Dynamic Asset Allocation with Hidden Volatility*

Felix Zhiyu Feng
University of Notre Dame

Mark M. Westerfield
University of Washington

May 2018

Preliminary and Incomplete

Abstract

We study a dynamic continuous-time principal-agent model with endogenous cash-flow volatility. The principal supplies the agent with capital for investment, but the agent can misallocate capital for private benefit and has private control over both the volatility of the project and the size of the investment. Depending on the curvature of the returns function, the optimal contract can yield either overly-risky or overly-prudent project selection; it can be implemented with a static two-part tariff on capital (a fixed cost plus a hurdle rate). Our model captures stylized facts about the use of hurdle rates in capital budgeting, particularly a zero price of volatility, and it helps reconcile mixed empirical evidence on risk choice and managerial compensation.

JEL Classification: D86, D82, G31
Key Words: dynamic agency, volatility control, capital budgeting, cost of capital, continuous time

*We thank George Georgiadis, Will Gornall, Raymond C. W. Leung, Bruno Strulovici, Curtis Taylor, Lucy White, Yao Zeng, seminar and conference participants at Boston University, Cheung Kong Graduate School of Business, Duke University, Tsinghua PBC School of Finance, University of British Columbia, University of Notre Dame, the Midwest Economic Theory conference and the Econometric Society North American Summer Meeting for helpful comments and suggestions.

Feng can be reached at ffeng@nd.edu; Westerfield can be reached at mwesterf@uw.edu

1 Introduction

Organizational settings in which project managers privately influence both the amount and riskiness of investment are ubiquitous. A manager may privately allocate capital across projects with different cash flow volatilities, or simultaneously discover and exploit opportunities with different risks. Nevertheless, the design of dynamic incentive plans that both allocate the appropriate amount of capital and induce managers to take the desired amount of volatility has challenged formal analysis.

In this paper, we build a continuous-time dynamic contracting model in which the agent has control over the size (capital invested) and the risk (cash flow volatility per unit of capital) of assets owned by the principal. By introducing separately controlled components of cash flow, we create a framework rich enough to capture meaningful risk choices. We derive the optimal contract and show that it can be implemented with a simple mechanism: the principal offers the agent two-part tariff, and the agent determines the amount of capital to manage. Importantly, the implementation can be made static; the principal does not need to dynamically adjust quantities nor keep track of the agent's performance history.

The contributions of this paper are three-fold: first, our findings rationalize the observed practice of capital budgeting documented by Jagannathan et al. (2016), Graham and Harvey (2001), and others[1]: firms impose capital rationing on managers through a deliberately high price of capital and that they most commonly do not adjust that cost of capital for risk. These stylized facts contradict conventional wisdom and textbook corporate finance instruction. In our model, the principal optimally imposes a high cost of capital to limit the scale of the agent's investment and the size of the agency problem. Moreover, because the principal can only assign incentives based on performance and the agent makes hidden choices over risk, the agent will optimally choose volatility such that the benefits and costs are equal, and the principal will not need to modify the cost of capital for risk, despite risk impacting performance. Secondly, we use our model to demonstrate and predict differences in the cross-section and the time-series of the cost of capital as well as the relationship between managerial risk-taking and pay-for-performance. Finally, we illustrate a modeling technique that allows the agent to privately determine *both* the capital intensity and cash-flow volatility, despite the principal being able to continuously monitor cash flows.[2]

---

[1]In Section 5, we summarize and condense the findings from Jagannathan et al. (2016), Graham and Harvey (2001), Graham and Harvey (2011), Graham and Harvey (2012), Jacobs and Sivdasani (2012), and Poterba and Summers (1995).

[2]Several recent papers such as Cvitanić et al. (2016b) and Leung (2017) have made attempts toward this

We begin with a basic dynamic moral hazard framework in which the principal (she) has assets or projects that she hires the agent (he) to manage. The agent takes hidden actions that determine the cash flow from these projects; the principal observes output and rewards the agent with consumption after high output and imposes termination after low output. Separation is costly, which makes the principal effectively risk averse. Our novel assumption is that the agent receives capital from the principal but privately decides both the risk of the project and the amount of capital that is actually invested. Any remaining capital is allocated to generate private benefits for the agent.[3] This joint modeling of capital intensity and risk circumvents the issue that volatility is observable for Brownian motion, which has been the primary obstacle precluding the analysis of dynamic risk choice in a tractable framework. Thus, in our setting the principal must provide incentives to generate both the desired risk choice and the desired capital intensity to avoid asset misallocation. The agent's choice is constrained by the observability of total risk/volatility – it is the *components* of risk/volatility that are unobservable and subject to agency manipulation.

We show that the resulting optimal contract can lead to both overly-risky or overly-prudent risk choice, relative to the first-best. To reduce the likelihood of costly separation following poor performance, the optimal contract reduces the volatility of the agent's continuation value, which has two components: the volatility of the project's cash flow, and the agent's exposure to it (his pay-performance sensitivity, or PPS). Exactly which component the principal reduces depends on the specific risk-return relationship, most critically, how much additional risk the agent takes as incentives are made less intense. When the risk taken by the agent is very sensitive to incentives, the principal offers more incentives, increasing PPS and reducing risk below the first-best level. In contrast, when the risk taken by the agent is relatively insensitive to incentives, the principal relaxes incentives, reducing PPS and resulting in overly-risky project choice. In contrast to the risk adjustment, capital intensity is always (weakly) less than the first-best because more intensive use of capital implies higher cash-flow volatility and requires higher pay-performance sensitivity.

We also demonstrate a generic and simple implementation for the optimal contract. In a standard recursive optimal contract, the principal would allocate capital to the agent and command a certain level of total cash flow risk. Our implementation allows the agent to

end. We discuss these studies below.

[3]Agency frictions of this kind are widespread: for example, a corporate manager may choose enjoyable but unproductive projects, netting private benefits but an inferior risk-return frontier. Similarly, an asset manager may not want to exert maximal effort to maintain all the available projects or investment options. He obtains private benefits (e.g. shirking) and the risk-return frontier is pulled down.

choose his own quantity of capital, the cash-flow risk target, and even his own compensation structure (his PPS or equity share). The principal simply supplies the capital to the agent at a certain cost calculated based on the agent's choices.[4] The cost of capital is subtracted from the project's output, and the remainder is split between the principal and agent based on the proposed equity share. Critically, the implementation is *static*: the principal does not need to make any dynamic adjustment to the formula used to calculate the cost of capital, and she does not need to track the agent's performance history either. In other words, the time-varying optimal policies of the dynamic optimally contract can be automatically carried out by offering the agent a static formula of hurdle rate.

Our novel implementation captures stylized facts about firms' use of hurdle rates in capital budgeting. As mentioned above, firms systematically use hurdle rates that are significantly higher than both the econometrician-estimated and firm-estimated cost of capital, passing up positive NPV projects. Our model explains capital rationing and the hurdle rate gap through an agency perspective in which we embed the cost of capital. Alternatively, the hurdle rate can also be interpreted as a preferred return to investors, which is standard in private equity contracts.[5]

Our implementation also rationalizes practices that deviate from textbook cost of capital usage, in particular failing to adjust for risk when determining hurdle rates. Intuitively, this results from the fact that the principal cannot provide separate incentives for capital and per-unit volatility in the optimal contract; our implementation gives the principal the ability to set an exact price of capital, which is sufficient to implement the optimal contract. Put differently, the agent's incentive compatibility condition equalizes the marginal product of capital in productive and unproductive projects. Thus, the cost of capital can be set equal to the marginal product of capital in unproductive projects without adjusting for the returns and risk in the productive projects, and the desired capital allocation will follow even so.

Finally, our results help reconcile empirical evidence regarding the correlation between investment risk and pay-performance sensitivity, which has been particularly ambiguous among existing studies.[6] Our model points out that there are actually two different mechanisms through which incentives are determined. The first is a "static" mechanism, which

[4] For a concrete example, imagine the principal tells the agent: "take however much capital you want. If you want a 10% equity share, your cost of capital is 8% per unit. If you want a 20% equity share, your cost of capital is 7%."

[5] See Metrick and Yasuda (2010) or Robinson and Sensoy (2013)

[6] Prendergast (2002) and Edmans et al. (2017) summarize the related theoretical and empirical studies. See Section 5 for more detailed discussion of this line of research

is described by the agent's incentive compatibility constraint, capturing the causal trade-off between incentives and risk. The second is a "dynamic" mechanism, which corresponds to the solution to the principal's maximization problem, where risk, size and PPS jointly evolve according to the agent's performance. Empirically, this means that the causal relationship between risk and PPS could be very different from their time-series correlations, which helps explain why existing studies failed to conclusively demonstrate how risk and PPS correlate with each other.

The basic framework of our model borrows from DeMarzo and Sannikov (2006) and Biais et al. (2007). Two other studies that examine volatility control in continuous-time are Cvitanić et al. (2016b) and Leung (2017). Cvitanić et al. (2016b) (and Cvitanić et al. (2016a)) assess optimal control over a multi-dimensional Brownian motion when the contract includes only a terminal payment and is sufficiently integrable. They show the principal can attain her optimal value (possibly in a limit) by maximizing over contracts that depend only on output and quadratic variation. The setup is similar to earlier work by Cadenillas et al. (2004) and more broadly the literature on delegated portfolio control such as Carpenter (2000), Ou-Yang (2003) and Lioui and Poncet (2013), which focus on exogenous compensation and/or information structures. Another contemporaneous work involving volatility control is Leung (2017), who, like us, assumes that cash flow is made of two components: agent's private choice of project risk and an exogenous market factor that is unobservable to the principal. However, because the market factor in Leung (2017) is exogenous, after the contract is in place, the agent has only limited ability to manipulate risk without being detected. Our paper is also broadly related to Biais et al. (2010), DeMarzo et al. (2013) and Li and Williams (2017) etc., who study optimal incentives when the agent's action generate discrete, verifiable jump risks. Finally, Epstein and Ji (2013) develop a volatility control model based on an ambiguity problem. In contrast, with standard preferences, we study the design of an optimal contract in an intuitive agency environment and show that our contract can be implemented with a simple structure largely resembling the practice of capital budgeting.[7]

[7] Other papers that investigate agency problems and capital usage in the same model include He (2011),DeMarzo et al. (2012), and Malenko (2018). We add to those papers by modeling an agency problem over capital intensity and the productivity of capital, as opposed to over mean cash flow or growth.

2 Model

In this section, we describe a principal-agent problem in which an agent is hired by a principal to manage an investment. The principal gives capital to the agent, and the agent privately allocates that capital among different projects or uses. The principal designs an incentive contract based on total output to induce the agent to choose the desired projects – the desired sources of cash flow and volatility.

2.1 The Basic Environment

Time is continuous. There is a principal that has access to capital and an agent that has access to projects. Both the principal and the agent are risk neutral. The principal has unlimited liability and a discount rate r, which is also her flow cost (rental rate) of liquid working capital. The agent has limited liability and a discount rate $\gamma > r$. The principal has outside option $L > 0$, and the agent has outside option $R = 0$, both of which are net of returning rented capital.[8] The agent cannot borrow or save.

The agent has access to a cash flow profile indexed by volatility σ, with $0 \leq \underline{\sigma} \leq \sigma \leq \overline{\sigma}$. Given a level of volatility and of invested capital $K_t \geq 0$, the agent's project choice generates a cumulative cash flow Y_t that evolves as

$$dY_t = f(K_t)\left[\mu(\sigma_t)dt + \sigma_t dZ_t\right], \qquad (1)$$

where Z_t is a standard Brownian motion. $\mu(\sigma)$ represents the agent's risk-efficient frontier: the best return that the agent can achieve given a level of volatility. Both σ_t and K_t can be instantaneously adjusted without cost; K_t represents *liquid working capital*, such as cash, machine-hours, etc.

We assume that $f(K)$ represents a standard decreasing-returns-to-scale technology: three-times differentiable with $f(0) = 0$; $f'(K) > 0$; $f''(K) < 0$; $\lim_{K\to 0} f'(K) = \infty$; $\lim_{K\to\infty} f'(K) = 0$. The agent has a limited selection of underlying projects, so each additional unit of capital is invested with less cash flow output.[9]

[8] $L > 0$ and $R = 0$ simplifies the exposition because the principal will not temporarily shut down production by granting the agent zero capital. In Section 6, we extent the model to include more general outside values.

[9] Allowing concavity in the production function can be more realistic than linearity (e.g. allowing for organizational frictions like a limited span of control), and the assumption generates a first-best with finite capital usage. Because our principal and agent are risk-neutral, the first-best will be achieved after some

We assume that $\mu(\sigma)$ represents the risk-efficient frontier given some set of underlying projects: $\mu''(\sigma) < 0$, $\mu(\sigma)$ attains its maximum inside $(\underline{\sigma}, \overline{\sigma})$, and $\mu(\overline{\sigma}) < 0$.[10] Together, they imply that some risk-taking is efficient, but at some level increasing volatility reduces the principal's average cash flows – there is negative "alpha" with infinite volatility.

In short, the assumptions on $f(K)$ and $\mu(\sigma)$ are flexible and allow a well-defined, interior first-best capital and volatility of investment $\{K^{FB}, \sigma^{FB}\}$, which are given by

$$\max_{K \geq 0;\, \sigma \geq \underline{\sigma}} [f(K)\mu(\sigma) - rK] \tag{2}$$

and characterized by the first-order conditions $0 = \mu'(\sigma^{FB})$ and $r = f'(K^{FB})\mu(\sigma^{FB})$.

2.2 The Agency Friction

The principal supplies capital K_t to the agent and a recommended level of volatility σ_t. The agent cannot obtain capital without the principal (an assumption that we drop in Section 4). The agent chooses two hidden actions: true volatility $\hat{\sigma}_t$ and the actual amount of investment $\hat{K}_t$ in productive, risky projects. In addition to productive projects, the agent has access to a project that produces zero cash flow but some private benefits. The agent allocates the remaining $K_t - \hat{K}_t$ capital to this zero-cash-flow project and receives a flow of private benefits $\lambda(K_t - \hat{K}_t)dt$, where $0 < \lambda \leq r$.[11] That is, the agent can mix between projects with productive cash flow and projects with private benefits. However, capital misallocation is (weakly) inefficient: capital cannot be used to generate private benefits in excess of its rental cost.

Our agency problem can be interpreted in several different ways:

---

histories (see Section 3). In contrast, a linear production function (e.g. $f(K) = K$) implies infinite first-best capital usage. To compensate, one would need to make the agent risk averse, as in Sannikov (2008). Risk aversion can generate a principal's value function that is strictly concave, so that the first-best level of capital is never implemented. However, the fundamental agency problem remains unchanged, and the comparative statics in Section 3 and the implementation in Section 4 remain substantively similar.

[10]For example, if the agent is choosing among underlying projects each with a normally distributed returns and a linear cost of volatility (e.g. something analogous to a repeated one-period CAPM), then we have $\mu(\sigma) = \underline{\mu} + C\sqrt{\sigma^2 - \underline{\sigma}^2} - b\sigma$. If instead the agent faces a constant return to risk minus a convex volatility penalty, then we may have $\mu(\sigma) = b\sigma - \sigma^{\alpha}$.

[11]We adopt a linear private benefits function to match the standard benchmark. In Section 6 we extend the model to allow general, non-linear private benefits $\Lambda(\hat{K}, K)$ from capital misallocation and discuss its implications on incentive compatibility and the optimal contract. We can also allow the agent's private benefits project to generate cash flow volatility, without qualitatively changing the model's outcome, as long as that volatility is not too large.

- Choosing $\hat{K}_t < K_t$ simply means shifting capital to enjoyable but unproductive projects. Thus, a manager with a desire for the quiet life (e.g. Bertrand and Mullainathan (2003)), or a manager who prefers not to travel to make site inspections (e.g. Giroud (2013)) would both qualify.

- In Section 4, we demonstrate an implementation of the optimal contract that allows the agent to request any amount of capital (K) from the principal and invest some of it ($\hat{K}_t < K_t$) to unproductive projects. In that case, the agency friction can also be naturally interpreted as empire-building, since the manager enjoys the control of more capital than it is needed for productive uses.

