Dynamic Multitasking and Managerial Investment Incentives*

Florian Hoffmann† Sebastian Pfeil‡

June 2018

Abstract

We study long-term investment in a dynamic agency model with multitasking. The manager's short-term task determines current performance which deteriorates if he invests in the firm's future profitability, his long-term task. The optimal contract dynamically balances incentives for short- and long-term performance such that investment is distorted upwards (downwards) relative to first-best in firms with high (low) technological returns to investment. These distortions decrease as good performance relaxes endogenous financial constraints arising from the agency problem, implying negative (positive) investment-cash flow sensitivities. Investment distortions and cash flow sensitivities increase in absolute terms with short-term performance pay and external financing costs.

Keywords: Continuous time contracting, multiple tasks, delegated investment, managerial compensation, endogenous financing frictions, investment dynamics.

JEL Classification: D86 (Economics of Contract: Theory), D92 (Intertemporal Firm Choice, Investment, Capacity, and Financing), J33 (Compensation Packages, Payment Methods).

*We thank Patrick Bolton, Wei Cui (CICF discussant), Guido Friebel, Sebastian Gryglewicz, Roman Inderst, Peter Kondor, Christian Leuz, Stephan Luck, Thomas Mariotti, Konstantin Milbradt (AFA discussant), Jean-Charles Rochet, Paul Schempp, Vladimir Vladimirov, Neng Wang, John Zhu, and conference and seminar participants at AFA 2015 in Boston, CICF 2013 in Shanghai, Columbia Business School, Erasmus School of Economics, Frankfurt School of Finance and Management, Stockholm School of Economics, Tilburg University, University of Amsterdam, University of Bonn, and University of Zürich for helpful comments. An earlier version of this paper has been circulated under the title "A Dynamic Agency Theory of Delegated Investment". Sebastian Pfeil acknowledges support by the University Research Priority Program FinReg of the University of Zurich.

†E-mail: fhoffmann@uni-bonn.de.

‡Erasmus School of Economics. E-mail: pfeil@ese.eur.nl.

1 Introduction

A manager responsible for a firm's operations usually has some form of discretion in running the day-to-day business, which relies on his specific skills or private information. Due to the separation of ownership and control, this gives rise to an agency problem, which has been the focus of much of the recent dynamic financial contracting literature.[1] Yet, in a dynamic world, firms also have to take strategic decisions and invest in order to maintain long-term profitability. Typically, this investment process also relies on information and skills of the same manager or is even delegated to him with discretion. In fact, many strategic investments, in particular those in intangibles such as investment in R&D, process innovation, product development, human capital, or distribution systems, share the following features: i) *actual* investment expenditures are not easily verifiable by firm owners as they are either not reported or reported expenses can be manipulated,[2] ii) their outcome is uncertain, and iii) they have persistent effects on the firm's profitability. The manager's hidden actions, thus, affect both the firm's current period payoffs as well as its long-term profitability, giving rise to a dynamic multitask problem.

To capture these ideas, we introduce delegated non-contractible investment in the firm's future profitability into a continuous time cash flow diversion model. We analyze how the need to incentivize the manager to meet both short-term as well as long-term targets affects the optimal compensation scheme as well as the efficiency of the investment process. Under the optimal long-term contract, investment depends on the entire history of past performance and is distorted away from the first-best level. We show that both the sign of the investment distortion as well as the comparative statics of investment (with respect to realized cash flows, corporate governance or financial frictions) crucially depend on the

[1] E.g., DeMarzo and Sannikov (2006), DeMarzo and Fishman (2007), Biais et al. (2007), or Hoffmann and Pfeil (2010). See also Biais et al. (2013) for an excellent survey.

[2] Indeed, several forms of intangible or "soft" investment expenditures are – unlike hard investments in plants, property or other equipment – not reported explicitly in firms' financial statements. For those intangible investments that are reported accounting rules often leave considerable discretion to managers "[...] to alter financial reports to either mislead some stakeholders about the underlying economic performance of the company, or to influence contractual outcomes that depend on reported accounting numbers," (Healy and Wahlen 1999) a practice referred to as *earnings management*. For instance, while R&D expenditures are reported in firms' income statements, "[...] the notion of what outlays are considered R&D [...] can be difficult to assess, and often represents the manager's discretionary choice" (Koh and Reeb 2015). Exploiting this discretion, e.g., by shifting core expenses to special items such as R&D, is referred to as *classification shifting* in the accounting literature (see, e.g., McVay 2006, Skaife et al. 2013, and Darrough et al. 2017, or, for related arguments also Bebchuk and Stole 1993, or Dutta and Reichelstein 2003).

technological returns to investment, i.e., the extend to which a firm's long-term profitability depends on successful investment. Investment is distorted upwards relative to the first-best benchmark when returns to investment are high and it is distorted downwards when they are low. These distortions decrease in *absolute terms* with financial slack and, thus, past performance. In line with recent empirical evidence (see, e.g., Peters and Taylor 2017 or Hovakimian 2009), this implies a positive relation between investment and realized cash flows in industries with low, and a negative relation in industries with high returns to intangible investment.[3] Further, if the costs of raising external funds increase, so will both the distortions in investment as well as its sensitivity to past performance. Our multitask model of investment, thus, provides a new perspective on investment-cash flow sensitivities as a measure of financial constraints when investment is non-contractible (cf., e.g., Brown and Petersen 2012 or Chen and Chen 2012 and, for the classical discussion related to capital investment, Fazzari et al. 1988 and Kaplan and Zingales 1997).

The key agency frictions in our dynamic multitask theory of investment are due to i) unobservable cash flows and ii) non-contractibility of actual investment expenditures. An agency problem with respect to investment then arises *endogenously* as a result of compensation for short-term performance which is necessary to deter the manager from diverting cash flows for private consumption: Since his compensation is tied to current earnings figures, the manager has an incentive to cut profitable investment in order to boost the firm's short-term profits at the expense of its long-term profitability.[4] Our model, thus, shares with the literature on managerial short-sightedness the notion that short-term incentive schemes can induce managers to focus excessively on current period outcomes.[5] However, in our model such unintended consequences of short-term incentive pay can be mitigated by tying compensation also to indicators of investment success such

[3]Clearly, *actual* investment expenditures are non-contractible for a large class of intangible investments, which is the key agency friction studied in this paper. However, in light of our model, *reported* (accounting) measures commonly used for empirical testing, such as R&D or SG&A spending, accurately reflect the respective actual expenditures under an incentive compatible contract.

[4]These incentives should be particularly strong for intangible investment expenditures, which, in contrast to investments in physical capital, are mainly expensed rather than capitalized. Accordingly, financial constraints, which in our model arise endogenously from the agency model, should be particularly strong for these investments due to their low collateral value (see Almeida and Campello 2007).

[5]This short-term bias has been attributed to career concerns (Narayanan 1985), takeover threats (Stein 1988), concerns about the firm's stock price over a near-term horizon (Stein 1989), mispricing of long-term assets (Shleifer and Vishny 1990), reputational herding (Zwiebel 1995) and short-term financing (von Thadden 1995).

as project milestones or the number of patents granted (see, e.g., Balkin et al. 2006).

Yet, as is well known from related (single-task) dynamic financial contracting models (see, e.g., DeMarzo and Sannikov 2006), exposing the manager to compensation risk in order to provide incentives is costly. In our model with bilateral risk neutrality this formally is a consequence of limited liability, which requires to terminate the contract if, after a series of bad outcomes, the manager is "too poor to be punished." In this case he must be replaced, which is costly ex-post, but part of the ex-ante optimal long-term contract. Hence, the need to tie the manager's pay to (imperfect) signals of investment success or failure, such as to incentivize investment, creates additional agency costs. The optimal investment schedule is then determined by trading-off the agency costs of investment with the potential efficiency gains as captured by the respective returns to investment. We show that the agency costs of investment are nonmonotonic in the implemented investment level. In particular, marginal agency costs are positive for low and negative for high levels of investment. As a consequence, investment is distorted downwards relative to the first-best (owner-manager) benchmark in industries with low returns to investment, i.e., industries in which first-best investment is low. By contrast, investment is distorted upwards relative to first-best when returns to investment and therefore also first-best investment is high.

To understand this result consider, first, the compensation policy that incentivizes a given investment level at minimal agency costs. Since firm owners are effectively risk averse with respect to variation in the *manager's* expected compensation, any contingent reward or punishment is costly. It is, thus, optimal to condition incentive pay mainly on the most informative performance signal (success/failure), i.e., the one that provides strongest incentives per unit of expected compensation (reward/punishment) as captured by a high absolute value of the likelihood ratio.

Since incentive costs are, thus, mainly incurred following the realization of the more informative investment outcome, it is optimal to distort investment relative to the first-best value such as to make this outcome less likely. Overinvestment may, hence, arise if incentives are predominantly given through punishment for bad outcomes (failure), which is the case if the investment technology is sufficiently profitable, corresponding to high first-best investment, such that an (unlikely) investment failure is particulary informative about the agent having deviated. Then, in order to reduce the probability of having to bear the high agency costs associated with an investment failure, it is optimal to distort investment upwards. Similarly, if the investment technology is rather unprofitable with low returns to investment, investment will be distorted downwards relative to the (low) first-best investment benchmark.

When the manager's track record improves, his stake in the firm under the optimal contract increases, mitigating the agency problem and, thus, reducing investment distortions. Accordingly, investment expenditures are negatively related to past performance (i.e., the entire history of cash flows) in overinvesting firms with high returns to investment, while in underinvesting firms with low returns to investment the relation is positive. Within any standard implementation (see, e.g., Biais et al. 2007), this history dependence of the optimal contract is captured by a measure of the firm's *financial slack*. For concreteness, consider an implementation in which the firm retains earnings to build up financial slack in the form of a cash buffer, that is used to cover investment expenditures, potential operating losses, and payments to investors. If these cash holdings increase, financial constraints are relaxed and investment distortions decrease, implying higher/lower investment in under-/overinvesting firms. When cash holdings fall to zero, the firm is no longer able to honor its payments to investors, triggering a costly restructuring of the firm, which involves replacing the incumbent manager and raising new external funds. We show that if the costs of raising external funds and, hence, financial constraints increase, investment distortions become more severe and investment becomes more sensitive to cash flows for any given level of financial slack. This amplification implies that investment *increases* and the relation between financial slack and investment becomes more *negative* in firms with high returns to investment, while investment *decreases* and its sensitivity to financial slack becomes more *positive* in firms with low returns to investment.

Our multitask theory of investment, thus, adds to the debate about the relationship between investment-cash flow sensitivities and financial constraints, which here arise endogenously from the agency problem. By focussing on investment in future profitability per unit of (physical) capital instead of standard capital investment, our model can be viewed as a first step towards accounting for the increasing importance of intangible investment, e.g., in knowledge or organizational capital (see, for instance, Brown and Petersen 2009). In this sense, recent empirical evidence in Hovakimian (2009) and Peters and Taylor (2017) indeed confirms our model's prediction that investment-cash flow sensitivities should be negative in industries where intangible investment is particularly important. Further, since we show that financially constrained firms may feature both positive as well as negative investment-cash flow sensitivities depending on industry-specific returns to investment, our model offers an explanation for the puzzling findings in Chen and Chen (2012) that on average (i.e., not controlling for intangible investment intensity) investment-cash flow sensitivities have disappeared despite financial constraints still being a first-order concern for firms.

The multitask nature of our dynamic optimal contracting model further allows to derive new testable implications regarding the relation between pay for short-term performance and long-term investment. If, in our model, the cash flow diversion problem becomes more severe, e.g., due to worse corporate governance, then i) the optimal contract has to be more sensitive with respect to short-term performance. This exacerbates the agency problem with respect to long-term investment, such that, ii) investment distortions increase in absolute terms. That is, steeper short-term incentives are associated with *higher* investment expenditures in firms with high returns to investment and with *lower* investment expenditures in firms with low returns to investment.

Related Literature. From a methodological perspective, our model is most closely related to Hoffmann and Pfeil (2010) who introduce exogenous (Poisson) shocks to firm profitability in a dynamic cash flow diversion model à la DeMarzo and Sannikov (2006). While the firm's long-term profitability in their model is purely determined by exogenous "lucky" shocks, it can be controlled by the manager in the present analysis via investment expenditures, giving rise to a fully dynamic multitask problem (see Holmstrom and Milgrom 1991). In that respect our analysis is related to Zhu (2018) who considers a model of persistent moral hazard in which the agent can choose between a short-term or a long-term action in every period. He characterizes the optimal contract that always induces the long-term action. Our focus is instead on determining the optimal interior investment level at every possible history while abstracting from problems of persistent private information.

Similarly, our approach is related to Varas (2017) who studies a multitasking model of project completion in which the agent can complete a project faster by reducing its quality which is a one-time action. In contrast, we consider a fully dynamic repeated problem in which the agent can divert cash flow for private consumption as well as move funds within the firm in order to meet either short- or long-term targets, where his discretion is restricted by the optimal contract. In this sense, our model is related to the capital budgeting literature starting with Harris and Raviv (1996) and extended to a dynamic setting in Malenko (2018).

Optimal investment with dynamic agency is also analyzed in the two-period model of Gertler (1992) as well as the multiperiod discrete time models in Quadrini (2004), Clementi and Hopenhayn (2006), and DeMarzo and Fishman (2007). Further, DeMarzo et al. (2012) analyze dynamic agency in a neoclassical investment setting using continuous time methods. All of these contributions consider capital investment which can be verified by firm owners. In our model, by contrast, the manager privately controls investment and, thus, has to be incentivized to invest in the firm owners' interest. Further, (intangible) investment in our model does not scale cash flows by increasing the capital stock as in neoclassical investment settings, but rather affects mean cash flows per unit of (constant) firm size.

The remainder of the paper is organized as follows. We introduce the model in Section 2. In Section 3 we derive the optimal dynamic contract and provide a detailed discussion of optimal compensation and investment. Section 4 provides comparative statics and derives empirical implications. Section 5 concludes. All proofs as well as some additional technical material are contained in the Appendix.

2 The Model

We consider an infinite horizon continuous time principal-agent model of a firm whose owners (the principal) hire a manager (the agent) to operate the business, which includes investment in the firms future profitability. We now, first, present the firm's cash flow process and investment technology. Second, we introduce the agency problem between firm owners and the manager, and formulate the optimal contracting problem.

