity share satisfy $\Psi_{it} \equiv \frac{\psi_{it}}{\sigma_i} = \frac{h}{\mu_i}$, then

$$\gamma W_t = \gamma \left(\sum_i \Psi_i M_{it} + A_t \right)$$

and the optimal contract is implemented.

The proof of Proposition 10 proceeds analogously and is omitted.

A.4.2 Proof of Proposition 11

The agent's continuation utility is assumed to satisfy $W_t = \Psi_t M_t + A_t$, and the personal account balance now satisfies

$$dA_t = \gamma A_t dt + dP_t + SC_t - SC_{t-}$$

where $SC_{t-} = \lim_{s \uparrow t} SC_t$.

---

[48]See He et al. (1992), p. 245, Th. 9.35

Analogously to the previous proof, the agent's HJB equation satisfies

$$\gamma W_t = \sup_a \Psi_t \left(\sum_i \left(\gamma M_{it} + \pi_t \mu_i a_i - \frac{dc_{it}}{dt} \right) \right) + \Psi_t \frac{dc_t}{dt} - \pi_t h \sum_i a_i +$$

$$+ \pi_t \sum_i \left(\Psi_{it} - \Psi_t \right) \mu_i a_{it} + dA_t + \frac{d\Psi_t}{dt} M_t$$

$$= \Psi_t \gamma M_t + \sup_a \left(\pi_t \sum_i \left(\Psi_t \mu_i - h + \left(\Psi_{it} - \Psi_t \right) \mu_i \right) a_i \right) + \gamma A_t$$

$$= \Psi_t \gamma M_t + \gamma A_t$$

and again the contract is implemented.

A.5 Proofs on the Model with Fixed Costs

A.5.1 Proof of Proposition 12

The proof consists of establishing the existence of a viscosity solution of Equation (25), subject to the appropriate boundary conditions. With the result in Proposition 23, a version of the verification argument in Proposition 17 implies that this solution equals the principal's optimal value function. Throughout, I use the viscosity solution approach.[49] In the following,

$$\mathcal{C}_a = \left\{ w : j_a\left(w\right) > \max_{a' \neq a} j_{a'}\left(w\right) - k\left(a, a'\right) \right\}$$

denotes the continuation region at which a certain project selection is optimal, while

$$\mathcal{S}_{a,a'} = \left\{ w : j_a\left(w\right) = j_{a'}\left(w\right) - k\left(a, a'\right) \right\}$$

denotes the region where the optimal project choice switches from a to a'. I use the abbreviation in (24) as well as

$$f_{a,I} = \sum_i \mu_i a_i - \kappa\left(I\right).$$

The continuous function $j_a\left(w\right)$ is a viscosity subsolution to (25), if for any $\phi \in C^2$ and w which is a local maximum of $j_a - \phi$,

$$\min \left\{ r j_a\left(w\right) - \max_I \mathcal{L}_{a,I} \phi\left(w\right) + f_{a,I} + \left(I - \delta\right) j_a\left(w\right), j_a\left(w\right) - \max_{a' \neq a} j_{a'}\left(w\right) - k\left(a, a'\right) \right\} \leq 0,$$

while it is a viscosity supersolution if for any $\phi \in C^2$ and w which is a local minimum of $j_a - \phi$,

$$\min \left\{ r j_a\left(w\right) - \max_I \mathcal{L}_{a,I} \phi\left(w\right) + f_{a,I} + \left(I - \delta\right) j_a\left(w\right), j_a\left(w\right) - \max_{a' \neq a} j_{a'}\left(w\right) - k\left(a, a'\right) \right\} \geq 0.$$

---

[49]Standard concepts and further references can be found in Crandall et al. (1992) and Fleming and Soner (2006).

Finally, j_a is a viscosity solution if it is both a sub- and supersolution.