- A manager might not want to spend the effort to maintain all possible opportunities. For example, a manager might watch a smaller number of potential investments. In doing so, he gains private benefits from shirking $\lambda\Delta_t$, and the efficient investment frontier is reduced to $f(K - \Delta_t)\,(\mu(\hat{\sigma}_t)dt + \hat{\sigma}_t dZ_t)$. Here, Δ_t plays the role of $K_t - \hat{K}_t$.

The cash-flow process Y is observable to the principal. Given the properties of Brownian motions, the principal can infer the true overall volatility, denoted Σ_t.[12] As a result, the principal can impose a particular level of overall volatility (e.g. by terminating the agent if the proper level is not observed). We make the more direct assumption that the principal simply controls the total cash-flow volatility from (1), labeled Σ_t, with

$$\Sigma_t \equiv f\left(K_t\right)\sigma_t = f(\hat{K}_t)\hat{\sigma}_t. \tag{3}$$

The second equality is the constraint that the agent must achieve the desired level of total volatility with his hidden choices.

The agency friction in our model comes from the fact that the principal does not observe the *source* of volatility – intensive capital use in productive projects or excessively risky projects. Put differently, the agent can generate the appearance of productive activity (cash flow volatility) while still putting capital to use generating private benefits. The agent can allocate $K_t - \hat{K}_t$ capital to the unproductive project while increasing the volatility in the productive project ($\hat{\sigma} > \sigma_t$), keeping aggregate volatility (Σ_t) constant. In so doing, the

[12] For concreteness, consider a heuristic example: a principal observes $dX_t = a_t dt + b_t dZ_t$ with X_0 known and $b_t > 0$, but the principal does not observe either a_t or b_t directly. The ability to observe the path of X implies the ability to observe the path of X^2. Since $d(X_t^2) - 2X_t dX_t = b_t^2 dt$, the principal is able to infer b_t along the path.

agent enjoys total private benefits $\lambda(K_t - \hat{K}_t)$. Thus, the principal provides an incentive contract to induce the agent to choose the desired *components* of volatility; the agent must be induced not to take bad risks that hide bad asset allocation.[13]

2.3 Objective Functions

Contracts in our model are characterized using the agent's continuation utility as the state variable. Denote the probability space as $(\Omega, \mathcal{F}, P)$, and the filtration as $\{\mathcal{F}_t\}_{t\geq 0}$ generated by the cash-flow history $\{Y_t\}_{t\geq 0}$. Contingent on the filtration, a contract specifies a payment process $\{C_t\}_{t\geq 0}$ to the agent, a stopping time τ when the contract is terminated, a sequence of capital $\{K_t\}_{t\geq 0}$ under the agent's management, and a sequence of recommended volatility levels $\{\sigma_t\}_{t\geq 0}$. $\{C_t\}_{t\geq 0}$ is non-decreasing because the agent is protected by limited liability. All quantities are assumed to be integrable and measurable under the usual conditions.

Given a contract, the agent chooses a given set of policy rules $\{\hat{K}_t, \hat{\sigma}_t\}_{t\geq 0}$. The agent's objective function is the expected discounted value of consumption plus private benefits

$$W_t^{\hat{K},\hat{\sigma}} = E_{\hat{K},\hat{\sigma}}\left[\int_t^{\tau} e^{-\gamma(s-t)}\left(dC_s + \lambda(K_s - \hat{K}_s)ds\right) + e^{-\gamma\tau}R \middle| \mathcal{F}_t\right], \tag{4}$$

while the principal's objective function is the expected discounted value of the cash flow, minus the rental cost of capital and payments to the agent

$$V_t^{\hat{K},\hat{\sigma}} = E_{\hat{K},\hat{\sigma}}\left[\int_t^{\tau} e^{-r(s-t)}\left(dY_s - rK_s ds - dC_s\right) + e^{-r\tau}L \middle| \mathcal{F}_t\right]. \tag{5}$$

where both expectations are taken under the probability measure associated with the agent's choices. The optimal contract is defined as:

Definition 1 *A contract is incentive compatible if the agent maximizes his objective function by choosing* $\{\hat{K}_t, \hat{\sigma}_t\}_{t\geq 0} = \{K_t, \sigma_t\}_{t\geq 0}$.

A contract is optimal if it maximizes the principal's objective function over the set of contracts that 1) are incentive compatible, 2) grant the agent his initial level of utility W_0,

[13] In our model, private benefits are linked to excess volatility at the project level, despite the fact that private benefits have no direct effect on cash-flow volatility. Instead, the effect is indirect: the agent allocates capital to the unproductive project and compensates with an excessively volatile productive project choice. Our setup contrasts with models that assume the agent generates risk directly from consuming private benefits – for example, shirking might mean increasing disaster risk, as in Biais et al. (2010) and Moreno-Bromberg and Roger (2016). Despite the direct/indirect distinction, both classes of models share the property that stronger incentives reduce project volatility (see Section 3.1).

and 3) give $W_t^{K,\sigma} \geq R$.

This definition restricts our analysis to contracts that involve no capital misallocation because we have defined incentive compatible contracts to mean $\hat{K}_t = K_t$. In developing the optimal contract in Section 3, we will restrict attention to contracts that implement zero misallocation. This is without loss of generality as long as misallocation is inefficient ($\lambda \leq r$), which we show in Proposition 3. We discuss generalizations to $\lambda > r$ in Section 6.

3 The Optimal Contract

In this section we derive the optimal contract. We begin by characterizing the properties of incentive compatible contracts and then proceed to the principal's Hamilton-Jacobi-Bellman (HJB) equation. We end with a categorization of contract types and some comparative statics. Our discussion in the text will be somewhat heuristic; proofs not immediately given in the text are in the Appendix.

3.1 Continuation Value and Incentive Compatibility

The following proposition summarizes the dynamics of the agent's continuation value W_t as well as the incentive compatibility condition:

Proposition 1 *Given any contract and any sequence of the agent's choices, there exists a predictable, finite process* β_t $(0 \leq t \leq \tau)$ *such that* W_t *evolves according to*

$$dW_t = \gamma W_t dt - \lambda(K_t - \hat{K}_t)dt - dC_t + \beta_t \left(dY_t - f(\hat{K}_t)\mu(\hat{\sigma}_t)dt\right) \tag{6}$$

The contract is incentive compatible if and only if

$$\{K_t, \sigma_t\} = \arg \max_{\hat{K}_t \in [0, K_t];\, f(\hat{K}_t)\hat{\sigma}_t = f(K_t)\sigma_t} \left[\beta_t f(\hat{K}_t)\mu(\hat{\sigma}_t) - \lambda \hat{K}_t\right] \tag{7}$$

If the contract is incentive compatible, then $\beta_t \geq 0$ *with*

$$\beta_t = \frac{\lambda}{f'(K_t)} \times \frac{1}{\mu(\sigma_t) - \mu'(\sigma_t)\sigma_t} \tag{8}$$

and (6) simplifies to

$$dW_t = \gamma W_t dt + \beta_t \Sigma_t dZ_t - dC_t. \tag{9}$$

The dynamics of W_t can be derived using standard martingale methods. The first three terms on the right hand side of (6) reflect the promise keeping constraint: because the agent has a positive discount rate, any utility not awarded today must be compensated with increased consumption in the future. The last term contains the agent's incentives, where β_t is the agent's pay-performance sensitivity (PPS): for every dollar of *excess* cash flow, the agent's continuation value changes by β_t dollars. Under an incentive compatible contract $\{\hat{K}_t, \hat{\sigma}_t\} = \{K_t, \sigma_t\}$, so (6) simplifies to (9).

The incentive compatibility condition (7) is a maximization over the discretionary part of the agent's instantaneous payoff. Given the evolution of the agent's continuation value (6), the agent chooses $\hat{K}_t$ and $\hat{\sigma}_t$ to maximize his flow utility:

$$\beta_t \mathrm{E}[dY_t] + \lambda(K_t - \hat{K}_t)dt = \underbrace{\beta_t f(\hat{K}_t)\mu(\hat{\sigma}_t)dt}_{\text{Cash flow paid to the agent}} + \underbrace{\lambda(K_t - \hat{K}_t)dt}_{\text{Agent's private benefit}} . \tag{10}$$

subject to two constraints. First, because the agent cannot borrow on his own, he must choose $\hat{K}_t \in [0, K_t]$. Second, because the principal controls aggregate volatility (3), the agent must choose his controls such that $f(\hat{K}_t)\hat{\sigma}_t = f(K_t)\sigma_t = \Sigma_t$. The resulting maximization problem is given in (7).[14]

We now discuss the important features of incentive compatible contracts in our setting. First, the agency friction does not prevent the principal from implementing the first-best outcome. In fact, the principal can do so even without giving the agent a full share of the project's cash flow:

Property 1 *By choosing $\beta_t = \frac{\lambda}{r} \leq 1$ and $K_t = K^{FB}$, the principal implements $\{K^{FB}, \sigma^{FB}\}$. We define $\beta^{FB} \equiv \frac{\lambda}{r} \leq 1$ to be the level of incentives which, when combined with K^{FB}, implement the first-best policies in the second-best problem.*

Intuitively, at $K_t = K^{FB}$, $\beta^{FB} = \frac{\lambda}{r}$ is the ratio of the marginal benefit from capital misallocation (λ) to the marginal value of productive capital (r), which is less than one. Substituting $\beta_t = \frac{\lambda}{r}$ into the agent's problem (7) produces the same outcome as the first-best optimization (2). No additional incentives are needed to implement the first-best level of volatility.

[14] The agent's problem in this model is effectively static: at any time t, the principal decides her optimal K_t and σ_t and uses β_t and Σ_t to implement such choices. This is a standard feature among dynamic contracting models in which the agent's action bears no persistent effect. The dynamics of the optimal contract come from the principal's side, which we describe in the next subsection.

Second, the agency friction requires the principal to be somewhat moderate in her risk-taking: the principal is precluded from implementing very-high or very-low volatility projects. This result will be especially important in Section 4 when we demonstrate an implementation through a cost of capital that has a small or zero adjustment for risk.

Property 2 *The principal cannot implement very-low volatility ($\sigma \leq \underline{\underline{\sigma}} \equiv \arg\max \frac{\mu(\sigma)}{\sigma}$).*
The principal will never implement very-high volatility ($\sigma \geq \overline{\overline{\sigma}} \equiv \max\{\sigma|\mu(\sigma) = 0\}$).

Intuitively, if the principal tries to implement very-low volatility, such that the Sharpe ratio ($\mu(\sigma)/\sigma$) can be increased by adding volatility, the agent always has a profitable deviation. This results from the fact that $\frac{\mu(\hat{\sigma}_t)}{\hat{\sigma}_t} = \frac{\mathrm{E}[dY_t]}{\Sigma_t}$, so if the principal fixes Σ_t, then increasing $\hat{\sigma}$ increases $\mathrm{E}[dY_t]$. At the same time, $\Sigma_t = f(\hat{K}_t)\sigma_t$ implies that increasing $\hat{\sigma}$ also frees up capital to be used for private benefits. Thus, very-low volatility projects are never incentive compatible, because the agent's deviation (increasing $\hat{\sigma}$) increases both average cash flows and private benefits.

Further, the principal will never choose to implement a value of $\sigma > 0$ that generates negative expected cash flow. This follows from the fact that the principal can always generate zero cash flow with zero volatility by giving the agent zero capital, and the principal's value function is concave, which we show in the next Section.[15]

Third, for moderate volatility projects ($\sigma_t \in \left(\underline{\underline{\sigma}}, \overline{\overline{\sigma}}\right)$), stronger incentives are used to increase capital intensity and decrease volatility:

Property 3 *We have*

$$\frac{\partial}{\partial\sigma}\beta(\sigma, K) = \frac{1}{\lambda}f'(K)\beta(\sigma, K)^2\sigma\mu''(\sigma) < 0 \tag{11}$$

$$\frac{\partial}{\partial K}\beta(\sigma, K) = -\frac{f''(K)}{f'(K)}\beta(\sigma, K) > 0 \tag{12}$$

There are two equivalent ways to understand why $\beta_\sigma < 0$. The first is to recall that the principal can only implement levels of volatility for which the Sharpe ratio ($\mu(\sigma)/\sigma$) is decreasing in σ. This implies that that average cash flow $\mathrm{E}[dY_t] = \Sigma_t \frac{\mu(\hat{\sigma}_t)}{\hat{\sigma}_t}$ is decreasing in σ over the intermediate-σ range. Thus, stronger incentives are needed to increase the *efficiency* of risk-taking and average cash flows.

[15]This result is different from Szydlowski (2016), in which the volatility of cash flow cannot be completely eliminated without terminating the contract. In that model, the principal may implement projects with negative cash flows if those projects are easy to incentivize.

The second way is to recall that the agent is tempted to take *excessive* risks to conceal his misallocation because of the total volatility constraint $f(\hat{K})\hat{\sigma} = f(K)\sigma$. However, the marginal return to productive projects is decreasing ($\mu''(\sigma) < 0$), which implies there is a cash flow cost to the agent for excess risk-taking. Furthermore, such cost is convex, so the impact of excessive risk-taking on the average cash flow is lower when volatility is lower. Consequently, stronger incentives are needed to prevent the agent from excessive risk-taking when volatility is lower.

In contrast, stronger incentives are needed to implement higher capital intensity ($\beta_K > 0$), because the marginal return to productive use of capital ($f'(K)$) is decreasing, while the marginal value of capital in generating private benefits (λ) is constant. In other words, there is a shortage of good projects but no shortage of bad projects. Thus, stronger incentives are needed to induce the agent to retain capital for productive purposes when projects are already large. Agency acts to exacerbate decreasing returns to scale, as it is more difficult to prevent capital misallocation when there is a large amount of capital in place.

We illustrate the solution to the agent's problem as well as the features we discussed above in Figures 1 and 2.

3.2 The Principal's Value Function

Given the results of Proposition 1, the principal's problem is to maximize her objective function (5), subject to the incentive compatibility constraint (7), the law of motion for W_t (9), and the agent's participation constraint ($W_t \geq R = 0$). The agent's continuation utility W is a sufficient state variable to characterize the principal's maximal payoff under the optimal contract. Let $F(W)$ be the principal's expected payoff (5) for the optimal contract given the agent's continuation value. $F(W)$ can be fully characterized with an ODE with boundary conditions summarized in the following Proposition:[16]

Proposition 2 *A solution to the principal's problem $F(W)$ exists, is unique and is concave on $W \in [R, W_C]$, where W_C is chosen so that $F'(W_C) = -1$ and $F''(W_C) = 0$. $F(W)$ solves*

$$rF(W) = \max_{K \in (k_0, \infty);\, \sigma \in (\underline{\sigma}, \overline{\overline{\sigma}})} \Big[f(K)\mu(\sigma) - rK + \gamma W F'(W) + \frac{1}{2}\beta^2(K,\sigma) f^2(K) \sigma^2 F''(W) \Big] \tag{13}$$

[16]To reduce the level of mathematical formalism needed and maintain focus on the economic insight, we make a few largely technical assumptions to simplify the proofs. See the Appendix for details.