2.1 Cash Flow Process and Investment Technology

The firm's cash flows, $\mathbf{Y} = \{Y_t : t \geq 0\}$, net of investment expenditures, $\mathbf{I} = \{I_t \in [0, \overline{I}] : t \geq 0\}$, evolve according to

$$dY_t = (\mu_t - I_t)\, dt + \sigma dZ_t, \tag{1}$$

where μ_t denotes the prevailing drift rate of cash flows, which we refer to as the firm's "profitability," σ the instantaneous volatility, and $\mathbf{Z}$ is a standard Brownian motion on a complete probability space. For notational simplicity we restrict attention in the main text to two possible profitability levels $\mu_t \in \{\mu^l, \mu^h\}$, where $0 < \mu^l < \mu^h$. It is straightforward to extend the analysis allowing for an arbitrary number of profitability levels (see Appendix B.2, where we also show that our main results continue to hold).

Investment affects the firm's future profitability by determining its ability to adopt and commercialize new technologies as they become available.[6] As illustrated in Figure 1, the industry is subject to (rare) exogenous *technology shocks* governed by a Poisson process $\mathbf{N}$ with intensity ν, indicating the availability of a new technology. If there is a technology shock ($dN_t = 1$), the firm is able to adopt the new technology with probability $p(I_t)$ and its future profitability will be high ($\mu_{t^+} := \lim_{s \downarrow t} \mu_s = \mu^h$). We refer to this event as an *investment success*, which can be interpreted as moving to or staying at the research frontier. With probability $1 - p(I_t)$ the firm is, however, unable to adopt the new technology sufficiently quickly to stay competitive, and its future profitability will, thus, be low ($\mu_{t^+} = \mu^l$). We refer to this as an *investment failure*, which can be interpreted as falling or staying behind the research frontier. For future reference, note that the occurrence of an investment success is, thus, governed by a Poisson process $\mathbf{N^g}$ with arrival rate $\nu p(I_t)$ and the occurrence of an investment failure by a Poisson process $\mathbf{N^b}$ with arrival rate $\nu (1 - p(I_t))$.[7] In between two technology shocks, profitability remains unchanged such that investment has persistent effects. Finally, we stipulate that the success probability $p(\cdot) \in [0, 1]$ is an increasing and strictly concave function of the investment amount I, satisfying $p(0) = 0$ and $p(\overline{I}) = 1$.

[6] Concretely, investment expenditures can be interpreted as a firm's choice of *absorptive capacity* in the sense of Cohen and Levinthal (1990), describing its capability of "assimilating new, external information and apply it to commercial ends." While we frame the investment technology within the context of technology adoption it is, hence, more widely applicable. E.g., one could similarly think of the industry being subject to unpredictable demand shocks, to which the firm can react quickly enough only if it has invested sufficient resources, e.g., into building capacity in market research or product design.

[7] Investment in our model, thus, depreciates instantly, reflecting the need to keep up with fast technological progress. Indeed, evidence suggests that, e.g., R&D investment depreciates much faster than physical capital (see, Bernstein and Mamuneas 2006). Nevertheless, the main drivers of our results do not depend on this assumption and our insights should extend also to the case where investment depreciates gradually over time, which however significantly complicates the analysis.

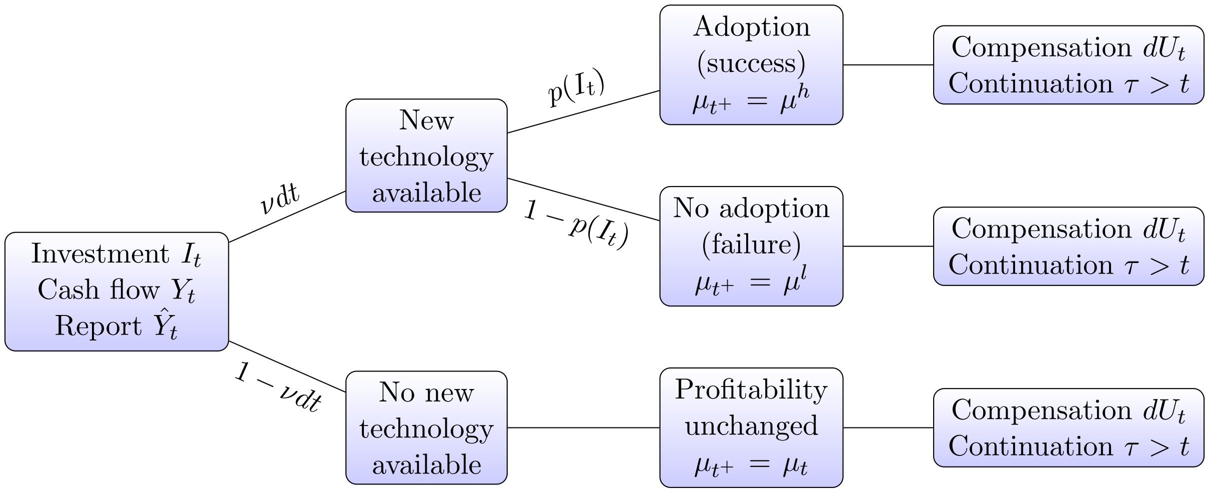

Figure 1: The figure illustrates the timing of events in a generic time interval $[t, t + dt]$: i) The agent decides on investment, current cash flows are realized, and reported to the principal. ii) A new technology becomes available with probability νdt. iii) The firm is able to adopt the new technology with probability $p(I_t)$, leading to high future profitability ($\mu_{t^+} = \mu^h$). With probability $1 - p(I_t)$, the firms fails to adopt the new technology and has low future profitability ($\mu_{t^+} = \mu^l$). If no new technology is available (which happens with probability $1 - \nu dt$) profitability remains unchanged. iv) The agent gets compensated and the continuation decision is taken.

2.2 Agency Problem

Firm owners have to hire a manager to (profitably) run the business which relies on the manager's specific skills or private information. Both parties are risk neutral and the manager is protected by limited liability and has no initial wealth. Hence, owners have to bear the costs of setting up the firm (normalized to zero) and cover operating losses. In order to capture liquidity needs on behalf of the manager, we assume, as is standard in dynamic contracting models (see, e.g., DeMarzo and Duffie 1999 or DeMarzo and Sannikov 2006), that he is relatively impatient, i.e., the discount rates of the manager (γ) and firm owners (r) satisfy $\gamma > r$.

Running the firm requires some form of discretion over the firm's cash flows, which we model as follows: Firm owners do not observe cash flows $\mathbf{Y}$ directly but only the manager's contractible reports $\hat{\mathbf{Y}}$. The difference between actual and reported cash flows is determined by the manager's hidden action, which is the source of the agency problem. In particular, we assume that the manager can divert cash flows for private consumption. To capture possible costs of concealing and taking funds out of the firm we stipulate that the manager can consume fraction $\lambda \in (0, 1]$ of each unit diverted.

As investment similarly requires a considerable understanding of the firm's internal processes as well as the market it is operating in, it is also delegated to the manager with discretion. To model this in a concise way we assume that, in addition to cash flows, also actual investment expenditures $\mathbf{I}$ are not contractible. Hence, the manager can freely use cash flows for investment purposes or, respectively, inflate net cash flows $\mathbf{Y}$ by deviating from the prescribed investment schedule (see (1)). While firm owners may observe (contractible) reports of $\mathbf{I}$, e.g., as documented in the firm's income statement, these reports do not have to coincide with actual investment expenditures as the manager can divert funds from his investment budget for private benefit, or misrepresent running expenses as investments.[8] Hence, independently of whether reports of investment expenditures are available or not, to provide incentives the contract has to condition on the (contractible) investment outcome, as governed by the Poisson processes $\mathbf{N^g}$ and $\mathbf{N^b}$. In fact, as the manager will always report to have followed the contractually specified investment schedule, providing incentives for investing according to contract trivially ensures truthful reporting.

The principal can commit to a long-term compensation contract $(\mathbf{U}, \tau)$ specifying – based on reported cash flows and investment outcomes – wage payments to the agent, dU_t, as well as the (random) time τ when the agent is fired and replaced, which incurs the principal fixed costs $k > 0$. Still, because limited liability requires the cumulative wage process $\mathbf{U} = \{U_t : 0 \leq t \leq \tau\}$ to be non-decreasing, firing and replacing the agent may be necessary for incentive provision and will indeed be part of the ex-ante optimal contract. We further assume that the agent cannot save privately, implying that $dY_t - d\hat{Y}_t \geq 0$, i.e., he can only underreport cash flows.[9] Hence, the agent's consumption flow at time t equals the sum of wage payments and the utility from consuming diverted funds (if any):

$$dC_t = dU_t + \lambda \left[dY_t - d\hat{Y}_t dt \right].$$

[8]Indeed, several relevant components of intangible investment have to be reported (such as R&D), however, accounting rules leave considerable discretion to managers over which expenses to classify in the respective categories. This allows managers, e.g., to shift core expenses to other expense categories (such as R&D) to improve core earnings, a form of earnings management referred to as classification shifting (see footnote 2). Within the context of our model, core earnings *reported* in the firm's income statement would then follow $d\hat{Y}_t = (\mu_t - a_t + b_t)dt + \sigma dZ_t$, where a_t represents diversion for private consumption and b_t denotes the amount of core expenses shifted to another expense category, captured as investment expenditures in our model. The difference of reported and actual investment is then given by b_t. While it is irrelevant for our theoretical analysis whether investment is reported or not, this becomes relevant when considering empirical testing. We defer further discussion to Section 4, where we also consider an alternative interpretation of our agency problem with respect to investment as an effort problem.

[9]Given that the agent is risk-neutral and relatively impatient, this is without loss of generality.

When the agent is fired he receives his outside option, which is normalized to zero. The incumbent agent's total expected wealth at the start of his employment, hence, is

$$w_0 = \mathbb{E}_0^S \left[\int_0^\tau e^{-\gamma t} dC_t \right], \tag{2}$$

where the expectation is conditional on time zero information and depends on the agent's reporting and investment strategy $\mathbf{S} = \left\{ \hat{Y}_t, I_t : 0 \leq t \leq \tau \right\}$.[10] The actual value of $w_0 > 0$ in (2) is determined by the two parties' relative bargaining power. For concreteness, in what follows, we will assume that the principal enjoys all bargaining power and the agent accepts any contract with $w_0 \geq 0$.[11]

If the agent is replaced, the principal's value is given by L_τ, which denotes the (endogenous) expected profit from the relationship with a new agent, net of replacement costs k. We assume that k is sufficiently small such that continuing the firm is optimal. Hence, at $t = 0$, the principal's total expected profit, delivering the agent an expected payoff of w_0 and given initial profitability μ_0, is

$$f_0 = \mathbb{E}_0 \left[\int_0^\tau e^{-rt} \left(d\hat{Y}_t - dU_t \right) + e^{-r\tau} L_\tau \right]. \tag{3}$$

Given a compensation contract $(\mathbf{U}, \tau)$, the agent chooses a feasible strategy $\mathbf{S}$ to maximize his initial expected payoff w_0 from (2). A strategy $\mathbf{S}$ is called *incentive compatible* if it solves the agent's maximization problem. Hence, an incentive compatible contract can be described by the triple $(\mathbf{S}^*, \mathbf{U}, \tau)$, where $\mathbf{S}^*$ is the (incentive compatible) strategy that the principal wants to induce. The associated (global) incentive constraint is given by

$$\mathbb{E}_0^{S^*} \left[\int_0^\tau e^{-\gamma t} dC_t \right] \geq \mathbb{E}_0^{\widetilde{S}} \left[\int_0^\tau e^{-\gamma t} dC_t \right], \quad \text{for any } \widetilde{\mathbf{S}} \neq \mathbf{S}^*. \tag{4}$$

We can simplify the analysis considerably by relying on a version of the revelation principle, which allows us to restrict attention to truth-telling contracts, implementing $\hat{\mathbf{Y}} = \mathbf{Y}$.[12] The contracting problem then is to find an incentive compatible truth-telling contract maximizing the principal's expected profit f_0 for given initial profitability μ_0, delivering expected payoff w_0 to the agent and satisfying limited liability. We refer to the solution

[10] Formally, $\mathbb{E}^S$ denotes the expectation under the probability measure Q^S induced by $\mathbf{S}$. If obvious, we will not state the measure associated with the expectation operator in the following.

[11] This is not crucial for our results. We just require relative bargaining power to be constant over time.

[12] For a formal argument compare Lemma 1 and Proposition 2 in DeMarzo and Sannikov (2006).

of this constrained maximization problem, which includes the optimal investment profile, as the *optimal contract*.

A heuristic outline of the timing of events taking place in any infinitesimal time interval $[t, t+dt]$ prior to replacement of the incumbent manager, is illustrated in Figure 1. Note, in particular, that the agent chooses investment prior to observing whether a technology shock occurred or not, which is the main important "sequentiality" in our continuous time model, while compensation and possible replacement of the manager occur thereafter.[13]

3 Model Solution

In this section we solve for the optimal contract using a dynamic programming approach. As we will show, the optimal contract can be written in terms of the agent's continuation payoff as the single state variable. We derive the dynamics of this key state variable along with local incentive compatibility constraints in Section 3.1. In Section 3.2 we set up the recursive formulation of the contracting problem and characterize the optimal contract. Finally, in Section 3.3, we discuss in detail the distortions in investment relative to the first-best (owner-manager) benchmark as well as its dynamics under the optimal contract.

3.1 Continuation Payoff and Local Incentive Compatibility

For any truth-telling contract, define the agent's continuation payoff, w_t, as his future expected discounted payoff at time t, given he will follow strategy $\mathbf{S}^*$ from t onwards, i.e.,

$$w_t = E_t^{S^*}\left[\int_t^\tau e^{-\gamma(s-t)} dC_s\right]. \tag{5}$$

While w_t is the continuation payoff after observing whether a technology shock occurs in t or not, investment expenditures can only condition on the respective value before this uncertainty is resolved, which we denote by $w_{t^-} := \lim_{s\uparrow t} w_s$.[14] This variable will serve, together with profitability μ_t, as the state variable in the recursive formulation of the optimal contracting problem. Applying standard martingale techniques, the evolution of w_{t^-} can be characterized as follows:

[13]Formally, $\mathbf{I}$ is predictable with respect to $\mathcal{F} = \{\mathcal{F}_t, t \geq 0\}$, the filtration generated by $(\mathbf{Z}, \mathbf{N^g}, \mathbf{N^b})$, while $\mathbf{U}$ is adapted to $\widehat{\mathcal{F}} = \{\widehat{\mathcal{F}}_t, t \geq 0\}$, the filtration generated by $(\hat{\mathbf{Y}}, \mathbf{N^g}, \mathbf{N^b})$ and τ is an $\widehat{\mathcal{F}}$-measurable stopping time.

