Incentives, Project Choice and Dynamic Multitasking*

Martin Szydlowski†

October 2, 2012

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 signal to noise ratio (SN). 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 signal to noise 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 rule resembles a hurdle rate criterion.

1 Introduction

The standard paradigm for firm investment posits a continuous investment decision. Firms choose investment as a means to regulate their capital stock, which, except for adjustment costs, is perfectly scalable. While for certain firms, this framework may be reasonable, for

*I am indebted to Jeffrey Ely, Michael Fishman, and Bruno Strulovici for their guidance. This paper has benefited from helpful discussions with Bruno Biais, Janice Eberly, Arvind Krishnamurthy, Alessandro Pavan, Mark Satterthwaite, and Yuliy Sannikov. I would also like to thank seminar participants at the Canadian Economic Theory Conference, the North American Summer Meeting of the Econometric Society, the Stony Brook Game Theory Festival, and the Transatlantic Doctoral Conference.

†Northwestern University, Department of Economics, 302 Arthur Andersen Hall, 2001 Sheridan Road Evanston, IL 60208-2600. Email: martin.szydlowski@u.northwestern.edu

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

others it is not. When a firm is the host of many disparate activities, we can instead think of the firm consisting of a portfolio of potential projects. When these projects are executed, they generate risky cash flows, which depend on each project's individual characteristics. Inherently, this choice of projects is discrete, i.e. the firm can either engage in a project at any given point in time or not. Hence, instead of having the possibility of continuously adjusting future capital according to its expectations, the firm faces a much more difficult problem - to determine the optimal portfolio of projects at any given time.

The goal of my paper is to characterize the optimal project selection policy when the firm relies on a manager to execute its projects. The manager has the opportunity to shirk, and the cost of executing projects is entirely driven by the agency friction. Even though both the firm and the manager are risk neutral in my setup, project choice along the optimal path of the dynamic contract is not determined by the NPV criterion alone, as would be the case in the first best, but instead by a project-specific markup over NPV. This markup is a function of the manager's promised continuation utility, as well as the project's signal to noise ratio (SN), which measures how difficult it is to discern whether the manager has been putting in effort.

There is both over- and underinvestment relative to the NPV criterion. While underinvestment is driven by the cost of incentives, which induce the firm to forgo positive NPV projects due to their risk, overinvestment is caused by its inability to punish the manager in the presence of a limited liability constraint.

Similar to DeMarzo et al. (2010), I identify the manager's continuation value with the firm's cash balances and the value of the manager's personal account.[1] Thus, holding the account value constant, when the firm's cash holdings are small, project choice becomes more distorted relative to the first best and the firm forgoes positive NPV projects if they have a low signal to noise ratio, or equivalently, high risk. When the cash holdings are sufficiently high, first-best efficiency in project execution is achieved.

The cutoffs for project selection are a function of the entire history of past projects, output, and managerial effort, as well as the noise embedded in the project cash flows. In particular, when the firm's cash holdings are sufficiently high, low NPV but high SN projects are gradually phased out in favor of high NPV projects, and firms with higher cash balances can afford a more risky and more lucrative project portfolio. This dynamic is entirely driven by the cost of incentives, which are in turn embodied in the projects' SN ratios. Since managers have limited liability, the firm is liquidated when the agent's continuation value reaches its lower boundary. Incentivizing projects necessitates an increased managerial risk exposure,

[1] The full derivation is given in Section 5.

which in turn increases the liquidation probability.

This finding is opposite to the standard risk shifting result found in Jensen and Meckling (1976), where the possibility of liquidation leads firms to take on excessive risk. Recently, in a study of pension funds, Rauh (2009) finds that firms with weak credit ratings, which may be interpreted as a proxy for default, allocate more resources towards safer investments, while financially sound firms do the opposite, which is in line with my predictions.

I study several extensions of this framework. First, when the shareholders can allocate internal funds between projects, the fund allocation is distorted away from the projects with the highest NPVs and towards projects with low signal to noise ratios. Intuitively, a low SN implies a relatively high cost of exposing the agent to risk. When internal funding can increase the effectiveness of managerial effort, the associated cost of incentives is lowered. Thus, firms with low cash holdings will distort their allocation of funds, while firms with high cash holdings will engage in winner picking, and allocate all their funds to the most profitable project.

Further, my model nests DeMarzo et al. (2010) as a special case. Therefore, I can study the relationship between project choice, aggregate firm investment, and growth. The agency friction has a similar effect on project choice and aggregate investment, and both are either comparatively efficient or inefficient, depending on whether the agent's continuation value is large relative to firm size.

As in any study involving multitasking, the question whether the optimal incentive scheme can be made contingent on total firm performance alone, as opposed to individual project payoffs, is important. In my setting, unless all projects have the same characteristics, incentives based on total output will fail to implement the second best allocation. Instead, they make the underinvestment problem more severe, and induce a fundamental change to the project selection policy. When restricted to output based incentives, optimal project choice resembles a hurdle rate. In particular, at any point in time, the NPV of each chosen project will be above the same threshold, which in turn is a function of the project with the lowest NPV in the portfolio. This hurdle rate allocation is not efficient, since by conditioning the manager's incentive contract on total output alone, the firm is unable to fine-tune the risk exposure of the manager towards individual projects, and hence the contract carries excessive risk.

Consequently, my model suggests that hurdle rates, which are widely observed in practice, are not the outcomes of an optimal contract. Instead, they arise when the firm is unable to condition the contract on individual projects, or unable to find incentive schemes which condition on this information.

Indeed, I show in Section 5 that if the managerial incentive contract is limited to equity, the hurdle rate allocation arises as the optimal contract. To implement the second best contract, it is necessary to introduce payments to the manager contingent on the individual project's performance. I show that these payments can be interpreted as bonus payments, and hence the optimal contract can be implemented via an equity stake and boni.