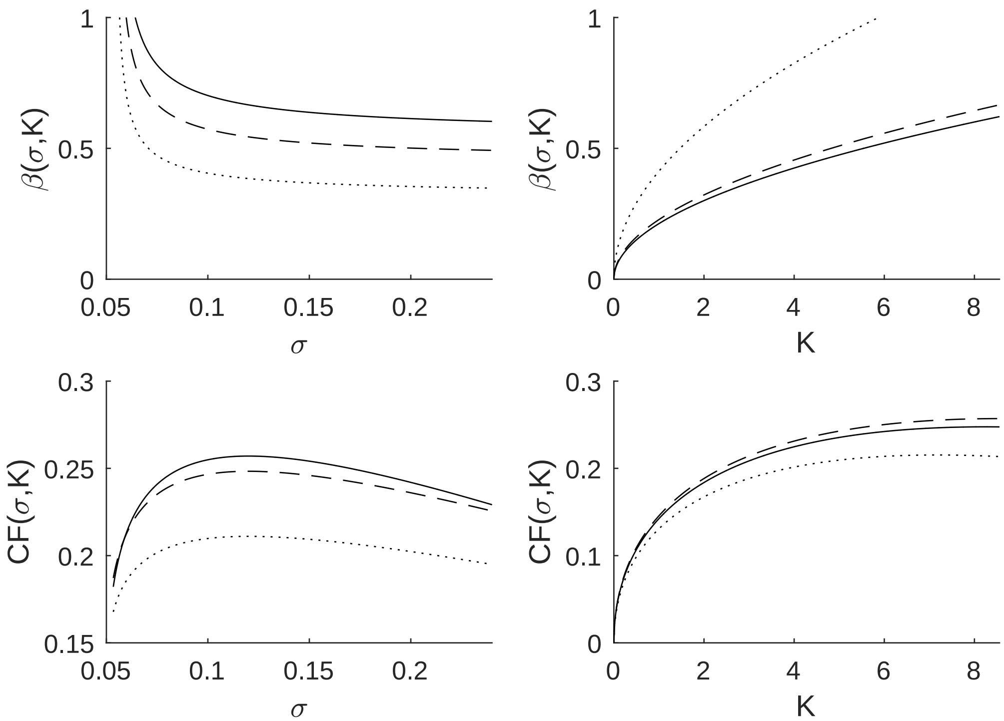

Figure 1: The top-left plot displays $\beta(\sigma, K)$ as a function of σ for $K = K^{FB} = 8.56$ (solid line), $K = \frac{2}{3}K^{FB}$ (dashed line), and $K = \frac{1}{3}K^{FB}$ (dotted line). The top-right plot displays $\beta(\sigma, K)$ as a function of K for $\sigma = \sigma^{FB} = 0.12$ (dashed line), $\sigma = \frac{1}{2}\sigma^{FB}$ (dotted line), and $\sigma = \frac{3}{2}\sigma^{FB}$ (solid line). The bottom two plots display average cash flows $CF(K, \sigma) = f(K)\mu(\sigma) - rK$ as a function of σ for each of the three values of K used in the top row, and as a function of K for each of the three values of σ used in the top row. These plots are generated using $\lambda = 0.02$, $r = 0.03$, $f(K) = 3K^{\frac{1}{2}}$, and $\mu(\sigma) = 0.07 + 0.5\left(\sigma^2 - 0.05^2\right)^{\frac{1}{2}} - 0.55\sigma$.

$F(W)$ is $\mathcal{C}^3$ for all $W \in [0, W_C)$, and has $F(0) = L$. $\{K, \sigma\}$ are the optimal policies, and $dC_t = max(W_t - W_C, 0)$. The agent's continuation utility evolves as in (9), which has a unique weak solution.

We now provide an intuitive derivation for Proposition 2. First, the principal will pay the agent only when the agent's continuation utility exceeds a given threshold, so that $dC_t = max(W_t - W_C, 0)$, with $F'(W_C) = -1$ and $F''(W_C) = 0$. This payment boundary exists because the agent is risk-neutral with respect to consumption and more impatient than the principal. W_C represents the point beyond which the cost of saving for the agent (due to his impatience) exceeds the benefit of avoiding contract termination after series of

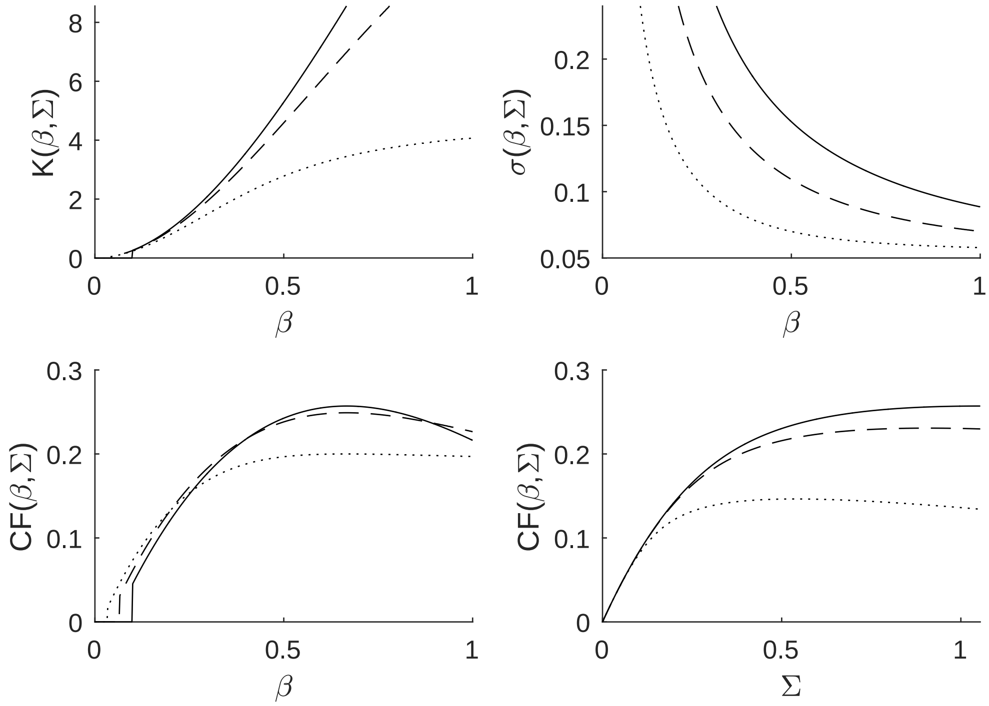

Figure 2: The top-left plot displays $K(\beta, \Sigma)$ as a function of β for $\Sigma = \Sigma^{FB} = 1.05$ (solid line), $\Sigma = \frac{2}{3}\Sigma^{FB}$ (dashed line), and $\Sigma = \frac{1}{3}\Sigma^{FB}$ (dotted line). The top-right plot displays $\sigma(\beta, \Sigma)$ as a function of β for $\Sigma = \Sigma^{FB} = 1.05$ (solid line), $\Sigma = \frac{2}{3}\Sigma^{FB}$ (dashed line), and $\Sigma = \frac{1}{3}\Sigma^{FB}$ (dotted line). The bottom two plots display average cash flows $CF(\beta, \Sigma)$ as a function of β for each of the three values of Σ used in the top row, and as a function of Σ for $\beta = \beta^{FB} = \frac{\lambda}{r} = 0.67$ (solid line), $\beta = \frac{2}{3}\beta^{FB}$ (dashed line), and $\beta = \frac{1}{3}\beta^{FB}$ (dotted line). These plots are generated using $\lambda = .02$, $r = .03$, $f(K) = 3K^{\frac{1}{2}}$, and $\mu(\sigma) = 0.07 + .5\left(\sigma^2 - 0.05^2\right)^{\frac{1}{2}} - 0.55\sigma$.

negative shocks.[17]

Second, the principal chooses $K > 0$ and has $F(0) = L$. This means that the project is always running and that if the agent's continuation utility drops to his outside value, the contract is terminated. This is a result of our parametric assumptions: $L > 0$ and the agent's

[17] These boundary conditions are the same as those in DeMarzo and Sannikov (2006), in which detailed argument can be found. The principal can always make a lump-sum payment of dC to the agent, moving the agent from W to $W - dC$. This transfer benefits the principal only if $F(W - dC) - dC \geq F(W)$, and so we have no transfers if $F'(W) \geq -1$. Thus, we define $W_C = \min\{W | F'(W) \leq -1\}$, and the smooth-pasting condition is $F'(W_C) = -1$. Since W_C is optimally chosen and the principal has linear utility, we have the super-contact condition $F''(W_C) = 0$. See Dumas (1991) for a general discussion of the smooth-pasting and super-contact conditions.

outside value, $R = 0$.[18]

Finally, for $W \in [R, W_C)$, applying Ito's Lemma to $dF(W)$, we obtain

$$dF(W_t) = \gamma W_t F'(W_t)dt + \frac{1}{2}\beta_t^2 f(K_t)^2 \sigma_t^2 F''(W_t)dt + \beta_t f(K_t)\sigma_t F'(W)dZ_t.$$

Combined with cash flows, this yields the principal's Hamilton-Jacobi-Bellman (HJB) equation (13). Because the principal's value function is concave, the $F''(W)$ term is always negative, which represents the principal's cost of providing incentives to the agent. We discuss properties of the HJB equation and the dynamics of the endogenous variables $K(W)$, $\beta(W)$, $\sigma(W)$ and $\Sigma(W)$ in the next subsection.[19]

3.3 Contract Description

The optimal contract and its key comparative statics can be summarized as the following:

Proposition 3 *The optimal contract has the following properties:*

1. *At $W = W_C$, the optimal contract uses $\beta(W) = \beta^{FB} = \frac{\lambda}{r}$ and implements K^{FB}, σ^{FB} and Σ^{FB}.*
2. *For all $W < W_C$, we have $K(W) < K^{FB}$, $\Sigma(W) < \Sigma^{FB}$, and $\beta(W) < \beta^{FB}$.*
3. *Define $\varepsilon_{\beta,\sigma} \equiv \frac{\partial \ln \beta(\sigma,K)}{\partial \ln \sigma} = \frac{\sigma^2 \mu''(\sigma)}{\mu(\sigma) - \sigma\mu'(\sigma)} < 0$ as the elasticity of incentives with respect to project volatility, which does not depend on K. If $\varepsilon_{\beta,\sigma} < -1$, then $\sigma_t \le \sigma^{FB}$ for all $W < W_C$. If $\varepsilon_{\beta,\sigma} > -1$, then $\sigma_t \ge \sigma^{FB}$ for all $W < W_C$.*
4. *If we generalize Definition 1 to allow for capital misallocation (private benefits) in optimal contracts, then all optimal contracts implement zero misallocation, except possibly at W_C.*

Proposition 3.1 can be found through direct evaluation of (13): because $F''(W_C) = 0$, the maximization problem in (13) is the same as the first-best maximization problem (2). Intuitively, $F''(W)$ is associated with the cost of incentive provision. The payment boundary

[18] In a more general model with $R > 0$, the principal may temporarily shut down the project by setting $K = 0$ and allowing the agent's continuation value (W) to reflect upwards. We discuss this generalization in Section 6. Our proof of Proposition 2 shows that this temporary shutdown will only ever occur at $W_t = R$, and only for $R > 0$.

[19] $\beta(W)$ is understood to mean $\beta(\sigma(W), K(W))$, with $\beta(\sigma, K)$ defined in (8). Similarly for $\Sigma(W)$ from (3).

exists because the cost of delaying payment (the agent's impatience) exactly equals the benefit (reduced possibility that future sequence of negative shocks moves W_t down to R), so the principal is effectively risk-neutral with respect to risk in the agent's continuation utility. This implies that the principal is temporarily able to implement the first-best project choice at $W = W_C$.

Between the left and right boundaries, $F''(W) < 0$, and the cost of incentive prevents the optimal contract from implementing the first-best project choice. There are two useful ways of understanding the principal's choices in this region. The first is to examine the cash-flow inputs, $\{K, \sigma\}$. These give us capital and risk choices at the investment level. The second is to examine the principal's volatility controls, $\{\Sigma, \beta\}$. These give us volatility and incentive choices at the relationship level. The mapping between $\{K, \sigma\}$ and $\{\Sigma, \beta\}$ is given by the formulas for Σ and β (3 and 8). Applying such mapping (with a slight abuse of notation), we can write

$$\mathrm{E}\left[dY - rKdt\right] = CF(K, \sigma) = CF(\Sigma, \beta). \tag{14}$$

We can then use $g(K) = \lambda \frac{f(K)}{f'(K)}$ and $h(\sigma) = \frac{\sigma}{\mu(\sigma) - \sigma\mu'(\sigma)}$ to write the the HJB equation (13) as:

$$rF(W) = \max_{K,\,\sigma} \left[CF(K, \sigma) + \gamma W F'(W) + \frac{1}{2} g(K)^2 h(\sigma)^2 F''(W) \right]; \tag{15}$$

$$rF(W) = \max_{\Sigma,\,\beta} \left[CF(\Sigma, \beta) + \gamma W F'(W) + \frac{1}{2} \Sigma^2 \beta^2 F''(W) \right]. \tag{16}$$

Because the expected cash flows ($CF(\Sigma, \beta)$) are concave and the costs ($\Sigma^2\beta^2$) are convex, the first-order conditions are sufficient to show that $\Sigma(W) < \Sigma^{FB}$ and $\beta(W) < \beta^{FB}$. A similar analysis shows that $\frac{\partial}{\partial K} CF(K, \sigma) > 0$ if $g'(K) > 0$ (guaranteed by $\beta_K > 0$ in Property 3) and $\frac{\partial}{\partial \sigma} CF(K, \sigma) > 0$ if $h'(\sigma) > 0$. Thus, we have $K(W) < K^{FB}$; $\sigma(W) < \sigma^{FB}$ if $h'(\sigma) > 0$, which is equivalent to $\varepsilon_{\beta,\sigma} < -1$.

The time series properties, which we will revisit in the implementation can be stated more strongly: as $F''(W)$ becomes more negative, both $\beta(W)$ and $\Sigma(W)$ decline, and as $F''(W)$ increases, both $\beta(W)$ and $\Sigma(W)$ increase. Thus, there is a strong time-series relationship, which is that the agent's share of cash flows increases at the same time as volatility limits increase. This is consistent with the empirical results from [ADD CITE].

The results in Proposition 3.2 and 3.3 show that the agency friction always causes the principal to reduce incentives below the level that would induce the first-best policies ($\beta_t \leq$

β^{FB}), and to do so in a way that reduces cash-flow volatility ($\Sigma_t \leq \Sigma^{FB}$). This is not a-priori obvious: the source of risk for the principal is volatility in the agent's continuation value, which can lead to a loss in default or near-default. Importantly, this risk is not driven by the volatility of the project's cash flow Σ alone, but by the volatility of the agent's continuation value, which equals the *product* $\Sigma\beta$, and so one can imagine that the principal might reduce the agent's share of volatility, imposing weaker incentives and allowing for more volatile cash flow. However, both Σ and β increase the expected cash flow, and they are complements in the cost term ($\frac{1}{2}\beta^2\Sigma^2 F''(W)$), so the principal reduces them both. This is not the case with project-level volatility σ, as we now describe.

A novel result of Proposition 3.3 is that the optimal contract may implement levels of project-based risk (σ) that are higher or lower than the first-best. The reason is that σ affects the volatility of the agent's continuation value through two opposing mechanisms: on the one hand, Property 3 shows that $\beta_\sigma < 0$. That is, implementing a smaller σ (which leads to a more risk-efficient cash flow) requires stronger incentives. On the other hand, cash-flow volatility is increasing in project-level volatility since $\Sigma = f(K)\sigma$. The volatility of the agent's continuation value is a product of these two effects, $\Sigma_\sigma > 0$ and $\beta_\sigma < 0$, and so whether the optimal contract implements a higher or lower σ relative to the first-best depends on which effect dominates.

The elasticity of incentives with respect to project volatility ($\varepsilon_{\beta,\sigma}$) captures the effect of σ on continuation value volatility. When $\varepsilon_{\beta,\sigma} < -1$, incentives can be made much weaker for high levels of project volatility; enough weaker that total volatility of the agent's continuation volatility ($\beta\Sigma$) is *lower* for high project-level volatility. When $\varepsilon_{\beta,\sigma} > -1$, incentives can be made only slightly weaker for high levels of project volatility; the dominant effect is that higher project-level volatility causes *higher* total continuation volatility. The critical distinction here is between cash-flow volatility and continuation-value volatility. The agency problem dictates that it is the risk of default and termination that generates losses – and therefore the agent's continuation value volatility that generates risk – but risky projects can be implemented by giving the agent a small share of those projects, and this creates low continuation-value volatility.

While capital K also affects continuation value volatility through both the incentives and project volatility channels, it does so in the same direction because Σ_K and β_K are both positive. More capital implies more total volatility, and stronger incentives are necessary to prevent shirking with high capital intensity because the marginal return to productive capital is lower. Thus the optimal contract always features *under-investment* ($K_t \leq K^{FB}$)

relative to the first-best.[20]

Last but not least, we note that because misallocation is assumed to be weakly inefficient ($\lambda \leq r$), the optimal contract is robust to considering positive capital misallocation in equilibrium, as shown in Proposition 3.4.

To summarize the results in this subsection, we illustrate an optimal contract in Figure 3. We label solutions for $\varepsilon_{\beta,\sigma} > -1$ and $\sigma_t \leq \sigma^{FB}$ as "Under-σ"; and solutions for $\varepsilon_{\beta,\sigma} < -1$ and $\sigma_t \geq \sigma^{FB}$ as "Over-σ".