Proposition 20. *There exists a family of viscosity solutions $\{j_a(w)\}_{a \in \mathcal{A}}$ to the system of Equations (25).*

Proof. By Ishii (1989), Proposition 3.4, for any particular a, Equation (25) has a continuous viscosity solution with boundary conditions $j_a(w_1) = j_1$ and $j_a(w_2) = j_2$, if there exist continuous super- and subsolutions $j_a^S(w)$ and $j_a^s(w)$ satisfying the boundary conditions, and $j_a^s \leq j_a^S$. The HJB equation with constant projects is given by

$$r j_a(w) = \max_I \mathcal{L}_{a,I} j_a(w) + f_{a,I} + (I - \delta) j_a(w), \tag{37}$$

and has a twice differentiable solution subject to the boundary conditions above.[50] Since the solution ignores possible switching, it is possible that $j_a(w) < \max_{a' \neq a} j_{a'}(w) - k(a, a')$. Therefore, it is a continuous viscosity subsolution to Equation (25). To obtain the supersolution, define $j_a^S(w) = j_a^s(w) + K(w)$ for some C^2 function $K(w)$ with $K(w_1) = K(w_2) = 0$, which satisfies

$$K(w) \geq \max_I \frac{1}{r - I + \delta} \mathcal{L}_{a,I} K(w),$$

so that j_a^S satisfies the HJB equation with the inequality

$$r j_a^S(w) \geq \max_I \mathcal{L}_{a,I} j_a^S(w) + f_{a,I} + (I - \delta) j_a^S(w).$$

As long as $j_1 > \max_{a' \neq a} j_{a'}(w) - k(a, a')$, and the same holds for j_2, a $K(w)$ can be found such that $j_a^S(w)$ is greater than the switching payoff, and $j_a^S(w)$ is a supersolution.

Define

$$j_{a*}(w) = \max_I \frac{\sum_i (\mu_i - h) a_i - (\gamma - I + \delta) w}{r - I + \delta},$$

and observe that $j_{a*}(w)$ and $j_{a'*}(w)$ can never cross for $a' \neq a$. For each a, whenever $j_{a*}(w) > \max_{a' \neq a} j_{a'*}(w) - k(a, a')$, define the boundary conditions

$$j_a(\bar{w}_a) = j_{a*}(\bar{w}_a)$$

and

$$j_a'(\bar{w}_a) = -1,$$

while if the opposite inequality holds, set

$$j_a(\bar{w}_a) = \max_{a' \neq a} j_{a'}(\bar{w}_a) - k(a, a').$$

---

[50]See Strulovici and Szydlowski (2012).

41

For all a, the boundary condition at zero is

$$j_a(0) = l.$$

Then, in the first case $\bar{w}_a \in \mathcal{C}_a$, and there exists a region such that Equation (37) holds. Then, a variant of the argument in Section A.2 establishes that there exists a solution which matches the boundary conditions, while in the second case, Ishii's result can be applied directly. $\qquad\square$

Lemma 21. *There exists an $\varepsilon > 0$ such that for any optimal contract, each continuation region $\mathcal{C}_a$ on which the project selection stays constant contains a subinterval of length greater than ε.*

Proof. This is immediate from the continuity of $j_a(w)$ for all $a \in \mathcal{A}$, and the switching cost being strictly positive. $\qquad\square$

Lemma 22. *For any project selection a and interval $[w_1, w_2]$ which satisfies $[w_1, w_2] \subset \mathcal{C}_a$ or $[w_1, w_2] \subset \mathcal{S}_{a,a'}$ for some $a' \neq a$, $j_a(w)$ is twice continuously differentiable.*

Proof. By the previous Lemma, a nonempty interval $[w_1, w_2] \subset \mathcal{C}_a$ is guaranteed to exist for all a. In this case, let a_1 be the optimal choice just left of w_1, and a_2 the optimal choice right of w_2.[51] Because project choice is constant on $[w_1, w_2]$, the HJB equation on this region reduces to (37), subject to the boundary conditions $j_a(w_1) = j_{a_1}(w_1) - k(a, a_1)$ and $j_a(w_2) = j_{a_2}(w_2) - k(a, a_2)$. This problem has a twice continuously differentiable solution,[52] and a standard verification argument implies that for $w \in [w_1, w_2]$, this solution equals the optimal value function.