[14]Technically, investment decisions have to be predictable with respect to investment outcomes.

Lemma 1. *For any contract, there exist some predictable processes* $\boldsymbol{\alpha} = \{\alpha_t : 0 \leq t \leq \tau\}$ *and* $\boldsymbol{\beta^j} = \left\{\beta_t^j : 0 \leq t \leq \tau\right\}$, $j \in \{g, b\}$, *such that, if the manager follows the recommended strategy* $\mathbf{S}^* = \{Y_t, I_t^* : 0 \leq t \leq \tau\}$, *his continuation payoff at any moment of time evolves according to:*

$$dw_{t^-} = \gamma w_{t^-} dt - dU_t + \alpha_t \big(d\hat{Y}_t - (\mu_t - I_t^*)\, dt\big) + \beta_t^g \big(dN_t^g - \nu p(I_t^*) dt\big) + \beta_t^b \big(dN_t^b - \nu\,(1 - p(I_t^*))\, dt\big), \tag{6}$$

where $d\hat{Y}_t - (\mu_t - I_t^*)\, dt$ *is the increment of a Brownian motion and* dN_t^g, dN_t^b *are increments of standard Poisson processes with arrival rates* $\nu p(I_t)$ *and* $\nu\,(1 - p(I_t))$ *respectively.*

To build intuition, let us discuss in more detail the evolution of w_{t^-} in (6). Due to promise keeping, the agent's promised wealth, w_{t^-}, has to grow at his discount rate, γ, while it must decrease with direct wage payments, dU_t. It also depends on his reporting strategy via sensitivity α_t and on the investment outcome via sensitivities β_t^g and β_t^b. So, if the agent underreports cash flows, such that he can immediately consume $\lambda(dY_t - d\hat{Y}_t)$, his continuation payoff is reduced by $\alpha_t(dY_t - d\hat{Y}_t)$. Incentive compatibility for truth-telling, thus, requires $\alpha_t \geq \lambda$. Next, consider a deviation from the recommended investment I_t^*. When the agent reduces investment I_t marginally below I_t^*, from (1), net current cash flows dY_t improve at the same rate, and the agent's continuation payoff w_{t^-} grows with sensitivity α_t. However, cutting investment implies that the probability of an investment success – triggering "reward" β_t^g – decreases at rate $\nu p'(I_t^*)$, while the probability of a failure – triggering "punishment" β_t^b – increases by $\nu p'(I_t^*)$. Hence, the agent has no incentives to invest less than the recommended level $I_t^* \in \left(0, \overline{I}\right]$ if and only if $\nu p'(I_t^*)\left(\beta_t^g - \beta_t^b\right) \geq \alpha_t$. Likewise, marginally increasing investment above I_t^* increases the probability of receiving bonus β^g by $\nu p'(I_t^*)$ and decreases the probability of punishment β^b by $\nu p'(I_t^*)$. However, as additional investment has to be financed out of current cash flows, the continuation payoff w_{t^-} is reduced by sensitivity α_t. Hence, the agent has no incentives to invest less than the recommended level $I_t^* \in \left[0, \overline{I}\right)$ if and only if $\nu p'(I_t^*)\left(\beta_t^g - \beta_t^b\right) \leq \alpha_t$. These observations lead to the local incentive compatibility conditions in Lemma 2 below.

Lemma 2. *The truth-telling contract* $\{\mathbf{S}^*, \mathbf{U}, \tau\}$ *with recommended investment* $I_t^* \in \left(0, \overline{I}\right)$ *is incentive compatible if and only if*

$$\alpha_t \geq \lambda \tag{7}$$

and

$$\nu p'(I_t^*)\left(\beta_t^g - \beta_t^b\right) = \alpha_t \tag{8}$$

holds for all $t \in [0, \tau)$, almost surely.[15] *Further, the limited liability constraint implies for all $t \in [0, \tau)$ that*

$$\beta_t^j \geq -w_{t^-},\ j \in \{g, b\}. \tag{9}$$

Incentive constraint (8) reflects the interaction between the problem of non-contractible investment and the cash flow diversion problem: As the agent's continuation value must be tied to reported cash flows ($\alpha_t > 0$), it also has to increase following an investment success and decrease following a failure to provide incentives for investment. More precisely, the sensitivity of the agent's continuation payoff with respect to the investment outcome, $\beta_t^g - \beta_t^b$, has to increase with the sensitivity to reported cash flows, α_t, for any given level of investment. Thus, a more severe cash flow diversion problem, reflected by greater diversion benefits λ, would require to impose more risk on the agent's income in two respects: First, from (7), the sensitivity to current cash flows, α_t, has to increase, because with higher diversion benefits the agent's income has to be linked more closely to reported cash flows in order to induce truthful reporting. Second, with his income being more sensitive to current cash flows, the agent has an incentive to inflate those by deviating from the recommended investment level. Hence, according to (8), his income has to be more responsive to the investment outcome, i.e., to future profitability as well.

3.2 Optimal Contract

The optimal contract can now be derived using the dynamic programming approach. Denote by $f^i(w)$, $i \in \{l, h\}$ the principal's value function, that is, the highest profit the principal can attain under any incentive compatible contract delivering expected payoff w to the agent and given the prevailing drift rate of cash flows, μ^i, $i \in \{l, h\}$, where we drop time subscripts for notational convenience.

The contracting problem can be greatly simplified by noting that, following any history, the optimal (continuation) contract is independent of the current profitability state μ^i. Intuitively this result follows from noting that, for a given level of investment, the

[15] If $I_t^* = 0$ (8) is replaced by $\nu p'(I_t)\left(\beta_t^g - \beta_t^b\right) \leq \alpha_t$, in case $I_t^* = \bar{I}$ we instead require $\nu p'(I_t)\left(\beta_t^g - \beta_t^b\right) \geq \alpha_t$.

probability of an investment success or an investment failure does not depend on current profitability and neither do the basic cash flow diversion problem nor the agency problem with respect to investment. As the optimal compensation and investment policies are, thus, independent of the prevailing profitability level, the principal's value in the two profitability states only differs by an additive constant, which captures the direct benefit from being at the research frontier, $\mu^h - \mu^l$, properly accounting for the Markov-switching structure.

Lemma 3. *The optimal contract* $(\mathbf{S}^*, \mathbf{U}, \tau)$ *is independent of current profitability* μ^i *and the principal's value functions in the high and in the low profitability state satisfy*

$$f(w) := f^l(w) = f^h(w) - \Delta, \tag{10}$$

with

$$\Delta := \frac{1}{r+\nu}\left(\mu^h - \mu^l\right). \tag{11}$$

From Lemma 3 we can now characterize the optimal contract based on the agent's continuation payoff w as the single state variable. Consider, first, the optimal compensation policy. Clearly, the principal can always compensate the agent directly by paying him a lump-sum of $dU > 0$ (at marginal costs of -1) and then move to the optimal contract with reduced continuation payoff $(w - dU)$. However, deferring compensation may be valuable: A higher promised wealth, w, relaxes future limited liability constraints as the agent can be punished for poor results by "clawing-back" previously deferred payments instead of having to inefficiently fire and replace him at cost k. In contrast to the costs of deferring compensation, which are due to the wedge in discount rates ($\gamma > r$), this benefit declines, however, as the agent's continuation payoff, w, increases and with it the probability of inefficient firing. This is reflected in the concavity of $f(w)$, which we show formally in the proof of Proposition 1 below. As a consequence, the principal is *effectively risk averse* with respect to variation in the agent's continuation value. One implication is that compensation is optimally deferred until a threshold $\overline{w}$ is reached, where $f'(\overline{w}) = -1$ and the agent is paid cash.

Next, since the principal discounts at rate r, his expected flow of value at time t must be $rf^i(w)dt$. This has to be equal to the expected instantaneous net cash flow $(\mu^i - I)\,dt$ plus the expected change in his value function, which can be computed using Itô's lemma and the change of variables formula for jump processes. Hence, on the interval without cash compensation, $w \in [0, \overline{w}]$, the principal's value function in the low profitability state must satisfy the following HJB equation ($f^h(w)$ then follows from (10)):

$$(r+\nu) f(w) = \max_{\alpha,\beta,I} \left\{ \begin{array}{l} \mu^l - I + \nu p(I)\Delta + \frac{1}{2}\sigma^2\alpha^2 f''(w) \\ + \left[\gamma w - \nu \left[\beta^g p(I) + \beta^b (1-p(I))\right]\right] f'(w) \\ +\nu p(I) f(w+\beta^g) + \nu (1-p(I)) f(w+\beta^b) \end{array} \right\} \tag{12}$$
$$s.t. \ (7),\ (8),\ (9).$$

Since cash transfers cause w to reflect at $\overline{w}$, the value function extends linearly to the right of the compensation threshold, i.e., $f(w) = f(\overline{w}) - (w - \overline{w})$ for $w \geq \overline{w}$. To solve for the optimal contract we have to pin down a solution to (12) and the compensation threshold $\overline{w}$. For this we require three boundary conditions: The first ("value matching") condition

$$f(0) = f(w^*) - k, \tag{13}$$

with $w^* \in \arg\max_w \{f(w)\}$, reflects that upon firing the incumbent agent at $w = 0$, the principal receives $f(w^*)$ from the relation with a new agent and bears replacement costs k. The second is a standard "smooth pasting" condition at the compensation threshold

$$f'(\overline{w}) = -1, \tag{14}$$

while the third boundary condition ("super contact") guarantees the optimal choice of $\overline{w}$

$$f''(\overline{w}) = 0. \tag{15}$$

The following Proposition characterizes the optimal contract, i.e., the solution to the boundary value problem in (12)-(15).

Proposition 1. *The optimal truth-telling contract with non-contractible investment takes the following form: Optimal investment $I(w)$ as well as the sensitivities $\alpha(w) = \lambda$ and $\beta^j(w)$, $j \in \{g, b\}$, are independent of μ^i, $i \in \{l, h\}$, and chosen as maximizers in (12). The incumbent agent's continuation payoff evolves according to (6) with $\alpha_t = \lambda$, $\beta_t^j = \beta^j(w_{t^-})$, $j \in \{g, b\}$ and $I_t = I(w_{t^-})$, $\forall t$. Optimal compensations satisfies $dU_t = \max\{w_{t^-} - \overline{w}, 0\}$. The incumbent agent is replaced when $w_{t^-} = 0$. The principal's expected payoff at any* *point in time is given by $f^i(w_t)$, $i \in \{h, l\}$, which satisfies $f(w) := f^l(w) = f^h(w) - \Delta$, where $f(w)$ is concave, strictly so for $0 \leq w < \overline{w}$ and solves, for $w \in [0, \overline{w}]$, the HJB equation in (12) subject to the boundary conditions (13) to (15).*

As is standard in dynamic cash flow diversion models (see, e.g., DeMarzo and Sannikov 2006), the agent's incentive constraint with respect to cash flow diversion (7) binds under the optimal contract. Formally, this result follows from the concavity of $f(w)$ and the observation that increasing α would increase the instantaneous volatility of w. Given that the basic cash flow diversion problem is by well now understood, in the following Section 3.3, we focus on the investment task and discuss in detail the optimal investment schedule, as well as the optimal choice of reward and punishment used to incentivize investment under the optimal contract of Proposition 1.

3.3 Optimal Investment

Incentive compatibility requires that the optimal contract contains a certain degree of reward following an investment success, $\beta^g(w) > 0$, and/or punishment after a failure, $\beta^b(w) < 0$. However, providing these incentives for investment is costly such that, in determining the optimal level of investment, the principal trades off the potential for higher profitability with the additional agency costs of providing incentives for investment. To see this formally, note that from (12) we can write the principal's problem of finding the optimal investment level as follows:

$$I(w) \in \arg\max_{I \in [0, \overline{I}]} \{\nu p(I)\Delta - I - \Phi(w, I)\}, \tag{16}$$

$$\text{with } \Phi(w, I) = \min_{\beta^g, \beta^b \ s.t. (8),(9)} \left\{ \begin{array}{r} \nu p(I)\left[f(w) + \beta^g f'(w) - f(w + \beta^g)\right] \\ +\nu\big(1 - p(I)\big)\left[f(w) + \beta^b f'(w) - f(w + \beta^b)\right] \end{array} \right\}. \tag{17}$$

From (16), optimal investment is chosen to maximize the (expected) technological returns to investment, $\nu p(I)\Delta$, net of investment expenditures, I, and the agency costs of providing incentives for investment denoted by Φ. Since firm owners are effectively risk-averse with respect to variation in the manager's compensation, as reflected in the concavity of $f(w)$, these agency costs are strictly positive whenever incentive compatible rewards for investment success, β^g, and punishments for investment failure, β^b, induce additional variation in the agent's continuation payoff w. Accordingly, the optimal combination of reward and punishment to implement a given level of investment are determined from (17), minimizing the agency costs of providing incentives for investment, Φ.

In the following, we now first study optimal compensation and investment distortions for a *given* value of the agent's continuation utility w. Then, we analyze the implied compensation and investment *dynamics* when w evolves as described in Proposition 1.

Optimal Incentives for Investment. Minimizing the agency costs of investment Φ in (17), with Lagrange multiplier $\eta(w)$ attached to the incentive compatibility constraint (8), implies that interior optimal values $\beta^g(w)$ and $\beta^b(w)$ have to satisfy[16]

$$\left[f'(w) - f'(w + \beta^i(w))\right] / LR^i(I(w)) = \eta(w), \tag{18}$$

for $i \in \{g, b\}$. Here $LR^i(I) := \frac{d}{dI} \log\left(\Pr\left(\mu = \mu^i | \, I\right)\right)$ denotes the *likelihood ratio* of the respective investment outcome, i.e., $LR^g(I) = p'(I)/p(I)$ for an investment success and $LR^b(I) = -p'(I)/\left(1 - p(I)\right)$ for an investment failure. As is well known $|LR^i(I)|$ measures the strength of incentive provision per unit of expected reward or punishment and, in that sense, how *informative* a performance signal - here, an investment outcome - is for providing incentives (see, e.g., Holmstrom 1979). Intuitively, according to (18), reward for investment success and punishment for failure are, hence, optimally chosen such as to equalize the respective expected cost of providing incentives (arising from the principal's "risk aversion") per unit of incentive provision.

