Incentives, Project Choice and Dynamic Multitasking*

Martin Szydlowski†

January 30, 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 execution which depends on both a project's NPV as well as it's signal to noise ratio (SN). The cutoff shifts dynamically depending on the past history of shocks, 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 will depend 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 others it is not. When a firm is the host of many disparate activities, we can instead think of the firm being comprised of a portfolio of potential projects. When these projects are

*I would like to thank Bruno Biais, Janice Eberly, Arvind Krishnamurthy, Alessandro Pavan, Mark Satterthwaite and Yuliy Sannikov for helpful discussions, and Jeffrey Ely, Mike Fishman and Bruno Strulovici for their guidance.

†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 executed, they provide the firm with risky cash flows, which will depend on the 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 it's expectations, the firm faces a much more difficult problem - to determine the optimal portfolio of projects at any given time.

Unsurprisingly, while the continuous investment framework has received considerable attention and has spawned a rich literature which attempts to generate realistic investment dynamics, [1] the literature on project choice has been sparse. This is due to the fact that to study project choice realistically, one has to consider both discreteness and dynamics.

The goal of my paper is to study the optimal project portfolio in a company which relies on a manager to execute the projects. I characterize the firm's optimal project selection policy which arises when the manager has an incentive to shirk, and show how the firm's portfolio of projects evolves over time. Even though both the shareholders and the manager are risk neutral in my setup, project choice along the optimal path of the dynamic contract will not be 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 will be a function of the manager's promised continuation utility, as well as the project's signal to noise ratio (SN), which can be interpreted as an indicator for how difficult it is for the shareholders to discern whether or not 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 shareholders to forgo positive NPV projects due to their risk, over investment is caused by the firm's 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.[2] 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 will forgo 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 will be 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,

[1] See Abel and Eberly (1994) for a canonical reference.

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

as the firm's cash holdings grow, 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. Interestingly, this dynamic is entirely driven by the cost of incentives, which are in turn embodied in the projects' SN ratios. The intuition is as follows.

Since managers have limited liability, the firm is terminated when the agent's continuation value reaches it's lower boundary. Thus, incentivizing projects comes with an increased managerial risk exposure, which in turn will increase the shutdown probability. The closer the agent's continuation value is to the lower bound, the more costly the additional risk becomes, to the point where a project's NPV cannot compensate the principal for the added risk.

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 positively affect the effectiveness of managerial effort, the associated cost of incentives is lowered. Thus, firms with low cash holdings will distort their allocation of funds.

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 will be 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 about 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 not 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, project choice will resemble 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 will be a function of the project with the lowest NPV in the portfolio. This hurdle rate allocation will not be 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 will carry 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 a certain condition is met.

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 choise 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 put forward by Schattler and Sung (1993) and Sannikov (2008). Recent contributions which share certain features with my setup include Biais et al. (2010), Fong (2007) and He (2009). For instance, Biais et al. (2010) study optimal investment and downsizing in a firm as a mean to implement a contract that causes the manager to exert effort in 'accident prevention', while He (2009) studies a model in which the agent's effort affects the firm size. While the first two focus on implementing a static effort decision, Fong (2007) studies a binary effort decision in a single task 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 distorts investment decisions. The framework in my model shares some of their features, such an investment-consumption decision by the principal, and linear utility. In contrast, I allow for multiple projects with different risk-return profiles which the firm can implement independently of their investment, and characterize the choice of projects as well as their implications for firm financing. Compared to DeMarzo et al. (2010), there are two decisions to make in the firm, the first being on scale and the second on a portfolio of projects to choose from, and the firm's project portfolio will vary over time, whereas it is comprised of a single, static project in DeMarzo et al. (2010). This also has implications for the implementation of the contract, since instead of a constant equity share, my model will feature equity transfers to and from the manager.

On the microeconomics side, the problem of multitasking has received significant attention since the seminal article of Holmstrom and Milgrom (1991). As the list of works it too long to be repeated here, I refer the reader to Bond and Gomes (2009) for a recent contribution and references. Due to the complex nature of the problem dynamic studies of multitasking are rare. Manso (2006) studies the trade off between exploration and exploitation of current discoveries in a two period setting. Miquel-Florensa (2007) considers a setup with two tasks in a discrete time setting and focuses on the question when, and whether, it is optimal to execute tasks sequentially or in parallel, which will depend on the strength of externalities between them. In my model the agent's effort does not affect project completion, but instead the cash flow of the project. Also, projects do not end as they do in Miquel-Florensa (2007). Instead, the principal in my model decides *which* and *how many* projects to execute in parallel at any given time, when to allocate more projects to the agent, and when to stop inducing effort in certain projects.

In a very recent paper, Hartman-Glaser et al. (2010) study a contract between a mortgage underwriter and secondary investors. They consider a multitasking model where the underwriter may exert effort in a subset of mortgages, and find that the agent either exerts effort in all mortgages or none so that the question of which tasks are allocated to the agent does not arise in equilibrium. They also find that bundling mortgages is optimal, which is reminiscent of Laux (2001), who reports a similar result in a static setting.

Finally, my model is related to the literature on optimal investment. Two strands are noteworthy. The real options literature, as summarized in Dixit et al. (1994), offers a complementary view on the issue of project choice. 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), only the choice of a single project is studied. This is due to the fact that characterizing the choice of multiple projects needs to take into account how taking one project changes the decision maker's value function with regard to the other projects, which is difficult to determine in the real options setup. In my model, the externality between projects is well behaved, and of second order only, allowing for the characterization of an entire project portfolio.

The literature on capital budgeting which has built on Harris and Raviv (1996) and Harris and Raviv (1998), studies the choice of projects when a division manager has superior 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 understood of as an allocation of a budget to the manager. In a similar setup, Berkovitch and Israel

(2004) derive an alternative 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 uni dimensional quality associated with them instead of risk and return, it is difficult to compare my results. If the average project payoff in my framework is interpreted as quality, and the relation between payoff and the SN ratio is positive and sufficiently large, then my model will imply that there are too many low quality projects and too few high quality projects in the firm's portfolio, in line with the above.

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

3 Model Setup

3.1 Projects and Investment Technology

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 it's 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 noise components 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. $a_{it} \in \{0, 1\}$. When $a_i = 1$, the cumulative project cash flow x_{it} evolves according to a Brownian Motion with drift μ_i and volatility σ_i, and the instantaneous cash flow dx_{it} is given by the diffusion

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

σ_i can be understood either as a measure of the riskiness of project i, or, equivalently, how difficult it is to infer the agent's effort a_i from observing the outcome path x_i of the project, 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. 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 project. 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 an 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 with filtration $\mathcal{F}_t$, which satisfies the usual conditions.[3] 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 $\gamma\pi$.[4] Let $\tau = \inf\{t \geq 0 : \ w_t = 0\}$ denote the random time at which shutdown occurs. The agent is remunerated with a non-negative consumption process $c = \{c_t \in \mathbb{R}_+ : 0 \leq t \leq \tau\}$ and the principal prescribes a vector-valued effort process[5] $a = \left\{(a_{it})_{i=1}^N \in \{0,1\} : 0 \leq t \leq \tau\right\}$.

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

[4] Note that this regime is not renegotiation proof as the value function of the principal will be 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.

[5] The discreteness of the effort process poses a potential problem. If the agent's effort a_{it} switches between zero and one for an infinite amount of times in a finite interval of time, the output process x_{it} will not be well behaved. Although this will not happen on the path of the optimal contract, it is not ruled out that there may be deviations where such behavior occurs. The problem can be solved by assuming a small positive switching cost that is incurred whenever effort changes. Then, it will never be optimal to switch effort infinitely many times in any finite time interval. For the sake of exposition, I omit switching costs from the

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

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

Here, dc_t is the consumption payment process the manager receives, and the manager's utility is the expected discounted consumption payments minus his expected effort cost, $h\sum_i a_{it}$, which is linear in the projects.

Note that on under the optimal contract, the agent's and the principal's information sets are the same. The agent's effort cost is linear and symmetric in each task effort a_{it}, and increases with firm size. The principal is risk neutral, and seeks to maximize the following expression

$$J(W,\pi) = \sup_{a,c} E\left(\int_0^\infty e^{-rt}\left(\left(\pi_t \sum_{i=1}^N \mu_i a_{it} - \pi_t \kappa\left(i_t\right)\right) dt - dc_t\right)\Bigg| W_0, \pi_0\right) \tag{4}$$

Here, the first term are the expected discounted cash flows from the projects, conditional on the manager's effort. The principal bears the expenses for investment in the firm $\pi_t\kappa\left(i_t\right)$, as well as for consumption payments ot 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 is useful for ruling out the case where the agent's consumption is postponed forever. Finally, I impose an upper bound on relative investment, $i < r + \delta$ to ensure that the principal's value function is bounded.[6]

3.3 Incentive Compatibility

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

$$W_t = E\left(\int_t^\infty e^{-\gamma(s-t)}\left(dc_s - \pi_t h \sum_i a_{it} ds\right) |\{a_s, c_s\}_{s\geq t}, \mathcal{F}_t\right) \tag{5}$$

initial analysis. Section 6 considers the relationship between the model with and without switching cost and shows that we obtain the current model as a limiting case when the switching costs go to zero.

[6] 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.

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

Then, the martingale representation theorem implies that the agent's continuation value W_t follows diffusion process in the multidimensional Brownian motion B_t, which yields a tractable representation for the agent's wealth process.

Intuitively, given any project selection rule a and consumption schedule c, the only source of uncertainty in the model will be 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. For Brownian Motion, this function takes a particularly simple form, as can be seen below.

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 $\left\{(\psi_{it})_{i=1}^{N} : 0 \le t \le \tau\right\}$, such that*

$$dW_t = \left(rW_t + \pi_t h \sum_i a_{it}\right) dt - dc_t + \pi_t \sum_{i=1}^{N} \psi_{it} dB_{it} \tag{6}$$

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

$$\psi_{it} \ge \frac{\sigma_i}{\mu_i} h \tag{7}$$

whenever $a_{it} = 1$.

In the same spirit as Sannikov (2008), 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 will be 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 will change 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.

Without exerting effort, his utility will rise by hdt. Because the principal would not know that the agent is shirking, she expects that $dB_{it} = \frac{1}{\sigma_i}\left(dx_{it} - \mu_i dt\right)$, while the true process is $dx_{it} = \sigma_i dB_{it}$. Hence, the principal's expectation of the noise process falls short by $-\frac{\mu_i}{\sigma_i} dt$, and by the representation, the manager loses $\psi_i \frac{\mu_i}{\sigma_i} dt$ in continuation utility. To induce effort, this loss must be larger than h, which leads to equation (7).

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}. For instance, when the project is relatively safe, and the ratio $\frac{\mu_i}{\sigma_i}$ is very large, observing a shortfall in output by roughly $\mu_i dt$ is a very unlikely event, and if the manager exerts effort this only happens when a large shock dB_{it} hits the project. Therefore, when observing a shortfall in output, the principal will conclude that tha manager has been shirking, and punish the manager severely.[7]

Finally, note that analogous to the discrete time contracting literature, equation (6) 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 will imply that the principal has to promise more to the agent at W_{t+dt}.