If $[w_1, w_2] \subset \mathcal{S}_{a,a'}$, since $\mathcal{S}_{a,a'} \subset \mathcal{C}_{a'}$,[53] Equation (37) with a replaced by a' holds for $j_{a'}$, and therefore

$$r j_a(w) = \max_I \mathcal{L}_{a',I} j_a(w) + f_{a'I}, + (I - \delta) j_a(w) - (r - I + \delta) k(a, a')$$

holds on $[w_1, w_2]$, with boundary conditions $j_a(w_1) = j_{a'}(w_1) - k(a, a')$ and $j_a(w_2) = j_{a'}(w_2) - k(a, a')$. Then the same argument implies that $j_a(w)$ is twice continuously differentiable. $\qquad\square$

There only remains to verify that the solution to the variational inequalities (25) is well behaved at the thresholds where project choice changes.

Proposition 23. *Let $\hat{w}$ be an optimal cutoff at which project selection changes. Then j_a is continuously differentiable at $\hat{w}$.*

Proof. Suppose that a is optimal on the left, and a' is optimal on the right of $\hat{w}$. By the previous Lemma, j_a is twice differentiable on a neighborhood left and right of $\hat{w}$ respectively. It therefore

---

[51]It is not necessary that $a_1 \neq a$ or $a_2 \neq a$.

[52]See Strulovici and Szydlowski (2012).

[53]This follows from the fact that $k(a, a') < k(a, a') + k(a', a'')$. Then, it is never optimal to switch from a to a', and then immediately to a'', since it incurs a higher cost that switching directly to a''.

has left and right limits at $\hat{w}$, which may not be equal. We have $j_a(w) = j_{a'}(w) - k(a, a')$ for $w > \hat{w}$ and

$$j_a(w) \geq j_{a'}(w) - k(a, a') \tag{38}$$

for $w < \hat{w}$.

Extending the region on which a is chosen to $[w_1, \hat{w}']$ for some $\hat{w}' > \hat{w}$ preserves continuity of $j_a(w)$ on that region. Therefore, if inequality (38) is strict at $\hat{w}$, we can extend the region to the right while preserving the inequality. This yields a higher payoff than switching at $\hat{w}$, so that the threshold cannot be optimal, which is a contradiction. Thus j_a is continuous at $\hat{w}$.

By the previous Lemma, the left and right derivatives $j'_{a,-}(\hat{w})$ and $j'_{a,+}(\hat{w})$ exist. Assume that $j'_{a',-}(\hat{w}) < j'_{a,+}(\hat{w})$,[54] take any number $x \in (j'_{a,-}(\hat{w}), j'_{a,+}(\hat{w}))$, and define

$$\phi_\varepsilon(w) = j_a(\hat{w}) + x(w - \hat{w}) + \frac{1}{2\varepsilon}(w - \hat{w})^2$$

for some $\varepsilon > 0$. We have $\phi_\varepsilon \in C^2$, and $\phi_\varepsilon(\hat{w}) = j_a(\hat{w})$. Since j_a is continuous at $\hat{w}$, this point is a local minimum of $j_a - \phi_\varepsilon$ for all ε. By the viscosity supersolution property, we have for some $w < \hat{w}$,

$$(r - I + \delta) \cdot j_a(\hat{w}) - x \cdot ((\gamma - I + \delta)\hat{w} + hn) - \frac{1}{2\varepsilon}\sum_i \psi_i^2 a_i \geq 0,$$

and sending $\varepsilon \to 0$ implies a contradiction. If $j'_{a,-}(\hat{w}) > j'_{a,+}(\hat{w})$, take some $x \in (j'_{a,+}(\hat{w}), j'_{a,-}(\hat{w}))$, and consider the function

$$\phi_\varepsilon(w) = rj_a(\hat{w}) + x(w - w_0) - \frac{1}{2\varepsilon}(w - \hat{w})^2.$$

$\hat{w}$ is a local maximum of $j_a(w) - \phi_\varepsilon(w)$ for all ε, and by the subsolution property we have for some $w > \hat{w}$

$$(r - I + \delta) \cdot j_a(\hat{w}) - x \cdot ((\gamma - I + \delta)\hat{w} + hn') + \frac{1}{2\varepsilon}\sum_i \psi_i^2 a'_i + k(a, a') \leq 0.$$

Here, $n' = \sum_i a'_i$ and the equation holds because $j_a(w) = j_{a'}(w) - k(a, a')$ for $w > \hat{w}$, and $w \in \mathcal{C}_{a'}$. Letting $\varepsilon \to 0$ again yields the contradiction, and hence, $j'_+(w) = j'_-(w)$. $\qquad \square$

The following Lemma establishes inequality (27).