To build further intuition, consider, first, the benchmark case with contractible investment (CI, a complete characterization can be found in Appendix B.1). Then, as the contract does not have to provide incentives for investment, optimal risk-sharing can be attained by setting $\beta^g_{CI}(w) = \beta^b_{CI}(w) = 0$. I.e., the principal, who is effectively risk-averse, bears all risk from the investment shock and perfectly insures the risk-neutral agent. This seemingly counterintuitive result follows from the fact that the principal is risk-averse with respect to variation not in his *own*, but in the *agent's* income, such that optimal risk sharing (corresponding to $\eta(w) = 0$ in (18)) prevails whenever the agent's compensation is not contingent on investment outcomes that occur with positive probability.

[16] If, for $i = b$, (18) has no solution $\beta^b(w) \geq -w$, the limited liability constraint (9) binds such that $\beta^b(w) = -w$ and the respective value of $\beta^g(w)$ follows directly from incentive compatibility (8). For a given level of investment, I, a sufficient condition for the limited liability constraint to be slack is that w is sufficiently large: $w \geq \lambda / \left(\nu p'(I)\right)$.

Instead, with non-contractible investment, the constrained optimal solution, subject to incentive compatibility, will usually have to deviate from optimal risk-sharing. According to (18), this is optimally done by shifting most of the required incentive costs to the more informative investment outcome, i.e., by setting[17]

$$\frac{f'(w+\beta^b(w))-f'(w)}{f'(w)-f'(w+\beta^g(w))}=\left|\frac{LR^b(I(w))}{LR^g(I(w))}\right|. \tag{19}$$

From $\left|LR^b(I)/LR^g(I)\right| = p(I)/(1-p(I))$ incentive costs are, thus, mainly incurred following the investment outcome that is less likely to occur.[18] Now, as the probability of successful investment increases in the implemented investment level, we obtain the following result comparing the composition of incentive costs for different sets of parameters $\psi \in \{\mu^h, \mu^l, \nu, \lambda, r, \gamma, \sigma, k\}$ which map into different values of $I(w)$.[19]

Lemma 4. *Fix $w \in (0, \overline{w})$ and consider two sets of parameter values ψ' and ψ'', such that $I(w)|_{\psi'} < I(w)|_{\psi''}$ under the optimal contract of Proposition 1. Then, as long as the limited liability constraint does not bind, rewards for investment success, $\beta^g(w)$, and punishments for investment failure, $\beta^b(w)$, are chosen such that the ratio of the costs of providing incentives through punishment relative to the costs of providing incentives through reward is strictly increasing in $I(w)|_\psi$, i.e.,*

$$\left.\frac{f'(w+\beta^b(w))-f'(w)}{f'(w)-f'(w+\beta^g(w))}\right|_{\psi'} < \left.\frac{f'(w+\beta^b(w))-f'(w)}{f'(w)-f'(w+\beta^g(w))}\right|_{\psi''}.$$

Further, $\beta^g(w) \to 0$ as $I(w)|_\psi \to \overline{I}$, while for $I(w)|_\psi \to 0$, as the agency problem vanishes, $-\beta^b(w)/\beta^g(w) \to 0$, i.e., the first unit of investment is incentivized with rewards only. For all $I(w)|_\psi \notin \{0, \overline{I}\}$, both reward and punishment are used, i.e., $\beta^g(w)|_\psi > 0 > \beta^b(w)|_\psi$.

This dependence of the relative costs of providing incentives through punishment rather than through reward on the implemented level of investment will be key in understanding

[17]In that sense, optimal incentive pay conditions more heavily on the more informative investment outcome. Note, however, that (19) does not necessarily imply that $\left|\beta^i(w)\right| > \left|\beta^j(w)\right|$ whenever $\left|LR^i(I(w))\right| > \left|LR^j(I(w))\right|$ as the curvature of f changes with w.

[18]Note that this simple description of the relative likelihood ratio is due to the fact that, with binary investment outcomes, the incentive effect of a payment conditioning on either investment outcome is given by $|p'(I)|$. In Appendix B.2 we consider a more general setup with an arbitrary number of investment outcomes for which (18) still holds and show how our main insights extend.

[19]As will be shown below, optimal investment $I(w)$ varies also with w such that the following comparative statics similarly hold fixing ψ and considering different levels of w.

the distortions in investment which we will analyze next.

Investment Distortions. In the first-best benchmark without agency frictions, a profit-maximizing risk-neutral owner-manager with discount rate r would choose the following investment level for all $t \geq 0$:[20]

$$I_{FB} \in \arg\max_{I \in [0,\overline{I}]} \{\nu p(I)\Delta - I\}. \tag{20}$$

This first-best investment schedule is constant and satisfies, if interior, the first-order condition $\nu p'(I)\Delta = 1$, which trades off the potential for higher profitability against the marginal investment expenditures.

In order to determine whether, with non-contractible investment expenditures, the firm invests too much or too little relative to this (first-best) benchmark, it is crucial to understand how the agency costs of investment in (17) change in I, i.e., to sign the marginal agency costs of investment $\partial\Phi\big(w, I(w)\big)/\partial I$. If these are positive, investment is distorted downwards relative to first-best, while it is distorted upwards if marginal agency costs are negative. We will now show that agency costs of delegated investment are non-monotonic in the incentivized investment level I such that indeed both under- as well as overinvestment can arise in equilibrium.

To see this, it is instructive to look, first, at the boundary cases where the problem in (20) has a corner solution. Clearly, when the investment technology is very unprofitable, such that, in the extreme case, $I_{FB} = 0$, the principal can trivially implement the first-best investment level without incurring any agency costs as there is no need to provide incentives and optimal risk-sharing can be attained. However, also when the investment technology is highly profitable, such that $I_{FB} = \overline{I}$, first-best investment can be incentivized without inducing (additional) volatility in w, i.e., maintaining optimal risk-sharing. Now, the shadow value on the incentive constraint in (18) is zero as an investment failure becomes extremely informative, in the sense of the respective likelihood ratio going to infinity. Intuitively, as the probability of success then is $p(\overline{I}) = 1$, from Lemma 4, all incentives are

[20] As we show in Appendix B.1, the same investment level would be implemented in the case where only the cash flow diversion problem is present and investment is contractible. This is due to the fact that both the returns to investment, captured by Δ, as well as its costs, I, are independent of the agency problem. Notably, this is different in the neoclassical investment model considered by DeMarzo et al. (2012), where the returns to investment (Tobin's Q) are reduced by a cash flow diversion problem. Thus, even though investment is contractible in their model, it is always distorted below first-best.

optimally provided "off-equilibrium" by relying on punishment only, as long as this does not violate limited liability. Since agency costs, as given in (17), are zero for $\beta^g(w) = 0$ and $p(\overline{I}) = 1$, optimal risk-sharing prevails and optimal investment again equals the first-best level.

Lemma 5. *Fix $w \in [0, \overline{w}]$. Then, as long as the limited liability constraint does not bind, investment under the optimal contract of Proposition 1 is equal to first-best whenever first-best investment attains a corner solution, i.e., $I(w) = I_{FB}$ for $I_{FB} \in \{0, \overline{I}\}$.*

The discussion above makes clear that first-best investment can be implemented without incurring any agency costs if $I_{FB} = 0$, or $I_{FB} = \overline{I}$. However, this is no longer true for interior values off I_{FB} which, from Lemma 4, are optimally implemented using both reward, $\beta^g > 0$, as well as punishment, $\beta^b < 0$, thus, implying a deviation from optimal risk-sharing. Since agency costs to incentivize first-best investment are, hence, strictly positive if I_{FB} is interior, and zero when I_{FB} attains a corner solution, it is easy to see that *marginal* agency costs are positive if first-best investment is sufficiently low and they are negative if first-best investment is sufficiently high. As a result, investment under the optimal contract will be distorted downwards in the former and upwards in the latter case.

Proposition 2. *Fix $w \in (0, \overline{w})$. Then, as long as the limited liability constraint does not bind, investment is distorted downwards, $I(w) < I_{FB}$, for sufficiently low values of $I_{FB} > 0$, while it is distorted upwards, $I(w) > I_{FB}$, for sufficiently high values of $I_{FB} < \overline{I}$.*

To build intuition, note that increasing investment has two basic effects on agency costs: First, higher investment requires stronger incentives (see (8)), and, second, it affects the outcome distribution, thus, increasing the probability of having to reward the agent for a favorable investment outcome and decreasing the probability of having to punish him for an investment failure. While the first effect unambiguously tends to increase agency costs pushing towards underinvestment, the second effect can have either sign depending on the relative costs of providing incentives through reward rather than punishment. As Proposition 2 shows, overinvestment may, hence, arise if incentives are predominantly given through punishment for bad outcomes, which from Lemma 4 is the case if the investment technology is sufficiently profitable such that an (unlikely) investment failure constitutes a very informative performance signal as measured by a high relative likelihood ratio. Then, in order to reduce the probability of having to bear the high agency costs associated with an investment failure, it is optimal to increase $I(w)$ above the first-best value.[21] The opposite case applies if the investment technology is rather unprofitable implying low first-best investment, which would be optimally incentivized with high rewards following an informative investment success. Then, to avoid the associated high incentive costs, investment $I(w)$ is optimally distorted downwards relative to first-best.

Compensation and Investment Dynamics. So far, we studied how investment distortions and the incentive scheme used to implement investment depend on the profitability of the investment technology as captured by I_{FB}. For this analysis we kept the agent's continuation value fixed. In the following we will now analyze the resulting dynamics of investment and compensation when w evolves as described in Proposition 1. In order to do so, it is useful to look, first, at extreme values of the state space, i.e., $w = \overline{w}$ and $w = 0$.

As the agent's continuation value reaches the compensation boundary ($w = \overline{w}$), he has accumulated so much wealth inside the firm that the agency problem is relaxed sufficiently for firm owners' effective risk-aversion to disappear. In fact, rewards are taken out as cash payments and do not induce costly variation in w (see Proposition 1). Hence, it is optimal not to punish the agent and to rely exclusively on rewards to incentivize investment. As a result, agency costs of investment are zero and investment equals first-best.[22]

For $w < \overline{w}$ the principal's value function is strictly concave such that it is optimal to use both rewards and punishments according to the optimal compensation policy in (19). However, for low values of the agent's continuation payoff, punishment is eventually restricted by the binding limited liability constraint (9). In fact, in the limiting case, as we approach the termination boundary ($w = 0$), only rewards can be used to incentivize any $I > 0$, as the agent is "too poor to be punished." Hence, incentives have to be provided mainly through rewards such that an investment success is the investment outcome that is particularly costly in terms of incentive provision. As a result, investment will be optimally distorted downwards for all $I_{FB} \in (0, \overline{I})$. These observations are summarized in the following Lemma.

[21] Note that this intuition is robust also beyond our simple binary state setting. In Appendix B.2 we allow for an arbitrary number of investment outcomes and derive, analogous to Proposition 2, conditions for when over- and underinvestment occurs. In particular, we show that, under standard assumptions, there is overinvestment whenever low investment outcomes (akin to an investment failure in the binary model) are sufficiently informative for providing incentives, as reflected in a high relative likelihood ratio.

[22] To see this formally, substitute $\overline{w}$ in (16) and note that $f(w)$ extends linearly to the right of $\overline{w}$.

Lemma 6. *Fix $I_{FB} \in (0, \overline{I})$. Then under the optimal contract of Proposition 1*

i) For $w \geq 0$ sufficiently small, the limited liability constraint for $\beta^b(w)$ binds such that the agent is instantly fired following a failure. Investment is distorted downwards, $I(w) < I_{FB}$.

ii) For $w < \overline{w}$ sufficiently large, incentives are provided through punishment and reward, which are optimally chosen according to (19). Investment may be distorted upwards or downwards as shown in Proposition 2.

iii) For $w \geq \overline{w}$, incentives are provided through rewards only and investment is equal to first-best, $I(w) = I_{FB}$.

Figure 2 illustrates the compensation and investment dynamics implied by Lemma 6 in an equilibrium with overinvestment (high I_{FB} in left panels), and in one with underinvestment (low I_{FB} in right panels). The two scenarios are generated by varying the "returns to investment" $\Delta\mu := \mu^h - \mu^l$, while all other parameters and functional forms are kept constant. Note that $\Delta\mu$ is a purely technological parameter in that it does not affect the agency problem for a given investment schedule.[23]

As shown in Lemma 6, there is no punishment ($\beta^b(w) = 0$) if the agent's track record is very poor ($w \to 0$) or if it is very good ($w \to \overline{w}$), while $\beta^b(w)$ is strictly negative in between. The upper panels of Figure 2 illustrate cases where the punishment policy $\beta^b(w) \leq 0$ is in fact U-shaped in the agent's continuation value w. The lower the degree of punishment, keeping all else constant, the more generous rewards are needed to satisfy incentive compatibility. Notably, this is true independently of whether the lack of punishment is imposed by the binding limited liability constraint (for low w), or if it is in fact optimal not to punish (for high w). As a result, also the reward policy $\beta^g(w) \geq 0$ in our numerical example is U-shaped in w.

Turning to the optimal investment policy $I(w)$ as depicted in the lower panels of Figure 2, note first that investment is distorted downwards relative to first-best for low returns to investment $\Delta\mu$ (right panel) and upwards for high $\Delta\mu$ (left panel), as long as w is sufficiently large, which corresponds to a slack limited liability constraint. In both cases investment $I(w)$ approaches the respective first-best value at the compensation boundary,

[23] I.e., if one had to implement in both cases the same (exogenously given) investment schedule, then the optimal contract would be identical, despite the difference in $\Delta\mu$.

i.e., for $w \to \overline{w}$. Hence, in the case with high returns to investment (left panel), investment is distorted upwards and decreases with the agent's stake in the firm w, while, for low returns to investment (right panel), it is distorted downwards and increases with w.