3.4 The Optimal Contract

The with the result of Lemma 1, the optimal contract can be expressed in 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 relevant state space will consist of continuation value and firm size pairs $(W, \pi) \in \mathbb{R}_+^2$, and the optimal consumption and effort processes will be functions of the state variables, i.e. $c = c(W, \pi)$ and $a = a(W, \pi)$.

The principal seeks to maximize the firm value, net of consumption payments and investment costs, subject to the promise keeping constraint (6), the law of motion for firm size and the incentive compatibility condition (7).

$$
\begin{aligned}
J(W,\pi) &= \max_{\{\psi_{it}, c_t, \tau, i_t\}} E\left(\int_0^{\tau} e^{-rt}\left(\pi_t \sum_{i=1}^{N} \mu_i a_{it} dt - dc_t - \pi_t \kappa\left(i_t\right) dt\right) + e^{-r\tau}\gamma\pi_t \,|\mathcal{F}_0\right) \quad (8)\\
s.t.\;\; dW_t &= \left(rW_t + \pi_t h \sum_i a_{it}\right) dt - dc_t + \pi_t \sum_{i=1}^{N} \psi_{it} dB_{it}\\
d\pi_t &= \pi_t\left(i-\delta\right) dt\\
\psi_{it} &\geq \frac{\sigma_i}{\mu_i} h \;\text{ if }\; a_i = 1
\end{aligned}
$$

Notice that except for the payout to the agent, dc_t, the principal's value function $J(W, \pi)$ is

[7]Formally, in equation (6), the shortfall in output is equivalent to a very large negative realization of dB_{it} and taking ψ_{it} as given will imply that W_t falls by a relatively large amount, while the opposite is true for when σ_i is large.

scalable by π_t. The same holds true for the agent's continuation value dW_t, except that now

$$\frac{dW_t}{\pi_t} = \left(r\frac{W_t}{\pi_t} + h\sum_i a_{it} \right) dt + \sum_{i=1}^{N} \psi_{it} dB_{it} \tag{9}$$

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, the principal's value function satisfies a scaled version of the HJB equation and is given by

$$rj\left(w\right) = \sup_{c,a,i} \sum_i \mu_i - \kappa\left(i\right) + j'\left(w\right)\left(\left(r - i + \delta\right) w + hn\right) + j''\left(w\right) \frac{1}{2} \sum_i \psi_i^2 - \left(i - \delta\right) j\left(w\right) \tag{10}$$

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

4 Properties of the Optimal Contract

4.1 Shape of the Value Function

The HJB equation (10) is key to characterizing the optimal contract. To determine the contract's properties at all points in the state space, we need to establish how j and it's first two derivatives depend on w. This is the purpose of this section.

The solution to equation (10) satisfies the familiar properties from DeMarzo et al. (2010). It features two boundaries, 0 and $\bar{w}$. At the first, the contract is terminated, while at the second, the agent is given a consumption payment dc_t so that the scaled continuation value process w_t reflects at $\bar{w}$. At this point the smooth pasting and super contact conditions

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

are satisfied, while at the shutdown boundary only the value matching condition $j\left(\bar{w}\right) = L$ holds. Throughout the region $(0, \bar{w})$, we have $j'\left(w\right) > -1$ and $j''\left(w\right) < 0$, that is, the firm is risk averse.

Hence, the qualitative properties of the value function, as far as it's shape and shutdown or payment conditions are concerned is unaffected by having different investment projects. Note however, that although these qualitative conditions hold, the value function will be

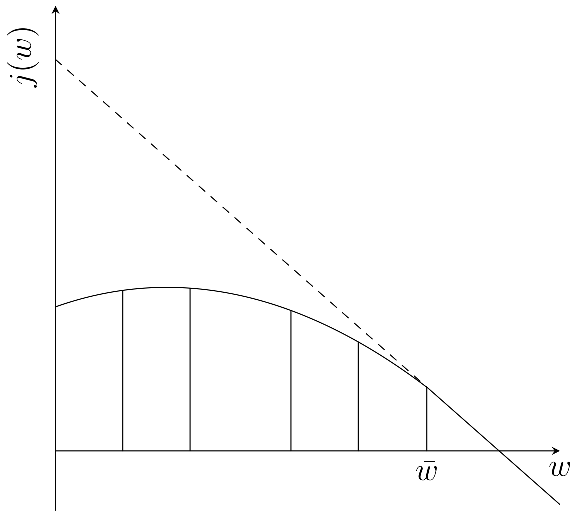

Figure 1: Shape of j(w) and continuation regions

different from the one featured in DeMarzo et al. (2010), since the HJB equation on the continuation region is different.

Proposition 2. *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 j is C^2 on the entire region $(0, \bar{w})$ and is C^3on $Int(\mathcal{C}_a)$ for all a with $j''' > 0$ wherever it exists and when project selection changes, $j'''(w)$ exhibits a jump.*

Figure 1 illustrates the results.

4.2 Project Choice

Having characterized the properties of the principal's value function, I now turn to the optimal choice of projects. In the first best benchmark the project is chosen if and only if $\mu_i > h$, i.e. whenever the payoff of the project is higher than the effort cost. Hence, in the first best, project choice follows the NPV criterion, and project assignment 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 (10). At each point w, the principal chooses the agent's consumption, investment and projects to maximize the right hand side, taking the function $j(w)$ as given. A convenient feature of this representation of the problem in (8) is that the function to be maximized is separable in the individual projects. Therefore, we can calculate the marginal benefit of each project which is given by

$$b_i(w) = \mu_i + j'(w)\,h + j''(w)\,\frac{1}{2}\psi_i^2 \tag{11}$$

and which is a function of both the project's characteristics and the scaled continuation value w. Due to this separability, projects are executed whenever $b_i(w) > 0$, and we can use the results in Proposition 2 to characterize each project's marginal benefit.

Further, we can relate expression (11) to the NPV rule rewriting it as

$$b_i(w) = \mu_i - h + (j'(w) + 1)\,h + \frac{1}{2}j''(w)\,\psi_i^2 \tag{12}$$

Since we have $j'(w) \geq -1$, the term $(j'(w) + 1)\,h$ will be positive, while $\frac{1}{2}j''(w)\,\psi_i^2$ will be negative. Intuitively, $\frac{1}{2}j''(w)\,\psi_i^2$ measures the cost of providing incentives for the agent, and $(j'(w) + 1)\,h \geq 0$ measures the benefit of of moving closer to the efficient 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

$$b_i(w) = \mathrm{NPV}_i + (j'(w) + 1)\,h + \frac{1}{2}j''(w)\,\frac{h^2}{\mathrm{SN}_i^2} \tag{13}$$

Hence, the marginal benefit of implementing a project will depend positively on both the net present value and the project's signal to noise ratio. While the effect of NPV is obvious, as higher NPV implies higher expected cash flows from the project, the signal to noise ratio works though the manager's incentives. The lower the signal to noise ratio, the harder it is to detect deviations by the agent, and thus, the stronger the incentives need to be to motivate the agent to work. Since $j''(w) < 0$, this is costly in the eyes of the principal and the firm is effectively risk averse, even though all involved parties are risk neutral.

The reason is that under moral hazard, the manager is fired as the continuation value W_t reaches zero. Since termination would not occur in the first best, it is inefficient, and the $j''(w)$ indicates exactly how this termination risk is weighted by the firm.

In line with this intuition, the third derivative of the value function $j'''(w)$ is positive wherever it exists, so that $j''(w)$ is a strictly increasing process in w. That is, as the manager's promised value moves away from the termination boundary, the probability of the firm being liquidated decreases, and so does the principal's risk aversion. When the manager takes on additional projects, the higher volatility in the law of motion for the continuation value (6) implies that at any W_t the probability of being terminated in the future is higher, which implies additional costs for the principal.

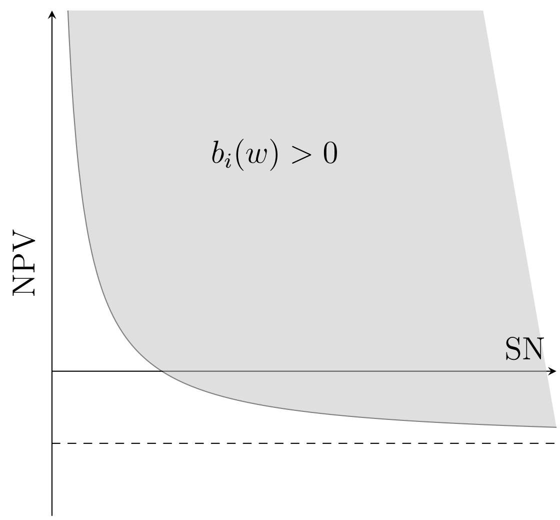

Figure 2: NPV vs. SN boundary

Setting $b_i(w) = 0$, we can derive the minimal NPV which the shareholders will require to implement a project.

$$\underline{\mathrm{NPV}}(\mathrm{SN}, w) = -\left(j'(w) + 1\right) h - \frac{1}{2} j''(w) \frac{h^2}{\mathrm{SN}^2} \tag{14}$$

Consequently, all projects with higher than the minimal NPV will be implemented, while all others will not. This threshold is be 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 will be chosen for a particular w.

4.3 Project Portfolio Dynamics

The choice of projects will evolve over time, as w changes. Since $j''' > 0$, i.e. the 'cost' of low SN is decreasing and because of $j'' < 0$, the cost of compensating effort is increasing. We have

$$\forall\, w \in (0, \bar{w})\;\; j''(w) > j''(0) \text{ and } j'(w) < j'(0) < 0 \tag{15}$$

therefore, projects with a sufficiently high NPV relative to their SN will always be chosen.

Proposition 3. *If $NPV_i > -j''(0)\,\frac{1}{2}\frac{h^2}{SN_i^2}$, we have $b_i(w) > 0$ for all $w \in (0, \bar{w})$ and there is no distortion in from moral hazard in selecting project i.*

For all other projects, implementation will be distorted by the agency problem and the trade off between NPR and SN will become significant. Positive NPV projects may not be implemented, unless their SN is sufficiently large so that it compensates the principal for the added risk. As w grows, the termination probability declines, and with it the cost of exposing the agent to risk, so that holding NPV constant, shareholders will accept a lower signal to noise ratio.

This effect is captured by the fact that $j'''(w) > 0$ in equation (13). At the same time, $(j'(w) + 1)\,h$ is decreasing, which implies that while the relative cost of risk exposure decreases, the cost of compensating the agent for effort is increasing. This generates a non-trivial dynamic for project selection.