Lemma 24. *Consider a cutoff $\hat{w}$ at which the optimal choice changes from a to a' as w crosses $\hat{w}$ from below. Then $j''_{a'}(\hat{w}) \leq j''_a(\hat{w})$.*

Proof. For any $w \in \mathcal{C}_{a'}$ we have

$$rj_{a'}(w) \geq \max_I \mathcal{L}_{a,I} j_{a'}(w) + f_{a,I} + (I - \delta)j_{a'}(w) - (r - I + \delta)k(a, a'), \tag{39}$$

---

[54]By the equality $j_a(w) = j_{a'}(w) - k(a, a')$ for $w \geq \hat{w}$ and the fact that $\hat{w} \in \text{int}\mathcal{C}_{a'}$, we actually have $j'_{a,+}(\hat{w}) = j'_{a'}(\hat{w})$.

which can be established using $\phi = j_{a'} - k(a, a')$ as a test function and verifying that w is a local minimum of $j_a - \phi$. The equation then follows from the supersolution property.

The proof of Proposition 23 shows that $\hat{w} \in \mathcal{C}_{a'}$, and therefore Equation (39) holds on some interval $[\hat{w}, w'] \subset \mathcal{C}_{a'}$, and can be rewritten as

$$
\begin{aligned}
j_{a'}''(w) \quad \leq \quad & \frac{1}{\frac{1}{2}\sum_i \psi_i^2 a_i} \Bigg((r - I + \delta)\left(j_{a'}(w) + k(a, a')\right) - \sum_i \mu_i a_i + \kappa(I) \\
& - j_{a'}'(w)\left((\gamma - I + \delta)w + h\sum_i a_i \right) \Bigg),
\end{aligned}
$$

where i is the optimal investment choice. The right hand side of this equation is continuous at $\hat{w}$, and in particular the left and right limits are equal. Since the HJB equation with a as optimal project choice must hold left of $\hat{w}$, we have

$$
\begin{aligned}
j_{a'}''(\hat{w}) \quad \leq \quad & \lim_{w \uparrow \hat{w}} \frac{1}{\frac{1}{2}\sum_i \psi_i^2 a_i} \Bigg((r - I + \delta)j_a(w) - \sum_i \mu_i a_i + \kappa(I) \\
& - j_a'(w)\left((\gamma - I + \delta)w + h\sum_i a_i \right) \Bigg) \\
= \quad & j_a''(\hat{w}),
\end{aligned}
$$

which is the relation to be proven. $\qquad\square$

A.5.2 Proof of Proposition 13

Lemma 25. *Let* $k = \max_{a,a'} k(a, a')$ *and denote with* $j_{a,k}(w)$ *the solution to the system of Equations* (12) *given action* a *and maximal switching cost* k. *For all* w *and* a, $j_{a,k}(w)$ *converges uniformly to a function* $\tilde{j}(w)$ *as* $k \to 0$.

Proof. For any $a \neq a'$ and w, $j_{a,k}(w) \geq j_{a',k}(w) - k(a, a')$ so that $k(a, a') \geq j_{a'}(w) - j_a(w)$. Since the same inequality holds with a and a' reversed, we have

$$
k(a', a) \leq j_{a',k}(w) - j_{a,k}(w) \leq k(a, a').
$$

Since the bounds are uniform in w, $j_{a,k}$ and $j_{a',k}$ converge uniformly to some function $\tilde{j}$ as $k \to 0$. $\quad\square$

The remainder of the proof consists of showing that the limit $\tilde{j}$ equals j, the unique solution to the HJB Equation (8). I index continuation $\mathcal{C}_a^k$ and switching regions $\mathcal{S}_{a,a'}^k$ with k, since they depend on the switching cost.