4 Implementation

In this section, we show that the optimal contract can be implemented with a simple two-part tariff on capital. The most critical features of the implementation are that the two-part tariff is a *static* function of the agent's visible choices, and the principal *does not keep track* of the agent's continuation utility. In other words, after setting up the two-part tariff, the principal's role is simply to apply the static functions.

The principal offers to rent capital to the agent as a two-part tariff:

- The fixed capital of production (the assets that have liquidation value L) is assigned a rental price ϕ.
- The variable capital of production, K, is assigned a unit price (i.e. a hurdle rate) θ.

The agent can freely request any level of capital and cash flow volatility, and even his pay-for-performance sensitivity (i.e. any combination of $\{\tilde{K}_t, \tilde{\Sigma}_t, \tilde{\beta}_t\}$. The tilde notation is used to indicate that those quantities are choices of the agent. The tariff is adjusted based on the agent's choices: $\theta = \theta(\tilde{\Sigma}_t, \tilde{\beta}_t)$ and $\phi = \phi(\tilde{\Sigma}_t, \tilde{\beta}_t)$, and the cost of capital is deducted from the project's cash flows:

$$dY_t^{NEW} \equiv dY_t - \tilde{K}_t\theta(\tilde{\Sigma}_t, \tilde{\beta}_t)dt - \phi(\tilde{\Sigma}_t, \tilde{\beta}_t)dt. \tag{17}$$

The agent retains $\tilde{\beta}_t dY^{NEW}$, which is placed into a cash account with balance M_t that the agent controls. This account grows at interest rate γ but the agent can freely withdraw

[20] In Section 6 we extend the model and allow general, non-linear private benefit from misallocation. There, it is possible that weaker incentives are required for larger investment (e.g. $\beta_K < 0$) and thus the optimal contract may reduce the volatility of the agent's continuation utility by increasing K, resulting in *over-investment* ($K_t \geq K^{FB}$) relative to the first-best.

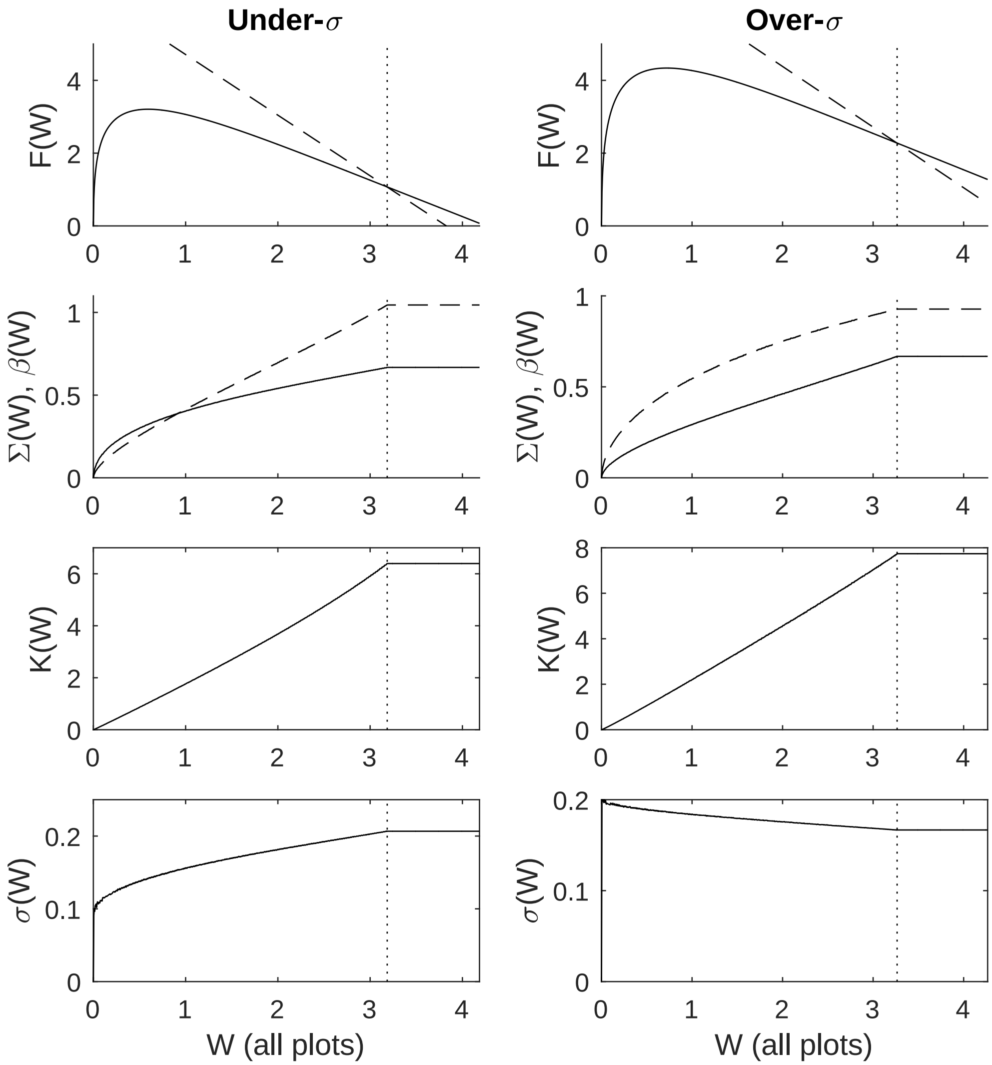

Figure 3: All plots are generated using $f(K) = 2K^{\frac{1}{2}}$, $L = 0$, $R = 0$, $\lambda = 0.02$, $r = 0.03$, and $\gamma = 0.05$. The 'Under-σ' column uses $\mu(\sigma) = \frac{1}{3}\sigma^{\frac{1}{2}} - 0.37\sigma$ and generates $W_C = 3.19$; the 'Over-σ' column uses $\mu(\sigma) = -3\sigma^2 + \sigma$ and generates $W_C = 3.27$. In the top row, the solid line is $F(W)$, and the dashed line is the right-boundary condition $rF(W_C) = \max[CF(K,\sigma)] - \gamma W_C$. In the second row, the solid line is $\beta(W)$ and the dashed line is $\Sigma(W)$.

consumption from the account as long as the balance is positive. That is, the account balance evolves according to:

$$dM_t = \gamma M_t dt + \tilde{\beta}_t dY_t^{NEW} - dC_t. \tag{18}$$

The original agency friction remains: the agent can still privately invest only a portion of the capital given ($\hat{K} \leq \tilde{K}$) into project $\hat{\sigma}$ and obtain the corresponding private benefit $\lambda(\tilde{K}-\hat{K})$, subject to the same volatility constraint as before: $f(\hat{K})\hat{\sigma} = \tilde{\Sigma}$.

The most critical property of the implementation is that, although the equilibrium hurdle rate θ and rental cost ϕ are time-varying, they are static functions of the agent's choices $\{\tilde{\Sigma}_t, \tilde{\beta}_t\}$. The principal does not need make any dynamic adjustment to those functions based on the agent's specific policy choices or his performance history. In fact, she does not even need to track the agent's continuation utility; the account M_t performs this role. That is, $W_t = M_t$ in equilibrium, despite the agent can withdraw consumption from the account at any time.

We formally summarize the implementation as follows:

Definition 2 *An implementation is a two-part tariff $\{\phi(\Sigma,\beta), \theta(\Sigma,\beta)\}$ such that for all $t \geq 0$, the agent chooses a set of actions $\{\tilde{\Sigma}_t, \tilde{\beta}_t, \tilde{K}_t, \hat{K}_t, dC_t\}$ to maximize his expected flow utility:*

$$\tilde{\beta} dY_t^{NEW} + \lambda(\tilde{K}_t - \hat{K}_t) \tag{19}$$

subject to (17) and (18).

An optimal implementation is two-part tariff $\{\phi(\Sigma,\beta), \theta(\Sigma,\beta)\}$ that induces the agent to choose $\{\tilde{\Sigma}_t, \tilde{\beta}_t, \tilde{K}_t, \hat{K}_t, dC_t\}$ as in the optimal contract, with $dC_t = \max(M_t - W^C, 0)$, and the agent quitting when $M_t = 0$.

To construct the optimal implementation, we will need the map $\{\Sigma, \beta\} \to \{K, \sigma\}$ consistent with the IC condition. This is the same map we used in Section 3.3 to define average cash flows as $CF(\Sigma, \beta)$. Together, the functional forms for $\beta(K,\sigma) = \frac{\lambda}{f'(K_t)} \times \frac{1}{\mu(\sigma_t) - \mu'(\sigma_t)\sigma_t}$ (as in 8) and $\Sigma(K,\sigma) = f(K)\sigma$ (3) create such a map. With a slight abuse of notation, define $K(\Sigma,\beta)$ and $\sigma(\Sigma,\beta)$ as the value of K and σ given from $\{\Sigma,\beta\}$ based on (8) and (3). That is, the values of K and σ are the incentive compatible values given the principal's controls. Then, we have

Proposition 4 *There exists an optimal implementation with the following properties:*

$$\theta(\tilde{\Sigma}, \tilde{\beta}) = \frac{\lambda}{\tilde{\beta}} \tag{20}$$

$$\phi(\tilde{\Sigma}, \tilde{\beta}) = f(K(\tilde{\Sigma}, \tilde{\beta}))\mu(\sigma(\tilde{\Sigma}, \tilde{\beta})) - \theta(\tilde{\Sigma}, \tilde{\beta})K(\tilde{\Sigma}, \tilde{\beta}) \tag{21}$$

where $K(\tilde{\Sigma}, \tilde{\beta}), \sigma(\tilde{\Sigma}, \tilde{\beta})$ *jointly solve*

$$\tilde{\beta} = \frac{\lambda}{f'(K)} \frac{1}{\mu(\sigma) - \sigma\mu'(\sigma)} \tag{22}$$

$$\tilde{\Sigma} = f(K)\sigma \tag{23}$$

Moreover, in equilibrium, $\mathrm{E}[dY_t^{NEW}] = 0$ *for all* t*.*

The optimal implementation is the result of several intuitions about how the agency problem is constructed. First, every element of the adjusted cash flow dY^{NEW} is observable to both the principal and the agent. Thus, the underlying moral hazard problem is the same in the implementation as in the optimal contract. Put differently, while it is usually assumed that the principal takes the agent's output from the underlying productive technology as the performance criterion, the principal can in fact look at any performance criterion that she likes. This intuition appears to be general: the principal can choose an augmented cash-flow process that induces the agent to choose the right level of capital, cash-flow volatility, and cash-flow share.

Second, the agent is indifferent across incentive compatible contracts, and so the agent can be induced to pick the principal's desired contract when given a choice. The functions $\phi(\tilde{\Sigma}, \tilde{\beta})$ and $\theta(\tilde{\Sigma}, \tilde{\beta})$ are constructed with two goals in mind:

1. Given the agent's choices of $\{\tilde{\Sigma}, \tilde{\beta}\}$, the agent's marginal value of capital is equal to $\theta(\tilde{\Sigma}, \tilde{\beta})$. The agent's hidden action optimality implies that the marginal value of capital in the productive project ($\mathrm{E}\left[\beta_t dY\right]$) will be the same as the marginal value of capital in private benefits (λ). Thus, we can set the public cost of capital to equal the agent's private marginal value of capital, and the agent will have no incentive to ask for more capital to obtain private benefits. Since the agent is receiving a β_t fraction of the project's cash flows, and the cost of capital is deducted from those cash flows, it must be the case that the private marginal value to the agent of new capital is $\frac{\lambda}{\beta_t}$. Setting $\theta(\tilde{\Sigma}, \tilde{\beta}) = \frac{\lambda}{\tilde{\beta}}$ ensures that the agent always picks the desired level of capital and puts it to productive use.

2. Given the agent's choices of $\{\tilde{\Sigma}, \tilde{\beta}\}$, and the value of θ that puts capital to productive use, the agent will choose $\{\tilde{K}, \hat{K}, \tilde{\sigma}\}$ so as to maximize $\mathrm{E}\left[dY^{NEW}\right]$. $\phi(\tilde{\Sigma}, \tilde{\beta})$ is a fixed rental fee, and it is set so that this maximum $\mathrm{E}\left[dY^{NEW}\right]$ is zero for whatever $\{\tilde{\Sigma}, \tilde{\beta}\}$ the agent chooses. Thus, the agent's expected flow utility is the same for any choice

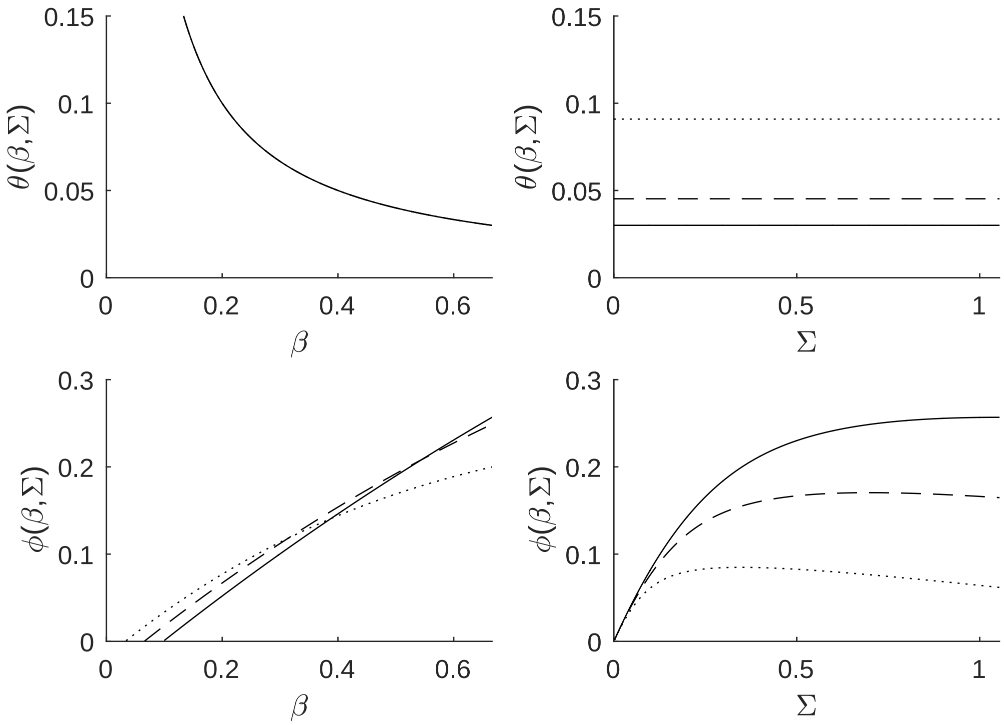

Figure 4: These plots are calculated using the values from Figure 2 and depict the functions in Proposition 4. The two left plots display $\theta(\beta, \Sigma)$ and $\phi(\beta, \Sigma)$ as functions of β for $\Sigma = \Sigma^{FB} = 1.05$ (solid line), $\Sigma = \frac{2}{3}\Sigma^{FB}$ (dashed line), and $\Sigma = \frac{1}{3}\Sigma^{FB}$ (dotted line). The two right plots display $\theta(\beta, \Sigma)$ and $\phi(\beta, \Sigma)$ as functions of Σ for $\beta = \beta^{FB} = \frac{\lambda}{r} = 0.67$ (solid line), $\beta = \frac{2}{3}\beta^{FB}$ (dashed line), and $\beta = \frac{1}{3}\beta^{FB}$ (dotted line). Note that $\theta(\beta, \Sigma)$ does not vary with Σ. These plots are generated using $\lambda = .02$, $r = .03$, $f(K) = 3K^{\frac{1}{2}}$, and $\mu(\sigma) = 0.07 + .5\left(\sigma^2 - 0.05^2\right)^{\frac{1}{2}} - 0.55\sigma$.

of $\{\tilde{\Sigma}, \tilde{\beta}\}$, and the agent is induced to make the principal's desired choice. In other words, the agent receives his full value from the contract in W_0 when it is signed, and the entire flow of expected producer surplus goes to the principal. Since the agent cannot do better than the principal's desired choices, the agent is induced to make the principal's desired choices.

We illustrate the shape of the two-part tariff in Figure 4.