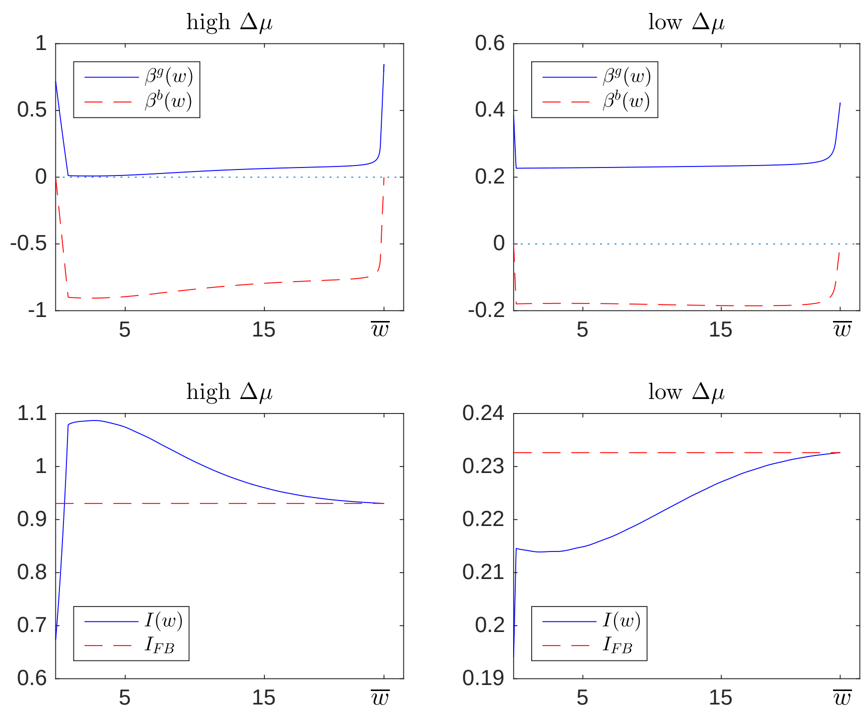

Figure 2: The upper panels illustrate the punishment and reward dynamics and the lower panels the investment dynamics under the optimal contract for low $\Delta\mu = 1.1$ and for high $\Delta\mu = 2.2$. We stipulate $p(I) = \phi\sqrt{I}$ and the parameter values used for calibration are: $\nu = 1.2$, $\sigma = 10$, $\lambda = 0.5$, $k = 15$, $\gamma = 0.15$, $r = 0.1$, $\phi = 0.95$.

4 Empirical Implications

In this section, we derive empirical implications regarding the relation between investment expenditures and available internal funds, how this relation is affected by financial frictions, the impact of corporate governance on investment, as well as the relation of investment to the structure of executive compensation. In order to derive these testable implications from our dynamic agency model, we have to map key model variables and parameters into observables. To do so, we will now, first, discuss how to measure the intangible investment expenditures we seek to model and how to interpret our key agency friction of non-contractibility of investment within the context of these applications. Then, second, we will show how to relate key features of the optimal contract to observables within an implementation.

Recall that by focussing on non-contractible investment in future profitability per unit of (physical) capital instead of standard capital investment, our model seeks to explain aspects of intangible investment such as investment in knowledge or organizational capital (including human capital, brands, customer relationships and distribution system). For the purpose of empirical testing, these types of intangible investment could be measured, e.g., by R&D spending, or by parts of selling, general and administrative (SG&A) spending respectively (see, e.g., Peters and Taylor 2017).[24] Within the context of our model, these *reported* (accounting) measures, correspond to non-contractible *actual* (or true) investment expenditures under the (optimal) incentive compatible contract from Section 3.

As mentioned previously, absent the appropriate incentives, actual and reported investment expenditures need not coincide perfectly, which can be motivated in a variety of ways. Within our main application to R&D note that, while R&D investments have to be expensed and reported when material (FASB 1974), accounting rules such as U.S. Generally Accepted Accounting Principles (GAAP) do leave considerable discretion to managers regarding which costs to classify as R&D.[25] Exploiting this discretion – e.g., to artificially improve short-term "core" earnings at the cost of long-term profitability – is referred to as "classification shifting" in the accounting literature.[26] Hence, actual investment expenditures are not fully verifiable, and have to be incentivized (see, e.g., Balkin et al. 2000 for evidence). This similarly seems to be relevant for investment in organizational capital for which only rather coarse measures, e.g., in the form of SG&A, are available (cf., Peters and Taylor 2017). However, even if actual investment expenditures can be verified by auditors or are directly observable, our simple agency model could be interpreted as capturing, in reduced form, settings in which the productivity of a given amount of intangible investment depends on complementary but *hidden* managerial effort. Then, the need to provide

[24]In the following, we do not consider intangible investment expenditures, for which no (accounting) measures are available, e.g., because they have to be kept secret from competitors, as in the context of product development or investment in product quality. Note, however, that our subsequent implications could still be tested by using observable investment successes or failures, the probability of which is monotonically related to investment, as a proxy for unobservable investment expenditures.

[25]The Financial Accounting Standards Board (1974) acknowledges that "the differences [in research and development] among enterprises and industries are so great that a detailed prescription of the activities and related costs includable in research and development, either for all companies or on an industry-by industry basis, is not a realistic undertaking for the FASB." See,also Horwitz and Kolodny (1980).

[26]Cf., relatedly, the notion of so-called "perk consumption" which also is a common interpretation of standard cash flow diversion (see, e.g., DeMarzo and Sannikov 2006).

incentives for managerial effort leads to qualitatively very similar implications as in our baseline model.[27]

To allow for empirical testing, in the following we will cast the predictions of our dynamic agency model within an implementation of the optimal contract. While it is well known that such an implementation is not unique (see, e.g., DeMarzo et al. 2012 for a discussion), the robust feature is that the key state variable, the agent's continuation payoff, is interpreted as a measure of the firm's *financial slack*. To see this relation, recall that when the agent's continuation payoff falls to zero, this triggers a restructuring of the firm, which is costly for its owners and involves the replacement of the incumbent management. As, for incentive reasons, the agent's continuation payoff w_t moves with cash flows with sensitivity λ, the maximum loss that a firm can sustain without triggering a restructuring, i.e., its financial slack, is equal to $M_t := w_t/\lambda$. While this insight can be formalized in a variety of ways, for illustrative purposes, consider one particular implementation of the optimal contract based on cash reserves as a measure of financial slack (see Appendix B.3 for details). If cash reserves fall to zero, which could be interpreted as insolvency, the firm needs to raise cash from the capital market, which involves a fixed cost of k. These cost of raising external funds, thus, capture the key financing friction in this implementation of our optimal contract.

In the following, we will express the firm's investment as a function of cash holdings,

$$INV_t := I(\lambda M_t) = I(w_t),$$

and, analogously, define the "investment-cash sensitivity" by

$$ICS_t := \frac{\partial \ln(INV_t)}{\partial \ln(M_t)},$$

which is the elasticity of investment to cash holdings.[28] Since the firm in our model

[27] To make this the most apparent, one could consider a variant of our model in which the probability for success $p(\cdot)$ and for failure $1 - p(\cdot)$ depend on the product of (contractible) investment I_t and effort $e_t \in [0, 1]$. When the manager deviates from the recommended effort level, i.e., $e_t < e_t^* \equiv 1$, he enjoys shirking benefits of $\lambda(1-e_t)I_t$. It is then straightforward to show that the necessary and sufficient condition under which $e_t^* \equiv 1$ is incentive compatible for a given (contractible) investment level $\tilde{I}$, is exactly the same condition under which a recommended investment level of $I_t^* = \tilde{I}$ is incentive compatible in our baseline model (see (8)). The two problems are, thus, isomorphic.

[28] ICS, thus, refers to the percentage change in investment induced by a percentage change in cash holdings, whereas the empirical literature usually considers investment sensitivities with respect to cash flows (i.e., the absolute change in cash holdings). While elasticities better capture the fact that under the

operates forever, i.e., within the implementation, it is always optimal to refinance when internal funds are depleted, cross-sectional ("ensemble") averages of INV and ICS can be obtained by computing the corresponding path-wise long-term averages,

$$\overline{INV} := \frac{1}{T}\sum_{t=0}^{T} INV_t, \tag{21}$$

$$\overline{ICS} := \frac{1}{T}\sum_{t=0}^{T} ICS_t, \tag{22}$$

with T sufficiently large.[29] Based on the analysis in Section 3, we consider the cross-sectional averages for subsets of firms (e.g., industries), which differ in the returns to investment as measured by the increase in expected cash flows due to successful investment, $\Delta\mu = \mu^h - \mu^l$. Empirically, one would expect firms for instance in the pharma or the high-tech sector to have particularly high returns to intangible investment (cf., Peters and Taylor 2017). We vary $\Delta\mu$ between 0 and an upper limit that guarantees an interior solution for the optimal investment level solving (16).

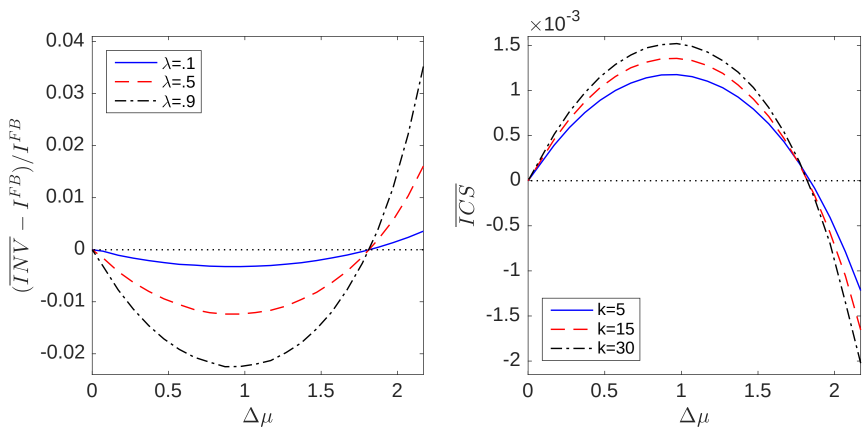

Figure 3: Plots average investment distortions (left panel), and average elasticity of investment with respect to cash (right panel) as a function of the industry returns to investment $\Delta\mu$. We stipulate $p(I) = \phi\sqrt{I}$ and the parameter values used for calibration are: $\nu = 1.2$, $\sigma = 10$, $\lambda = 0.5$, $k = 15$, $\gamma = 0.15$, $r = 0.1$, $\phi = 0.95$.

optimal contract investment (and its derivative) depends on financial slack, i.e., *cash holdings,* and not only on *cash flows*, our qualitative results still hold true if cash flows would be considered instead.

[29]This ergodicity property is exploited also by Brunnermeier and Sannikov (2014) or Klimenko et al. (2016). We compute (21) and (22) by Monte-Carlo simulation with a total number of $T = 10^7$ periods with period length $h = 10^{-3}$. We further specify $p(I) = \phi\sqrt{I}$, with $\phi > 0$ such that $I \in [0, 1/\phi^2]$. For more details and an outline of the employed algorithm, we refer to Appendix B.4.

Investment Distortions. The left panel of Figure 3 shows how *average* industry investment changes with the returns to investment in the industry, $\Delta\mu$. In particular, the investment distortion relative to the "owner-manager" (or contractible investment) benchmark is negative for the average firm in a low returns to investment industry while it is positive when returns to investment in the respective industry are high.

Implication 1. *Compared to average investment in otherwise similar owner-manager firms, average delegated investment is*

i) distorted downwards in industries with low returns to investment, and

ii) distorted upwards in industries with high returns to investment.

Notably, in our model these investment distortions arise endogenously via the optimal contractual response to a multitask problem, without assuming any direct costs or benefits of investment on the manager's side (such as, e.g., empire building preferences or pressure from the stock-market).

Compensation and Corporate Governance. A common presumption in much of the literature on managerial short-termism is that short-term oriented compensation biases managers towards short-term performance, leading them to neglect (hard to verify) long-term investment (see, e.g., Bebchuk and Fried 2004 or Edmans et al. 2017). The left panel of Figure 3 illustrates that, under a long-term contract that optimally balances incentives over different horizons, the relation between short-term compensation and long-term investment can also be positive. In particular, as the agent's diversion benefit λ increases, more compensation for short-term performance is needed to induce truth-telling according to incentive constraint (7). In our multitask setting, then also the exposure to the investment outcome needs to increase to incentivize a given level of investment according to incentive constraint (8). As a consequence, investment distortions go up, which implies more severe underinvestment (overinvestment) in industries with low (high) returns to investment, $\Delta\mu$.

Implication 2. *The relation between compensation for short-term performance and long-term investment is*

i) negative in industries with low returns to investment, and *ii) positive in industries with high returns to investment.*

Given that better corporate governance reduces the manager's potential to divert funds for private consumption (lower λ), Implication 2 can be interpreted also as a statement on the relation between corporate governance and long-term investment.[30] Our multitask model therefore describes a novel channel relating the impact of corporate governance on investment to industry characteristics, which does not assume any direct costs or benefits of investment on the manager's side.

Investment-Cash Sensitivities. The right panel of Figure 3 plots the average investment-cash sensitivity $\overline{ICS}$ implied by the optimal contract as a function of the returns to investment $\Delta\mu$. Since the agency problem is mitigated when the manager's "stake" in the firm grows, investment distortions relative to the constant first-best level decrease with positive firm performance and, thus, the level of financial slack (cf., the discussion following Lemma 6). As a consequence, the sign of the average investment-cash sensitivity $\overline{ICS}$ is determined by the sign of the average investment distortion as reported in Implication 1.

Implication 3. *The relation between investment and internal funds (cash) is*

i) positive in industries with low returns to investment, and

ii) negative in industries with high returns to investment.