Lemma 26. *Let $\bar{w}_{a,k} \in \mathcal{C}_a^k$ be a family of thresholds such that $j'_{a,k}(\bar{w}_{a,k}) = -1$ and $j''_{a,k}(\bar{w}_{a,k}) = 0$. Then for k sufficiently small, $\bar{w}_{a,k} \in \mathcal{S}^k_{a,a_{fb}}$ for all a, and there exists a unique finite threshold $\bar{w}^0$ such that $\bar{w}_{a,k} \to \bar{w}^0$.*

Proof. At $\bar{w}_{a,k}$, the HJB Equation (37) implies

$$r j_{a,k}(\bar{w}_{a,k}) = \max_I \sum_i (\mu_i - h) a_i - \kappa(I) - (\gamma - I + \delta) \bar{w}_{a,k} + (I - \delta) j_{a,k}(\bar{w}_{a,k}).$$

Let $i_{a,k}$ be the optimal choice of investment, define

$$j_*(w) = \max_I \frac{\sum_i (\mu_i - h)^+ - \kappa(I) - (\gamma - I + \delta) w}{r - I + \delta}$$

and suppose that $\bar{w}_{a,k} \in \mathcal{C}_a^k$ as $k \to 0$. Then,

$$
\begin{aligned}
j_{a,k}(\bar{w}_{a,k}) - j_*(\bar{w}_{a,k}) &\leq j_{a,k}(\bar{w}_{a,k}) - \frac{\sum_i (\mu_i - h)^+ - \kappa(I_{a,k}) - (\gamma - I_{a,k} + \delta) w}{r - I_{a,k} + \delta} \\
&\leq \frac{\sum_i (\mu_i - h) a_i - \sum_i (\mu_i - h)^+}{r - I_{a,k} + \delta}
\end{aligned}
$$

where $I_{a,k}$ is the optimal investment given project choice a and cost k at $\bar{w}_{a,k}$. Whenever $a \neq a_{fb}$, this expression is uniformly bounded below zero for all k, since $I_{a,k} \leq \bar{I} < r + \delta$. But the definition of $\mathcal{C}_a^k$ implies that

$$k(a, a_{fb}) \geq j_*(\bar{w}_{a,k}) - j_{a,k}(\bar{w}_{a,k}),$$

which yields a contradiction as $k \to 0$. Therefore, for all a and k sufficiently small, $\bar{w}_{a,k} \in \mathcal{S}^k_{a,a_{fb}}$, and $\mathcal{S}^k_{a,a_{fb}} \subset \mathcal{C}^k_{a_{fb}}$, we have $\bar{w}_{a,k} = \bar{w}_{a_{fb},k}$. Let w' be the point closes to $\bar{w}_{a_{fb},k}$ at which it is optimal to switch from a_{fb} to some other project allocation a'. Then $[w', \bar{w}_{a_{fb},k}] \subset \mathcal{C}^k_{a_{fb}}$, and $j_{a_{fb},k}(w)$ satisfies Equation (37) with $a = a_{fb}$, and boundary conditions $j_{a_{fb},k}(w') = j_{a',k}(w') - k(a_{fb}, a')$, $j'_{a_{fb},k}(\bar{w}_{a_{fb},k}) = -1$ and $j''_{a_{fb},k}(\bar{w}_{a_{fb},k}) = 0$. The argument in Lemma 15 can be applied to $j_{a_{fb},k}(w)$ on this region, and the threshold $\bar{w}_{a_{fb},k}$ is unique. Therefore there exists some $\bar{w}^0$ such that $\bar{w}_{a,k} \to \bar{w}^0$ for all a and k. The value $\bar{w}^0$ must be finite, since for all k, $j_{a,k}(w)$ is bounded above $-w - k(a, a_0)$, where a_0 is the allocation where no projects are executed, and $j'_*(w) < -1$. $\square$

By the existence of a finite limit $\bar{w}^0$, we can restrict attention to analyzing the viscosity solutions to (25) on some finite interval $[0, w_{max}]$ with $w_{max} > \bar{w}^0$ as $k \to 0$.