In this implementation, the hurdle rate θ_t has several properties worth noting. First, in the equilibrium, $\theta_t = \lambda/\beta_t$ implies $\theta_t \geq r$ because $\beta_t \leq \lambda/r$, with the equality holds only at $W = W^C$. That is, the principal always sets the hurdle rate above the true cost at which she obtains capital from external markets. Instead, the cost of capital is determined by the agency problem.

Secondly, $\partial\theta/\partial\tilde{\Sigma} = 0$. In other words, the hurdle rate is adjusted for the agent's choice of cash flow share, but it is not additionally adjusted for risk. This is exactly because the agent is already optimizing – the marginal product of capital in productive use must equal the marginal value of capital in private benefits, and the private value is constant.

Third, the agent can choose his own compensation timing. This is standard (e.g. DeMarzo and Sannikov (2006) or Biais et al. (2007)), and it results from the fact that giving the agent access to a hidden savings technology with interest rate γ (equal to the agent's discount rate) does not make the risk-neutral agent better off.

Fourth, the agent can choose his own compensation structure – his own pay-performance sensitivity $\tilde{\beta}$. This is surprising because the private benefit from capital misallocation, λ, is fixed. For example, in DeMarzo and Sannikov (2006) the desired pay-performance sensitivity is implemented with inside equity that is chosen so that the marginal benefit from reporting additional cash flow is equal to the marginal benefit from capital misallocation. Both are constant. However, in our implementation, θ is designed such that the agent has to trade off two different controls, $\tilde{\beta}_t$ and $\hat{K}_t$. If $\tilde{\beta}_t$ is chosen to be very small, the cost of capital will be very high, and so the agent has to use the capital productively to avoid a loss of continuation value. If $\tilde{\beta}_t$ is chosen to be high, then capital is cheap, but the gains to capital misallocation are lower than the agent's chosen cash-flow residual (pay-performance sensitivity).

Finally, we have written the implementation so that the cost of variable capital (θ_t) is deducted from the project cash flow instead of the continuation value. This is not required; we could deduct the cost of variable capital from the agent's account M directly. The only difference is whether we interpret the cost of capital as being paid by the agent or by the project (e.g. a preferred return to outside investors, or not). We can also take either β_t or Σ_t out of the agent's choice set; i.e. we have presented the most decentralized implementation by giving the agent the choice over both β_t and Σ_t as well as K_t, but that is not required.

5 Empirical Discussion

5.1 Capital Budgeting and the Cost of Capital

There is broad empirical agreement on the basic stylized facts surrounding the use of hurdle rates for capital budgeting:[21]

[21] This list is a summary of results in Jagannathan et al. (2016), Graham and Harvey (2001), Graham and Harvey (2011), Graham and Harvey (2012), Jacobs and Sivdasani (2012), and Poterba and Summers (1995).

- Most or almost all firms use DCF methods with a hurdle rate. That hurdle rate is substantially above *both* the econometrician-estimated and firm-estimated cost of capital. For example, Jagannathan et al. (2016) find an average hurdle rate of 15% compared to an average cost of capital of 8%. They find that this is not likely to be caused by behavioral biases or driven by managerial exaggeration.

- Firms engage in deliberate capital rationing. This rationing is often a response to non-financial constraints; more that half of firms report that they pass up apparently positive NPV projects because of constraints on managerial time and expertise (55.3%, Jagannathan et al. (2016)).

Our model is consistent with these results on hurdle rates and capital rationing: first and foremost, we demonstrate in Section 4 that the equilibrium hurdle rate is indeed higher than the firm's true cost of capital (i.e. $\theta_t = \lambda/\beta_t \geq r$). This is optimal because the agency problem imposes a constraint on the use of managerial time and expertise. The principal must offer the agent a portion of residual cash flow in order to induce the desired project choice and capital usage. This portion, combined with limited liability on the part of the manager, creates the possibility of termination, which entails the loss of a high NPV project. To avoid the larger loss, the principal accepts the smaller loss of reducing the scale of the agent's activity in order to lower the volatility of the agent's residual claim. In short, our model suggests that extracting the full value of the "time and expertise" of managers is an agency problem that requires the principal to limit the scale of the agent's production or investment activity through capital rationing, created by a high hurdle rate.[22]

[To Be Added: Lack of risk adjustment, company-wide versus divisional hurdle rates, and volatility budgets.]

5.2 Risk-Taking and Pay-Performance Sensitivity

Dynamic adjustment in risk-taking (σ) can lead to overly-risky or overly-prudent investments. The source of this result is the difference between the volatility of the agent's inside value – which is what drives the principal's termination/default risk – and the volatility of the cash flow. The volatility of the agent's continuation value is the product of project

[22] Our model thus offers predictions based on time-series variation in the cost of capital. First, capital rationing should decrease after success and increase after failure. In particular, the gap between the hurdle rate and the cost of capital should be smaller after success than after failure. Second, adjustments of the hurdle rate for idiosyncratic risk should be conditional: larger after failure than after success.

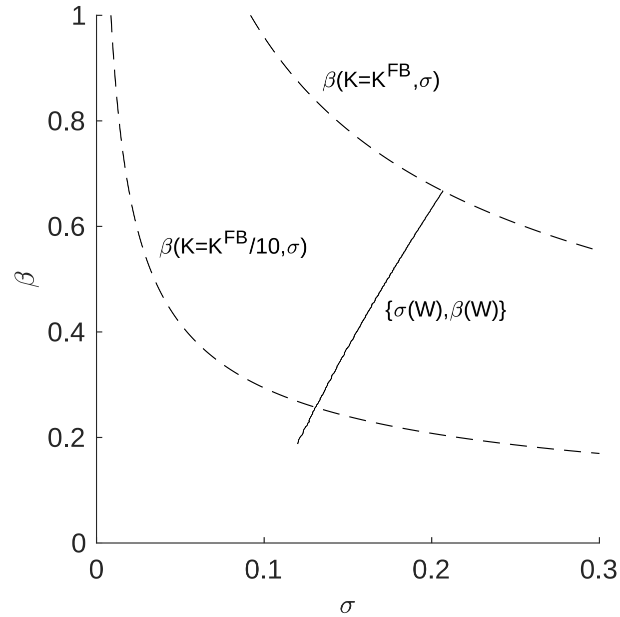

Figure 5: These plots depict β as a function of σ (equation 8) and the locus of points $\{\beta(W), \sigma(W)\}$ (the solution to 13). The plots are generated using the 'under-σ' specification from Figure 3: $f(K) = 2K^{\frac{1}{2}}$, $L = 0$, $R = 0$, $\lambda = 0.02$, $r = 0.03$, $\gamma = 0.05$, and $\mu(\sigma) = \frac{1}{3}\sigma^{\frac{1}{2}} - 0.37\sigma$.

cash-flow volatility and the intensity of incentives. Thus, depending on model specification, risk reduction can mean reducing project volatility or allowing project volatility to increase in order to reduce incentive intensity. One result is that actions that look like risk-shifting can actually be risk-reducing.

The dynamic risk adjustment can generate strong empirical differences between the static and dynamic responses to incentives – the time series and the cross section might look very different because of different sources of variation. Consider, for example, the static $\beta(K, \sigma)$ function in (8): we have $\beta_\sigma < 0$, meaning that for any given state of the world, stronger incentives *cause* lower project risk-taking. However, when we consider the path of incentives and risk-taking over time (see, e.g. Figure 3), the correlation between incentives and total risk (β and Σ) is always *positive*. At the same time, the correlation between incentives and project risk (β and σ) can be positive (in the 'under-σ' specification) or negative (in the 'over-σ' specification). Thus, the causal results of incentives and the dynamic correlation can have opposite signs. We illustrate this in Figure 5, where the solid line is how the economy evolves as a function of W, and the dashed lines are the function $\beta(K, \sigma)$ for two different values of K.

Empirically, this result means that our model predicts different results for time series and cross-sectional tests in the 'under-σ' specification, because the source of variation is different.[23] If the dominant source of variation is the history of success or failure (which changes over time), then we should see a positive relationship between the agent's share and volatility. This is movement along the dynamic curve, and may be more likely in time series variation. If the dominant source of variation is firm characteristics that are stable and not the history of success or failure, then we might see a negative relationship between the agent's share and volatility. Further, an experiment or identification that correctly uncovers the causal mechanism between the agent's share and volatility uncovers movement along the static curve, which generates a negative relationship.

The 'over-σ' specification does deliver another empirical test. In settings with easily scalable investment and in which the relevant risk is agent-separation, rather than default, increasing project risk after failure should be more common. Investment funds, especially mutual funds and hedge funds, would seem to be good examples. They have no explicit risk of default, and to the extent that fund manager skill is real, the primary danger to fund value is that the high-skill manager leaves. In fact, many empirical results, (e.g. Chevalier and Ellison (1997), Aragon and Nanda (2011), Huang et al. (2011)) find increasing project risk after failure to be the case. However, those studies often attribute the increase in risk to convex incentive schemes. Our mechanism is different: increasing project risk actually decreases termination risk. A useful empirical test would be to distinguish changes in inside and outside values, and to see to what extent that difference impacts investment risk.

[23]The empirical literature has indeed found contradictory results across different settings. while some studies find that firms increase risk-taking following poor performance (e.g., Eisdorfer (2008)), others find no such effect (e.g., Andrade and Kaplan (1998)) or a reduction of risk (e.g. Gilje (2016)). The asset management literature paints a similarly mixed picture (e.g., Rauh (2008), Huang et al. (2011), Aragon and Nanda (2011), and their discussions of the literature). Note that while these empirical studies address different institutional backgrounds and may follow different assumptions, our model suggests that more sophisticated controls will be needed to assess the relationship between risk choice and performance. The relationship should be consistent in the time series for each firm, but differ in the cross section based on unobservable characteristics, like investment opportunities.

6 Further Discussion

6.1 $R > 0$ and Temporary Shutdown

So far, we have assumed that $L > 0$ and $R = 0$, with the result that principal always chooses $K > 0$. If we generalize our assumptions to $L \geq 0$ and $R \geq 0$, the principal may choose $K = 0$ for some histories. This represents a temporary shutdown of the firm by the principal. Proposition 2 becomes

Proposition 5 *A solution to the principal's problem $F(W)$ exists, is unique and is concave on $W \in [R, W_C]$, where W_C is chosen so that $F'(W_C) = -1$ and $F''(W_C) = 0$. $F(W)$ solves*

$$rF(W) = \max_{K \in \{0 \cup (k_0, \infty)\}; \sigma \in (\underline{\underline{\sigma}}, \overline{\overline{\sigma}})} \Big[f(K)\mu(\sigma) - rK + \gamma W F'(W) + \frac{1}{2}\beta^2(K, \sigma) f^2(K) \sigma^2 F''(W) \Big] \tag{24}$$

$F(W)$ is $\mathcal{C}$ for all W and $\mathcal{C}^3$ for all $W \in (R, W_C)$, and has $F(R) = L > L^*$ if $K(R) > 0$ and $F(R) = L^*$ if $K(R) = 0$. Further, $K(W) > 0$ for $W \in (R, W_C]$. The agent's continuation utility evolves as in (9), which has a unique weak solution.

This Proposition is what we prove in the appendix. Proposition 2 is a corollary.

The key result is that termination is optional for the principal: if the principal chooses $K(\tilde{W}) = 0$, then the law of motion for W_t (9, with $\Sigma = 0$) implies that W_t reflects upwards at $\tilde{W}$. Thus, shutdown is temporary. The principal will only optimally choose $K = 0$ at $W = R$, and only when $L \leq L^*$, for some constant L^*.

The fact that the principal only chooses $K = 0$ at $W = R$ arises because setting $K = 0$ is costly, and so the principal wishes to delay paying that cost as long as possible. There are two costs to setting $K = 0$. One is an opportunity cost because any time with $K = 0$ is time the principal might otherwise have positive expected cash flow. The second cost is that by causing W_t to reflect early, the principal is causing W_t to reflect upward at a level that is closer to the agent's consumption boundary, so the agent will be awarded consumption sooner.

The fact that the principal only chooses $K = 0$ if L is low is economically straightforward: the principal only avoids default and termination if her value in default is low. If L is high enough, the principal simply accepts default rather than pay the opportunity cost associated with $K = 0$. Note that because of the costs associated with termination, the principal's value function is concave regardless of whether termination occurs in equilibrium. If $L > L^*$, there is a direct cost associated with termination as long as L is less than the discounted, first-best cash flow. This cost makes volatility undesirable, and the principal's value function is concave. If $L < L^*$, there is a direct cost to termination that the principal avoids in equilibrium, choosing instead to pay the opportunity cost of shutting down the project. The principal forgoes the project's cash flow, allowing the agent's continuation value to reflect upwards.

Our model's temporary shutdown is different from that of Zhu (2013), who shows that the principal can relax incentives to either pay the agent with private benefits instead of cash or relax incentives to prevent termination/default after bad cash-flow realizations. Zhu (2013) has fixed project scale, and so when incentives are relaxed the principal cannot also restrict private benefits to the agent. This makes the principal worse off because of an inability to control the timing of private benefits separately from the timing of incentives. Further, the principal might have to shut down incentives before termination would otherwise happen (incentives can be relaxed at $W > R$ in Zhu (2013)) to make sure that the agent's continuation value is high enough to support both private benefits and the continued project.

While our model does feature an incentives shutdown, the mechanism is very different because the project's size is endogenous. In our model, the principal has to pay an ongoing rental cost of capital (rK_t) in order to fund the project; the amount of capital determines the project's scale and the potential private benefits available to the agent. Since paying the agent with private benefits delivers benefits that are less than the rental cost of capital ($\lambda \leq r$), the principal will always couple zero incentives with zero project size. This prevents the agent from receiving additional negative cash-flow shocks, so the agent's continuation value drifts upward, and the project can continue. Because this shutdown is costly, the principal delays as much as possible, and our firm does not shut down until the last moment, $W_t = R$.

6.2 A General Private Benefits Function

We can also extend the model and consider general, non-linear specifications for the agent's private benefits from capital misallocation, which we currently assumed to be $\lambda(K - \hat{K})$. This assumption implies that the agent derives a constant marginal benefit from capital misallocation which, combined with the decreasing returns to scale of $f(K)$, can be interpreted as "there is a shortage of good projects but no shortage of bad projects". While constant marginal private benefit is consistent with many existing models of dynamic moral hazard problems and allows an easy verification of no-misallocation in equilibrium (as long as $\lambda < r$), relaxing such assumption not only affects little of the derivation and implementation of the optimal contract, but also expands the scope of analytical predictions generated by our model.

We can still apply the argument used in deriving (8) in Property 3 to derive the incentive compatibility condition under generic private benefit. We summarize the new condition and its implication into the following proposition:

Proposition 6 *Let $\Lambda(\hat{K}, K)$ denote the generic private benefit from capital misallocation. Assuming $\Lambda(\hat{K}, K)$ is twice-differentiable, the optimal incentive compatibility condition is*

$$\beta(\sigma_t, K_t) = \frac{\Lambda_{\hat{K}}(K_t, K_t)}{f'(K_t)} \times \frac{1}{\mu(\sigma_t) - \mu'(\sigma_t)\sigma_t} \tag{25}$$

where $\Lambda_{\hat{K}}(K, K)$ represents $\Lambda_{\hat{K}}(\hat{K}, K)|_{\hat{K}=K}$. The optimal contract, following Definition 1, has the following property that is different from that described in Proposition 3: define $g(K) = \Lambda_{\hat{K}}(K, K)\frac{f(K)}{f'(K_t)}$. For all $W < W_C$, if $g'(K) > 0$, the principal chooses $K \leq K^{FB}$. If $g'(K) < 0$, then the principal may choose $K > K^{FB}$ under some parameter values.