Proposition 4. *For any w, and project allocation a, there exist a cutoff $\overline{SN}(w)$ such that for all $SN_i > \overline{SN}(w)$, $b_i'(w) < 0$ and for all $SN_i < \overline{SN}(w)$, $b_i'(w) > 0$. Equivalently, we have,*

$$b_i'(w) > 0 \text{ if } \psi_i^2 > \frac{1}{n} \sum_{j:a_j=1} \psi_j^2 \tag{16}$$

and, for two projects i and j,

$$b_i - b_j > 0 \text{ whenever } \Delta NPV > -j''(w) \frac{h^2}{2} \Delta \frac{1}{SN^2} \tag{17}$$

and

$$b_i' - b_j' = j'''(w) \frac{h^2}{2} \Delta \frac{1}{SN^2} \tag{18}$$

Hence, at any w, marginally raising the scaled continuation value implies that $b_i(w)$, the value of the project to the principal, rises for projects with sufficiently low signal to noise ratio. Alternatively, in line with the intuition just outlined, we should expect that as w rises, the benefit of more risky projects rises as well. The Proposition shows that this will be the case when a project required risk exposure ψ_i is above the mean risk exposure of the currently chosen portfolio.

Interestingly, the model provides a rationale for why, and when, negative NPV projects are executed. To see this effect notice that if $\sigma_i = 0$, i.e. the signal to noise ratio is infinite, the break even NPV satisfies

$$\overline{\text{NPV}} = -(j'(w) + 1)\,h < 0 \tag{19}$$

Standard explanations for this phenomenon include 'empire building' as in Jensen (1986), and private benefits to the manager. In my setup, the reason is rather different and can be explained due to the interaction of two forces.

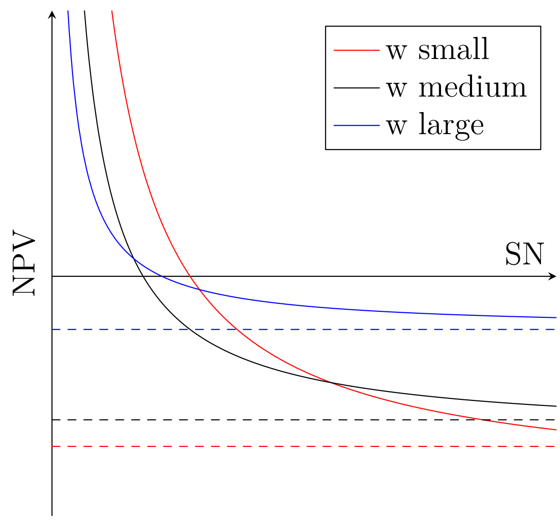

Figure 3: Project Cutoffs as a function of w

If w is relatively close to zero, i.e. $j'(w) > 0$ and the termination probability is high, taking on a very safe negative NPV project will serve to push the agent's continuation value up in expectation. Hence the firm benefits from lowering the future expected termination probability, in exchange for current losses. At the same time, a low continuation value implies that the effective cost of compensating the agent for effort is lower than the first best. This is due to the fact that a low w will imply that effectively the principal does not need to reimburse the entire effort cost to the agent. Instead, payment is made only at the boundary $\bar{w}$, and the farther away w is from the payment boundary, the lower the expected present value of the payment, and hence the lower the expected cost for the principal. Indeed, as w rises, the minimal NPV will rise as well. Eventually, when w is close to $\bar{w}$, the relative importance of the SN criterion will vanish, and only positive NPV projects will be executed, irrespective of their riskiness. 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. Together with the results from Proposition 4, this implies that as w grows, the composition of the firm's optimal portfolio will shift towards more high-NPV, low-SN projects, while more and more negative NPV projects are sorted out. Finally, as $w \to \bar{w}$, the break-even NPV line will become successively flatter and approach the x-axis.

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.[8] 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. *Whenever another project is added, $b_i'(w)$ jumps down for all projects and whenever a project is dropped, $b_i'(w)$ jumps up for all projects.*

4.4 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

$$dW_t = (rW_t + h(a_t))\,dt + \bar{\psi}_t \pi_t \sum_{i=1}^{N} \sigma_i dB_{it} \tag{20}$$

and the IC constraint becomes

$$a_{it} = 1 \Rightarrow \bar{\psi}_t \quad \geq \quad \frac{h}{\mu_i} \tag{21}$$

Hence, the firm will choose a certain portfolio of projects, and the project ι with the lowest NPV will determine $\bar{\psi}$, 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 outlied 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 foregos $\mu_i dt$ in payoff, and by the representation (20), will suffer a reduction in continuation value by $-\mu_i\bar{\psi}$. The deviation is not profitable when $\bar{\psi} \geq \frac{h}{\mu_i}$,

[8] 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 will cause a discrete jump in j'', and hence in b_i for all projects.

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 will thus have

$$\bar{\psi}_t = \max_{i:a_{it}=1} \frac{h}{\mu_i} \tag{22}$$

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

$$\mathrm{NPV}_i \geq h \frac{\bar{\psi} - 1}{\bar{\psi}} \tag{23}$$

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

Since the hurdle NPV will depend on both the current scaled continuation value w as well as the current project selection, it will shift 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 will now be given by

$$rj(w) = \sup_{c,a,i} \sum_i \mu_i - \kappa(i) + j'(w)\left((r - i + \delta)w + hn\right) + j''(w)\frac{1}{2}\bar{\psi}^2 \sum_i \sigma_i^2 - (i - \delta)j(w) \tag{24}$$

The main difference to equation (10) 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.

4.5 Investment in Projects

I now consider the question how agency conflicts affect the allocation of funds between competing 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

$$\sum_i \pi_i \leq k\pi \tag{25}$$

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

$$dx_{it} = \left(\pi_{it}\mu_i a_{it} dt + \sigma_i \pi_t dB_{it}\right) \tag{26}$$

Total instantaneous cash flow will hence follow

$$dx_t = \pi_t \left(\sum_i \tilde{\pi}_{it}\mu_i a_{it} dt + \sigma_i dB_{it}\right) \tag{27}$$

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

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

$$\psi_i \geq h\frac{\sigma_i}{\mu_i}\frac{1}{\tilde{\pi}_i} \tag{28}$$

Hence, 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}, and 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 (10) will now change to

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

where λ is the Lagrange multiplier associated with resource constraint (25). Given project i is implemented, it's capital allocation will now solve the FOC

$$\mu_i - \lambda - j''(w)\frac{h^2}{\mathrm{SN}_i^2}\frac{1}{\tilde{\pi}_i^3} = 0 \tag{29}$$

which implies that

$$\tilde{\pi}_i = \left(\frac{-j''(w)\, h^2}{\mathrm{SN}_i^2 (\lambda - \mu_i)} \right)^{\frac{1}{3}} \tag{30}$$

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

As the following Lemma shows, project funding will increase 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. *We have $\lambda'(w) < 0$ and project funding will be increasing in w whenever $\mu_i - \lambda > \frac{-\lambda'(w)}{j'''(w)} j''(w)\, h$ and decreasing otherwise.*

Proof. We have

$$\frac{\partial \tilde{\pi}_i}{\partial w} = \frac{1}{3} \left(\frac{j''(w)\, h^2}{\mathrm{SN}_i^2 (\mu_i - \lambda)} \right)^{-\frac{2}{3}} h^2 \mathrm{SN}_i^2 \frac{j'''(w)(\mu_i - \lambda) + \lambda'(w) \cdot j''(w)}{(\mu_i - \lambda)^2 \mathrm{SN}_i^4} \tag{31}$$

which is positive whenever the condition holds. To see that $\lambda'(w) < 0$ note that by the resource constraint

$$\sum_i \frac{\partial \tilde{\pi}_i}{\partial w} = 0 \tag{32}$$

and if $\lambda'(w) > 0$, $\frac{\partial \tilde{\pi}_i}{\partial w} < 0$ for all w and i, so that the equation above cannot hold. □

5 Implementation

In this section I discuss how the optimal contract can be implemented. I consider two setups, one where the firm can issue equity on individual projects, and one where it cannot. In the first case, the firm will hold cash balances and give the manager equity shares in the projects. These shares are vested, in the sense that insufficient performance of the manager will lead to him losing shares, while the manager will be granted new shares for opening projects if past performance was high.

When only shares in the firm and not the projects can be issued however, equity is not sufficient to implement the optimal contract. The intuition for this is analogous to Section

4.4, and I show that the hurdle rate allocation from Section 4.4 will be implemented when equity is issued on the firm level. To achieve the second best, it will be necessary to introduce a measure which takes managerial performance in the individual projects into account, and I discuss how this measure corresponds to bonus structures seen in practice.

5.1 Project-Specific Equity

As a benchmark, I describe an implementation when the firm can issue equity on the individual projects, and uses cash holdings to finance it's operations. Let M_t denote the total stock of cash, which can be allocated among the projects so that $M_t = \sum_i M_{it}$ where M_{it} is the stock of cash associated with project i. We have

$$dM_{it} = rM_{it}dt + dX_{it} - d\mathrm{Div}_{it} - dc_{it} - \alpha_i \kappa (i) \, dt \tag{33}$$

The first term is the interest earned on the current stock of cash at rate r. The cash stock has inflows in form of the project's random cash flow dX_{it} each period and outflows in terms of the dividends paid on the equity issued for project i, as well as the payout for the agent. The terms dc_{it} are for accounting purposes only. In fact, since we know that the agent must receive a payout dc_t when w hits $\bar{w}$, any assignment of payouts to the projects such that $\sum_i dc_{it} = dc_t$ will yield the same result. The same holds true for the assignment of investment costs towards projects, which are split according to share α_i, with $\sum_{i:a_i=1} \alpha_i = 1$. Suppose that the equity holders require a minimal dividend which must satisfy

$$d\mathrm{Div}_{it} = (r - \gamma) \, M_{it}dt - \alpha_i \kappa (i) \, dt \tag{34}$$

In addition, the manager is endowed with a personal account[9] with balance A_t, which pays interest at rate γ and has the following features. At the beginning of the contract, $A_0 = 0$ and the manager is endowed with an equity share Ψ_{it} in each project which is executed at $t = 0$. Whenever a new project is executed, the manager buys equity in the project at a pre-determined price and whenever a project is halted, the manager sells the equity share back to the shareholders at price a certain price. Proceeds from buying and selling shares are deposited into the personal account.[10] Finally, assume that the manager may not access funds inside the account, except for when a dividend dc_t is paid.

[9]We can think of the account either in terms of cash or incentive points - the distinction turns out to be irrelevant.

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

Then we have the following Proposition.

Proposition 7. *Suppose that the firm holds a cash balance M_t which satisfies $\sum_i M_{it} = M_t$ as well as equation (33), and that the minimal dividend process satisfies (34). Further, at time t the manager holds $\Psi_{it} = \frac{\psi_{it}}{\sigma_i}$ equity shares in project i, is endowed with an personal account with balance A_t which pays interest at rate γ and decides on the implementation of projects as well as the payout policy.*

Whenever the sum of personal account and the managerial share of the firm's cash holdings $\sum_i \Psi_{it} M_{it} + A_t$ equals zero, the agent is fired and the firm is shut down, and whenever a project is started the manager buys shares according to a pre-determined price which corresponds exactly to M_{it} and whenever the project is stopped the manager sells his shares at the same price. Proceeds from the purchase and sales of shares are going towards the personal account.

Then the contract from Section 4.4 will be implemented and the agent's continuation value will satisfy $W_t = \sum_i \Psi_{it} M_{it} + A_t$.