Lemma 27. *$\tilde{j}(w)$ is continuously differentiable on $[0, \bar{w}^0]$.*

Proof. If $w \in \mathcal{C}_a^k$, $j''_{a,k}(w)$ is bounded, since $j_{a,k}$ satisfies Equation (37). If $w \in \mathcal{S}^k_{a,a'}$, we have

$$r j_{a,k}(w) = \max_I \mathcal{L}_{a',i} j_{a,k}(w) + f_{a',i} - r k(a, a'),$$

45

because $\mathcal{S}_{a,a'}^k \subset \mathcal{C}_{a'}^k$ and $j_{a,k}(w) = j_{a',k}(w) - rk(a,a')$, so that $j_{a,k}''(w)$ is again bounded. Since the number of switching points is finite, $j_{a,k}'(w)$ is differentiable almost everywhere with bounded derivative, and therefore Lipschitz for all w. The family of functions $\left\{ j_{a,k}'(w) \right\}_k$ is equicontinuous on $\left[0, \bar{w}^0\right]$, and by the Arzelà-Ascoli Theorem, there exists a subsequence which converges to a continuously differentiable function. Repeating the argument for all a implies that $\tilde{j}$ is C^1, since the limits have to be equal. $\qquad\square$

Proposition 28. *$\tilde{j}$ is a viscosity solution to the HJB Equation 8.*

Proof. The proof is a variant of the argument in Dolcetta and Evans (1984), who study optimal switching in a deterministic setting. Take a C^2 function ϕ such that some $w_0 \in [0, w_{max}]$ is a strict local minimum of $\tilde{j}(w) - \phi(w)$. By the uniform convergence of $j_{a,k}(w)$, for each $a \in \mathcal{A}$ there exists a point $w_{a,k}$ which is a local minimum of $j_{a,k}(w) - \phi(w)$, and which converges to w_0 as $k \to 0$. Since $j_{a,k}$ is a viscosity solution to (25), this implies the inequality

$$\min \left\{ r j_{a,k}(w_{a,k}) - \max_I f_{a,I} + \mathcal{L}_{a,I}\phi(w_{a,k}) + (I - \delta) j_{a,k}(w_{a,k}) , \right.$$
$$\left. j_{a,k}(w_{a,k}) - \max_{a' \neq a} j_{a',k}(w_{a,k}) - k(a,a') \right\} \geq 0.$$

By definition of $j_{a,k}$, $j_{a,k}(w_{a,k}) \geq \max_{a' \neq a} j_{a',k}(w_{a,k}) - k(a,a')$, which implies

$$r j_{a,k}(w_{a,k}) \geq \max_I \left\{ f_{a,I} + \mathcal{L}_{a,I}\phi(w_{a,k}) + (I - \delta) j_{a,k}(w_{a,k}) \right\}.$$

As $k \to 0$, this yields

$$r\tilde{j}(w_0) \geq \max_I f_{a,I} + \mathcal{L}_{a,I}\phi(w_0) + (I - \delta)\tilde{j}(w_0),$$

and repeating the argument for all $a \in \mathcal{A}$ implies

$$r\tilde{j}(w_0) \geq \max_{a,I} f_{a,I} + \mathcal{L}_{a,I}\phi(w_0) + (I - \delta)\tilde{j}(w_0).$$

Thus, $\tilde{j}$ is a viscosity supersolution of Equation (8).

Taking another C^2 function ϕ such that w_0 is a strict local maximum of $\tilde{j}(w) - \phi(w)$, for k sufficiently small there exist points $w_{a,k}$ which are strict local maxima of $j_{a,k}(w) - \phi(w)$. To save notation define $j_k(w) = j_{a(k),k}(w)$ and $w_k = w_{a(k),k}$ where $a(k)$ is given by

$$j_{a(k),k}\left(w_{a(k),k}\right) - \phi\left(w_{a(k),k}\right) = \max_{a \in \mathcal{A}} \left\{ j_{a,k}(w_{a,k}) - \phi(w_{a,k}) \right\}.$$

Since for each a, $w_{a,k}$ is a strict local maximum of $j_{a,k}(w) - \phi(w)$, this definition ensures that

$$j_k(w_k) - \phi(w_k) \geq \max_{a \neq a(k)} j_{a,k}(w_{a,k}) - \phi(w_{a,k}) \geq \max_{a \neq a(k)} j_{a,k}(w_k) - \phi(w_k),$$