There is a large empirical literature (starting with Fazzari et al. 1988) analyzing the relationship between firms' investment and cash flows. While initially, the literature has considered mainly capital investment, recent papers have explicitly considered also intangible investment, such as spending on R&D (see, e.g., Chen and Chen 2012 or Peters and Taylor 2017), documenting a quantitatively small or insignificant relationship in the cross section. In light of our theory, the sign of the relationship should, however, depend on the returns to these intangible investments, which is expected to differ considerably across firms and industries. This provides a possible explanation for why, looking at subsamples, Peters and Taylor (2017) do find that investment-cash flow sensitivities are negative for

[30] In line with this implication, Brav et al. (2017) find a negative relation between hedge fund activism and R&D expenditures in their sample of high-tech firms, while Aghion et al. (2013) report a strong positive relation between institutional ownership and R&D. The role of incentive pay as a substitute for corporate governance is documented, e.g., in Core et al. (1999) or, more recently, Fahlenbrach (2009).

industries with a high intensity of intangible investment, such as the high-tech and health sector, while it is positive in the consumer and manufacturing sector with low intangible investment.[31]

Financial Frictions and ICS. Much of the empirical investment-cash flow sensitivity literature (staring with Fazzari et al. 1988) has evolved around the question of whether cash flows are a good proxy for financial constraints. That investment in our model changes with financial slack and, thus, also with cash flows, is a consequence of *endogenous* financial constraints arising from the agency problem. In particular, a firm's financial slack reduces the probability with which the firm has to raise costly external funds and, thus, relaxes financial constraints for given costs of external financing k. Varying k then allows us to conduct comparative statics with respect to the severity of financial constraints for any given level of cash holdings. We find that, if the costs of external financing increase, investment becomes more sensitive to cash flows. As illustrated in the right panel of Figure 3, investment-cash sensitivities increase *in absolute terms* with financing frictions, but the sign and strength of this relation depends on the industry returns to investment. Our model, thus, provides a novel perspective on how *intangible* investment-cash flow sensitivities can be used as a measure for financial constraints.

Implication 4. *The relation between the costs of external financing and investment-cash sensitivities is*

i) positive in industries with low returns to investment, and

ii) negative in industries with high returns to investment.

Given the growing importance of intangible capital, Implication 4 may also provide a first step in explaining the puzzling decline of total investment-cash flow sensitivities in the cross-sectional average over the last decades (see, e.g., Brown and Petersen 2009 and Chen and Chen 2012). However, in light of our model, these result should not be interpreted as evidence against the validity of using investment-cash flow sensitivities as a measure of financial constraints.

[31] Similarly, Hovakimian (2009) finds that firms with negative investment-cash flow sensitivities have higher R&D expenditures, higher market-to-book ratios, and lower asset tangibility than firms with positive investment-cash flow sensitivities.

5 Conclusion

In this paper, we analyze the dynamics of corporate investment under endogenous financial constraints in a dynamic agency model with multitasking. The manager privately observes the firm's cash flows which he can divert for private consumption and, in addition, use to finance non-contractible investment in the firm's future profitability. Thus, when reported cash flows are low (or even negative), the principal does not know whether this is because of stealing or because of investment. To ensure truth-telling and investment in the interest of firm owners, the optimal compensation scheme ties the manager's compensation both to reported cash flows as well as to imperfect performance signals indicating investment success or failure. Exposing the manager to compensation risk is, however, costly such that incentivizing investment causes additional agency costs.

The optimal investment profile trades off these agency costs of investment with the potential efficiency gains. As a result, investment will be history dependent and distorted away from its first-best level. Investment distortions decrease (in absolute terms) with good performance, as the manager's stake in the firm increases mitigating the agency problem. Our model predicts overinvestment and negative investment-cash flow sensitivities in industries where returns to non-contractible intangible investment such as R&D are high (e.g., as being in a position of technological leadership is important like in the pharmaceutical industry). By contrast, in industries where this is not the case, investment will be distorted downwards and increasing in cash flows.

If raising external funds is more expensive, investment will be more severely distorted and more sensitive to past performance; i.e., in firms with high returns to investment, investment increases and the relation between financial slack and investment becomes more negative, while investment decreases and its sensitivity to financial slack becomes more positive in firms with low returns to investment. Our multitask theory, thus, offers a new perspective on the interpretation of investment-cash flow sensitivities as a measure of financial constraints, taking into account the increased importance of alternative forms of investment such as process innovation, R&D, or intangible investment in general.

References

Aghion, P., J. Van Reenen, and L. Zingales, 2013, "Innovation and Institutional Ownership," *American Economic Review* 103:277–304.

Almeida, H., and M. Campello, 2007, "Financial Constraints, Asset Tangibility, and Corporate Investment," *Review of Financial Studies* 20:1429–1460.

Balkin, D., G. Markman, and L. Gomez-Mejia, 2000, "Is CEO Pay in High-Technology Firms Related to Innovation?" *The Academy of Management Journal* 43:1118–1129.

Bebchuk, L. A., and J. M. Fried, 2004, "Paying without Performance,"Harvard University Press.

Bebchuk, L. A., and L. A. Stole, 1993, "Do Short-Term Objectives Lead to Under- or Overinvestment in Long-Term Projects?" *Journal of Finance* 48:719–729.

Bernstein, J., and T. Mamuneas, 2006, "R&D Depreciation, Stocks, User Costs and Productivity Growth for US R&D Intensive Industries," *Structural Change and Economic Dynamics* 17:70–98.

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.

Biais, B., T. Mariotti, and J.-C. Rochet, 2013, "Dynamic Financial Contracting," in Advances in Economics and Econometrics, Tenth World Congress, D. Acemoglu, M. Arellano, E. Dekel, Cambridge University Press, vol. 1:125–171.

Brav, A., W. Jiang, S. Ma, and X. Tian, 2017, "How Does Hedge Fund Activism Reshape Corporate Innovation?" *Journal of Financial Economics* forthcoming.

Brown, J., and B. Petersen, 2009, "Why Has the Investment-Cash Flow Sensitivity Declined so Sharply? Rising R&D and Equity Market Developments," *Journal of Banking & Finance* 33:971–984.

Brunnermeier, M., and Y. Sannikov, 2014, "A Macroeconomic Model with a Financial Sector," *American Economic Review* 104:379–421.

Chen, H., and S. Chen, 2012, "Investment-Cash Flow Sensitivity Cannot Be a Good Measure of Financial Constraints: Evidence from Time series," *Journal of Financial Economics* 103:393–410.

Clementi, G., and H. Hopenhayn, 2006, "A Theory of Financing Constraints and Firm Dynamics", *Quarterly Journal of Economics* 121:229–265.

Cohen, W., and D. Levinthal, 1990, "Absorptive Capacity: A New Perspective on Learning and Innovation," *Administrative Science Quarterly* 35:128–152.

Core, J., R. Holthausen, and D. Larcker, 1999, "Corporate Governance, Chief Executive Officer Compensation, and Firm Performance," *Journal of Financial Economics* 51:371–406.

Darrough, M., J. Suh, and Y. Shen, 2017, "Product-Market Competition and Firms Disclosure Strategy of R&D Expenditures," Working Paper Zicklin School of Business.

DeMarzo, P., and D. Duffie, 1999, "A Liquidity-Based Model of Security Design," *Econometrica* 67:65–99.

DeMarzo, P., and M. Fishman, 2007, "Agency and Optimal Investment Dynamics," *Review of Financial Studies* 20:151–188.

DeMarzo, P., M. Fishman, Z. He, and N. Wang, 2012, "Dynamic Agency and the q Theory of Investment," *Journal of Finance* 67:2295–2340.

DeMarzo, P., and Y. Sannikov, 2006, "Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model," *Journal of Finance* 61:2681–2724.

Dutta, S., and S. Reichelstein, 2003, "Leading Indicator Variables, Performance Measurement, and Long-Term Versus Short-Term Contracts," *Journal of Accounting Research* 41:837–866.

Edmans, A., V. Fang, and A. Huang, 2017, "The Long-Term Consequences of Short-Term Incentives," CEPR Discussion Paper 12305.

Fahlenbrach, R., 2009, "Shareholder Rights, Boards, and CEO Compensation," *Review of Finance* 13:81–113.

Fazzari, S., G. Hubbard, and B. Petersen, 1988, "Financing Constraint and Corporate Investment," *Brookings Papers on Economic Activity* 141–195.

Gertler, M., 1992, “Financial Capacity and Output Fluctuations in an Economy with Multi-Period Financial Relationships,” *Review of Economic Studies* 59:455–472.

Harris, M., and A. Raviv, 1996, “The Capital Budgeting Process: Incentives and Information,” *The Journal of Finance* 51:1139–74.

Hoffmann, F., and S. Pfeil, 2010, “Reward for Luck in a Dynamic Agency Model,” *Review of Financial Studies* 23:3329–3345.

Holmstrom, B., 1979, “Moral Hazard and Observability,” *Bell Journal of Economics* 10:74–91.

Holmstrom, B., and P. Milgrom, 1991, “Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design,” *Journal of Law, Economics & Organization* 7:24–52.

Horwitz, B., and R. Kolodny, 1980, “The Economic Effects of Involuntary Uniformity in the Financial Reporting of R&D Expenditures.” *Journal of Accounting Research* 18:38–74.

Hovakimian, G., 2009, “Determinants of Investment Cash Flow Sensitivity,” *Financial Management* 38:161–183.

Jacod, J., and A. Shiryaev, 2003, “Limit Theorems for Stochastic Processes,” 2nd edn, Springer, Berlin.

Kaplan, S. N., and L. Zingales, 1997, “Do Financing Constraints Explain why Investment is Correlated with Cash Flow?” *Quarterly Journal of Economics* 112:169–215.

Klimenko, N., Pfeil, S., Rochet, J.-C., and G. DeNicolò, 2016, “Aggregate Bank Capital and Credit Dynamics,” Swiss Finance Institute Research Paper No. 16-42.

Koh, P. S., and D. M. Reeb, 2015, “Missing R&D,” *Journal of Accounting and Economics* 60:73–94.

Malenko, A., 2018, “Optimal Dynamic Capital Budgeting,” Working Paper MIT Sloan School of Management.

McVay, S. E., 2006, “Earnings Management Using Classification Shifting: An Examination of Core Earnings and Special Items,” *The Accounting Review* 81:501–531.

Narayanan, M., 1985, "Managerial Incentives for Short-Term Results," *Journal of Finance* 40:1469–1484.

Peters, R., and L. Taylor, 2017, "Intangible Capital and the Investment-q Relation," *Journal of Financial Economics* 123:251–272.

Quadrini, V., 2004, "Investment and Liquidation in Renegotiation-Proof Contracts with Moral Hazard," *Journal of Monetary Economics* 51:713–751.

Seru, A., 2014, "Firm Boundaries Matter: Evidence from Conglomerates and R&D Activity," *Journal of Financial Economics* 111:381–405.

Shleifer, A., and R. Vishny, 1990, "Equilibrium Short Horizons of Investors and Firms," *American Economic Review Papers and Proceedings* 80:148–153.

Skaife, H. A., L. A. Swenson, and D. D. Wangerin, 2013, "Classification Shifting of R&D Expenses," Working Paper University of Wisconsin-Madison.

Stein, J., 1988, "Takeover Threats and Managerial Myopia," *Journal of Political Economy* 96:61–80.

Stein, J., 1989, "Efficient Capital Markets, Inefficient Firms: A Model of Myopic Corporate Behavior," *Quarterly Journal of Economics* 104:655–669.

Varas, F., 2017, "Managerial Short-Termism, Turnover Policy, and the Dynamics of Incentives," *Review of Financial Studies* forthcoming.

von Thadden, E., 1995, "Long-Term Contracts, Short-Term Investment and Monitoring," *Review of Economic Studies* 62:557–575.

Zhu, J. Y., 2018, "Myopic Agency," *Review of Economic Studies* 85:1352–1388.

Zwiebel, J., 1995, "Corporate Conservatism and Relative Compensation," *Journal of Political Economy* 103:1–25.

Appendix A Proofs

Proof of Lemma 1. For $S = S^*$, we define the manager's expected lifetime utility evaluated conditional on time t information by:

$$v_t = \int_0^t e^{-\gamma s} dU_s + e^{-\gamma t} w_t, \tag{A.1}$$

which, by construction, is a martingale with respect to the filtration generated by (Z, N^j), $j \in \{g, b\}$ under the probability measure Q^{S^*}. By the martingale representation theorem (cf., Theorem III 4.34 in Jacod and Shiryaev 2003), we can express v_t for some predictable processes α and β^j, $j \in \{g, b\}$ as

$$\begin{aligned} v_t &= v_0 + \int_0^t e^{-\gamma s} \alpha_s \left(d\hat{Y}_s - (\mu_s - I_s^*)\, ds \right) + \int_0^t e^{-\gamma s} \beta_s^g \left(dN_s^g - \nu p(I_s^*) ds \right) \\ &\quad + \int_0^t e^{-\gamma s} \beta_s^b \left(dN_s^b - \nu \left(1 - p(I_s^*)\right) ds \right), \end{aligned}$$

where, under Q^{S^*}, $\hat{Y}_t - \int_0^t (\mu_s - I_s^*)\, ds$ is a standard Brownian motion and $N_t^g - \int_0^t \nu p(I_s^*) ds$ as well as $N_t^b - \int_0^t \nu \left(1 - p(I_s^*)\right) ds$ are compensated Poisson processes. Differentiating both this representation and the definition of v_t in (A.1) yields (6). **Q.E.D.**

Proof of Lemma 2. Consider any feasible policy of the agent, $S = \left\{ \hat{Y}_t, I_t : 0 \le t \le \tau \right\}$, with $dY_t - d\hat{Y}_t \ge 0$ and $I_t \ge 0$. The associated expected lifetime utility is given by

$$\begin{aligned} v_t &= v_0 + \int_0^t e^{-\gamma s} \alpha_s \left(d\hat{Y}_s - (\mu_s - I_s^*)\, ds \right) + \int_0^t e^{-\gamma s} \lambda \left(dY_s - d\hat{Y}_s \right) \\ &\quad + \int_0^t e^{-\gamma s} \beta_s^g \left(dN_s^g - \nu p(I_s^*) ds \right) + \int_0^t e^{-\gamma s} \beta_s^b \left(dN_s^b - \nu \left(1 - p(I_s^*)\right) ds \right), \end{aligned} \tag{A.2}$$

where $d\hat{Y}_t - (\mu_t - I_t^*)\, dt = \sigma dZ_t$ for $S = S^*$, and dN_t^g (dN_t^b) is a Poisson process with intensity $\nu p(I_t)$ $(\nu (1 - p(I_t)))$. Differentiating (A.2) and taking expectations gives

$$e^{\gamma t} E\left[dv_t\right] = (\lambda - \alpha_t)\, E\left[dY_t - d\hat{Y}_t \right] + \alpha_t \left(I_t^* - I_t\right) dt + \nu \left(\beta_t^g - \beta_t^b \right) \left(p(I_t) - p(I_t^*) \right) dt.$$