There are two important differences brought by allowing a generic private benefit function $\Lambda(\hat{K}, K)$. The first is how incentive varies with project size, which is captured by:

$$\frac{\partial}{\partial K}\beta(\sigma, K) = \left(\frac{\Lambda_{\hat{K}K}}{\Lambda_{\hat{K}}} - \frac{f''(K)}{f'(K)}\right)\beta(\sigma, K) \tag{26}$$

Unlike (12), equation (26) can be either positive or negative depending on the specific functional form of $\Lambda(\hat{K}, K)$. In our baseline model, we assume the agent's marginal private benefit from misallocating an additional dollar of capital is constant (λ). In general though, it can depend on the project size. If the marginal private benefit of misallocating each dollar declines with project size, and the decline is faster than the speed of marginal productivity ($f'(K)$) declines, then the agent's tendency for misallocation is weaker when managing a larger project. Consequently, the required incentive for no misallocation is lower as project size grows.[24]

[24]An example of such case is when $\Lambda(\hat{K}, K) = \lambda(K - \hat{K})/K$, (and $f(K) = K^{\alpha}, 0 < \alpha < 1$), which implies $\beta_K = -\frac{\lambda}{\alpha}(1+\alpha)K^{-(2+\alpha)}\beta < 0$. One can interpret this case as the agent's private benefit depends on the *fraction* – instead of quantity – of capital misallocated. If one re-write the private benefit as a result of effort

The fact that β may decrease in K implies the second important difference: the possibility of "over-investment" in the optimal contract, or $K(W) > K^{FB}$. This is in contrast to the baseline model with linear private benefit, which only features "under-investment". The intuition is the same as that behind why the optimal contract could feature either "overly-risky" and "overly-prudent" project choice: as W moves away from W_C, the principal seeks to lower the volatility of W. Under generic, non-linear private benefit, β could be decreasing in K. If such decrease is fast enough, the principal may achieve a less volatile W with a high K and thus low β.[25] This result complements other studies featuring over-investment under the optimal contract in which other market frictions and elements are introduced on top a dynamic moral hazard problem.[26]

We can even extend the model further by allowing private benefit to be a function of project risk or total risk (e.g. $\Lambda(\hat{K}, K, \sigma)$ or $\Lambda(\hat{K}, K, \Sigma)$. We can also consider the case in which capital misallocation generates no return but some volatility to the cash flow, instead of no return and no volatility as currently assumed. Nevertheless, the basic agency friction and how it determines the optimal contract remains the same: the agent has the tendency to allocate capital towards activities that generate personal benefit and must take sub-optimally high risk in order to hide such behavior. The strength of the incentive required to prevent misallocation depends on the marginal private benefit of misallocation and the gain/loss of marginal return from the productive use of capital due to the increased risk. Consequently, when the principal wants to lower the degree of the agency friction by lowering the volatility of the agent's continuation utility, she may find it optimal to either decrease total cash flow volatility (lower σ and/or lower K) or decrease incentives (higher σ and/or higher K), resulting in different combinations of overly-prudent/overly-risky project choices and under/over-investment in the equilibrium.

Finally, we have been focusing on optimal contracts following Definition 1, i.e. a contract that eliminates capital misallocation in equilibrium because misallocation is inefficient. This is standard in the literature such that an incentive compatibility contract is also socially

(as discussed in the next subsection) then it represents the case in which the agent's effort has a *multiplicative* – instead of additive – effect on output.

[25]In Proposition 6 we note that while $g'(K) > 0$ is a sufficient condition for under-investment, $g'(K) < 0$ is only necessary but not sufficient for over-investment. When $g'(K) < 0$, the first order condition for K from the HJB equation (15) implies $f'(K)\mu(\sigma) - r = -h(\sigma)^2 F''(W)g(K)g'(K) > 0$, or $f'(K) < r/\mu(\sigma)$, which is not sufficient to deduce $K(W) > K^{FB}$ since $\mu(\sigma) < \mu(\sigma^{FB}) = r/f'(K^{FB})$ for all $\sigma \neq \sigma^{FB}$. Nevertheless, we present numerical examples in Figure X that illustrates the existence of over-investment.

[26]For example, Ai and Li (2015) and Bolton et al. (2017) with limited commitment, Szydlowski (2016) and Gryglewicz et al. (2018) with multi-tasking, Gryglewicz and Hartman-Glaser (2017) with real options.

desirable. We can consider the case in which misallocation is efficient: for example when $\Lambda(\hat{K}, K) = \lambda(\hat{K} - K)$ as in the baseline model but $\lambda > r$. However, this changes the right boundary of $F(W)$ in an uninteresting way. If capital used for the agent's private benefit is $\delta \leq \overline{\delta}$, the principal's HJB equation becomes

$$rF(W) = \max_{K,\sigma,\delta} \Big[f(K)\mu(\sigma) - rK - r\delta \ + (\gamma W - \lambda\delta)F'(W) + \frac{1}{2}\beta^2(K,\sigma)f^2(K)\sigma^2 F''(W) \Big]$$

The principal would pay private benefits whenever the agent's continuation utility exceeded W_S, defined by $F'(W_S) = -\frac{r}{\lambda}$ (instead of the current boundary condition, $F'(W_C) = -1$.)

6.3 Dimensionality

Our model uses two-dimensional incentives (β and Σ) to implement a two-dimensional agent's choice (K and σ). This raises a potentially interesting question: what would happen if the agent had more choices, particularly an effort choice. We have assumed that the agent implements a portfolio of projects with a particular risk-return tradeoff, but what if the agent could work to improve the tradeoff? What if the relationship between K and σ was not separable? We might assume that, instead of (1), we have

$$dY_t = \alpha(K_t, \sigma_t, e_t)dt + \varphi(K_t, \sigma_t)dZ_t \tag{27}$$

where e_t is effort. The agent's private benefits function might then be $B(K_t - \hat{K}_t, e_t)$. In this setup, the principal then has two controls (β and $\Sigma = \varphi(K_t, \sigma_t)$) to implement a three-dimensional choice by the agent. The incentive compatibility condition (7) becomes

$$\{K, \sigma, e\} = \arg\max_{\hat{K},\hat{\sigma},\hat{e}} \left[\beta\alpha(\hat{K}, \hat{\sigma}, \hat{e}) + B(K - \hat{K}, \hat{e})\right] \tag{28}$$

under the constraint that $\Sigma = \varphi(\hat{K}, \hat{\sigma})$.

Because the principal only has two controls, but the agent has three choices, the set of choices the principal can actually implement is a two-dimensional curve in a three-dimensional space. Depending on the functional forms of α and φ, this implementable set might be very far from the first-best. Informally, the principal can condition on output to provide incentives that capital be used efficiently and that total risk is as desired, but her controls are not more granular than that; adding an effort choice need not increase the size of the set of implementable actions.

A key insight of our model of volatility is that it represents a unique setting that cannot be captured by an economically reasonable hidden effort model, such as a simple variation of DeMarzo and Sannikov (2006). One might think that our model could be moved in the opposite direction by giving the principal control over volatility and capital and giving the agent a hidden effort choice over drift.[27] However, this change of variables would put all the economics into the private benefits function in a completely uninterpretable way.[28] Instead, we give the agent a second choice over per-unit-of-capital volatility, give the principal a second control over total volatility, and impose economically reasonable assumptions on the production function $f(K)\mu(\sigma)$. These choices are structured so that equilibrium output could be obtained with positive probability under both the agent's true action and under the principal's desired action, thus allowing us to use the martingale methods of Sannikov (2008). In sum, our model demonstrates that agency problems over the composition of volatility are possible and interesting – total cash-flow volatility (Σ), project volatility (σ), and agent's continuation value volatility ($\beta\sigma$) are all meaningfully different.

7 Conclusion

Continuous-time principal-agent models have developed rapidly with applications to expanding areas of economic research. Despite the progress, almost all of the existing models involve the agent controlling the drift of the output/cash-flow process. We deviate from the literature by considering the optimal contract when the agent controls the volatility. In the model, overall cash flow is made up of two components: the individual risk of the project and capital intensity, both of which are observable only to the agent. The principal must incentivize the agent to choose the desirable level of project risk and capital intensity. Such setting represents a unique and different environment than a drift-control model such as modified

[27]That is, define $dY_t = f(K_t)e_t dt + \Sigma_t dZ_t$ instead of (1), and use $e_t = \frac{f(\hat{K}_t)}{f(K_t)}\mu\left(\frac{\Sigma_t}{f(\hat{K}_t)}\right)$ as the hidden action instead of $\hat{K}_t$. Then, the agent's private benefits function becomes $B(e, K, \Sigma) = \lambda\left(K - \hat{K}(e, K, \Sigma)\right)$, where $\hat{K}(e_t, K_t, \Sigma_t)$ is an inversion of $e(\hat{K}, K, \Sigma)$.

[28]First, $B(e, K, \Sigma)$ has unsigned derivatives (i.e. B_K and B_Σ must change signs at arbitrary point over the relevant range of the problem). Second, even the level of private benefits is difficult to assess because the constraint $B(e(K, K, \Sigma), K, \Sigma) = 0$ is a technological constraint that must be imposed exogenously and has no clear meaning. These issues are especially prominent under generic, non-linear private benefit $\Lambda(\hat{K}, K)$.

versions of DeMarzo and Sannikov (2006) or Biais et al. (2007).

Interestingly, we find the optimal incentives can be implemented without loss of generality with a simple hurdle rate against which the agent's performance – the realized cash flow – is measured. The agent is allowed to propose the amount of capital and the level of risk he wants, even his own incentive power. Moreover, the principal is freed from the need to keep track of the agent's performance. A natural direction of future research is to explore the implications for asset pricing and portfolio intermediation (e.g. Buffa et al. (2015) with an optimal contract). The hurdle rate becomes the required return of limited partners or outside investors, and the joint determination of contracts and equilibrium asset prices is a natural question to explore.

Appendix

In this appendix, we provide proofs of propositions included in the main text. We will assume only that $L \geq 0$ and $R \geq 0$. The restriction to $L > 0$ and $R = 0$, as discussed in the text of Section 2, is a corollary.

To reduce the level of mathematical formalism in the proofs of the following propositions, we maintain the following assumptions

Assumption A.1 *$f(K)$ and $\mu(\sigma)$ satisfy the following assumptions in addition to those mentioned in Section 2.1:*

1. $\frac{d^2}{dK^2}\left(\frac{f(K)}{f'(K)}\right)^2 \geq 0$ *for all* $K > 0$.
2. $\frac{d^2}{d\sigma^2}\left(\frac{\sigma}{\mu(\sigma)-\sigma\mu'(\sigma)}\right)^2 \geq 0$ *for all* $\sigma \geq \underline{\sigma}$ *with* $\mu(\sigma) > 0$.
3. *There is a minimum positive amount of capital that the principal can grant the agent:* $K \in \{0 \cup [k_0, \infty)\}$ *for some* $k_0 > 0$ *very small.*

The first line ensures that decreasing returns occurs smoothly enough for the principal's problem to be strictly concave, so that there are no jumps in K_t. The second ensures that the variance of the agent's continuation value is convex in the standard deviation of the project's cash flow. Altogether, these restrictions are economically innocuous and consistent with most commonly used production functions and risk-return relationships. Examples of $f(K)$ that meets the conditions include $f(K) = \ln(1+K)$ and $f(K) = K^\alpha$, $\alpha \in (0,1)$. Examples of $\mu(\sigma)$ include $\mu(\sigma) = \sigma^a - b\sigma$, $a \in \left(0, \frac{1}{2}\right)$; $\mu(\sigma) = b\sigma - \sigma^a$, $a > 1$; and $\mu(\sigma) = \underline{\mu} + C\sqrt{\sigma^2 - \underline{\sigma}^2} - b\sigma$, which is the mean-variance efficient frontier if the agent has access to several projects with normally distributed cash flows. In that case we assume C and $\underline{\mu}$ are not too large.

The final condition simplifies the proof of the existence and uniqueness of the principal's Hamilton-Jacobi-Bellman ODE. The proceeding analysis is valid with or without this assumption. One should think of k_0 as being very small: e.g. the principal cannot allocate less than one penny of capital without allocating zero capital. See Piskorski and Westerfield (2016) for such a proof when the principal's control can go continuously to zero. Our assumption is a restriction on the principal rather than on the agent: incentive compatibility conditions will still be required at $K = k_0$.

Proof of Proposition 1

First[29], we define the agent's total expected utility received under a contract conditional on his information at time t as:

$$U_t = E^{\hat{K},\hat{\sigma}}\left[\int_0^\tau e^{-\gamma u} dC_u + \int_0^\tau e^{-\gamma u}\lambda\hat{\Delta}_u du + e^{-\gamma\tau}R\,|\mathcal{F}_t\right],$$

[29] This proof a slightly modified version of a proof in Piskorski and Westerfield (2016), which in turn is based on a similar proof in Sannikov (2008).

where $\hat{\Delta}_u = K_u - \hat{K}_u$. We note that the process $U = \{U_t, \mathcal{F}_t; 0 \le t < \tau\}$ is an $\mathcal{F}_t$-martingale. The expectation is taken with respect to the probability measure induced by $\{\hat{K}, \hat{\sigma}\}$, such that $\int_0^T \Sigma_t dZ_t^{\hat{K},\hat{\sigma}} dt = \int_0^T dY_t - \int_0^T f(\hat{K}_t)\mu(\hat{\sigma}_t)dt$ is a Martingale for all $T > 0$. Recall here that σ_t, $\hat{\sigma}_t$, K_t, and $\hat{K}_t$ are all bounded from below by zero and from above by positive constants, so Σ_t is also bounded below by zero and above by a positive constant. Then, by the martingale representation theorem for Lévy processes, there exists a $\mathcal{F}_t$-predictable, integrable process β such that

$$U_t = U_0 + \int_0^t e^{-\gamma u}\beta_u \Sigma_u dZ_u^{\hat{K},\hat{\sigma}}. \tag{29}$$

Recall the agent's continuation value $W_t^{\hat{K},\hat{\sigma}}$ defined in (4). We have

$$U_t = \int_0^t e^{-\gamma u} dC_u + \int_0^t e^{-\gamma u}\lambda \hat{\Delta}_u du + e^{-\gamma t} W_t^{\hat{K},\hat{\sigma}}. \tag{30}$$

for $t \le \tau$. Differentiating (29) and (30), we obtain

$$dU_t = e^{-\gamma t}\beta_t \Sigma_t dZ_t^{\hat{K},\hat{\sigma}} = e^{-\gamma t} dC_t + e^{-\gamma t}\lambda\hat{\Delta}_t dt - \gamma e^{-\gamma t} W_t^{\hat{K},\hat{\sigma}} dt + e^{-\gamma t} dW_t^{\hat{K},\hat{\sigma}},$$

therefore

$$dW_t^{\hat{K},\hat{\sigma}} = \gamma W_t^{\hat{K},\hat{\sigma}} dt - dC_t - \lambda \hat{\Delta}_t dt + \beta_t \Sigma_t dZ_t^{\hat{K},\hat{\sigma}}.$$

This equation also implies the evolution of promised value given in (6), and the evolution given in (9) for $\{\hat{K}, \hat{\sigma}\} = \{K, \sigma\}$.

Next, define $\tilde{U}_t$ to be the payoff to a strategy $\{\tilde{K}, \tilde{\sigma}\}$ that consists of following an arbitrary strategy until time $t < \tau$ and then $\{K, \sigma\}$ thereafter, then

$$\tilde{U}_t = \int_0^t e^{-\gamma u} dC_u + \int_0^t e^{-\gamma u}\lambda \tilde{\Delta}_u du + e^{-\gamma t} W_t^{K,\sigma}.$$

Differentiating $\tilde{U}_t$ and combining terms yields

$$\begin{aligned} e^{\gamma t} d\tilde{U}_t &= \lambda \left(K_t - \tilde{K}_t\right) dt + \beta_t \Sigma_t dZ_t^{K,\sigma} \\ &= \lambda \left(K_t - \tilde{K}_t\right) dt + \beta_t (f(\tilde{K}_t)\mu(\tilde{\sigma}_t) - f(K_t)\mu(\sigma_t))dt + \beta_t \Sigma_t dZ_t^{\tilde{K}_t,\tilde{\sigma}_t} \end{aligned}$$

where the second equality reflects a change in the probability measure from the one induced by $\{K, \sigma\}$ to the one induced by $\{\tilde{K}, \tilde{\sigma}\}$. Recall here that both $\{K, \sigma\}$ and $\{\tilde{K}, \tilde{\sigma}\}$ must generate the same Σ, per the agent's constraints. If (7) does not hold on a set of positive measure, then the agent could chose $\{\tilde{K}, \tilde{\sigma}\}$ such that

$$\beta_t f(\tilde{K}_t)\mu(\tilde{\sigma}_t) - \lambda \tilde{K}_t > \beta_t f(K_t)\mu(\sigma_t) - \lambda K_t$$

that is, the drift of $\tilde{U}$ is always nonnegative and strictly positive on a set of positive measure, which implies

$$E^{\tilde{K},\tilde{\sigma}}\left[\tilde{U}_t\right] > \tilde{U}_0 = W_0^{K,\sigma},$$

and so the strategy $\{K, \sigma\}$ would not be optimal for the agent. If (7) does hold for the strategy $\{K, \sigma\}$ then $\tilde{U}_t$ is a super-martingale (under measure induced by $\{\tilde{K}, \tilde{\sigma}\}$) for any strategy $\{\tilde{K}, \tilde{\sigma}\}$, that is,

$$E^{\tilde{K},\tilde{\sigma}}\left[\tilde{U}_t\right] \leq \tilde{U}_0 = W_0^{\tilde{K},\tilde{\sigma}}.$$

which proves that choosing $\{K, \sigma\}$ is optimal for the agent if and only if (7) holds for the strategy $\{K, \sigma\}$.