Since the number of projects will change over time, the risk exposure for the manager, and therefore the optimal equity share will not be static, as in DeMarzo and Sannikov (2006) or DeMarzo et al. (2010), and it may be necessary to adjust the manager's equity share when the project selection changes. However, these equity transfers may also distort incentives, since if the manager expects to be stripped of shares in the future, he may be less likely to put in effort. To counter this effect, I show that the implementation features the manager buying and selling equity at ex-ante agreed on transfer prices, which exactly offset the adversarial incentive effect from equity purchases and sales. Proceeds from these transactions as well as the manager's bonus payments will be escrowed in a personal account, from which the manager will be paid once his performance history is sufficiently good.

The contract I derive shares many features with contracts found in reality. As Murphy (1999) documents, the vast majority of CEO incentive contracts consist of a wage, which is normalized to zero in my setup, equity holdings and bonus payments. The latter are set by shareholders ex ante, and provide payments to the manager depending on his performance in different categories. The total bonus payment is then a linear function of the boni of the individual categories. The results in Murphy (1999) suggest that while the equity stake is needed to provide the manager with a baseline level of incentives, bonus payments are used to fine-tune the incentive plan, and make sure that the manager puts in the desired amount of effort into the different projects. I show in Section 5 how this intuition translates into my setup.

When projects choice is a binary decision, and associated with fixed costs, we are in a real options framework. My results share many features with the real options literature, although they are the consequence of very different mechanisms. I explain the connection in Section 6 in detail, and I also provide a discussion on how my model can be viewed as an approximation to a model which is more in line with the real options literature.

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. The implementation outlined in the paragraph above is derived in Section 5. Finally, Section 6 provides a discussion how my setup relates to the real options framework, which also deals with binary investment decisions, 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 manager's effort 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. (2010), who study a firm's investment decisions based on a continuous time moral hazard framework, and find that the agency friction opens a wedge between average and marginal Q, and induces underinvestment. My framework features multiple projects with varying risk-return profiles, as well as a continuous investment variable, and nests the model in DeMarzo et al. (2010). This allows me to study both the choice of the optimal project portfolio and its interaction with investment. Instead of a constant equity share, the optimal contract in my setup requires the manager to purchase and sell equity at transfer prices set by the firm, and the implementation with a constant equity share held by the manager ceases to be optimal.

The problem of multitasking has received significant attention since the seminal article of Holmstrom and Milgrom (1991).[2] 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

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

will either exert effort in all mortgages or none.

Finally, my model is related to the literature on optimal investment. The real options literature[3] 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 manager 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 optimal contract can be implemented by allocating a fixed budget to the manager. 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] See e.g. Dixit et al. (1994) for a comprehensive overview.

3 Model Setup

3.1 Projects and Investment Technology

Consider a long term contract between the manager of a firm, the agent, and shareholders who act as the principal. The firm is equipped with a portfolio of N potential projects, indexed by $i \in \{1, ..., N\}$. Time t is continuous and infinite, and each project i is characterized by its risk-return profile (μ_i, σ_i). Projects contribute to the firm's cash flow, and their output depends on the agent's effort decision as well as a Brownian noise component B_{it}. The Brownian motions are mutually independent, i.e. $B_{it} \perp B_{js}$ for all $i \neq j$ and times $t, s \geq 0$. The agent's effort decision in project i at time t is denoted as a_{it}. To capture the discrete nature of project implementation, a_{it} is binary, i.e. $a_{it} \in \{0, 1\}$. I denote with a_t the vector of choices for each project $(a_{it})_{i=1}^{N}$, and with $\mathcal{A} = \{0, 1\}^N$ the set of possible project allocations. When $a_i = 1$, the cumulative project cash flow x_{it} evolves according to a diffusion process with drift μ_i and volatility σ_i, and the instantaneous cash flow dx_{it} is given by

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

σ_i can be understood either as a measure of the riskiness of project i, while μ_i measures the payoff of the agent's effort in the task. As shall be seen, the inverse signal to noise ratio, $\frac{1}{SN_i} = \frac{\mu_i}{\sigma_i}$, is proportional to the cost of exposing the manager to the necessary risk to motivate effort, and measures how difficult it is to infer the agent's effort from observing the outcome path x_{it}. The event $a_{it} = 1$ shall be interpreted as project i being *assigned* to the manager, or alternatively project i being *implemented* at time t.

Total cash flow depends on both firm size π_t and the total output from all implemented projects. Shareholders receive a total cash flow of $\pi_t \sum_i dx_{it}$, and decide how much to either pay out as dividend, leave to the agent as consumption, or use for investment. Given the investment decision I_t, firm size is deterministic and follows the law of motion

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

The principal bears the adjustment cost in investment, $\pi\kappa\left(\frac{I}{\pi}\right)$, which is deducted from the firm's cash flows. Letting $i = \frac{I}{\pi}$ be the ratio of investment to current firm size, I assume that the investment cost $\kappa(i)$ is increasing and convex with $\kappa(0) = 0$.

3.2 Utility Functions and the Contract Space

The vector of project-specific Brownian noise $B_t := (B_{1t}, ..., B_{Nt})$ is defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with filtration $\mathcal{F}_t$, which satisfies the usual conditions.[4] Each project's output can be fully observed by the principal and contracted upon, while effort is unobservable. The agent has limited liability so that for all t, $W_t \geq 0$. When $W_t = 0$, the agent is fired, the firm is shut down and the principal receives a salvage value of $\gamma\pi$.[5] Let $\tau = \inf\{t \geq 0 : \ w_t = 0\}$ denote the random time at which shutdown occurs. The agent is remunerated with a cumulative consumption process $c = \{c_t \in \mathbb{R}_+ : 0 \leq t \leq \tau\}$ with non-negative increments,[6] and the principal prescribes a vector-valued effort process[7] $a = \left\{(a_{it})_{i=1}^N \in \{0,1\} : 0 \leq t \leq \tau\right\}$. Effort and consumption are both progressively measurable with respect to $\mathcal{F}_t$.