Clearly, v_t is a martingale for $S = S^*$. For $S = S^*$ to be incentive compatible, the drift of v_t has to be non-positive for all possible deviations, i.e., v_t has to be a supermartingale for any feasible $S \neq S^*$. Consider, first, a deviation from truth-telling, i.e., underreporting for consumption $dY_t - d\hat{Y}_t > 0$. This deviation is suboptimal for the agent if $\alpha_t \ge \lambda$. Second, increasing I_t marginally above I_t^*, is suboptimal for the agent if $\alpha_t \ge \nu p'(I_t) \left(\beta_t^g - \beta_t^b \right)$, while, third, decreasing I_t marginally below I_t^*, is suboptimal if $\alpha_t \leq \nu p'(I_t)\left(\beta_t^g - \beta_t^b\right)$. So, incentive compatibility requires (7) and (8) to hold. **Q.E.D.**

Proof of Lemma 3. Consider the process κ counting the number of replacements and the associated time points $\tau(\kappa)$ where we normalize $\tau(0) = 0$. Then we can write the principal's value as

$$f^i(w) = \max_{\mathbf{I},\mathbf{U},\tau}\left\{E\left[\int_0^\infty e^{-rs}\left(dY_s - dU_s\right) - k\sum_{\kappa=1}^{\infty} e^{-r\tau(\kappa)}\middle| w, \mu^i\right]\right\},$$

where maximization is subject to incentive compatibility in (4), limited liability as well as promise-keeping $w_0 = w$. Now denote the (stochastic) time of the next technology shock by T and recall that technology shocks arrive with exogenous rate, such that T is clearly beyond both the agent's and the principal's control. Thus, substituting from (1), we have

$$f^i(w) = \frac{\mu^i}{r+\nu} + \max_{\mathbf{I},\mathbf{U},\tau}\left\{E\left[\int_T^\infty e^{-rs}\mu_s ds - \int_0^\infty e^{-rs}\left(dU_s + I_s ds\right) - k\sum_{\kappa=1}^{\infty} e^{-r\tau(\kappa)}\middle| w\right]\right\},$$

where the maximization problem determining the optimal contract is clearly independent of $\mu_0 = \mu^i$ for any w. **Q.E.D.**

Proof of Proposition 1. We first show that the function $f(w)$ is concave in w and then verify that the contract of Proposition 1 maximizes the principal's expected payoff. **Concavity.** Consider the function $f(w)$ that, for a given w and μ, solves

$$(r+\nu)\, f(w) = \begin{array}{c} \mu + \nu p(I\,(w))\Delta - I\,(w) + \frac{1}{2}(\alpha\,(w))^2\sigma^2 f''(w) \\ + \left[\gamma w - \nu\left(\beta^g\,(w)\,p(I\,(w)) + \beta^b\,(w)\,(1 - p(I\,(w)))\right)\right] f'(w) \\ + \nu p(I\,(w)) f\,(w + \beta^g\,(w)) + \nu\,(1 - p(I\,(w)))\, f(w + \beta^b\,(w)) \end{array} \tag{A.3}$$

for $w \in [0, \overline{w}]$ and $f\,(w) = f(\overline{w}) - (w - \overline{w})$ for $w > \overline{w}$, with boundary conditions $f\,(0) = \max_w \{f\,(w) - k\}$, $f'\,(\overline{w}) = -1$, and $f''\,(\overline{w}) = 0$.

Investment $I\,(w)$ satisfies the first order condition

$$1 = \left[\begin{array}{c} \nu p'(I\,(w))\left(\beta^g\,(w) - \beta^b\,(w)\right) f_w(w) \\ + \nu p'(I\,(w))\left[f(w + \beta^g\,(w)) + \Delta - f(w + \beta^b\,(w))\right] \\ + \alpha\,(w)\, p(I\,(w)) \frac{p''(I(w))}{p'(I(w))^2}\left[f'(w) - f'(w + \beta^g\,(w))\right] \end{array}\right], \tag{A.4}$$

and $\beta^i(w), i \in \{g, b\}$ is determined by

$$f'(w) = p(I(w))f'(w + \beta^g(w)) + (1 - p(I(w)))f'(w + \beta^b(w)) \tag{A.5}$$

for $\beta^i(w) \geq -w$ while $\beta^i(w) = -w$ otherwise.

Differentiating (A.3) and using (A.4) and (A.5) yields

$$-(\gamma - r) f'(w) = \left[\gamma w - \nu\left(\frac{\alpha(w) p(I(w))}{\nu p'(I(w))} + \beta^b(w)\right)\right] f''(w) + \frac{1}{2}\sigma^2 (\alpha(w))^2 f'''(w).$$

From the boundary conditions at $\overline{w}$ we get $f'''(\overline{w}) = 2\frac{\gamma - r}{\sigma^2 \alpha(\overline{w})^2} > 0$, such that $\exists \varepsilon > 0$ with $f'''(\overline{w} - \varepsilon) > 0$ and $f''(\overline{w} - \varepsilon) < 0$.

The proof then is by contradiction. So assume that $\exists \breve{w} := \sup\{w < \overline{w} : f''(w) \geq 0\}$, where it holds by continuity that $f''(\breve{w}) = 0$ and $f'''(\breve{w}) < 0$, implying that $f'(\breve{w}) = -f'''(\breve{w})\frac{1}{2}\frac{\sigma^2\alpha(\breve{w})^2}{\gamma - r} > 0$. Now, consider two points $w^1 < \breve{w} < w^2$ close to $\breve{w}$, such that $f''(w^1) > 0 > f''(w^2)$ and $w^1 f'(w^1) = w^2 f'(w^2)$ and observe that $f(w)$ can be written as

$$rf(w) = \gamma w f'(w) + \frac{1}{2}\sigma^2 (\alpha(w))^2 f''(w) + g(w),$$

with

$$g(w) = \mu - I(w) + \nu \left[\begin{array}{c} p(I(w))\left(f(w + \beta^b(w) + \frac{\alpha(w)}{\nu p'(I(w))}) + \Delta\right) \\ +(1 - p(I(w)))f(w + \beta^b(w)) - f(w) \\ -\left(\frac{\alpha(w)p(I(w))}{\nu p'(I(w))} + \beta^b(w)\right) f'(w) \end{array}\right].$$

Next, compute the differential of $g(w)$ around $\breve{w}$, $dg(w)|_{w=\breve{w}} = g'(\breve{w})dw$, and observe that

$$\begin{aligned} g'(\breve{w}) \quad = \quad & I_w(w, \mu) \left[\begin{array}{c} -1 + \nu p'(I(w))\left[f(\breve{w} + \beta^g(\breve{w})) + \Delta - f(\breve{w} + \beta^b(w))\right] \\ -\alpha(w) f'(\breve{w}) \\ +\frac{\alpha(w)p(I(w))p''(I(w))}{p'(I(w))^2}\left[f'(\breve{w}) - f'(\breve{w} + \beta^g(w))\right] \end{array}\right] \\ & +\nu\left(1 + \beta^b_w(w)\right)\left[\begin{array}{c} p(I(w))f'(\breve{w} + \beta^g(w)) \\ +(1 - p(I(w)))f'(\breve{w} + \beta^b(w)) - f'(\breve{w}) \end{array}\right] \\ & -\nu\left(\frac{\alpha(w)p(I(w))}{\nu p'(I(w))} + \beta^b(w)\right) f''(\breve{w}) \\ = 0, & \end{aligned}$$

which follows from $f''(\breve{w}) = 0$ together with (A.4) and either (A.5), or, in case the limited liability constraint binds, $\beta^b_w(w) = \partial\beta^b(w)/\partial w = -1$. Thus, evaluating $f(w)$ in w^1 and

w^2, we get

$$r\left[f(w^1)-f(w^2)\right]=\frac{1}{2}\sigma^2\alpha\left(w\right)^2\left[f''(w^1)-f''(w^2)\right]>0,$$

where we have used that the effect of a change in $g(w)$ around $\breve{w}$ is of second order and will thus be dominated by the change in $f''(w)$. However, this directly contradicts $f'(\breve{w})>0$.

Verification. For any incentive compatible contract $(\boldsymbol{S},\boldsymbol{U},\tau)$, define

$$G_t=\int_0^t e^{-rs}\left(dY_s-dU_s\right)+e^{-rt}f\left(w_{t^-},\mu_t\right)$$

and recall that w_{t^-} evolves according to

$$dw_{t^-}=\gamma w_{t^-}dt-dU_t+\alpha_t\sigma dZ_t+\beta_t^g\left(dN_t^g-\nu p(I_t)dt\right)+\beta_t^b\left(dN_t^b-\nu\left(1-p(I_t)\right)dt\right),$$

where $\alpha_t\geq\lambda$ and $\beta_t^g-\beta_t^b=\frac{\alpha_t}{\nu p'(I_t)}$. Recall that $f\left(w\right)=f\left(w,\mu^l\right)=f\left(w,\mu^h\right)-\Delta$ so that from differentiating G using Ito's lemma and the change in variables formula for point processes as well as (A.3), we get

$$\begin{aligned}
e^{rt}dG_t &= \nu\left(\begin{array}{c} p(I_t)\left[f\left(w_{t^-}+\beta_t^g\right)+\Delta\right]+\left(1-p(I_t)\right)f(w_{t^-}+\beta_t^b) \\ -\left[\beta_t^g p\left(I_t\right)+\beta_t^b\left(1-p\left(I_t\right)\right)\right]f'\left(w_{t^-}\right)-I_t\end{array}\right)dt \\
&-\nu\left(\begin{array}{c} p(I\left(w\right))\left[f\left(w+\beta^g\left(w\right)\right)+\Delta\right]+\left(1-p(I\left(w\right))\right)f(w+\beta^b\left(w\right)) \\ -\left[\beta^g\left(w\right)p(I\left(w\right))+\beta^b\left(w\right)\left(1-p(I\left(w\right))\right)\right]f'\left(w\right)-I\left(w\right)\end{array}\right)dt \\
&+\frac{1}{2}\left(\alpha_t^2-\lambda^2\right)\sigma^2 f''\left(w_{t^-}\right)dt-\left[1+f'\left(w_{t^-}\right)\right]dU_t+\left[\sigma+\alpha_t\sigma f'\left(w_{t^-}\right)\right]dZ_t \\
&+\left[f\left(w_{t^-}+\beta_t^g\right)+\Delta\right]dM_t^g+f\left(w_{t^-}+\beta_t^b\right)dM_t^b-f\left(w_{t^-}\right)\left(dM_t^g+dM_t^b\right)
\end{aligned} \tag{A.6}$$

for $w=w_{t^-}\in[0,\overline{w}]$, where M^i denote the compensated point processes associated with N^i, $i\in\{g,b\}$. Now observe that the sum of the first two lines is less or equal to zero, because $\beta^i\left(w\right),i\in\{g,b\}$ and $I\left(w\right)$ are the solution to

$$\max_{\beta^i\geq-w,I}\left[\begin{array}{c} p(I)\left[f\left(w+\beta^g\right)+\Delta\right]+\left(1-p(I)\right)f(w+\beta^b) \\ -\left(\beta^g p\left(I\right)+\beta^b\left(1-p\left(I\right)\right)\right)f'\left(w\right)-I\end{array}\right].$$

Now turn to the third line of (A.6). The first term is non positive as $f''\leq 0$ and $\alpha_t\geq\lambda$ for any $t\geq 0$. The second term is non positive as $f'\geq -1$ and $dU\geq 0$ while Z and M^i are martingales. Hence, G_t is a supermartingale and a martingale if and only if for $t>0$, $\beta_t^i=\beta^i\left(w\right)$, $I_t=I\left(w\right)$, $\alpha_t=\lambda$ and $dU_t>0$ only when $w>\overline{w}$.

Now consider the principal's expected payoff under any incentive compatible contract

$(\boldsymbol{S}, \boldsymbol{U}, \tau)$:

$$E\left[\int_0^\tau e^{-rs}\left(dY_s - dU_s\right) + e^{r\tau} L_\tau,\right]$$

where τ denotes the time when the incumbent manager is replaced and L_τ the principal's expected profits from restarting with a new agent, net of replacement costs k. We then have that

$$\begin{aligned}
& E\left[\int_0^\tau e^{-rt}\left(dY_s - dU_s\right) + e^{r\tau} L_\tau,\right] \\
= \; & E\left[G_{t\wedge\tau}\right] + E\left[\mathbf{1}_{t\leq\tau}\left(\int_t^\tau e^{-rs}\left(dY_s - dU_s\right) + e^{-r\tau} L_\tau - e^{-rt} f\left(w_{t^-}\right)\right)\right] \\
\leq \; & f\left(w_{0^-}\right) + E\left[\mathbf{1}_{t\leq\tau}\left(\int_t^\tau e^{-rs}\left(dY_s - dU_s\right) + e^{-r\tau} L_\tau - e^{-rt} f\left(w_{t^-}\right)\right)\right] \\
= \; & f\left(w_{0^-}\right) + e^{-rt} E\left[\mathbf{1}_{t\leq\tau}\left(E\left[\int_t^\tau e^{-r(s-t)}\left(dY_s - dU_s\right) + e^{-r(\tau-t)} L_\tau | \mathcal{F}_t\right] - f\left(w_{t^-}\right)\right)\right],
\end{aligned}$$

where the inequality follows since $G_{t\wedge\tau}$ is a supermartingale and $G_0 = f\left(w_{0^-}\right)$. Now note that

$$E\left[\int_t^\tau e^{-r(s-t)}\left(dY_s - dU_s\right) + e^{-r(\tau-t)} L_\tau | \mathcal{F}_t\right] < \frac{\mu^h}{r} - w_{t^-},$$

and, as $f' \geq -1$, we have that $f\left(w_{t^-}\right) + w_t \geq L_\tau$, which yields

$$E\left[\int_0^\tau e^{-rt}\left(dY_s - dU_s\right) + e^{r\tau} L_\tau,\right] \leq f\left(w_{0^-}\right) + e^{-rt} E\left[\mathbf{1}_{t\leq\tau}\left(\frac{\mu^h}{r} - L_\tau\right)\right].$$

Taking $t \to \infty$ yields

$$E\left[\int_0^\tau e^{-rt}\left(dY_s - dU_s\right) + e^{r\tau} L_\tau,\right] \leq f\left(w_{0^-}\right). \tag{A.7}$$

Finally, under the contract stated in Proposition 1, $G_{t\wedge\tau}$ is a martingale and, hence, (A.7) holds with equality. **Q.E.D.**

Proof of Lemma 4. Whenever the limited liability constraint (9) does not bind, for which it is sufficient that $w \geq \frac{\lambda}{\nu p'(\overline{I})}$, first order condition (19) holds. The first claim then follows immediately from rewriting (19) as

$$\frac{f'(w + \beta^b(w)) - f'(w)}{f'(w) - f'(w + \beta^g(w))} = \frac{p(I(w))}{1 - p(I(w))}, \tag{A.8}$$

and observing that $p(I(w))|_{\psi'} < p(I(w))|_{\psi''}$ since $I(w)|_{\psi'} < I(w)|_{\psi''}$ and $p'(\cdot) > 0$. Next, when $I(w)|_{\psi} \to \overline{I}$ and the limited liability constraint is slack, strict concavity of $f(w)$ implies that (19) can only be satisfied for $\beta^g(w) \to 0$. In case $p(I) \to 0$, (19) similarly requires that $\beta^b(w) \to 0$, while, from incentive compatibility, $\beta^g(w) \to \frac{\lambda}{\nu p'(0)}$. For bounded $p'(0)$, it then trivially follows that $-\beta^b(w)/\beta^g(w)$ goes to zero as $I(w) \to 0$. Now, if $p'(I) \to \infty$ as $I \to 0$, both $\beta^g(w)$ and $\beta^b(w)$ go to zero as $I \to 0$. Then, up to a first-order approximation, we have $f'(w+\beta^b(w)) = f'(w)+\beta^b(w)f''(w)$, and $f'(w+\beta^g(w)) = f'(w)+\beta^g(w)f''(w)$. Hence, there exists an $\varepsilon > 0$ such that, from (A.8), we have $\frac{p(\varepsilon)}{1-p(\varepsilon)} = \frac{-\beta^b(w)}{\beta^g(w)}$.