Finally, incentive compatibility requires that $\beta \geq 0$ because otherwise it is impossible for (7) to hold at $\hat{K} = K > 0$. If $K_t = 0$, then $\hat{K}_t = 0$ and no deviation is possible, so the value of β_t is irrelevant and we say $\beta_t = 0$.

Proof of Property 3

Substituting $\hat{\sigma} = \frac{\Sigma}{f(\hat{K}_t)}$ into (7) yields a new maximization problem

$$K = \arg \max_{0 \leq \hat{K} \leq K} \beta f(\hat{K}) \mu\left(\frac{\Sigma}{f(\hat{K})}\right) - \lambda \hat{K} \tag{31}$$

Taking the first order condition of the objective function and setting it to zero yields

$$\beta f'(\hat{K})\mu\left(\hat{\sigma}\right) - \beta f'(\hat{K})\mu'\left(\hat{\sigma}\right)\hat{\sigma} - \lambda = 0 \tag{32}$$

which implies (8). Meanwhile the second order condition of the objective function is given by

$$\beta f''(\hat{K})\left(\mu\left(\hat{\sigma}\right) - \mu'\left(\hat{\sigma}\right)\hat{\sigma}\right) + \beta f(\hat{K})\mu''(\hat{\sigma})\hat{\sigma}^2 < 0 \tag{33}$$

where the inequality follows from $f''(K) < 0$ and $\mu''(\sigma) < 0$. Therefore, there is a unique maximum described by the first-order condition.

The derivatives β_K and β_σ are direct calculations from (8).

Proof of Propositions 2 and 5

We prove Proposition 5, of which Proposition 2 is a corollary. We start with four preliminary results:

Lemma 7 *Any solution to the principal's value function problem must have consumption awarded to the agent at* W_C *such that* $W_t \leq W_C$*, with* $F'(W_C) = -1$ *and* $F''(W_C) = 0$*.*

These are the value matching and super-contact conditions.

Lemma 8 *Let* $\hat{F}$ *solve (13). Assume that the boundary conditions in Lemma 7 are met at some* $\hat{W}_C$*, but ignore the boundary condition at* $W = R$*. Then* $\hat{F}''(W) < 0$ *for all* $W < \hat{W}_C$ *on which* $\hat{F}$ *is defined.*

Proof. Taking the derivative (from the left or the right) of the HJB equation (13) with respect to W and using the envelope theorem yields

$$0 = (\gamma - r)\, F'(W) + \gamma W F''(W) + \frac{1}{2}\Sigma^2(W)\beta^2(W)F'''(W) \tag{34}$$

If $F'' = 0$ and $F''' < 0$, then (34) implies $F' > 0$. Moreover, from (13), $F'' = 0$ and $F' > 0$ together imply $rF > \max[f(K)\mu(\sigma) - rK]$, which is impossible since $\max[f(K)\mu(\sigma) - rK]$ characterizes the solution under the first-best scenario. Therefore $F'' = 0$ and $F''' < 0$ are jointly impossible: if $F'' = 0$, then it must be that $F''' > 0$ (i.e. F'' can only cross zero from below). Therefore if $F''(\hat{W}_C) = 0$ for some $\hat{W}_C$ then $F'' < 0$ for $W < \hat{W}_C$. ■

Lemma 9 *Let* $\hat{F}$ *and* $\tilde{F}$ *both solve (13) Assume that the boundary conditions in Lemma 7) are met at* $\hat{W}_C$ *and* $\tilde{W}_C$*, respectively, but ignore the boundary conditions at* R*. Then the following four statements are equivalent:*

- $\tilde{W}_C < \hat{W}_C$
- $\tilde{F}(W) > \hat{F}(W)$ *for all* $W \leq \tilde{W}_C$ *such that* $\hat{F}$ *and* $\tilde{F}$ *both exist.*
- $\tilde{F}'(W) < \hat{F}'(W)$ *for all* $W \leq \tilde{W}_C$ *such that* $\hat{F}$ *and* $\tilde{F}$ *both exist.*
- $\tilde{F}''(W) > \hat{F}''(W)$ *for all* $W \leq \tilde{W}_C$ *such that* $\hat{F}$ *and* $\tilde{F}$ *both exist.*

The arguments in Piskorski and Westerfield (2016), Lemma 8, are sufficient.

Lemma 10 *Let* $\hat{F}$ *solve (13). Assume that the boundary conditions in Lemma 7) are met at some* $\hat{W}_C$*, but ignore the boundary condition at* $W = R$*. Then, if the principal chooses* $K(\tilde{W}) = 0$*, the law of motion for* W_t *(9, with* $\Sigma = 0$*) implies that* W_t *reflects upward at* $\tilde{W}$*. The principal will optimally choose* $K = 0$ *only at* $W = R$*.*

Proof. First, any solution to the ODE characterized by (13) with $K(\tilde{W}) = 0$ for any $\tilde{W} > R$ can be improved by moving the $K = 0$ transition point to the left. The principal of optimality[30] implies that at $\tilde{W}$ for which $K(\tilde{W}) = 0$, the HJB equation must follow (13) with continuous $F(W)$ (the value-matching condition) and $F'(W)$ (the smooth pasting condition). In addition, if the highest $\tilde{W}$ for which $K(\tilde{W}) = 0$ lies in the interior of W (i.e. $\tilde{W} > R$), then $F''(W)$ must be continuous as well (the super-contact condition). Finally, there can be at most one value of $\tilde{W}$ for which a transition exists.

Our problem makes the continuity of $F''(\tilde{W})$ impossible. At $\tilde{W} > R$ for which $K(\tilde{W}) = 0$, (13) becomes $0 = -rF(W) + \gamma W F'(W)$ (which generates an analytical solution, $F(W) = const \times W^{\frac{r}{g}}$). Taking the derivative of the HJB equation, solving for $F'(W)$ and substituting that back into $0 = -rF(W) + \gamma W F'(W)$ yields $rF(W) = -\frac{\gamma^2}{\gamma - r} W^2 F''(W)$. Since $F(W)$ is bounded above by the first-best, it must be the case that $F''(W)$ is bounded from below by a constant and this bound does not depend on k_0. However, inspection of the first-order condition of (13) for K, remembering that σ and β are bounded, implies that $F''(W)$ is arbitrarily negative if k_0 is arbitrarily small. This is a contradiction, so the second derivative cannot be continuous, and the transition from $K = k_0$ to $K = 0$ cannot be optimal on the interior of W.

In addition, there cannot be a transition from $K > k_0$ to $K = 0$ on the interior of W. Taking the derivative of the HJB equation on either the left or right side of $\tilde{W}$ yields:

$$0 = (\gamma - r) F'(W) + \gamma W F''(W) + \frac{1}{2}\Sigma^2(W)\beta^2(W)F'''(W). \tag{35}$$

Since F' and F'' are continuous at $\tilde{W}$ we have $0 = (\gamma - r) F'(\tilde{W}) + \gamma \tilde{W} F''(\tilde{W})$ (with the derivative taken on the left-hand side). Since $K(W > \tilde{W}) > 0$, it must be the case that $\lim_{W \downarrow \tilde{W}} F'''(W) = 0$. In addition, concavity plus examination of the analytical solution for $W < \tilde{W}$ shows $\lim_{W \uparrow \tilde{W}} F'''(W) > 0$. This jump in $F'''(W)$ at $\tilde{W}$ has the wrong sign for an optimal jump in K. Thus, there cannot be a transition from $K > k_0$ to $K = 0$ on the interior of W. ■

We now proceed to analyze the HJB equation.[31] A necessary condition for optimality (Lemma 10) is that $K \geq k_0$ on $W \in (R, W_C]$, so we will consider that region first. Re-writing (13), we have

$$F''(W) = \min_{K \geq k_0;\, \sigma \in (\underline{\sigma}, \overline{\overline{\sigma}})} \frac{rF(W) - f(K)\mu(\sigma) - rK - \gamma W F'(W)}{\frac{1}{2}\beta(K,\sigma)^2 K^2 \sigma^2} \tag{36}$$

The right-hand-side can be written as the function $H_{K,\sigma}(W, F(W), F'(W))$, which is differentiable in all of its arguments. Since K and β are bounded away from zero and

[30] See Dumas (1991) for a detailed theoretical discussion or Piskorski and Westerfield (2016) or Zhu (2013) for applications in a similar setting.

[31] This part of the proof is based on a similar proof in Sannikov (2008), altered to this setting and extended to include the possibility that the agent's volatility might be zero.

infinity, $H_{K,\sigma}$ has uniformly bounded derivatives in $F'(W)$ and $F(W)$, and $H_{K,\sigma}$ is Lipschitz continuous in $F(W)$ and $F'(W)$. It follows that solutions to (36) exist, are $\mathcal{C}^2$ and are unique and continuous is initial conditions.[32] Inspection of (36) and (34) shows that the solution is at least $\mathcal{C}^3$, rather than just $\mathcal{C}^2$ on $W \in (R, W_C)$.

To complete our solution, we need to show that our remaining condition can be met: that for any L, either there exists a W_C such that $F(R; W_C) = L$ and $K(R) > 0$ or there exists a W_C such that $F(R; W_C) \geq L$ and $K(R) = 0$. We need only consider $K = 0$ at R from Lemma 10.

Lemma 9 shows that proposed solutions to (13) that obey the right boundary condition (at W_C) can be ranked by the proposed value $\hat{W}_C$. For $\hat{W}_C = R$, we have the highest solution with $rF(R = \hat{W}_C) = \max\left[f(K)\mu(\sigma) - rK\right] - \gamma R$. As we increase $\hat{W}_C$, $F(W; \hat{W}_C)$ declines for all W; thus, $\lim_{W \to R} F(W; \hat{W}_C)$ declines as $\hat{W}_C$ increases.

Within the set of solutions that obey the right boundary condition, we consider the subset with $K(R; \hat{W}_C) > 0$. Label the infimum value of $F(R; \hat{W}_C) > 0$ as L^*, and the corresponding value of $\hat{W}_C$ as W_C^*. Any solution with $K(R; \hat{W}_C) > 0$ must be better than the principal's value with $K = 0$ and immediate payment (flow $\gamma R dt$), which is finite. Thus, L^* and W_C^* are both finite.

Because solutions are continuous in initial conditions, L^* is also the maximum value of $F(R; \hat{W}_C)$ among all proposed solutions with $K(R) = 0$. Then, consider the proposed solution with V_C^* such that $K(R, W_C^*) = 0$ and $F(W; W_C^*) = L^*$. This solution can be implemented for any value of L: because $K(R; W_C^*) = 0$, there is no termination and L is never realized. This solution is preferred to all other no-termination solutions by construction.

If, instead, $L > L^*$, then the principal prefers to allow termination, with $K(R) > 0$. Since L is greater than the infimum value of $F(R; \hat{W}_C)$ among all proposed solutions with $K(R) > 0$, and solutions to the ODE are continuous in initial conditions and ordered (Lemma 9), there exists exactly one solution with $\hat{W}_C < W_C^*$ such that $F(R; \hat{W}_C) = L$.

This shows existence and uniqueness of solutions to the HJB equation with a given liquidation value L: If $L > L^*$, then $K(W) > 0$ and solutions exist and are unique by the arguments given above. If $L < L^*$, then we use the solution that generates $K(R) = 0$ and $F(R) = L^*$, which exists and is unique.

Standard existence and uniqueness results (see e.g. Karatzas and Shreve (1998)), expanded to include Sticky Brownian Motions (see arguments from e.g. Harrison and Lemoine (1981) or Engelbert and Peskir (2014)) are sufficient to show that (9) has a unique solution and that W_t is a Sticky Brownian Motion near W_R if $K(W_R) = 0$.

The proof is completed with a standard dynamic-programming verification argument. We illustrate $F(W)$ under various parameter choices in Figure 6.

[32]To see these conditions directly, use one of the first order conditions (K or σ) to solve out F'', and observe that for regions in which $\{K, \sigma\}$ are interior, (36) is a first-order ODE that can be solved through direct integration, taking $K(W)$ and $\sigma(W)$ as unknown, bounded functions. Similarly, in regions in which $K = k_0$, we have constant or bounded coefficients and standard results imply existence and uniqueness there (see e.g. Piskorski and Westerfield (2016) or Zhu (2013)). We use the more powerful Lipschitz continuity to show the solution is $\mathcal{C}^2$.

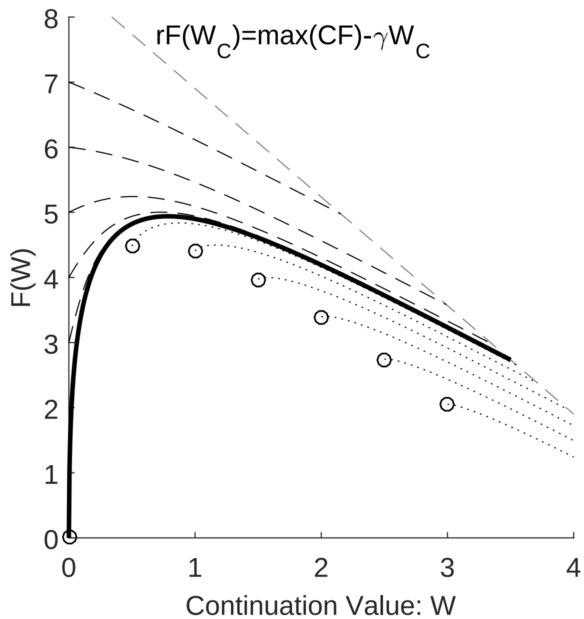

Figure 6: We generate solutions to the HJB equation by varying the right boundary point W_C along the line described by $rF(W_C) = \max(CF) - \gamma W_C$ (dashed grey line), which has $F'(W_C) = -1$ and $F''(W_C) = 0$. Our parameter choice implies $L^* = 0$. The solid line is the solution with $F(R = 0) = 0 = L^*$. The small circles indicate values of W with $K(W) = 0$, at which point W reflects upward. The dotted lines connecting W with $K(W) = 0$ (small circles) to W_C also solutions to the HJB, but they have $K(W > R) = 0$ and achieve lower values for the principal that the solid black line. The dashed lines above the solid line are solutions with default in equilibrium, $F(R) = L > L^*$ and $K(R = 0) > 0$. The plot uses $\mu(\sigma) = 0.07 + .5\left(\sigma^2 - 0.05^2\right)^{\frac{1}{2}} - 0.55\sigma$ (e.g. the efficient frontier for a mixture of normally distributed payoffs) and $f(K) = 3K^{\frac{1}{2}}$.

Corollary 11 *If $R = 0$, then $L^* = 0$. $L > 0$ implies the principal chooses $K > 0$ for all W, and the boundary value is $F(R = 0) = L$, and $F(W)$ is $\mathcal{C}^3$.*

Proof. If $R = 0$ and $K = 0$, then $dW = 0dt$, and no further consumption is realized for either party. Thus, the principal chooses to obtain L if $L \geq 0$, i.e. $L^* = 0$. ■

Proof of Proposition 3

Part 1:

Direct substitution of the boundary conditions at W_C into (13).