The agent seeks to maximize his discounted lifetime utility W_0, which is given by

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

Here, dc_t is the increment in the consumption process the manager receives, and the manager's utility consists of the expected discounted value of consumption payments minus his expected effort cost, $h\sum_i a_{it}$, which is linear and symmetric in each task effort a_{it}, and increases with firm size.

Under any incentive compatible contract, the agent's and the principal's information sets are identical. The principal is risk neutral, and seeks to maximize the following expression

$$J_0 = E\left[\int_0^\tau e^{-rt}\left(\left(\pi_t \sum_{i=1}^N \mu_i a_{it} - \pi_t \kappa\left(i_t\right)\right) dt - dc_t\right) |W_0, \pi_0\right]. \tag{4}$$

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

[5] Note that this regime is not renegotiation proof as the value function of the principal is upward sloping in the agent's continuation value when the latter is close to zero. Then, renegotiation would be beneficial since instead of shutting the firm down, the principal would benefit from giving the agent a higher continuation value. While allowing renegotiation diminishes the principal's ability to incentivize the agent, it does not alter the qualitative properties of the contract.

[6] Precisely, we have $c_t - \lim_{t' \uparrow t} c_{t'} \geq 0$ almost surely for all t.

[7] 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 for the contract. 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.

Here, the first term are consists of the expected discounted cash flows from the projects, conditional on the manager's effort. The principal bears the expenses for investment in the firm $\pi_t \kappa(i_t)$, as well as for consumption payments to the manager.

I assume that principal and agent have different discount factors, and that the principal is more patient, i.e. $r < \gamma$. As noted in DeMarzo and Sannikov (2006), this assumption prevents the principal from postponing the agent's consumption forever. Finally, I impose an upper bound on relative investment, $i \leq \bar{i} < r + \delta$ to ensure that the principal's value function is bounded.[8]

3.3 Incentive Compatibility

Given effort and consumption schedules (a, c) and firing time τ, the manager's continuation utility at time t is given by

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

which is the analog of expression (3) at time $t > 0$.

Then, the martingale representation theorem[9] implies that the agent's continuation value W_t follows a diffusion process in the multidimensional 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 of the realizations of this uncertainty.

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*

$$dW_t = \left(\gamma W_t + \pi_t h \sum_i a_{it}\right) dt - dc_t + \pi_t \sum_{i=1}^N \psi_{it} dB_{it}. \tag{5}$$

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

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

[8] If $i = r + \delta$ the shareholders' value of the firm might be infinite, since the firm would grow at a fast enough rate to negate any discounting.

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

whenever $a_{it} = 1$.

I interpret ψ_{it} as the sensitivity of the agent's continuation value with respect to the risk of project i. Since the principal can control both consumption and effort, she is able to implicitly determine how much the agent's continuation utility responds to uncertainty, and we can think of ψ_{it} being chosen directly by the principal. 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 the agent's decision, consider a deviation for a short period of time dt, during which the manager 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 of the noise process falls short by $-\pi_t \frac{\mu_i}{\sigma_i} dt$, and by the representation, the manager 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 is important for providing incentives. When the manager shirks, he affects the principal's beliefs about the realization of dB_{it}. When the project is relatively safe, and the ratio $SN_i = \frac{\mu_i}{\sigma_i}$ is large, observing a shortfall in output by $\mu_i dt$ while the manager 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 .[10]

Analogously to the discrete time contracting literature, equation (5) should be interpreted as a promise keeping constraint. Given a continuation value W_t, higher consumption dc_t implies that ceteris paribus, the manager'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

The 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 by the principal. The principal seeks to maximize the firm value (4), subject to the promise keeping constraint (5), the law

[10] 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.

of motion for firm size, and the incentive compatibility condition (6).

$$\begin{aligned}
J(W_0,\pi_0) &= \max_{\{\psi_{it},c_t,\tau,i_t\}} E\left[\int_0^\tau e^{-rt}\left(\pi_t\sum_{i=1}^N \mu_i a_{it}dt - dc_t - \pi_t\kappa\left(i_t\right)dt\right) + e^{-r\tau}\gamma\pi_t\left|\mathcal{F}_0\right.\right] \quad (7)\\
s.t.\ dW_t &= \left(\gamma W_t + \pi_t h\sum_i a_{it}\right)dt - dc_t + \pi_t\sum_{i=1}^N \psi_{it}dB_{it}\\
d\pi_t &= \pi_t\left(i-\delta\right)dt\\
\psi_{it} &\geq \frac{\sigma_i}{\mu_i}h \text{ if } a_i = 1
\end{aligned}$$

Notice that except for the payout to the agent, dc_t, the principal's value function $J\left(W,\pi\right)$ is scalable by π_t. The same holds true for the agent's continuation value dW_t, and

$$\frac{dW_t}{\pi_t} = \left(\gamma\frac{W_t}{\pi_t} + h\sum_i a_{it}\right)dt + \sum_{i=1}^N \psi_{it}dB_{it}.$$

This suggests that if we take $w_t = \frac{W_t}{\pi_t}$ as the relevant state variable, the principal's value function can be scaled by π_t and expressed in terms of w_t alone. Then, as I verify in Proposition 2, the principal's value function satisfies a scaled version of the HJB equation, and is given by

$$rj\left(w\right) = \sup_{c,a,i}\sum_i \mu_i - \kappa\left(i\right) + j'\left(w\right)\left(\left(\gamma - i + \delta\right)w + hn\right) + j''\left(w\right)\frac{1}{2}\sum_i \psi_i^2 + \left(i-\delta\right)j\left(w\right), \quad (8)$$

where $n = \sum_{i=1}^N a_i$ is the number of projects.[11]

4 Properties of the Optimal Contract

4.1 Shape of the Value Function

The HJB equation (8) is key to characterizing the optimal contract. The proposition below verifies that the solution to this equation equals the principal's optimal value function. This solution is then used in the following sections to characterize the choice of projects, and investment.