Intuitively, while the equity shares serve to expose the manager to just enough risk to incentivize effort, the cash balances and dividend policy are ensuring that the manager's continuation value grows at exactly the right rate.

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 would be one where if the manager performs badly, and a project is halted, he gets stripped of the equity share without a compensation. Then, however, given the contract is implemented with cash and equity only, the continuation value could not be continuous, since the total value of his equity stake would jump downward. More importantly, the manager would know that his continuation value will undergo a jump in the future, which in turn may distort his incentives right before being stripped of his share.

To counter this effect, it is necessary to compensate the manager whenever shares are either awarded or taken away. This leads to the interpretation that the manager will instead be required to either buy or sell shares to the shareholders at pre-determined transfer prices, which are set precisely ot the amount at which the transfer does not distort incentives while ensuring the optimal level of risk exposure to the manager.[11]

Further, the contract will now depend not only on the cash holdings of the firm, as in

[11] Note that 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 works, 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.

DeMarzo et al. (2010), but 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 firm is terminated.[12] 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.

Finally, note that at the optimal contract the equity shares satisfy $\Psi_{it} = \frac{h}{\mu_i}$. This is because the process dM_{it} already carries a Brownian noise component of with volatility σ_i from the output process dX_{it}. Therefore, the manager will hold less shares in projects with higher NPV, independent of their risk.

5.2 Firm Level Equity

In reality, corporations issue stock based on the entire firm's performance, instead of individual subdivisions, or even plants or R&D projects and hence the previous discussion is best understood as a benchmark case. 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.

Formally, suppose the agent is granted an equity share $\bar{\Psi}_t$, which grants him a fraction of ownership in the firm's total cash holdings $M_t = \sum_i M_{it}$, as well as any dividend payments, should they arise. Now, the dividend process satisfies simply

$$d\text{Div}_t = (r - \gamma) M_t dt - \kappa(i) dt \tag{35}$$

while the cash holdings process follows

$$dM_t = rM_t dt + dX_t - dc_t - d\text{Div}_t - \kappa(i) dt \tag{36}$$

and again the personal account A_t is used to escrow proceedings from the manager's equity transactions. Since $\bar{\Psi}_t$ effectively conditions the managerial equity share on the total output of the projects, $dX_t = \sum_{it} dX_{it}$ it comes at no surprise that the allocation implemented is not the optimal contract, but the hurdle-rate contract in Section (4.4).

Proposition 8. *Suppose that the firm holds a cash balance M_t which satisfies (36), and that*

[12]Note that I interpret either $M_{it} < 0$ or $A_t < 0$ as the firm taking on short term debt.

the minimal dividend process satisfies (35)*. Further, at time t the manager holds Ψ_t equity shares, is endowed with an personal account with balance A_t which pays interest at rate γ and decides on the implementation of projects as well as the payout policy.*

Whenever $\Psi_t M_t + A_t$ reaches zero, the agent is fired and the firm is shut down, and whenever a project is started the manager buys shares according to a pre-determined price which corresponds exactly to M_t and whenever the project is stopped the manager sells his shares according to the same price. Proceeds from the purchase and sales of shares are going towards the personal account.

Then the contract from Section 4 will be implemented and the agent's continuation value will satisfy $W_t = \Psi_t M_t + A_t$.

Building on Proposition 7, we can show intuitively why only the hurdle rate contract will be implemented. Given an equity share Ψ_t, the managerial share of the firm's cash holdings responds to the manager's effort decision in project i according to $\Psi_t \pi_t \mu_i a_{it} dt$, while at the same time the manager incurs an effort cost of $\pi_t h a_{it} dt$. Consequently, effort is only optimal if $\Psi_t \geq \frac{h}{\mu_i}$. Since this must hold for each project which is implemented, we have that

$$\Psi_t = \max_{i:a_i=1} \frac{h}{\mu_i} = \max_{i:a_i=1} \Psi_{it} \tag{37}$$

and the risk exposure is the same as in Section 4.

Proposition 8 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, my results suggest that the widespread use of hurdle rates may not be optimal, as for example Berkovitch and Israel (2004) suggest, but instead the result of flawed incentive contracts.

5.3 Firm Level Equity and Bonus Contracts

In order to implement the optimal contract, it is necessary introduce a project-dependent component into the implementation. There are of course are many ways to achieve this. For instance, the firm could set up an incentive point account which is designed to exactly mimic the evolution of the manager's continuation value in equation (6).

My implementation features an equity share and bonus payments to the manager, which are made contingent on his performance in the different projects. In a survey on managerial compensation Murphy (1999) finds that indeed the majority of managerial incentive contracts feature a mix of equity and boni, and that the total bonus is a sum of boni in individual categories, which are generally specified by the shareholders. Hence, my model serves as a rationalization of these practices.

Suppose that in addition to the instruments in Proposition 8, the firm makes payments dP_t to the manager's personal account according to the performance of the individual projects, and that the amount of equity issued Ψ_t satisfies $\Psi_t = \min_{i:a_i=1} \Psi_{it}$.

To implement the optimal contract, suppose now that the investor can pay a bonus directly to the agent contingent on his performance in each project. In particular, the payment process follows

$$dP_t = \sum_i (\Psi_{it} - \Psi_t)\, dX_{it} \tag{38}$$

which is dependent on Ψ_{it}, the equity stake in project i that would be optimal if equity on the individual projects could be issued. In this implementation, the bonus payments thus act towards providing the manager with additional project dependent incentives in excess of those already provided by the equity stake.

Proposition 9. *Suppose that the managerial incentive contract is just as in Proposition 8 with the exception of a bonus payment process dP_t which satisfies equation (38) and is escrowed in the manager's personal account, and the managerial equity share satisfies $\Psi_t = \min_{i:a_i=1} \Psi_{it}$. 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 not provide unnecessary risk exposure, 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

It is worth comparing my model and it's predictions with the theory of real options[13], which next to NPV analysis is the canonical framework used to evaluate investment decisions. In

[13] See, e.g. Dixit et al. (1994) or Abel and Eberly (1994).

the real options framework, an investor needs to decide whether, and when, to undertake a project that carries a fixed cost and whose value changes stochastically over time. The main insight of real options theory is that simply following the NPV criterion, i.e. investing whenever NPV>0 is not optimal. This is due to the fact that there exists an 'option value' to investment, and exercising that option entails a loss of option value. Similarly, if starting a project gives access to follow-up projects, investing may be associated with acquiring an option value. Since the NPV criterion does not take this into account, investment decisions will be suboptimal.

Superficially, my model provides similar results to the ones obtained by the real options literature. For instance, projects are not executed if NPV>0 but if it is above a threshold depending on the project's riskiness, and there will be 'delay' in the sense that positive NPV projects will not be started. Note however, that the mechanisms which drive the results in both frameworks are entirely different. With real options, not exerting a positive NPV project simply means that the current NPV is not enough to compensate for the loss of option value, and hence the delay is in fact efficient, whereas in my model, delay relative to the NPV criterion stems from the fact that moral hazard induces an additional cost the principal has to take into account when implementing projects, and would not occur in the first best. My model generates this pattern entirely without relying on irreversibility or sunk costs, which is crucial in the real options literature, and without which the option value would necessarily be zero.

Further, in a real options framework, it is well possible that a negative NPV project is executed, a fact that is generally attributed to issues such as empire building and private benefits to managers which cannot be remedied by contractual arrangements. This happens exactly when starting the project gives access to another option, p.e. for follow up projects, which compensates for both the negative NPV and the loss of the original option, which is shown by Baldwin (1982) and Roberts and Weitzman (1981) in the contexts of product market competition and sequential investment.

My framework provides a different rationalization for this phenomenon. Due to limited liability, if the firm is close to termination, the principal cannot drive the continuation value up, and thereby effectively punish the agent for past failures, by demanding cash payments. In this situation, the only way to lower the agent's current utility is to mandate more projects. For the principal, this will come at a benefit since it will push the agent's continuation value away from the termination boundary and thereby lower the termination probability. Alternatively, a low w implies that in expectation, the principal does not need to reimburse the agent for his entire effort cost. In this sense, the principal values the cost of effort differently under moral hazard, and will undertake projects whose NPV is negative, since he does not expect to pay the full amount of the accrued effort costs.

Finally, real options models carry an inherent difficulty when it comes to characterizing the simultaneous choice of projects. This is because the fixed costs of starting, and potentially terminating projects will in general imply that the marginal benefit of each project depends on the entire chosen portfolio in ways which are difficult to characterize unless an explicit solution to the decision maker's value function can be obtained.

The following Section illustrates how my model can serve as an approximation to a real-options framework with multiple projects.

6.2 Project Choice under Fixed Costs

To study the relation between a real options framework and my setup, I introduce fixed costs of project execution. In particular, assume that projects can be started and shut down at will as before, but executing, or re-starting, a project carries a fixed cost $k_e > 0$, whereas shutting the project down entails cost $k_s > 0$. Therefore, whenever a project is not being executed, the firm holds a real option to start the project, while when the project is implemented, the firm holds an option to halt it.[14]

Let $\{\tau_n^{i,e}\}_{n=1}^{\infty}$ denote the sequence of stopping times at which project i is started or re-started, and let $\{\tau_n^{i,s}\}_{n=1}^{\infty}$ denote the analog for when project i is halted. Finally, the total fixed cost of switching from portfolio a to portfolio a' is

$$k\left(a, a'\right) = \sum_{i: a_i = 1 \wedge a_{i'} = 0} k_s + \sum_{i: a_{i=0} \wedge a_i = 1} k_e \tag{39}$$

Under the optimal policy, the principal's value function will be now given by

$$\begin{aligned} J(W, \pi) &= E\left(\int_0^\tau e^{-rt}\left(\pi_t \sum_{i=1}^N \mu_i a_{it} dt - dc_t - \pi_t \kappa\left(i_t\right) dt\right) + \right. \\ & \left. + e^{-r\tau}\gamma\pi_t - \sum_{i=1}^N \sum_{n=1}^\infty e^{-r\tau_n^{i,s}} k_s - \sum_{i=1}^N \sum_{n=1}^\infty e^{-r\tau_n^{i,e}} k_e \right| \mathcal{F}_0 \Bigg) \end{aligned} \tag{40}$$

and the principal will face the same constraints as in (8) of Section 3.

[14] Note that none of my proofs rely on this specification, and the results will hold for any cost function $k\left(a, a'\right) > 0$ for $a \neq a'$.

Introducing fixed costs affects the project selection policy in two important ways. First, project selection will not be linearly independent anymore. Thus, the currently chosen project portfolio will have 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.

Proposition 10 describes the structure of the optimal contract in the real options case.

Proposition 10. *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 $\phi \in C^2$*

$$\mathcal{L}_{i,a}\phi\left(w\right) = \left(\left(r - i + \delta\right)w + h\sum_i a_i\right)\phi'\left(w\right) + \phi''\left(w\right)\frac{1}{2}\sum_i \psi_i^2 \tag{41}$$

The solution to problem (40) is determined by the following system of quasi variational inequalities for all a and w.

$$\min\left\{rj_a - \mathcal{L}_{i,a}j_a - \sum_i \mu_i + \kappa\left(i\right) - \left(i - \delta\right)j_a, j_a - \max_{a' \neq a} j_{a'} - k\left(a, a'\right)\right\} = 0 \tag{42}$$

For any w and a, $j_a \in C^1$ will satisfy the above equation in a viscosity sense.