46

and therefore

$$j_k(w_k) > j_{a,k}(w_k) - k(a(k), a) \tag{40}$$

for all $a \neq a(k)$. By the viscosity subsolution property of j_k, we have

$$\min \left\{ rj_k(w_k) - \max_I \left\{ f_{a(k),I} + \mathcal{L}_{a(k),I} \phi(w_k) + (I - \delta) j_k(w_k) \right\}, \right.$$
$$\left. j_k(w_k) - \max_{a' \neq a} j_k(w_k) - k(a, a') \right\} \leq 0,$$

which combined with Equation (40) implies that

$$rj_k(w_k) \leq \max_I \left\{ f_{a(k),I} + \mathcal{L}_{a(k),I} \phi(w_k) + (I - \delta) j_k(w_k) \right\}.$$

Since $\mathcal{A}$ is finite, there exists a subsequence $\{k_n\} \to 0$ such that $a(k_n)$ converges to some $a_0 \in \mathcal{A}$. Since for all $a \in \mathcal{A}$, $w_{a,k} \to w_0$, we have

$$r\tilde{j}(w_0) \leq \max_I \left\{ f_{a_0,I} + \mathcal{L}_{a_0,I} \phi(w_0) + (I - \delta) \tilde{j}(w_0) \right\} \leq \max_{a,I} \left\{ f_{a,I} + \mathcal{L}_{a,I} \phi(w_0) + (I - \delta) \tilde{j}(w_0) \right\}.$$

Hence, $\tilde{j}$ is a viscosity subsolution of the HJB Equation (8). Lemmas 26 and 27 imply that $\tilde{j}$ satisfies the boundary conditions $\tilde{j}(\bar{w}) = j_*(\bar{w})$ and $\tilde{j}'(\bar{w}) = -1$. Since for each k, $j_{a,k}(0) = l$, the condition $\tilde{j}(0) = l$ holds as well. $\square$

Proposition 29. *The viscosity solution of the HJB Equation* (8) *with boundary conditions* $\tilde{j}(0) = l$, $\tilde{j}(\bar{w}) = j_*(\bar{w})$ *and* $\tilde{j}'(\bar{w}) = -1$ *is unique and equals* j, *the solution of the HJB equation.*

Proof. The HJB equation satisfies all assumptions of Theorem 3.3 in Ishii (1989), which implies that it has a unique viscosity solution which satisfies the boundary conditions $\tilde{j}(0) = l$ and $\tilde{j}(\hat{w}) = j_*(\hat{w})$ for arbitrary $\hat{w}$. Choosing $\hat{w} = \bar{w}$, so that $\tilde{j}'(\bar{w}) = -1$, which is feasible by the previous Proposition, implies that the HJB equation has a unique viscosity solution. The argument in Section A.2 shows hat the equation also has a unique twice differentiable solution with the same boundary conditions. Since any such solution is a fortiori also a viscosity solution,[55] uniqueness implies that $\tilde{j} = j$ for all $w \in [0, \bar{w}]$. $\square$

Since $\tilde{j}$ is twice continuously differentiable, the jumps in the second derivatives of the value function with switching cost must converge to zero, which can be used to show that the regions of w where a certain project portfolio is chosen at maximal cost k must converge to the region from the HJB equation, without switching cost, as k goes to zero.

---

[55]This is verified by taking the twice differentiable solution j instead of the test functions ϕ in the definition of the viscosity super- and subsolution properties, which are trivially satisfied in this case.

Corollary 30. *As $k \to 0$, for all $a \in \mathcal{A}$ and $w \in [0, \bar{w}]$,*

$$\left| j''_{a,k}(w) - j''_{a',k}(w) \right| \to 0,$$

and

$$\mathcal{C}^k_a \to \mathcal{C}_a$$

in the Hausdorff distance.

References

Aliprantis, C. D. and K. C. Border (2006). *Infinite dimensional analysis: a hitchhiker's guide.* Springer Verlag.

Baldwin, C. Y. (1982). Optimal sequential investment when capital is not readily reversible. *Journal of Finance 37*(3), 763–782.

Berkovitch, E. and R. Israel (2004). Why the NPV criterion does not maximize NPV. *Review of Financial Studies 17*(1), 239.