[11] The argument is detailed in Section A.1 of the Appendix.

Proposition 2. *The HJB equation (8) with the boundary conditions*

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

has a unique twice continuously differentiable solution on the interval $[0, \bar{w}]$, and equals the principal's optimal value function. The region $(0, \bar{w})$ is partitioned into continuation regions $\mathcal{C}_a$ on which a particular project selection a is optimal, and cutoffs $w(a, a')$ on which the project selection changes. 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. If the principal pays the agent, her value changes by $-(1 - j'(w_t))\, dc_t$.[12] Therefore, $dc_t > 0$ if $j'(w_t) \leq -1$, and $dc_t = 0$ otherwise. If the principal's value function is strictly concave, the point $\bar{w}$ at which the agent is paid is unique, since $j'(w)$ is decreasing and by the promise keeping constraint (5), w_t is reflected downwards at $\bar{w}$. This is captured by 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.[13] When the continuation value reaches zero the agent is fired, because the limited liability condition prevents the principal from punishing low outputs, which makes it impossible to incentivize effort.

4.2 Project Choice

In the first best benchmark, 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 or current firm size π.

Under moral hazard, the choice of projects is determined from the scaled HJB equation (8).

[12] 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.

[13] 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 $-(1 - j'(w'))\, dc_t > 0$. The optimal payment threshold $\bar{w}$ is the one at which this is not profitable.

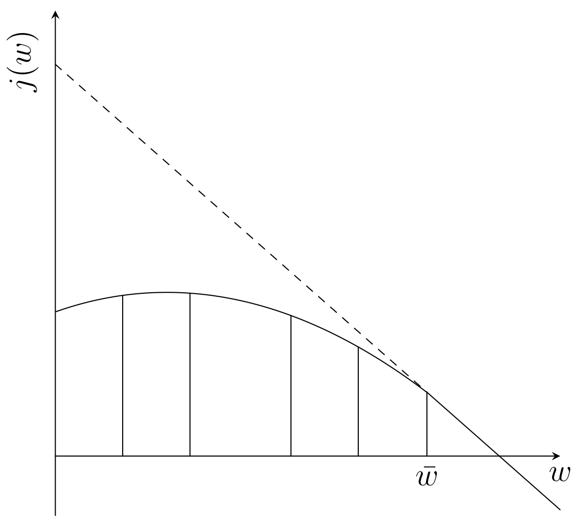

Figure 1: Shape of j(w) and continuation regions

A convenient feature of this representation of the principal's problem is that the function to be maximized is separable in the individual projects. Therefore, conditional on the agent's scaled continuation value, we can calculate the marginal benefit of each project separately, which is given by

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

Projects are executed whenever $b_i\left(w\right) > 0$, and the shape of the value function obtained in Proposition 2 can be used to characterize the marginal benefit function. To see how project choice and NPV relate, we can rewrite equation (9) as

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

Since $j'\left(w\right) \geq -1$, the term $\left(j'\left(w\right)+1\right)h$ is positive for all w, while $\frac{1}{2}j''\left(w\right)\psi_i^2$ is negative. Intuitively, $\frac{1}{2}j''\left(w\right)\psi_i^2$ measures the cost of providing additional incentives for the agent, and $\left(j'\left(w\right)+1\right)h \geq 0$ measures the benefit of of moving closer to the boundary $\bar{w}$.

Since $\psi_i^2 = h^2\left(\frac{\sigma_i}{\mu_i}\right)^2 = \frac{h^2}{SN_i^2}$, we have

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

The marginal benefit of implementing a project depends positively on both the net present value and the project's signal to noise ratio. While a higher NPV implies higher expected payoffs from the project, the signal to noise ratio works though the manager's incentives.

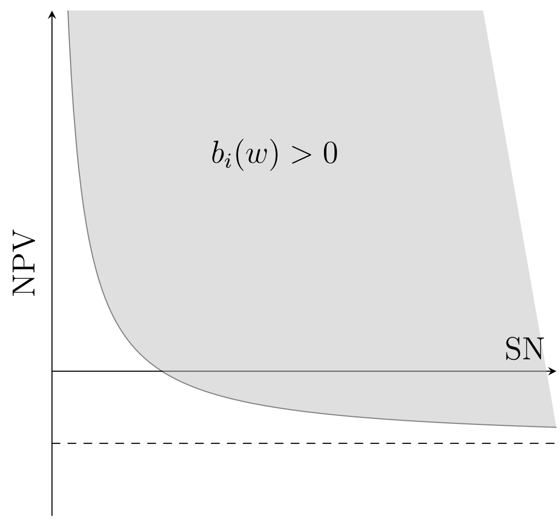

Figure 2: NPV vs. SN boundary

The lower the ratio, the harder it is to detect shirking, and the stronger the incentives have to be to motivate the agent. By the representation in equation (5), this is equivalent to a higher volatility of w_t, which induces a higher likelihood of hitting the boundary at which the agent is fired in the future. Since termination does not occur in the first best, it is inefficient, and the cost of increasing the termination probability is represented by $j''(w)$.