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

$$\begin{aligned} j_a\left(w\left(a, a'\right)\right) &= j_{a'}\left(w\left(a, a'\right)\right) \\ j'_a\left(w\left(a, a'\right)\right) &= j'_{a'}\left(w\left(a, a'\right)\right) \end{aligned} \tag{43}$$

And

$$j''_a\left(w\left(a, a'\right)\right) \geq j''_{a'}\left(w\left(a, a'\right)\right) \tag{44}$$

Due to the fixed costs of setting up and scrapping projects, the value function is dependent on the current project portfolio a, which is expressed by the notation $j_a(w)$. Equation (42) encodes the optimal choice of a as a function of w. A particular project selection a will be optimal for some range of w, in which case

$$j_a\left(w\right) > \max_{a' \neq a} j_{a'}\left(w\right) - k\left(a, a'\right) \tag{45}$$

In line with the previous notation, I label this open set the continuation region $\mathcal{C}_a$, and it's complement the switching region $\mathcal{S}_a$. Equation (42) implies that on $\mathcal{C}_a$, the analog of the

HJB equation without fixed costs, equation (10), will hold, albeit under a different set of boundary conditions, which are given by (43).

Note that equation (45) does not imply

$$j_a(w) > \max_{a' \neq a} j_{a'}(w) \tag{46}$$

so that we have a hysteresis effect. Hence, the firm may stick to a locally suboptimal project portfolio, which happens exactly when the potential benefit of changing the portfolio does not outweigh the fixed costs.

An important feature of the contract under switching costs is that the principal's value function will not be C^2 at the thresholds $w(a, a')$. Recall that in the case without fixed costs, the project portfolio had a simple structure. Only projects with positive marginal benefit $b_i(w)$ were chosen, and for a particular w, the level of the benefit for each i is independent of whether or not other projects are chosen. Equation (13) implies that the marginal benefit of each project will be in part determined by the amount of risk exposure ψ_i, which is needed to incentivize effort. In the eyes of the principal, the additional risk exposure is weighted by $j''(w)$, which now may be discontinuous, and in particular

$$j_a''(w(a, a')) \geq j_{a'}''(w(a, a')) \tag{47}$$

Whenever the inequality is strict, changing the project portfolio implies a spillover effect. Immediately after the change, the firm becomes more risk averse, which implies a downward jump in the valuation of all projects, whether executed or not. Therefore, in the real options case, the optimal portfolio of projects has to be determined jointly and the linear decomposition that underlied equation (13) breaks down.

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

Proposition 11. *Let $\hat{w}(a, a')$ and $w(a, a')$ denote the threshold at which the optimal action changes from a to a' in the case without*[15] *and with fixed costs respectively, and let j and j_a denote the value functions for both cases. Then as $\max\{k_e, k_s\} \to 0$ we have $\hat{w}(a, a') \to w(a, a')$, $j_a(w) \to j(w)$ $\forall w$ and $\left|j_{a+}''(w) - j_{a-}''(w)\right| \to 0$ $\forall w$.*

Proposition (11) specifies in which sense we may take the model in Section 3 as an approximation to a real-options style model with fixed costs. When these costs are small, the value

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

functions will converge towards j, the value function without costs and the marginal benefit function b_i will be 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 will coincide. Therefore, when the fixed costs are small enough we can ignore the direct spill-over effects between projects and determine the project portfolio as in the previous case.

Finally, it is noteworthy that Proposition 10 implies an interesting caveat. Suppose that the new portfolio has strictly more projects than the old one, that is $\forall i\ a_i = 1 \Rightarrow a_{i'} = 1$ and $\exists j$ such that $a_j = 0$ and $a_i = 1$. Then the optimality conditions around the threshold $w(a, a')$ imply

$$\begin{aligned}\sum_{i:\text{ added}} \left(\text{NPV}_i + \left(j'_{a'}(w) + 1\right) + j''_{a'}(w) \frac{1}{2} \frac{h^2}{\text{SN}_i^2} \right) &= \left(j''_a(w) - j''_{a'}(w)\right) \frac{1}{2} \sum_{i:a_i = a_{i'} = 1} \frac{h^2}{\text{SN}_i^2} \\ &\quad + \sum_{i:\text{ added}} k_e\end{aligned}$$

which we can interpret as

$$\sum_{i:\text{ added}} b_{i,a'}(w) = \text{Risk Aversion Spillover} + \text{ Fixed Costs} \tag{48}$$

Thus, whenever the firm adds more projects, the total marginal benefit of the projects added must, from the viewpoint after the addition, exceed not only the fixed costs, but also the spillover effect on existing projects.

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 in a firm under moral hazard, as well as the distribution of funds among projects and the persistence of bonus contracts in CEO compensation. Further, it explains the use of different criteria to evaluate projects aside from the NPV criterion, which is broad practice in companies, as Graham and Harvey (2001) shows.

The optimal project selection policy implies that projects are selected whenever their NPV is above a cutoff depending on the firm's current free cash flow, as well as the project's risk-return ratio. This cutoff will change stochastically over time, and depend on the agent's cumulative past performance. Firms with a large free cash flow relative to firm size will feature a relatively efficient investment portfolio, comprised of high-risk, positive return projects, whereas firms with a low cash flow will suffer from an inefficient choice of investment projects, passing up positive NPV projects when their risk is too high. At any given point in time, the absolute benefit of projects with above-average risk increases with the free cash flow, and decreases whenever projects are relatively safe. The first best project selection schedule is attained whenever the free cash flow is large enough. The manager is assigned more and riskier projects after a history of sufficiently good performance, while a poorly performing manager will see the number of projects assigned to him diminished. 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 grows. 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

Given the combined stopping and control problem in (8), the principal's HJB equation satisfies the following HJB variational inequality.

$$0 = \max\left\{\sup_{(c,a,i)} -rJ + \pi\sum_i \mu_i a_i - dc + J_w\left(rW + \pi h\sum_i a_{it}\right) + \frac{1}{2}J_{ww}\pi^2\sum_i \psi_i^2 \right.$$
$$\left. + J_\pi \pi (i-\delta) \quad , \quad \gamma\pi - J(W,\pi)\right\} \tag{49}$$

That is, there will be a continuation region $\mathcal{C} \subset \mathbb{R}_+^2$ such that for values $(W,\pi) \in \mathcal{C}$, the firm will continue to operate, and shutdown will occur on the boundary of $\mathcal{C}$. When the firm is shut down, the principal's value satisfies $\gamma\pi = J(W,\pi)$, and on $\mathcal{C}$, we have the HJB equation

$$rJ(W,\pi) = \sup_{(c,a,i)} \pi\sum_i \mu_i a_i - \pi\kappa(i) - dc + J_w\left(rW + \pi h\sum_i a_{it}\right) + \frac{1}{2}J_{ww}\pi^2\sum_i \psi_i^2 + J_\pi \pi(i-\delta) \tag{50}$$

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 $J_{ww} = \frac{1}{\pi}j''(w)$, the HJB equation (50) can be converted to equation (10). Since both laws of motion (2) and (6), as well as the termination value $\gamma\pi$ obey the same scaling, this implies that the control problem in (8) is equivalent to the scaled control problem in w alone. Finally, since dc_t is an impulse control, there will be a region on $(0,\infty)$ such that $dc_t = 0$, and w will reflect on the boundary of this region.

A.2 Properties of the Value Function

In this section I will show the main properties of the value function. Since the smooth pasting and value matching properties of the value function at $\bar{w}$ are well known, I shall focus on establishing these properties at the cutoffs at which the project choice changes. That the value function is twice continuously differentiable for stochastic control problems with a continuous action space is well established, and the standard sources in control theory Oksendal (2003), generally make no distinction between control with continuous or discrete actions. My proof here is general, in the sense that the same argument can be applied to any control problem with discrete actions, without having to recourse to arguments specific to the problem at hand.. I use the viscosity solution approach, and the first two Lemmata are based on the literature on optimal switching, see, p.e. Pham et al. (2009).

Lemma 12. *Assume that for all admissible investment policies i, there exists an $\varepsilon > 0$ such that for all t, $i_t < r + \delta - \varepsilon$. Then, for all (W, π),$J(W, \pi) \in C$.*

Proof. I actually prove a stronger condition, Lipschitz continuity. Consider an arbitrary policy (a, i, τ) consisting of project choice a, investment i and termination time τ.

Consider the policy a. We can equivalently express project choice as an impulse control process. In particular, each policy a will induce a sequence of stopping times $\{\tau_i\}_{i=0}^{\infty}$ with $\tau_0 = 0$ which index the times at which project choice changes as well as a sequence of vector stopping times $\{\eta_i\}_{i=0}^{\infty}$ with $\eta_i \in \{0, 1\}^N$ and $\eta_0 = a(w_0)$ which indicates the project allocation policy just after τ_i. That is, for any t, $a_t = \sum_{i=1}^{\infty} 1\{\tau_{i-1} \leq t < \tau_i\}\eta_i$.

Then we can express the process w_t as

$$w_t = w_0 + \int_0^t \left(\gamma w_s + h \sum_{i=1}^{N} a_s\right) dt + \int_0^t \sum_{i=1}^{N} \frac{\sigma_i^2 h^2}{\mu_i^2} a_{it} dB_{it} \tag{51}$$

and for W_t and π_t we have

$$\pi_t = \pi_0 + \int_0^t \pi_s (i_s - \delta) ds \tag{52}$$

and

$$W_t = W_0 + \int_0^t \left(\gamma W_s + h\pi_s \sum_{i=1}^{N} a_s\right) ds + \int_0^t \pi_s \sum_{i=1}^{N} \frac{\sigma_i^2 h^2}{\mu_i^2} a_{it} dB_s \tag{53}$$

And under the policy, the value function is

$$J(W, \pi) = E\left(\int_0^{\tau} e^{-rt} \pi_t \left(\sum_i \mu_i a_{it} - \kappa(i_t)\right) dt\right) \tag{54}$$

Take two pairs (W, π) and (W', π') under the same policy, but where the Laws (53) and (52) exhibit $W_0 \neq W_0'$ and $\pi_0 \neq \pi_0'$. Then we have

$$\begin{aligned}
\left|J\left(W,\pi\right)-J\left(W',\pi'\right)\right| &\leq \sup_{i,a,\tau} E\left(\int_0^\tau e^{-rt}\left|\left(\pi_t-\pi_t'\right)\cdot\left(\sum_i \mu_i a_{it}-\kappa\left(i_t\right)\right)\right| dt\right) \\
&\leq \sup_{i,a,\tau} E\left(\int_0^\tau e^{-rt}\left|\pi_t-\pi_t'\right|\cdot\left|\sum_i \mu_i a_{it}-\kappa\left(i_t\right)\right| dt\right) \\
&\leq \sup_{i,a,\tau} E\left(\int_0^\tau \exp\left(-rt+\int_0^t \left(i_s-\delta\right) ds\right)\cdot\left|\sum_i \mu_i a_{it}-\kappa\left(i_t\right)\right| dt\right)\left|\pi-\pi_0\right|
\end{aligned}$$