Biais, B., T. Mariotti, J. C. Rochet, and S. Villeneuve (2010). Large risks, limited liability, and dynamic moral hazard. *Econometrica 78*(1), 73–118.

Bond, P. and A. Gomes (2009). Multitask principal-agent problems: Optimal contracts, fragility, and effort misallocation. *Journal of Economic Theory 144*(1), 175–211.

Cadenillas, A., J. Cvitanic, and F. Zapatero (2007). Optimal risk-sharing with effort and project choice. *Journal of Economic Theory 133*(1), 403–440.

Crandall, M. G., H. Ishii, and P. L. Lions (1992). User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc 27*(1), 1–67.

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

DeMarzo, P. M. and Y. Sannikov (2006). Optimal security design and dynamic capital structure in a continuous-time agency model. *Journal of Finance 61*(6), 2681–2724.

Dixit, A. K., R. S. Pindyck, and G. A. Davis (1994). *Investment under uncertainty*, Volume 15. Princeton University Press Princeton, NJ.

Dolcetta, I. C. and L. C. Evans (1984). Optimal switching for ordinary differential equations. *SIAM journal on control and optimization 22*(1), 143–161.

Fleming, W. H. and H. M. Soner (2006). *Controlled Markov processes and viscosity solutions*, Volume 25. Springer Verlag.

Fong, K. (2007). Evaluating skilled experts: Optimal scoring rules for surgeons.

Friedman, A. (1975). *Stochastic differential equations and applications.* New York: Academic Press.

Graham, J. R. and C. R. Harvey (2001). The theory and practice of corporate finance: evidence from the field* 1. *Journal of financial economics 60*(2-3), 187–243.

Grenadier, S. R. and A. Malenko (2010). Real options signaling games with applications to corporate finance.

Grenadier, S. R. and N. Wang (2005). Investment timing, agency, and information. *Journal of Financial Economics 75*(3), 493–533.

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. *Society for Industrial and Applied Mathematics, Philadelphia.*

Hartman-Glaser, B., T. Piskorski, and A. Tchistyi (2010). Optimal securitization with moral hazard.

He, S., C. Wang, and J. Yan (1992). *Semimartingale theory and stochastic calculus.* Taylor & Francis US.

He, Z. (2009). Optimal executive compensation when firm size follows geometric brownian motion. *Review of Financial Studies 22*(2), 859.

He, Z. and W. Xiong (2008). Multi-market delegated asset management. *NBER Working Paper.*

Holmstrom, 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.

Ishii, H. (1989). On uniqueness and existence of viscosity solutions of fully nonlinear second-order elliptic PDE's. *Communications on Pure and Applied Mathematics 42*(1), 15—45.

Jensen, M. C. (1986). Agency costs of free cash flow, corporate finance, and takeovers. *The American Economic Review 76*(2), 323–329.

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.

Laux, C. (2001). Limited-liability and incentive contracting with multiple projects. *The RAND Journal of Economics 32*(3), 514–526.

Makarov, I. and G. Plantin (2010). Rewarding trading skills without inducing gambling.

Malenko, A. (2011). Optimal design of internal capital markets.

Manso, G. (2006). Motivating innovation.

Miquel-Florensa, J. (2007). Optimal incentives in dynamic multiple project contracts. *Working Papers.*

Morellec, E. and N. Schürhoff (2010). Corporate investment and financing under asymmetric information. *Journal of Financial Economics.*

Murphy, K. J. (1999). Executive compensation. *Handbook of labor economics 3*, 2485–2563.

Ou-Yang, H. (2003). Optimal contracts in a continuous-time delegated portfolio management problem. *Review of Financial Studies 16*(1), 173.

Rauh, J. D. (2009). Risk shifting versus risk management: Investment policy in corporate pension plans. *Review of Financial Studies 22*(7), 2687.

Roberts, K. and M. L. Weitzman (1981). Funding criteria for research, development, and exploration projects. *Econometrica: Journal of the Econometric Society 49*(5), 1261–1288.

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. *The Review of Economic Studies 54*(4), 599–617.

Strulovici, B. and M. Szydlowski (2012). On the smoothness of value functions and the existence of optimal strategies. Discussion Paper 1542, Northwestern University.