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

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

Consequently, all projects with higher than the minimal NPV are implemented, while all others are not. The threshold is a function of both the current scaled continuation value, as well as the project's signal to noise ratio. Figure 2 illustrates the non-linear relationship between NPV and SN and 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 positive NPV is not sufficient to compensate for the increase in termination probability, while overinvestment is due to the manager's limited liability constraint.[14] The principal cannot demand monetary compensation from the manager after bad performance, and must terminate all projects once w_t hits zero. To avert this, the principal can allocate more projects to the manager, which

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

increases his disutility of effort, and serves as a less inefficient punishment device.[15] This effect may compensate for the negative NPV of a project, and firm benefits from lowering the future expected termination probability, in exchange for current losses. The projects given as a punishment cannot be too difficult to incentivize, since otherwise the principal would have to increase the volatility of the manager's continuation value, which in turn makes firing more likely. 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.[16]

4.3 Project Portfolio Dynamics

The choice of projects evolves over time as w_t changes, and the marginal benefit of each project is non-monotonic in w. A higher continuation value implies a higher cost of compensating the agent for effort, which is captured by $j'(w)$ being decreasing. When $j'(w) > 0$, the cost of incentives may be increasing, because the principal's value increases with w, so she stands to lose more if a shock to the project outputs drives the contract into termination. When w is sufficiently large, and $j(w)$ is decreasing, the cost of incentives is decreasing in w as well, since termination becomes less likely, and the potential loss decreases with w. 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. Thus, the marginal benefit function at $\bar{w}$ is

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

and the project allocation converges to the first best when w_t approaches $\bar{w}$. 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.

As the continuation value grows, and the costs of incentives decline, it is intuitive to think

[15] 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.

[16] 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 $w < \bar{w}$.

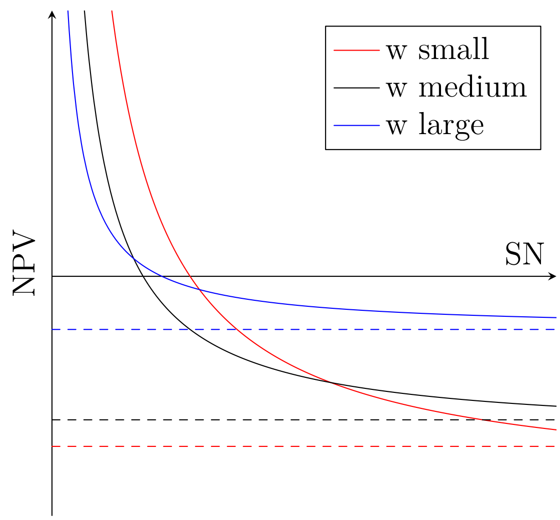

Figure 3: Project Cutoffs as a function of w

that the firm's optimal portfolio shifts towards high-NPV, low-SN projects, while projects with negative NPV are sorted out. Proposition 3 details the sense in which this intuition holds. When the cost of incentives is decreasing, the marginal benefit of any project which is sufficiently difficult to incentivize increases with w. 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 of every project is linear, the result in Proposition 3 can be sharpened, and for w sufficiently large, an increase implies that the firm adds project 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, we can 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, at each point in the state space, the choice of projects is independent. 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.[17] Intuitively, there is no hysteresis effect as in the real options literature and the firm can freely switch between projects.

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, $j_+'''(\hat{w}) < j_-'''(\hat{w})$.*

4.4 Project Choice and Investment

In the contract under moral hazard, optimal investment $i(w)$ is strictly below the first best level of investment. 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, $i(w)$ is increasing in w. 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 implies that for all w, $i(w) < i_{fb}$.

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.

[17] 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.

4.5 Output- vs. Project-Based Incentives

In this subsection, I illustrate how the criterion for project choice changes when the agent cannot be offered incentives based on individual project outputs. Suppose that the contract can only be conditioned on total output $dx_t = \sum_i dx_{it}$. With these output-based incentives, repeating the argument from Lemma 1 shows that the agent's continuation value satisfies

$$dW_t = (\gamma W_t + h(a_t))\, dt + \bar{\psi}_t \pi_t \sum_{i=1}^{N} \sigma_i dB_{it}, \tag{11}$$

and the IC constraint becomes

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

Hence, the firm chooses a certain portfolio of projects, and the project with the lowest NPV determines $\bar{\psi}$, the risk exposure required to incentivize all projects in the portfolio. To see why this needs to be the case, consider a variant of the intuition outlined in Section 3.3. If the manager 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 projects in the portfolio. Since $j''(w) < 0$, we have

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

Equivalently, given a portfolio and a value for $\bar{\psi}$, all projects in the portfolio necessarily satisfy

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

and therefore the portfolio appears to have been chosen using a hurdle rate or minimum NPV criterion.

Since the hurdle NPV depends on both the current scaled continuation value w as well as the current project selection, it shifts non-monotonically as w changes. It will however converge to the NPV > 0 criterion when w approaches $\bar{w}$.

We can deduce further properties of the contract by examining the principal's value function, which is given by

$$rj(w) = \sup_{c,a,i} \sum_i \mu_i - \kappa(i) + j'(w)((\gamma - i + \delta)w + hn) + j''(w)\frac{1}{2}\bar{\psi}^2 \sum_i \sigma_i^2 + (i-\delta)j(w).$$

The main difference to equation (8) is the coefficient of $j''(w)$. Under project based incentives the total risk inherent in the contract is $\sum_i \psi_i^2 = \sum_i \sigma_i^2 \left(\frac{h}{\mu_i}\right)^2$, while output based incentives raise the coefficient 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 under output than under project based incentives, as long as $\frac{h}{\mu_i} \neq \frac{h}{\mu_i}$ for some projects i and j with $a_i = a_j = 1$, and in particular the risk exposure on any project is at least as high as under project based incentives. Therefore, conditioning the incentive contract on total output alone cannot be efficient, since it exposes the agent to excessive risk.