By $i_s < r+\delta-\varepsilon$ for all policies, we have

$$\left|J\left(W,\pi\right)-J\left(W',\pi'\right)\right| \leq \sup_{i,a,\tau} E\left(\int_0^\tau \exp\left(-\varepsilon t\right)\cdot\left|\sum_i \mu_i a_{it}-\kappa\left(i_t\right)\right| dt\right) \tag{55}$$

Similarly, since $\kappa'\left(i\right)>0$

$$\int_0^\tau e^{-\varepsilon t}\kappa\left(i_t\right) dt < \int_0^\tau e^{-\varepsilon t}\kappa\left(r+\delta-\varepsilon\right) dt \tag{56}$$

It is then easy to see that the transversality condition holds for equation (55), and hence

$$\left|J\left(W,\pi\right)-J\left(W',\pi'\right)\right| \leq C\cdot\left|\pi-\pi_0\right| \tag{57}$$

for some constant $C<\infty$. Since $j\left(w\right)=E\int_0^\tau e^{-rt}\sum_i \mu_i a_i-\kappa\left(i_t\right) dt$, we can apply the same logic go get Lipschitz continuity in j. $\square$

Lemma 13. *$j\in C^1$ on $\left(0,\bar{w}\right)$.*

Proof. Take two continuation regions $\mathcal{C}_a$ and $\mathcal{C}_{a'}$ with $\sup\left\{w\in\mathcal{C}_a\right\}=\inf\left\{w\in\mathcal{C}_{a'}\right\}:=w_{n+1}$. We know that on both $\mathcal{C}_a$ and $\mathcal{C}_{a'}$, j solves the HJB equation (10), and is C^2 on these regions.

Suppose that $j_-'\left(w_{n+1}\right)\neq j_+'\left(w_{n+1}\right)$ and assume $j_-'\left(w_{n+1}\right)<j_+'\left(w_{n+1}\right)$. Take any number $x\in\left(j'_-\left(w_{n+1}\right),j_+'\left(w_{n+1}\right)\right)$ and let

$$\phi_\varepsilon\left(w\right)=rj\left(w_{n+1}\right)+x\left(w-w_{n+1}\right)+\frac{1}{2\varepsilon}\left(w-w_{n+1}\right)^2 \tag{58}$$

Note that $\phi_\varepsilon\in C^2$, $\phi_\varepsilon\left(w_{n+1}\right)=j\left(w_{n+1}\right)$ and that for all ε there exists a neighborhood $B_\delta\left(w_{n+1}\right)$ such that $j>\phi_\varepsilon$ on $B_\delta\left(w_{n+1}\right)$. Then by the viscosity supersolution property of j, it must be the case that for $w \in B_\delta(w_{n+1})$ and $w < w_{n+1}$

$$(r-i+\delta)\cdot j(w_{n+1}) - x\cdot((\gamma-i+\delta)w_{n+1}+hn) - \frac{1}{2\varepsilon}\sum_i \psi_i^2 a_i \geq 0 \tag{59}$$

But sending $\varepsilon \to 0$ yields the contradiction. Hence $j'_-(w_{n+1}) \geq j'_+(w_{n+1})$.

Now suppose $j'_-(w_{n+1}) > j'_+(w_{n+1})$ and take again $x \in \left(j'_+(w_{n+1}), j'_-(w_{n+1})\right)$ and consider the function

$$\phi_\varepsilon(w) = rj(w_{n+1}) + x(w-w_0) - \frac{1}{2\varepsilon}(w-w_{n+1})^2 \tag{60}$$

For any $\varepsilon > 0$ there exists a neighborhood $B_\delta(w_{n+1})$ such that $\phi_\varepsilon > j$ on $B_\delta(w_{n+1})$ and again $\phi_\varepsilon(w_{n+1}) = j(w_{n+1})$. Hence, by the subsolution property of j, we must have

$$(r-i+\delta)\cdot j(w_{n+1}) - x\cdot((\gamma-i+\delta)w_{n+1}+hn) + \frac{1}{2\varepsilon}\sum_i \psi_i^2 a_i \leq 0 \tag{61}$$

and letting $\varepsilon \to 0$ yields the contradiction again. Hence, $j'_+(w) = j'_-(w)$ for all $w \in (0,\bar{w})$. $\square$

Lemma 14. *$j \in C^2$ on $(0,\bar{w})$.*

Proof. I only show that j'' is continuous at points where the project choice shifts. Continuity on $\bigcap_a \operatorname{Int}(\mathcal{C}_a)$ is shown in Lemma 15 and is independent of the results obtained here. Let

$$H_a(w) := \sup_i -(r-i+\delta)j(w) + \sum_i \mu_i - \kappa(i) + j'(w)((r-i+\delta)w+hn) + j''(w)\frac{1}{2}\sum_i \psi_i^2 \tag{62}$$

and by the HJB equation, $H_a(w) \leq 0\ \forall w < \bar{w}$. Consider wlog two continuation regions $\mathcal{C}_a$ and $\mathcal{C}_{a'}$ with $\sup\{w \in \mathcal{C}_a\} = \inf\{w \in \mathcal{C}_{a'}\} := w_{n+1}$.

We have

$$\sup_i -(r-i+\delta)j(w) + \sum_i \mu_i - \kappa(i) + j'(w)((\gamma-i+\delta)w+hn) + j''(w)\frac{1}{2}\sum_i \psi_i^2 = 0 \tag{63}$$

if and only if

$$j''(w) = \inf_{a,i} \frac{1}{\frac{1}{2}\sum_i \psi_i^2}\left((r-i+\delta)j(w) - \sum_i \mu_i a_i + \kappa(i) - j'(w)((\gamma-i+\delta)w+hn)\right) \tag{64}$$

since $\sum_i \psi_i^2 > 0$ on all paths.

Let

$$\tilde{j}_a(w) := \inf_i \frac{1}{\frac{1}{2}\sum_i \psi_i^2}\left((r-i+\delta)\,j(w) - \sum_i \mu_i a_i + \kappa(i) - j'(w)\left((\gamma-i+\delta)\,w + hn\right)\right) \tag{65}$$

Clearly, $j''(w) = \tilde{j}_a(w)$ on $\mathcal{C}_a$ and $j''(w) = \tilde{j}_{a'}(w)$ on $\mathcal{C}_{a'}$. Also, by the properties of the inf-operator, $\tilde{j}_a(w) < \tilde{j}_{a'}(w)$ on $\mathcal{C}_a$ and $\tilde{j}_{a'}(w) < \tilde{j}_a(w)$ on $\mathcal{C}_{a'}$. Taking limits again we have

$$\tilde{j}_{a,-}(w_{n+1}) \leq \tilde{j}_{a',-}\left(w_{n+1}\right) \tag{66}$$

and

$$\tilde{j}_{a,+}(w_{n+1}) \geq \tilde{j}_{a',+}\left(w_{n+1}\right) \tag{67}$$

But note that for fixed a, $\tilde{j}_a(w)$ is a continuous function in w and hence $\tilde{j}_{a,+}(w_{n+1}) = \tilde{j}_{a,-}(w_{n+1})$. Then we have the following series of inequalities

$$j''_-(w_{n+1}) = \tilde{j}_{a,-}(w_{n+1}) = \tilde{j}_{a,+}(w_{n+1}) \geq \tilde{j}_{a',+}\left(w_{n+1}\right) = j''_+(w_{n+1}) \tag{68}$$

and hence

$$j''_-(w_{n+1}) \geq j''_+(w_{n+1}) \tag{69}$$

To obtain the reverse inequality, use equation (66) to get

$$j''_-(w_{n+1}) = \tilde{j}_{a,-}(w_{n+1}) \leq \tilde{j}_{a',-}\left(w_{n+1}\right) = \tilde{j}_{a',+}\left(w_{n+1}\right) = j''_+(w_{n+1}) \tag{70}$$

Thus it must be the case that $j''_+(w_{n+1}) = j''_-(w_{n+1})$ and $j \in C^2$. Note that the case when one less project is executed is analogous. □

Lemma 15. *$j''(w)$ is negative for all $w \in (0, \bar{w})$.*

Proof. The proof is exactly analogous to DeMarzo et al. (2010), with the additional complication of having continuation regions, which is easliy resolved. First we note that for w close to $\bar{w}$, we have that w will be in some continuation region and hence

$$j'''(w) = \frac{1}{2}\frac{1}{\sum_i \psi_i^2}(\gamma - r) > 0 \tag{71}$$

Therefore, since $j''(\bar{w}) = 0$ we have $j''(w) < 0$. Now consider a $\tilde{w}$ such that $j''(\tilde{w}) = 0$ and $j''(\tilde{w} + \varepsilon) < 0$. We notice that $\tilde{w}$ is again in some continuation region and that j'' is continuous at $\tilde{w}$, so that a version of the HJB equation will hold. Then the remainder of the argument in DeMarzo et al. (2010) goes through, and we have $j''(w) < 0$ everywhere. $\square$

Lemma 16. *$j'''(w)$ exists on $int(\mathcal{C}_a)$ for every a and if $\gamma - r$ is small, we have $j'''(w) > 0$ $\forall w \in (0, \bar{w})$.*

Proof. Rewriting the scaled HJB equation (10),

$$j''(w) = \inf_{c,a,i} \frac{1}{\frac{1}{2}\sum_i \psi_i^2} \left((r - i + \delta) j(w) - \sum_i \mu_i a_i + \kappa(i) - j'(w)((\gamma - i + \delta) w + hn) \right) \tag{72}$$

By the envelope theorem,

$$j'''(w) = \frac{1}{\frac{1}{2}\sum_i \psi_i^2} \left((r - \gamma) j'(w) - j''(w)((\gamma - i + \delta) w + hn) \right) \tag{73}$$

and we can see that $j'''(w) \lesseqgtr 0$ whenever

$$(r - \gamma) j'(w) - j''(w)((\gamma - i + \delta) w + hn) \lesseqgtr 0 \tag{74}$$

Note that whenever $j'(w) < 0$ it must necessarily be the case that $j'''(w) > 0$, since $j'' < 0$ and $r < \gamma$. Therefore, the only case of concern is when $j'(w) > 0$.

Suppose that $j'''(w) < 0$. For any γ, w will be in some continuation region, so that some version of the HJB equation will hold, and therefore

$$j''(w) \left(1 - \frac{(\gamma - i + \delta) w + hn}{(\gamma - r)\sum_i \psi_i^2 \frac{1}{2}} \right) < \frac{(r - i + \delta) j(w) + k(i) - \sum_i \mu_i a_i}{\sum_i \psi_i^2 \frac{1}{2}} \tag{75}$$

Since for $\gamma \to r$, the term multiplying $j''(w)$ will tend to negative infinity, it must be negative for γ close enough to r. In this case,

$$j''(w) > \frac{(r - i + \delta) j(w) + k(i) - \sum_i \mu_i a_i}{\sum_i \psi_i^2 \frac{1}{2}} \cdot \left(1 - \frac{(\gamma - i + \delta) w + hn}{(\gamma - r)\sum_i \psi_i^2 \frac{1}{2}} \right)^{-1} \tag{76}$$

Note that as $\gamma \to r$, the RHS goes to zero.