Parts 2 and 3:

The principal's HJB equations (15 and 16) and the accompanying first-order conditions are

$$\begin{aligned}
rF(W) = &\max_{K,\sigma}\left[CF(K,\sigma) + \gamma W F'(W) + \frac{1}{2}g(K)^2 h(\sigma)^2 F''(W)\right] \\
FOC(K): &\quad CF_K(K,\sigma) + g'(K)g(K)h^2(\sigma)F''(W) \\
FOC(\sigma): &\quad CF_\sigma(K,\sigma) + g^2(K)h'(\sigma)h(\sigma)F''(W)
\end{aligned}$$

and

$$\begin{aligned}
rF(W) = &\max_{\Sigma,\beta}\left[CF(\Sigma,\beta) + \gamma W F'(W) + \frac{1}{2}\Sigma^2\beta^2 F''(W)\right] \\
FOC(\Sigma): &\quad CF_\Sigma(\Sigma,\beta) + \Sigma\beta^2 F''(W) \\
FOC(\beta): &\quad CF_\beta(\Sigma,\beta) + \Sigma^2\beta F''(W)
\end{aligned}$$

where $CF(K,\sigma) = f(K)\mu(\sigma) - rK$, $g(K) = f(K)/f'(K)$, $h(\sigma) = \lambda\sigma/(\mu(\sigma) - \sigma\mu'(\sigma))$. Assumption A.1 is sufficient to ensure strict concavity of the HJB equation with respect to K and σ. Signing $g'(K)$ and $h'(\sigma)$, remembering that $F''(W) \leq 0$, and evaluating the first-order conditions directly, are sufficient to demonstrate the statement of the property with respect to K and σ.

Next, we rewrite the cash flow as $CF(\beta,\Sigma) = CF(K(\beta,\Sigma),\sigma(\beta,\Sigma))$. Because the first-best has $CF_K(K^{FB},\sigma^{FB}) = CF_\sigma(K^{FB},\sigma^{FB}) = 0$, there is a corresponding value of (β^{FB},Σ^{FB}) which has $CF_K(K(\beta^{FB},\Sigma^{FB}),\sigma(\beta^{FB},\Sigma^{FB})) = CF_\sigma(K(\beta^{FB},\Sigma^{FB}),\sigma(\beta^{FB},\Sigma^{FB})) = 0$. The second order conditions for the HJB equation (16), evaluated at $\{\beta^{FB},\Sigma^{FB}\}$ are

$$SOC(\beta)|_{(\beta^{FB},\Sigma^{FB})}: \quad CF_{KK}\cdot(K_\beta)^2 + CF_{\sigma\sigma}\cdot(\sigma_\beta)^2 < 0 \tag{37}$$

$$SOC(\Sigma)|_{(\beta^{FB},\Sigma^{FB})}: \quad CF_{KK}\cdot(K_\Sigma)^2 + CF_{\sigma\sigma}\cdot(\sigma_\Sigma)^2 < 0, \tag{38}$$

where the inequalities follow from $CF_{KK} < 0$ and $CF_{\sigma\sigma} < 0$. Because the first-best is unique, with $CF_\beta = 0$ and $CF_\Sigma = 0$, we also have that $CF_\beta > 0$ implies $\beta < \beta^{FB}$ and $CF_\Sigma > 0$ implies $\Sigma < \Sigma^{FB}$. Combining this with the first-order conditions (16) yields the statement of the property.

Part 4:

Assume that the principal offers the agent a recommended level of capital misallocation, δ_t, to go with the assigned level of capital, $K_t + \delta_t$. Then the contract is incentive compatible if it implements $\hat{K}_t = K_t$ and $\hat{\delta}_t = \delta_t$. The incentive compatibility condition (7) becomes

$$\{K_t,\delta_t,\sigma_t\} = \arg\max_{\hat{K}_t+\hat{\delta}_t = K_t+\delta_t;\, f(\hat{K}_t)\hat{\sigma}_t = f(K_t)\sigma_t}\left[\beta_t f(\hat{K}_t)\mu(\hat{\sigma}_t) + \lambda\left(K_t + \delta_t - \hat{K}_t\right)\right]. \tag{39}$$

Note that the agent-controlled portion of the right-hand side is the same as in the original

IC condition (7). The first order conditions imply that β_t becomes

$$\beta_t = \beta(\sigma_t, K_t) = \frac{\lambda}{f'(K_t)} \times \frac{1}{\mu(\sigma_t) - \mu'(\sigma_t)\sigma_t} \; . \tag{40}$$

The evolution of the agent's continuation value becomes

$$dW_t = \gamma W_t dt + \beta_t \Sigma_t dZ_t - \lambda \delta_t dt - dC_t \; . \tag{41}$$

The principal's HJB equation (13) becomes

$$\begin{aligned} rF(W) = \max_{K+\delta \in \{0 \cup [k_0,\infty)\};\, \sigma \in (\underline{\underline{\sigma}}, \overline{\overline{\sigma}})} \Big[& f(K)\mu(\sigma) - rK - r\delta + (\gamma W - \lambda\delta)F'(W) \\ & + \frac{1}{2}\beta^2(K,\sigma) f^2(K)\sigma^2 F''(W) \Big] \end{aligned} \tag{42}$$

For any given choice of K, the objective function of (42) is linear in δ with coefficient $-r - \lambda F'(W)$. Since $F'(W) \geq -1$, we have $-r - \lambda F'(W) < 0$. Therefore, $\delta = 0$ is (weakly) optimal.

Proof of Proposition 4

Consider the implementation in two stages. First, we solve the agent's problem conditional on an arbitrary choice of $\{\tilde{\beta}, \tilde{\Sigma}\}$. The principal offers the agent the adjusted technology with

$$d\tilde{Y}_t^{NEW} = f(\hat{K}_t)\mu\left(\frac{\tilde{\Sigma}_t}{f(\hat{K}_t)}\right) dt - \tilde{K}_t \theta(\tilde{\beta}_t, \tilde{\Sigma}_t) dt - \phi(\tilde{\beta}_t, \tilde{\Sigma}_t) dt + \tilde{\Sigma}_t dZ_t. \tag{43}$$

Then, the probability measure induced by $\{\hat{K}, \hat{\sigma}, \tilde{K}\}$ is such that

$$\tilde{\Sigma}_t dZ_t^{\hat{K},\hat{\sigma},\tilde{K}} = \left(d\tilde{Y}_t - f(\hat{K})\mu\left(\frac{\tilde{\Sigma}}{f(\hat{K})}\right) dt + \tilde{K}_t \theta(\tilde{\beta}_t, \tilde{\Sigma}_t) dt + \phi(\tilde{\beta}_t, \tilde{\Sigma}_t) dt \right)$$

is an an $\mathcal{F}_t$-martingale (following the construction in Proposition 1). The incentive compatibility arguments in Proposition 1 also follow, with the modification that the agent's incentive compatibility condition is

$$\{\tilde{K}, \hat{K}\} = \arg\max \left[\tilde{\beta}_t f(\hat{K})\mu\left(\frac{\tilde{\Sigma}}{f(\hat{K})}\right) - \tilde{\beta}_t \tilde{K}_t \theta(\tilde{\beta}_t, \tilde{\Sigma}_t) + \lambda(\tilde{K}_t - \hat{K}_t) \right] \tag{44}$$

Plugging in $\theta = \frac{\lambda}{\tilde{\beta}_t}$ (20) yields

$$\{\tilde{K}, \hat{K}\} = \arg\max \left[\tilde{\beta}_t f(\hat{K}) \mu \left(\frac{\tilde{\Sigma}}{f(\hat{K})} \right) - \lambda \hat{K}_t \right] \tag{45}$$

and the observation that the agent is indifferent across choices of $\tilde{K}_t \geq \hat{K}_t$, and so will pick $\tilde{K}_t = \hat{K}_t$ (the principal's desired level, which is strictly preferred to $\tilde{K}_t < \hat{K}_t$). In addition, a comparison of (45) and (7), or direct calculation, shows that the agent's choice of $\hat{K}_t$ will be given by (8), with $\hat{\sigma}_t = \frac{\tilde{\Sigma}_t}{f(\hat{K}_t)}$.

Second, writing $\tilde{K} = \hat{K} = K(\tilde{\Sigma}, \tilde{\beta})$ and $\hat{\sigma}_t = \sigma(\tilde{\Sigma}, \tilde{\beta})$ (from 22 and 23), we obtain

$$d\tilde{Y}_t^{NEW} = f(K(\tilde{\Sigma}, \tilde{\beta}))\mu \left(\frac{\tilde{\Sigma}_t}{f(K(\tilde{\Sigma}, \tilde{\beta}))} \right) dt - K(\tilde{\Sigma}, \tilde{\beta})\theta(\tilde{\beta}_t, \tilde{\Sigma}_t)dt - \phi(\tilde{\beta}_t, \tilde{\Sigma}_t)dt + \tilde{\Sigma}_t dZ_t. \tag{46}$$

Using (21) and (20), we see that (46) reduces to $dY_t^{NEW} = \tilde{\Sigma}_t dZ_t$: the drift is identically equal to zero for any $\{\tilde{\beta}, \tilde{\Sigma}\}$. Thus we have $dW_t = \gamma W_t dt + \tilde{\beta}\tilde{\Sigma} dZ_t - dC_t$ from the arguments in Proposition 1 and $dM_t = \gamma M_t dt + \tilde{\beta}\tilde{\Sigma} dZ_t - dC_t$ from (18) and (46). Standard arguments (see DeMarzo and Sannikov (2006) or Biais et al. (2007)) imply that this is sufficient for $W_t = M_t$ and $dC_t = \max(M_t - W^C, 0)$. The indifference across $\{\tilde{\beta}, \tilde{\Sigma}\}$ allows the agent to pick the principal's desired level of $\{\tilde{\beta}, \tilde{\Sigma}\}$.

References

Ai, H. and R. Li (2015). Investment and ceo compensation under limited commitment. *Journal of Financial Economics 116*(3), 452–472.

Andrade, G. and S. N. Kaplan (1998). How costly is financial (not economic) distress? evidence from highly leveraged transactions that became distressed. *The Journal of Finance 53*(5), 1443–1493.

Aragon, G. O. and V. Nanda (2011). Tournament behavior in hedge funds: High-water marks, fund liquidation, and managerial stake. *Review of Financial Studies 25*(3), 937–974.

Bertrand, M. and S. Mullainathan (2003). Enjoying the quiet life? Corporate governance and managerial preferences. *Journal of Political Economy 111*(5), 1043–1075.

Biais, B., T. Mariotti, G. Plantin, and J.-C. Rochet (2007). Dynamic security design: Convergence to continuous time and asset pricing implications. *Review of Economic Studies 74*, 345–390.

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

Bolton, P., N. Wang, and J. Yang (2017). Liquidity and risk management: Coordinating investment and compensation policies. *Journal of Finance*. Working Paper.

Buffa, A. M., D. Vayanos, and P. Woolley (2015). Asset management contracts and equilibrium prices. NBER working paper.

Cadenillas, A., J. Cvitanić, and F. Zapatero (2004). Leverage decision and manager compensation with choice of effort and volatility. *Journal of Financial Economics 73*(1), 71–92.

Carpenter, J. N. (2000). Does option compensation increase managerial risk appetite? *Journal of Finance 55*(5), 2311–2331.

Chevalier, J. and G. Ellison (1997). Risk taking by mutual funds as a response to incentives. *The Journal of Political Economy 105*(6), 1167–1200.

Cvitanić, J., D. Possamaï, and N. Touzi (2016a). Dynamic programming approach to principal-agent problems. *Finance and Stochastics*. forthcoming.

Cvitanić, J., D. Possamaï, and N. Touzi (2016b). Moral hazard in dynamic risk management. *Management Science*. forthcoming.

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

DeMarzo, P. M., M. J. Fishman, Z. He, and N. Wang (2012). Dynamic agency and the q theory of investment. *Journal of Finance 67*(6), 2295–2340.

DeMarzo, P. M., D. Livdan, and A. Tchistyi (2013). Risking other people's money: Gambling, limited liability, and optimal incentives. Working Paper, Stanford University, University of California – Berkeley, and UIUC.

Dumas, B. (1991). Super contact and related optimality conditions. *Journal of Economic Dynamics and Control 75*(4), 675–685.

Edmans, A., X. Gabaix, and D. Jenter (2017). Executive compensation: A survey of theory and evidence. Technical report, National Bureau of Economic Research.

Eisdorfer, A. (2008). Empirical evidence of risk shifting in financially distressed firms. *The Journal of Finance 63*(2), 609–637.

Engelbert, H.-J. and G. Peskir (2014). Stochastic differential equations for sticky brownian motion. *Stochastics 86*(6), 993–1021.

Epstein, L. G. and S. Ji (2013). Ambiguous volatility and asset pricing in continuous time. *Review of Financial Studies 26*(7), 1740–1786.

Gilje, E. P. (2016). Do firms engage in risk-shifting? empirical evidence. *Review of Financial Studies 29*(11), 2925–2954.

Giroud, X. (2013). Proximity and investment: Evidence from plant-level data. *Quarterly Journal of Economics 128*(2), 861–915.

Graham, J. and C. Harvey (2011). Duke/CFO magazine global business outlook. U.S. topline tables, March 2011, www.cfosurvey.org/pastresults.htm.

Graham, J. and C. Harvey (2012). Duke/CFO magazine global business outlook. U.S. topline tables, June 2012, www.cfosurvey.org/pastresults.htm.

Graham, J. R. and C. R. Harvey (2001). The theory and practice of corporate finance: Evidence from the field. *Journal of Financial Economics 60*(2), 187–243.

Gryglewicz, S. and B. Hartman-Glaser (2017). Dynamic agency and real options. Working Paper, Erasmus University Rotterdam and UCLA.

Gryglewicz, S., S. Mayer, and E. Morellec (2018). Agency conflicts over the short and long run: Short-termism, long-termism, and pay-for-luck. Working Paper.

Harrison, J. M. and A. J. Lemoine (1981). Sticky brownian motion as the limit of storage processes. *Journal of Applied Probability 18*(1), 216–226.

He, Z. (2011). A model of dynamic compensation and capital structure. *Journal of Financial Economics 100*(2), 351–366.

Huang, J., C. Sialm, and H. Zhang (2011). Risk shifting and mutual fund performance. *Review of Financial Studies 24*(8), 2575–2616.

Jacobs, M. T. and A. Sivdasani (2012). Do you know your cost of capital? *Harvard Business Review* (7/8), 118–124.

Jagannathan, R., D. A. Matsa, I. Meier, and V. Tarhan (2016). Why do firms use high discount rates? *Journal of Financial Economics 120*, 445–463.

Karatzas, I. and S. Shreve (1998). *Brownian Motion and Stochastic Calculus* (2nd ed.). Springer-Verlag New York.

Leung, R. (2017). Continuous-time principal-agent problem with drift and stochastic volatility control: With applications to delegated portfolio management. Working paper. CKGSB.

Li, R. and N. Williams (2017). Optimal contracts with hidden risk. Working Paper, University of Wisconsin – Madison and University of Massachusetts – Boston.

Lioui, A. and P. Poncet (2013). Optimal benchmarking for active portfolio managers. *European Journal of Operational Research 226*(2), 268–276.

Malenko, A. (2018). Optimal dynamic capital budgeting. Working Paper, MIT.

Metrick, A. and A. Yasuda (2010). The economics of private equity funds. *Review of Financial Studies 23*(6), 2303–2341.

Moreno-Bromberg, S. and G. Roger (2016). Leverage and risk taking. Working Paper, University of Zurich and University of Sydney.

Ou-Yang, H. (2003). Optimal contracts in a continuous-time delegated portfolio management problem. *Review of Financial Studies 16*(1), 173–208.

Piskorski, T. and M. M. Westerfield (2016). Optimal dynamic contracts with moral hazard and costly monitoring. *Journal of Economic Theory 166*, 242–281.

Poterba, J. M. and L. H. Summers (1995). A CEO survey of U.S. companies' time horizons and hurdle rates. *MIT Sloan Management Review* (Fall), 43–53.

Prendergast, C. (2002). The tenuous trade-off between risk and incentives. *Journal of Political Economy 110*(5), 1071–1102.

Rauh, J. D. (2008). Risk shifting versus risk management: Investment policy in corporate pension plans. *Review of Financial Studies 22*(7), 2687–2733.

Robinson, D. T. and B. A. Sensoy (2013). Do private equity fund managers earn their fees? Compensation, ownership, and cash flow performance. *Review of Financial Studies 26*(11), 2760–2797.

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

Szydlowski, M. (2016). Incentives, project choice and dynamic multitasking. Working Paper, University of Minnesota.

Zhu, J. Y. (2013). Optimal contracts with shirking. *Review of Economic Studies 80*, 812–839.