4.6 Allocation of Funds in Projects

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

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

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

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

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 risk exposure on the agent's effort, 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

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

where λ is the Lagrange multiplier associated with resource constraint (13). Given project i is implemented, its capital allocation solves the FOC[18]

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

which implies that

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

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

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 $\lambda'(w)$. □

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

4.7 Project Dynamics with Stealing

In DeMarzo and Sannikov (2006) and DeMarzo et al. (2010), instead of putting in effort, the agent decides whether or not to steal from the principal. If he does, he receives an additional payoff of ϕdt, while the principal's expected increment in profit from a project is $\mu_i (1-\phi)\, dt$, where $\phi \in (0,1)$ measures the loss of social value from stealing. The main difference between this setting and mine is that in the optimal contract, the principal does not have to compensate the agent for effort when she implements the no-stealing outcome. Therefore, the overinvestment result of Section 4.2 cannot arise, since giving the agent additional projects has no impact on his utility in equilibrium.[19]

A benefit of the stealing setup is that the marginal benefit of each project is exclusively driven by the tradeoff of additional payoff versus the cost of providing incentives. With stealing and no investment, the HJB equation (8) becomes

$$rj(w) = \max_a \sum_i \mu_i + j'(w)\,\gamma w + j''(w)\,\frac{1}{2}\sum_i \psi_i^2, \tag{14}$$

and each project's marginal benefit function is simply $b_i(w) = \mu_i + j''(w)\,\frac{1}{2}\psi_i^2$. This allows for simpler dynamics of the optimal project portfolio.

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 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

[19] The projects which are not taken in the first best are those whose payoff μ_i is negative. Because of the cost of incentives, these projects continue to have a negative payoff under moral hazard. Therefore, there can only be underinvestment relative to the first best.

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 cash balances and assigns an equity share in every active project to the manager. The shares are vested, meaning that the manager will lose shares in current projects, or gain shares in new ones depending on his performance.

When only shares on the total output of the firm, and not the individual projects, can be issued 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 when the implementation consists of firm level equity only. To achieve the second best, the implementation must feature a measure of the manager'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. M_t denotes 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 stock of cash associated with project i. We have

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

The cash stock 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 $\kappa(i)$, and the payout to the agent.[20]

Equity holders are assumed to require a minimal dividend payoff which satisfies

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

The manager is endowed with a personal account[21] with balance A_t, which pays interest at rate γ, and is used in part to pay the manager. At any point in time, the manager receives an equity share Ψ_{it} in any active project. Whenever a new project is executed, the manager buys equity in that project at a pre-determined price, and when a project is halted, the manager sells off the equity. Proceeds from these sales and purchases are deposited in the personal account.[22] Finally, the manager 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 manager's consumption paid from the account, p_{it} the transfer price on share 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 $d\Psi_{it} < 0$.

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, the optimal contract can

[20] 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 $\sum_{i:a_i=1} \alpha_i = 1$.

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

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

be implemented.

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 manager is fired whenever $\sum_i \Psi_{it} M_{it} + A_t = 0$, for each i and t, the transfer price equals $p_{it} = M_{it}$. Then the contract from Section 4.5 will be 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 manager's continuation value W_t is a continuous function of time. An intuitive, but wrong, implementation of the contract is to simply strip the manager 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 manager loses the claim to future dividends, and distort incentives right before the jump. The transfer prices p_{it} are chosen to exactly offset the change in the manager's value, which is given by M_{it}.[23]

The implementation depends not only on the cash holdings of the firm, as in DeMarzo et al. (2010), 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 manager 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 manager from the equity stake, or an equivalent payout from the personal account. Hence the personal account may serve as another way to reward the manager 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 manager 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

[23] In DeMarzo et al. (2010), 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 the agent's equity share are not an issue.

satisfies simply

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

while the cash holdings process follows

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

and again the personal account is used to escrow proceedings from the manager'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) makes clear that holding a single equity stake $\bar{\Psi}_t$ conditions the manager'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 manager 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 will be 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 manager shirks. Given equity share Ψ_t, the growth in the value of the manager'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 manager's contract consists predominantly of an equity share, the inefficient hurdle rate contract is the only one that is implementable. Therefore, 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 boni, with the latter being a linearly weighted function of the manager's performance in individual categories set by shareholders. 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 manager receives a constant equity share Ψ_t, and bonus payments P_t, which are linear in individual project outputs

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

with positive weights β_{it}. Since in the optimal contract, the manager 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 manager'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 manager would hold if project specific shares could be issued. Although the manager 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 theory of real options, 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.[24] 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 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 $a' \neq a$,

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

and

$$k(a, a'') < k(a, a') + k(a', a'')$$

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.[25]

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

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

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

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

The solution to problem (23) is determined by the following system of quasi variational

[25] 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.

inequalities for all a and w

$$\min\left\{rj_a(w)-\mathcal{L}_{i,a}j_a(w)-\sum_i \mu_i+\kappa(i)+(i-\delta)j_a(w)\,,\right. \qquad (25)$$
$$\left. j_a(w)-\max_{a'\neq a} j_{a'}(w)-k(a,a')\right\} \;=\; 0.$$

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

$$\begin{aligned} j_a(w(a,a')) &= j_{a'}(w(a,a')) \\ j_a'(w(a,a')) &= j_{a'}'(w(a,a'))\,, \end{aligned} \qquad (26)$$

and

$$j_a''(w(a,a'))\geq j_{a'}''(w(a,a'))\,. \qquad (27)$$

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

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

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

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

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 a simple marginal benefit criterion. In particular, if the new portfolio has strictly more projects than the old one, then the optimality conditions around the threshold $w(a, a')$ imply

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

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*[26] *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 will be approximately continuous. This implies that 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. Moreover, the measure of the subset of w on which a particular project choice is optimal with fixed costs, but not chosen when $k > 0$ converges to zero.