By a similar agrument we have

$$j'(w) < \frac{(r-i+\delta)\,j(w)+k(i)-\sum_i \mu_i a_i}{\sum_i \psi_i^2 \frac{1}{2}} \left(\frac{(\gamma-i+\delta)\,w+hn}{\sum_i \psi_i^2 \frac{1}{2}} + \frac{r-\gamma}{(\gamma-i+\delta)\,w+hn} \right)^{-1} \tag{77}$$

which will converge to

$$j'(w) \leq \frac{(r-i+\delta)\,j(w)+k(i)-\sum_i \mu_i a_i}{(\gamma-i+\delta)\,w+hn} \tag{78}$$

as $\gamma \to r$. Let $K = (r-i+\delta)\,j(w)+k(i)-\sum_i \mu_i a_i$. Hence, if $j'''(w) < 0$ and $K > 0$ we have that for small γ, $j''(w) > 0$, which is a contradiction. Similarly, for $K \leq 0$ we must have that $j'(w) \leq 0$. But then since $j''(w) < 0$ we can't have $j'''(w) < 0$. Therefore we must have $j'''(w) \geq 0$ everywhere. □

Lemma 17. *The contract is optimal in the class of contracts with square integrable payout policies and firm sizes.*

Proof. Take any arbitrary contract which satisfies the agent's IC constraint. Then, define the auxiliary gain process G_t as

$$G_t = \int_0^t e^{-rs} \left(\sum_i dX_{is} - \pi_s \kappa(i_s)\,dt - dc_s \right) + e^{-rt} J(W_t, \pi_t) \tag{79}$$

By Ito's Lemma,

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

By the scaled HJB equation (10), it has to be the case that for any contract,

$$-(r-i+\delta)\,j(w) + \sum_i \mu_i - \kappa(i) + j'(w)\left((\gamma-i+\delta)\,w+hn\right) + j''(w)\frac{1}{2}\sum_i \psi_i^2 - (i-\delta)\,j(w) \leq 0 \tag{80}$$

Similarly, for any consumption payout policy, $-dc_t\left(1+j'(w)\right)\leq 0$ on $(0,\bar{w})$. Hence, the drift of the auxiliary gain process is non positive and the principal's expected time-zero profit is less than $J\left(W_0,\pi_0\right)$. $\square$

A.3 Proofs On Project Choice

The proof of Proposition 3 follows directly from equation (13) and the arguments outlined in the text. The proof for Proposition 4 is below.

Proposition 18. *For any w, and project allocation a, there exist a cutoff $\overline{SN}(w)$ such that for all $SN_i>\overline{SN}(w)$, $b_i'(w)<0$ and $SN_i<\overline{SN}(w)$, $b_i'(w)>0$.*

Proof. I first establish the cutoff $\overline{\mathrm{SN}}(w)$. We have $b_i'(w)=j''(w)\,h+j'''(w)\,\frac{1}{2}\frac{1}{\mathrm{SN}_i^2}$, hence $\overline{\mathrm{SN}}(w)=\sqrt{-\frac{j'''(w)\frac{1}{2}}{j''(w)h}}$. Note that, $-\frac{j'''(w)\frac{1}{2}}{j''(w)h}=\frac{1}{2}\frac{(\gamma-i+\delta)w+hn}{h\sum_i\psi_i^2}$ so that

$$\frac{d\overline{\mathrm{SN}}(w)}{dw}\propto(\gamma-i+\delta)-i'(w)\cdot w=(\gamma-i+\delta)+\frac{j''(w)\,w^2}{\kappa''(i)} \tag{81}$$

so $\frac{d\overline{\mathrm{SN}}(w)}{dw}>0$ when $i'(w)$ is relatively small.

Further, we have

$$\begin{aligned} b_i'(w) &= j''(w)\left(h-\frac{1}{2}\frac{1}{\sum_j\psi_j^2}\psi_i^2\left((\gamma-i+\delta)\,w+hn\right)\right)\\ &\propto j''(w)\left(h\left(\sum_j\psi_j^2-\psi_i^2 n\right)-(\gamma-i+\delta)\,w\right)\end{aligned}$$

using (**??**) and hence a sufficient condition is $\frac{1}{n}\sum_j\psi_j^2-\psi_i^2<0$. A sharp sufficient condition can be obtained as

$$\psi_i^2>\frac{1}{n}\sum_j\psi_j^2-\frac{1}{hn}(\gamma-i+\delta)\,w \tag{82}$$

$\square$

Proposition 19. *Whenever another project is added, $b_i'(w)$ jumps down for all projects and whenever a project is dropped, $b_i'(w)$ jumps up for all projects.*

Proof. Let $\Delta j'''(w)=j_+'''(w)-j_-'''(w)$ denote the jump in j'''. Suppose one more project is implemented right of the cutoff w, and denote that project by $i+1$. Then

$$\begin{aligned}\Delta j'''(w) &= \frac{-j''(w)\left((r-i+\delta)w+h(n+1)\right)}{\frac{1}{2}\left(\sum_i \psi_i^2+\psi_{i+1}^2\right)}+\frac{j''(w)\left((r-i+\delta)w+hn\right)}{\frac{1}{2}\sum_i \psi_i^2} \\ &= -\left(j'''_{+}(w)\frac{1}{2}\psi_{i+1}^2+j''(w)h\right)\frac{1}{\sum_i \psi_i^2}\end{aligned}$$

Hence, $\Delta j'''(w)<0$ whenever $j'''_{+}(w)\frac{1}{2}\psi_{i+1}^2+j''(w)h>0$. Equivalently,

$$\Delta j'''(w)=-\frac{1}{\frac{1}{2}\left(\sum_i \psi_i^2+\psi_{i+1}^2\right)}\left(j'''_{-}(w)\frac{1}{2}\psi_i^2+j''(w)h\right) \tag{83}$$

and $\Delta j'''(w)<0$ whenever $j'''_{-}(w)\frac{1}{2}\psi_{i+1}^2+j''(w)h>0$.

Now suppose $\Delta j'''(w)>0$, which is equivalent to

$$j'''_{\pm}(w)<-\frac{j''(w)h}{\frac{1}{2}\psi_{i+1}^2} \tag{84}$$

and note that this implies $b'_{i+1,\pm}(w)<0$. But if $i+1$ is added, it must be the case that b_{i+1} crosses 0 at w and hence $b'_{i+1,\pm}(w)>0$. Hence, it must be the case that

$$j'''_{\pm}(w)>\frac{j''(w)h}{\frac{1}{2}\psi_i^2}=\frac{j''(w)}{\frac{1}{2}h}\mathrm{SN}_{i+1}^2 \tag{85}$$

and $\Delta j'''(w)<0$. Hence, for all projects i, $\Delta b'_i=\frac{1}{2}\psi_i^2\Delta j'''<0$.

Similarly whenever a project is dropped, we have

$$\begin{aligned}\Delta j'''(w) &= \frac{1}{\frac{1}{2}\sum_i \psi_i^2}\left(j'''_{-}(w)\frac{1}{2}\psi_{i+1}^2+j''(w)h\right) \\ &= \frac{1}{\frac{1}{2}\left(\sum_i \psi_i^2+\psi_{i+1}^2\right)}\left(j'''_{+}(w)\frac{1}{2}\psi_{i+1}^2+j''(w)h\right)\end{aligned}$$

and if $\Delta j'''(w)<0$, we have

$$j'''_{\pm}(w)>\frac{-j''(w)h}{\frac{1}{2}\psi_{i+1}^2}=\frac{-j''(w)}{h\frac{1}{2}}\mathrm{SN}_{i+1}^2 \tag{86}$$

which implies $b'_{i+1,\pm}>0$ at w. But then b_{i+1} cannot cross zero from above. Hence, $\Delta b'_i(w)=\frac{1}{2}\psi_i^2\Delta j'''(w)>0$ whenever a project is dropped. □

A.4 Proofs on Implementation

The proof of Proposition 7 is below. The proof of Proposition 8 proceeds is exactly analogous and is therefore omitted.

Proof. I use a guess and verify technique. First consider the process $SC_t \equiv \sum_i \Psi_{it} M_{it}$, which I interpret as the share of the firm's cash holdings the manager has at time t. M_{it} follows the process in equation (33) and I assume that some policy Ψ_{it} with finite variation is implemented.

I posit that the decomposition $W_t = SC_t + A_t$ holds, where A_t is the current balance in the manager's personal account. At the optimal contract, Lemma 1 implies that W_t is a diffusion and hence continuous in t. From equation (33) it is clear that SC_t will be continuous whenever no change is made in the project portfolio, and exhibit a jump of $\Psi_{it} M_{it}$ when project i is added, and $-\Psi_{it} M_{it}$ when project i is dropped.[16] Now, define $dA_t = \gamma A_t dt$ whenever SC_t is continuous and $dA_t = \gamma A_t dt - \Psi_{it} M_{it}$ whenever project i is implemented and $dA_t = \gamma A_t dt + \Psi_{it} M_{it}$ whenever it is stopped. It is clear that then W_t is indeed continuous. Hence 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 indeed satisfies $W_t = SC_t + A_t$, by Ito's Lemma for semi martingales[17] we have

$$dW_t = \sum_i \left(\Psi_{it} dM_{it} + d\Psi_{it} M_{it}\right) + dA_t \tag{87}$$

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

$$\begin{aligned}
\gamma W_t &= \sup_{a_i} \sum_i \Psi_i \left(r M_{it} + \pi_t \mu_i a_i - \alpha_i \kappa(i) - (r-\gamma) M_{it} + \alpha_i \kappa(i) - \frac{dc}{dt} \right) \quad (88)\\
&= + \sum_i \Psi_i \frac{dc_{it}}{dt} - \pi_t h \sum_i a_i + \sum_i \frac{d\Psi_i}{dt} M_{it} + \frac{dA_t}{dt} \\
&= \sup_{a_i} \sum_i \Psi_i \left(r M_{it} + \pi_t \mu_i a_i - \alpha_i \kappa(i) - (r-\gamma) M_{it} + \alpha_i \kappa(i) - \frac{dc}{dt} \right) \\
&\quad + \sum_i \Psi_i \frac{dc_{it}}{dt} - \pi_t h \sum_i a_i + \gamma A_t
\end{aligned}$$

[16] Thus, the process SC_t is a semi martingale.
[17] See He et al. (1992), p. 245, Th. 9.35

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

$$\gamma W_t = \gamma \left(\sum_i \Psi_i M_{it} + A_t \right) \tag{89}$$

and the optimal contract is indeed implemented. $\square$

The proof of Proposition (9) is below.

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

$$dA_t = \gamma A_t dt + dP_t + SC_t - SC_{t-} \tag{90}$$

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

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

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

$\square$

A.5 Proofs on the Model with Fixed Costs

The proof of Proposition 10 consists of proving the validity of equation (42) together with the conditions at the cutoffs (43). It is easy to show that Lemma 12 and Lemma 13 still hold We can therefore follow a similar procedure as in A.2 to prove the representation in Proposition 10.