7 Conclusion

I analyze a continuous time moral hazard problem in which the manager's effort is distributed among different projects. Unlike past studies, project choice is simultaneous, and the possible feedback effect between projects is explicitly considered. The model sheds light on the optimal choice of projects under moral hazard, as well as the distribution of funds among projects and the persistence of bonus contracts in CEO compensation. Further, it explains

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

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 free 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 free cash flow when the cost of incentives is decreasing, while the benefit projects which are relatively safe decreases. The first best project selection schedule is attained whenever the free cash flow is large enough. The manager is given riskier projects after a history of sufficiently good performance, while a poorly performing manager 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 free cash flow 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' risk 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

$$rJ = \max_{a,i} \pi \sum_i \mu_i a_i - \pi\kappa(i) + J_w \left(\gamma W + \pi h \sum_i a_{it} \right) + \frac{1}{2} J_{ww} \pi^2 \sum_i \psi_i^2 + J_\pi \pi (i - \delta) \quad (29)$$

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

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

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_* 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

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

By Berge's Maximum Theorem,[27] $H(w, u, p)$ 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.[28] 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[29]

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

Finally, let $\bar{w}(s) = \inf\{w : j_s(w) = j_*(w)\}$ be the first point at which j_s hits j_*.[30] 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

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

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

[29] 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.

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

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 $\bar{w}$,[31] and is given by

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

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 $j'(\hat{w}) \geq 0$, we have

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

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

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

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

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

Since for all fixed i,

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

we have

$$j_{fb}(\hat{w}) < \frac{\sum_i \mu_i - \kappa(i_{fb})}{r - i_{fb} + \delta} - w \leq \frac{\sum_i \mu_i - \kappa(i^*)}{r - i^* + \delta} - w \leq j(\hat{w}),$$

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

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

which makes it impossible for j'' to cross zero from above, as required by the definition of $\hat{w}$. If $\hat{w}$ does not lie on the interior of any continuation region, there exists a project i such that $b_i(\hat{w}) = \mu_i + j'(\hat{w})\, h = 0$. Take some $w > \hat{w}$. Because $j'' < 0$ on $(\hat{w}, w)$, j' is decreasing on this region, and $b_i(w) < b_i(\hat{w}) = 0$. Thus, the project cannot be taken again on $(\hat{w}, w)$, and

[31] 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.

the project choice must stay constant on this region. This implies that the right derivative $j_+'''(\hat{w})$ exists, and it is positive since j' and j'' are negative. But then again $j''(w)$ cannot cross zero from above at $\hat{w}$. Therefore, $j''(w) < 0$ for all $w < \bar{w}$. $\square$

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

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

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

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

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

$$-1 = j'_{s'}(\bar{w}(s')) > j_s(\bar{w}(s')),$$

and by Proposition 14, j_s is strictly concave, and therefore $j_s(\bar{w}(s)) < j_s(\bar{w}(s')) < -1$. $\square$

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

Lemma 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,

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

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

$$-\left(r-i+\delta\right)j\left(w\right)+\sum_{i}\mu_{i}-\kappa\left(i\right)+j'\left(w\right)\left(\left(\gamma-i+\delta\right)w+hn\right)+j''\left(w\right)\frac{1}{2}\sum_{i}\psi_{i}^{2}\leq0, \quad (32)$$

and for any consumption payout policy, $-\pi_t\left(1+j'\left(w\right)\right)dc_t \leq 0$ on $\left(0,\bar{w}\right)$. Since $j'\left(w\right)$ 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

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

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

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

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

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

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 $\left(0,\bar{w}\right)$ 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 $\hat{a}(w)$ 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 $L(w, a(w), \hat{a}(w))$ 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

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

Here, the expectation operator is taken under the action $\hat{a}(w)$, and τ_ε denotes the first exit time of w_t from the interval. As $\varepsilon \to 0$, this expression converges to zero.[32]

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

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

A.3 Properties of the Value function

A.3.1 Proof of Proposition 3

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

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

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

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*

[32]Precisely, since $\hat{a}(w)$ is constant, w_t satisfies $dw_t = (\gamma w_t + h\hat{n})\,dt + \sum_i \psi_i \hat{a}_i dB_{it}$, and $\lambda(w)$ solves the differential equation $r\lambda(w) = \bar{L} + \lambda'(w)(\gamma w + h\hat{n}) + \lambda''(w)\frac{1}{2}\hat{\psi}^2$ 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 $\lambda'(w)$ which is uniform in ε. This implies that $|\lambda(w)| \le M\varepsilon$ 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.

$j'(w) < 0$ 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 $\psi_i^2 > -\frac{2j''(w)}{j'''(w)}$.

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 $\psi_j^2 > \frac{-2j''(w)}{j'''(w)}$, which is equivalent to

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

$\square$

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

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

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

$$j_{+}^{\prime\prime\prime}(\hat{w})-j_{-}^{\prime\prime\prime}(\hat{w})=\frac{1}{\frac{1}{2}\sum_{i}\psi_{i}^{2}a_{i}}b_{n+1,+}^{\prime}(\hat{w})<0.$$

The jump in $j^{\prime\prime\prime}(w)$ can be equivalently expressed as

$$j_{+}^{\prime\prime\prime}(\hat{w})-j_{-}^{\prime\prime\prime}(\hat{w})=-\frac{1}{\frac{1}{2}\sum_{i}\psi_{i}^{2}a_{i}+\frac{1}{2}\psi_{n+1}^{2}}b_{n+1,-}^{\prime}(\hat{w})$$

in case a project is added, and

$$j_{+}^{\prime\prime\prime}(\hat{w})-j_{-}^{\prime\prime\prime}(\hat{w})=\frac{1}{\frac{1}{2}\sum_{i}\psi_{i}^{2}a_{i}-\frac{1}{2}\psi_{n+1}^{2}}b_{n+1,-}^{\prime}(\hat{w})$$

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

$$b_{n+1,-}^{\prime}(\hat{w})=\left(1+\frac{\psi_{n+1}^{2}}{\sum_{i}\psi_{i}^{2}a_{i}}\right)b_{n+1,+}^{\prime}(\hat{w})$$

if a project is added, and

$$b_{n+1,-}^{\prime}(\hat{w})=\left(1-\frac{\psi_{n+1}^{2}}{\sum_{i}\psi_{i}^{2}a_{i}}\right)b_{n+1,+}^{\prime}(\hat{w})$$

if it is removed.

A.3.3 Proof of Proposition 7

The principal's HJB equation is given by

$$rj(w)=\max_{a}\sum_{i}\mu_{i}+j^{\prime}(w)\gamma w+j^{\prime\prime}(w)\frac{1}{2}\sum_{i}\psi_{i}^{2},$$

with boundary conditions $j(0)=l$, $j^{\prime}(\bar{w})=-1$ and $j^{\prime\prime}(\bar{w})=0$. 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 has the same qualitative features.

On any region where project choice is constant, we have

$$j^{\prime\prime\prime}(w)=\frac{-(\gamma-r)j^{\prime}(w)-j^{\prime\prime}(w)\gamma w}{\frac{1}{2}\sum_{i}\psi_{i}^{2}a_{i}}.$$

For w sufficiently close to zero, $j^{\prime\prime\prime}(w)$ 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 manager 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 manager'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.[33] 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 manager either sells or buys shares at the price of M_{it} per unit, which corresponds exactly to their naive value in terms of the firm's cash holdings.

To verify that the manager's continuation value satisfies (35), by Itô's Lemma for semimartingales[34] 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

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

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.

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

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

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

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

$$\begin{aligned}
\gamma W_t &= \sup_a \Psi_t \left(\sum_i \left(\gamma M_{it} + \pi_t \mu_i a_i - \frac{dc_{it}}{dt} \right) \right) + \Psi_t \frac{dc_t}{dt} - \pi_t h \sum_i a_i + \\
&\quad + \pi_t \sum_i \left(\Psi_{it} - \Psi_t \right) \mu_i a_{it} + dA_t + \frac{d\Psi_t}{dt} M_t \\
&= \Psi_t \gamma M_t + \sup_a \left(\pi_t \sum_i \left(\Psi_t \mu_i - h + \left(\Psi_{it} - \Psi_t \right) \mu_i \right) a_i \right) + \gamma A_t \\
&= \Psi_t \gamma M_t + \gamma A_t
\end{aligned}$$

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.[35] In the following,

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

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

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

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

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

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

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

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

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

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

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, the 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.[36] 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

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

[36] See Strulovici and Szydlowski (2011).

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

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

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. $\square$

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. $\square$

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.[37] 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,[38] 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 $\mathcal{S}_{a,a'} \subset \mathcal{C}_{a'}$,[39] equation (37) with a replaced by a' holds for $j_{a'}$, and therefore

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

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. □

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 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 $\hat{w}$.

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

[38] See Strulovici and Szydlowski (2011).

[39] 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''.

By the previous Lemma, the left and right derivatives $j'_{a,-}(\hat{w})$ and $j'_{a,+}(\hat{w})$ exist. Assume that $j'_{a',-}(\hat{w}) < j'_{a,+}(\hat{w})$,[40] take any number $x \in \left(j'_{a,-}(\hat{w}), j'_{a,+}(\hat{w})\right)$, 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 $x \in \left(j'_{a,+}(\hat{w}), j'_{a,-}(\hat{w})\right)$, 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, $j'_+(w) = j'_-(w)$. $\square$

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

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

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.

[40] 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 $j'_{a,+}(\hat{w}) = j'_{a'}(\hat{w})$.

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

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

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. $\square$

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

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

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

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

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

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,

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

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

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

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 $\left[w', \bar{w}_{a_{fb,k}}\right] \subset \mathcal{C}^k_{a_{fb}}$, and $j_{a_{fb},k}(w)$ satisfies equation (37) with $a = a_{fb}$, and boundary conditions $j_{a_{fb},k}(w') = j_{a',k}(w') - k(a_{fb}, a')$, $j'_{a_{fb},k}\left(\bar{w}_{a_{fb},k}\right) = -1$ and $j''_{a_{fb},k}\left(\bar{w}_{a_{fb},k}\right) = 0$. 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 $w \in \mathcal{S}_{a,a'}^k$, we have

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

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 $[0, \bar{w}^0]$, 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. □

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

$$\begin{aligned}\min\Big\{ rj_{a,k}(w_{a,k}) - \max_i f_{a,i} + \mathcal{L}_{a,i}\phi(w_{a,k}) + (i-\delta) j_{a,k}(w_{a,k}), \\ j_{a,k}(w_{a,k}) - \max_{a' \neq a} j_{a',k}(w_{a,k}) - k(a, a')\Big\} \quad \geq \quad 0.\end{aligned}$$

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

$$rj_{a,k}(w_{a,k}) \geq \max_i \left\{ f_{a,i} + \mathcal{L}_{a,i}\phi(w_{a,k}) + (i-\delta) j_{a,k}(w_{a,k}) \right\}.$$

As $k \to 0$, this yields

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

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

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

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

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

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

$$\begin{aligned} \min\Big\{ & r j_k(w_k) - \max_i \left\{\mathcal{L}_{a(k),i}\phi(w_k) + (i-\delta) j_k(w_k)\right\}, \\ & j_k(w_k) - \max_{a'\neq a} j_k(w_k) - k(a, a')\Big\} \leq 0, \end{aligned}$$

which combined with equation (40) implies that

$$r j_k(w_k) \leq \max_i \left\{f_{a(k),i} + \mathcal{L}_{a(k),i}\phi(w_k) + (i-\delta) j_k(w_k)\right\}.$$

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

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

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. □

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,[41] 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.

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

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

in the Hausdorff distance.

[41] 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.

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, January). 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 (2010). Dynamic agency and the q theory of investment.

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. *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.

Strulovici, B. and M. Szydlowski (2011). On the smoothness of value functions.