In general, the part of Lemma 14 which establishes the C^2 property at the cutoffs where project selection changes will not hold, since it relies on the same HJB equation being valid on both sides of the threshold.

The following Lemma establishes equation (44).

Lemma 20. *Consider a cutoff $w(a,a')$ such that wlog for some $\varepsilon > 0$, $(w-\varepsilon, w) \subset \mathcal{C}_a$ and $(w, w+\varepsilon) \subset \mathcal{S}_a$. Then $j''_{a'}(w(a,a')) \geq j''_a(w(a,a'))$.*

Proof. For notational convenience define

$$f_a(w) = \sum_i \mu_i + \kappa(i) - (i-\delta) j_a \tag{91}$$

First we note that for any $x \in \mathcal{C}_{a'}$ we have

$$rj_{a'} - \mathcal{L}_a j_{a'} - f_a - rk(a,a') \geq 0 \tag{92}$$

which can be established using $\phi_{a'} = j_a - k(a,a')$ as a test function and verifying that $\phi_{a'}$ is a supersolution to j_a for any $w \in \mathcal{C}_{a'}$. Then, the supersolution property immediately implies the above equation.

Since we have $k(a,a') > 0$ and $k(a',a) > 0$ it must be the case that actually $w(a,a') \in \mathcal{C}_{a'}$ and therefore $j''_{a'} \in \mathcal{C}$ on some neighborhood of $w(a,a')$ and in particular on some region of $\mathcal{C}_i$. By the above equation,

$$rj_{a'}(w) - f_a(w) - \left((\gamma - i + \delta) w + h \sum_i a_i\right) j'_{a'}(w) - \frac{1}{2} j''_{a'}(w) \sum_i \psi_i^2 a_i - k(a,a') \geq 0 \tag{93}$$

for all $\{w : \; w \geq w(a,a'), \; w \in \mathcal{C}_{a'}\}$ which can be rewritten as

$$j''_{a'}(w) \;\; \leq \;\; \frac{1}{\frac{1}{2}\sum_i \psi_i^2 a_i} \left((r - i + \delta) j_{a'}(w) - \sum_i \mu_i a_i + \kappa(i) - j'_{a'}(w) \left((\gamma - i + \delta) w + h \sum_i a_i\right)\right)$$

Notice that because the conditions in equation (43) which are the value matching and the smooth pasting conditions hold the RHS of the equation is continuous around $w(a,a')$. Therefore, the right limit of the right hand side will correspond to the left limit. But since for all $\varepsilon > 0$, $w(a,a') - \varepsilon \in \mathcal{C}_a$ the HJB equation for j_a must hold with equality left of $w(a,a')$ and therefore

$$\begin{aligned} j''_{a'}\left(w\left(a,a'\right)\right) &\leq \lim_{w\uparrow w(a,a')} \frac{1}{\frac{1}{2}\sum_i \psi_i^2 a_i} \left(\left(r-i+\delta\right) j_a\left(w\right) - \sum_i \mu_i a_i + \kappa\left(i\right) \right. \\ &\qquad \left. - j'_a\left(w\right)\left(\left(\gamma - i + \delta\right) w + h \sum_i a_i \right) \right) \\ &= j''_{a-}\left(w\left(a,a'\right)\right) \end{aligned}$$

which is the relation to be proven. □

The proof of Proposition 11 relies mostly on the HJBQVI in equation (42).

Proof. Take any a and note that for any w either $j_a > \max_{a' \neq a} j_{a'} - k\left(a,a'\right)$, or there exists an a' such that $j_a = j_{a'} - k\left(a,a'\right)$. Therefore, we have that for all $a \neq a'$ and all w, $j_a \geq j_{a'} - k\left(a,a'\right)$. Then by analogy

$$k\left(a',a\right) \leq j_a - j_{a'} \leq k\left(a,a'\right) \tag{94}$$

for all w and especially

$$\sup_w \left|j_a\left(w\right) - j_{a'}\left(w\right)\right| \leq \max\left\{k\left(a',a\right), k\left(a,a'\right)\right\} \tag{95}$$

and as $k\left(a,a'\right) \to 0$ it must be the case that $j_a \to j_{a'}$ for all a, a'. Therefore the functions j_a converge uniformly to a continuous function $\tilde{j}$, and the remainder of the proof consists of verifying that indeed $j = \tilde{j}$.

First, denote with $\mathcal{C}_a^k$ the closure of the continuation regions given cost function $k\left(.,.\right)$ and recall that on $\mathcal{C}_a^k$, both $j_a \geq \max_{a' \neq a} j_{a'} - k\left(a,a'\right)$ and the HJB equation

$$rj_a - f_a - \mathcal{L}_a j_a = 0 \tag{96}$$

hold. As $k\left(a,a'\right) \to 0$ we have that $j_a \geq \max_{a' \neq a} j_{a'} - k\left(a,a'\right)$ holds for all w and therefore each set $\mathcal{C}_a^k$ will converge in the Hausdorff distance to a set $\tilde{\mathcal{C}}_a$, on which the above HJB equation must hold for the limit of j_a, which is exactly $\tilde{j}$. Therefore

$$\tilde{\mathcal{C}}_a = \left\{w : \, r\tilde{j} - f_a - \mathcal{L}_a \tilde{j} = 0\right\} \tag{97}$$

Notice that for all k we must have that $\bigcup_a \mathcal{C}_a$ covers the entire range of w, since by optimality of the set of functions j_a there must be always some action that will be optimally chosen for all w. In particular, this implies that whenever $w \in \mathcal{C}_a^k$ but $w \notin \mathcal{C}_a^{k'}$ for some $k' < k$ then there exists an a' such that $w \in \mathcal{C}_{a'}^{k'}$. Therefore, for all w $\exists a$ such that $w \in \tilde{\mathcal{C}}_a$. Note that the fact that $\tilde{j} \in C$ implies that the value matching condition must hold whenever there is a cutoff $\tilde{w}\left(a, a'\right)$ where the project choice switches from a to a'.

Further, it is easy to show that since $\tilde{j}$ is a viscosity solution to the HJB equation on $\tilde{\mathcal{C}}_a$, as well as value matching conditions on the boundaries of $\tilde{\mathcal{C}}_a$, it must indeed be C^2 on $\tilde{\mathcal{C}}_a$.[18]

With continuity of $\tilde{j}$ and the C^2 property on the continuation regions, we can then repeat the steps of Lemma 13, which continues to hold in the presence of fixed costs.

To conclude, note that if $\tilde{j} \in C^2$ everywhere, it will satisfy exactly the same conditions as j, and then $j = \tilde{j}$ by the uniqueness of solutions to the HJB equation (10).

To show this, consider some a and a' so that the cutoffs $w\left(a, a'\right)$ and $w\left(a', a\right)$ converge to a cutoff $\tilde{w}\left(a, a'\right)$. Assume wlog that for $k > 0$ we have that a is optimal left of $w\left(a, a'\right)$ and a' is optimal right of $w\left(a', a\right)$.

For $k > 0$ we have on $\mathcal{S}_{a,a'}$,

$$r j_a - \mathcal{L}_{a'} j_{a'} - f_{a'} + rk\left(a, a'\right) = 0 \tag{98}$$

and hence taking $w \downarrow w\left(a, a'\right)$ and using the value matching and smooth pasting conditions in (43) we have

$$rk\left(a, a'\right) + \left(h\sum_i a_i - h\sum_i a_i'\right) j_a' + \frac{1}{2}\left(\sum_i \psi_i^2 a_i j_{a-}'' - \sum_i \psi_i^2 a_i' j_{a'+}''\right) + f_a - f_{a'} = 0 \tag{99}$$

at $w_+\left(a, a'\right)$. Rearranging terms and expanding this implies

$$\begin{aligned} \frac{1}{2}\sum_i \psi_i^2 a_i \left(j_{a-}'' - j_{a'+}''\right) &= -rk\left(a, a'\right) - \left(f_a - f_{a'}\right) \\ &\quad - j_{a'}'\left(h\sum_i a_i - h\sum_i a_i'\right) - \frac{1}{2}\left(\sum_i \psi_i^2 a_i - \sum_i \psi_i^2 a_i'\right) j_{a+}'' \end{aligned}$$

By equation (44), the LHS is greater or equal to zero. By analogy, we can obtain a similar

[18] The result is an immediate application of Friedman (1975), p. 134, Theorem 2.4, since the HJB equation on $\tilde{\mathcal{C}}_a$ together with the value matching conditions will define a Dirichlet problem and in my setup the volatility coefficient in the HJB equation is bounded strictly above zero.

expression at the cutoff $w\left(a',a\right)$

$$\frac{1}{2}\sum_i \psi_i^2 a_i\left(j''_{a-}-j''_{a'+}\right) = rk\left(a,a'\right)-\left(f_a-f_{a'}\right)$$
$$-j'_{a'}\left(h\sum_i a_i - h\sum_i a'_i\right)-\frac{1}{2}\left(\sum_i \psi_i^2 a_i - \sum_i \psi_i^2 a'_i\right) j''_{a+}$$

and from equation (44), we have that the LHS is non positive. Suppose that indeed $\exists\varepsilon>0$ such that $j''_{a-}\left(w\left(a,a'\right)\right)-j''_{a+}\left(w\left(a,a'\right)\right)\geq\varepsilon$ and $j''_{a-}\left(w\left(a',a\right)\right)-j''_{a+}\left(w\left(a',a\right)\right)\leq\varepsilon$ even as $k\to 0$. Then we have

$$\lim_{|k|\to 0} j''_{a-}\left(w\left(a,a'\right)\right)-j''_{a+}\left(w\left(a,a'\right)\right) > \lim_{|k|\to 0} j''_{a-}\left(w\left(a',a\right)\right)-j''_{a+}\left(w\left(a',a\right)\right) \tag{100}$$

which is a contradiction to $\left|w\left(a,a'\right)-w\left(a',a\right)\right|\to 0$.

Therefore, $\tilde{j}\in C^2$ everywhere and hence the boundary conditions on the$\tilde{\mathcal{C}}_a$ are exactly the same as on the $\mathcal{C}_a$ in the case without fixed costs, and the uniqueness of solutions to HJB equations implies that indeed $\tilde{j}=j$. $\square$

References

Abel, A. B. and J. C. Eberly (1994). A unified model of investment under uncertainty. *The American Economic Review 84*(5), 1369–1384.

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

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

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

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

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

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.

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-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.

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

Jensen, M. C. and W. H. Meckling (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of financial economics 3*(4), 305–360.

Karatzas, I. and S. E. Shreve (1991). *Brownian motion and stochastic calculus*. Springer.

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

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

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

Manso, G. (2006). Motivating innovation.

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

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

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

Oksendal, B. K. (2003). *Stochastic differential equations: an introduction with applications*. Springer Verlag.

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

Pham, H., V. L. Vath, and X. Y. Zhou (2009). Optimal switching over multiple regimes. *SIAM Journal on Control and Optimization 48*(4), 2217–2253.

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.