Finally, the last claim follows from strict concavity of $f(w)$ for $w < \overline{w}$, implying that (19) can only be satisfied for $\beta^g > 0 > \beta^b$. **Q.E.D.**

Proof of Lemma 5. See main text. **Q.E.D.**

Proof of Proposition 2. Denote the right-hand side of the HJB in (12) by $\mathcal{L}$ and take the first derivative with respect to I, noting that $\beta^g = \beta^b + \frac{\lambda}{\nu p'(I)}$ from incentive constraint (8), to obtain

$$\frac{\partial \mathcal{L}}{\partial I} = \nu p'(I\left(w\right))\Delta - 1 - \left[\left(\frac{\lambda}{\nu}\frac{-p''(I(w))}{\left[p'(I(w))\right]^2}\right)\delta^{gb}\left(w\right) + \nu p'\left(I\left(w\right)\right)\left(\delta^g\left(w\right) - \delta^b\left(w\right)\right)\right], \tag{A.9}$$

with

$$\delta^{gb}\left(w\right) := p(I(w))\left(1-p(I(w))\right)\nu\left[f'(w+\beta^b(w)) - f'(w+\beta^g\left(w\right))\right] \geq 0, \tag{A.10}$$

$$\delta^j\left(w\right) := f(w) + \beta^j\left(w\right)f'(w) - f(w+\beta^j\left(w\right)) \geq 0,\ j \in \{g,b\}, \tag{A.11}$$

where we have used (19), which holds whenever the limited liability constraint does not bind, and the inequalities in (A.10) and (A.11) follow from strict concavity of $f(w)$ for $w < \overline{w}$. The term in square brackets in (A.9) are the marginal agency costs of investment which we denote by $\phi(w)$.

To show the overinvestment result for sufficiently high levels of I_{FB} it is sufficient to show that the marginal agency costs of investment are strictly negative as I_{FB} approaches $\overline{I}$, such that $\partial\mathcal{L}/\partial I|_{I(w)=I_{FB}} > 0$ for $I_{FB} \nearrow \overline{I}$. The result then follows from continuity and the fact that $I(w) = I_{FB}$ is the uniquely optimal investment level for $I_{FB} = \overline{I}$. So, note that $p(\overline{I}) = 1$ implies that $\delta^{gb}(w) \to 0$ and, by Lemma 5, $\beta^g(w), \delta^g(w) \to 0$, such that $\phi(w)|_{I(w)=I_{FB}} \to \nu p'(I_{FB})\delta^b(w) < 0$, where the inequality follows again from the concavity of $f(w)$ and the fact that $\beta^b(w) > 0$ by incentive constraint (8).[32]

[32]Note also that from $\nu p'(I_{FB})\Delta - 1 = 0$ we must have that $p'(I_{FB}) > 0$, which is also a necessary condition for limited liability not to bind.

As for the remaining case, underinvestment if I_{FB} is sufficiently low, we will show that $\partial\mathcal{L}/\partial I|_{I(w)=I_{FB}} \to 0$ from below as $I_{FB} \searrow 0$. To see this note that, in the limit, δ^g dominates δ^b from Lemma 5, such that $\phi(w)$ is strictly positive. The result then follows from continuity and the fact that $I(w) = I_{FB}$ is the uniquely optimal investment level for $I_{FB} = 0$. **Q.E.D.**

Proof of Lemma 6. For part i) we first show that there exists $\widehat{w}_\beta \in (0, \overline{w})$ such that $\beta^b(w) = -w$ for $w \in [0, \widehat{w}_\beta]$. To see this, note that, for any $I(w) > 0$, we must have $\beta^g(w) - \beta^b(w) > 0$, which together with the strict concavity of $f(w)$ for $w \in [0, \overline{w})$ implies that the first-order condition in (19) cannot be satisfied in a neighborhood of $w = 0$. Thus, we get the stated result, with $\beta^b(w) = -w$ and $\beta^g(w) = \beta^b + \frac{\lambda}{\nu p'(I(w))} > 0$ on $w \in [0, \widehat{w}_\beta]$, where

$$\widehat{w}_\beta := \min\left\{w > 0 : p(I(w))f'\left(\frac{\lambda}{\nu p'(I(w))}\right) + \big(1 - p(I(w))\big)f'(0) = f'(w)\right\}.$$

Next, let us show the underinvestment result for small w. With the limited liability constraint binding, marginal agency costs of investment $\phi(w)$ in ($A.9$) can be rewritten to obtain

$$\begin{aligned}\phi(w) = \nu p'(I(w))&\left[f(\frac{\lambda}{\nu p'(I(w))}) - \left(f(0) + \frac{\lambda}{\nu p'(I(w))}f'(w)\right)\right]\\ &+ \lambda\frac{p(I(w))p''(I(w))}{(p'(I(w)))^2}\left[f'(w) - f'(w + \beta^g(w))\right].\end{aligned}$$

It remains to show that this expression is negative for small w. To see this, note first, that the expression in the second line is negative for all w by strict concavity of $f(w)$. As for the remaining right-hand-side expression, it also follows from concavity of $f(w)$ that this is negative for w small enough.

For part iii), note that as $f(w)$ extends linearly for $w > \overline{w}$, we have that $f'(\overline{w}+\beta^g(\overline{w})) = f'(\overline{w})$ and (19) can only be satisfied if $\beta^b(\overline{w}) = 0$. It remains to show that $I(\overline{w}) = I^{FB}$. Since $\beta^g(\overline{w}) > 0$, but $f'(\overline{w} + \beta^g(\overline{w})) = f'(\overline{w})$, and $\beta^b(\overline{w}) = 0$, the marginal agency costs of investment $\phi(w)$ in ($A.9$) are equal to zero, implying first-best investment.

Part ii) then is immediate as $\beta^b(\overline{w}) = 0$ and continuity imply that the limited liability constraint (9) is slack for w close to $\overline{w}$ such that Lemma 4 and Proposition 2 apply. **Q.E.D.**

Appendix B Additional Material (Online)

This Appendix provides some supplementary material, including a complete characterization of the optimal contract in the benchmark case with contractible investment expenditures (Appendix B.1), as well as an analysis of the extension of the baseline model to an arbitrary number of investment outcomes (Appendix B.2). Further, in Appendix B.3 we provide an implementation of the optimal contract, while Appendix B.4 contains an outline of the algorithm used for solving for the optimal contract numerically.

B.1 Contractible Investment Benchmark

When investment expenditures are contractible, the problem reduces to a standard single-task cash flow diversion problem in which the principal directly controls investment. Formally, he then also solves the boundary value problem in (12)-(15), but does not have to respect the incentive constraint for investment (8). We denote the respective value function with contractible investment by $f_{CI}(w)$ and index contractual parameters, such as the compensation boundary $\overline{w}_{CI}$. The solution to this benchmark case is summarized as follows:

Proposition B.1. *Assume investment expenditures are contractible, then, under the optimal truth-telling contract, investment is given by the constant first-best investment level I_{FB} in (20) $\forall t$. The incumbent agent's continuation payoff evolves according to (6) with $\alpha_t = \lambda$, $\beta_t^g = \beta_t^b = 0$ and $I_t = I_{FB}$, $\forall t$. When $w_{t^-} \in [0, \overline{w}_{CI})$, $dU_t = 0$; when $w_{t^-} \geq \overline{w}_{CI}$ payments dU_t cause w_{t^-} to reflect at $\overline{w}_{CI}$. The incumbent agent is replaced when $w_{t^-} = 0$. The principal's expected payoff at any point in time is given by $f_{CI}^i(w_t)$, $i \in \{h, l\}$, which satisfies $f_{CI}(w) := f_{CI}^l(w) = f_{CI}^h(w) - \Delta$, where $f_{CI}(w)$ is concave, strictly so for $0 \leq w < \overline{w}_{CI}$ and solves, for $w \in [0, \overline{w}_{CI}]$, the HJB equation*

$$rf_{CI}(w) = \mu^l - I_{FB} + \nu p(I_{FB})\Delta + \gamma w f'_{CI}(w) + \frac{1}{2}\sigma^2\lambda^2 f''_{CI}(w) \tag{B.1}$$

with boundary conditions $f_{CI}(0) = f_{CI}(w^*_{CI}) - k$, where $w^*_{CI} \in \arg\max_w \{f_{CI}(w)\}$, $f'_{CI}(\overline{w}_{CI}) = -1$ and $f''_{CI}(\overline{w}_{CI}) = 0$.

Proof of Proposition B.1. This result is a straightforward extension of Hoffmann and Pfeil (2010), who study a dynamic cash-flow diversion models with exogenous shocks to profitability; we therefore will be brief. Note, first, that $f_{CI}(w)$ is strictly concave, which follows from the same arguments as in Hoffmann and Pfeil (2010). Further, the incentive constraint binds, i.e., $\alpha = \lambda$. Interior solutions for I are then given by

$$\Delta + \left[f_{CI}(w+\beta^g) - f_{CI}(w+\beta^b)\right] - \left(\beta^g - \beta^b\right) f'_{CI}(w) = \frac{1}{\nu p'(I)}, \tag{B.2}$$

while the β^j, $j \in \{g, b\}$, are determined from $f'_{CI}(w+\beta^g) = f'_{CI}(w) = f'_{CI}(w+\beta^b)$. Due to the strict concavity of f_{CI} this immediately implies that $\beta^g = 0$ and $\beta^b = 0$, and, plugging into (B.2), we find that optimal investment is equal to the first-best level as characterized in (20), such that the proposed solution achieves efficient investment. Verification is then standard (c.f., Hoffmann and Pfeil 2010) and therefore omitted. **Q.E.D.**

As in the case with non-contractible investment, the agent's incentive constraint with respect to cash flow diversion (7) binds under the optimal contract because it is costly to provide incentives (formally the instantaneous volatility of w increases in α and $f_{CI}(w)$ is concave). A similar argument implies that, when investment expenditures are contractible and there is, thus, no need to provide incentives based on the investment outcome, it is optimal to choose $\beta^g_{CI} = \beta^b_{CI} = 0$.[33] As a consequence, the agent's continuation value is insensitive to the investment outcome and the principal's investment problem is, thus, independent of the cash-flow diversion problem. Hence, the optimal investment policy is equal to first-best.

B.2 General Model

In this Appendix we show how our analysis can be extended beyond the binary state structure allowing firm profitability to take on values $\{\mu^i\}_{i=1}^M$ for any $M \geq 2$, where $\mu^1 < \mu^2 < ... < \mu^M$.[34] As in our baseline model the industry is subject to (rare) exogenous *technology shocks* governed by a Poisson process $\mathbf{N}$ with intensity ν, indicating the availability of a new technology. Investment determines the probability distribution over future profitability in the event of a technology shock, now according to $p_i(I_t) := \Pr(\mu_{t^+} = \mu^i | I_t)$ for $i = 1, ..., M$.[35] It is then again convenient to define Poisson processes $\mathbf{N}^i$ with arrival rate $\nu p_i(I_t)$ for $i = 1, ..., M$ that capture the investment outcome in the event of a tech-

[33] Formally, this follows from the fact that the costs of compensating the agent, as reflected in the slope $f'_{CI}(w)$, are independent of current profitability. If the costs of compensating the agent differed across states $i \in \{h, l\}$, e.g., due to state-dependent hiring costs, it would also be efficiency enhancing to specify $\beta^g_{CI}, \beta^b_{CI} \neq 0$ (cf., Hoffmann and Pfeil (2010) for a formal model of this "reward for luck" effect).

[34] For expositional clarity we assume that it is profitable to run the firm in each state, which requires that μ^1 is sufficiently large.

[35] Clearly, we must have $p_i(I) \geq 0$ as well as $\sum_{i=1}^M p_i(I) = 1$. Further, to ensure identifiability, the probability distribution $p_i(I)$ must vary with investment I, in particular, assuming differentiability of $p_i(I)$ for $I \in (0, \bar{I})$, there must for each I exist an i such that $\frac{d}{dI} p_i(I) \neq 0$.

nology shock with $dN_t^i = 1$ if $\mu_{t+} = \mu^i$ and zero else. In between two technology shocks, profitability remains unchanged. Hence, first-best investment, which corresponds to optimal investment when investment expenditures are contractible (cf., Proposition B.1), solves

$$I_{FB} = \arg\max \left\{ \nu \sum_{i=1}^{M} p_i(I)\Delta^i - I \right\},$$

where $\Delta^i := \frac{\mu^i - \mu^1}{r+\nu}$ captures the gain from increasing profitability above the minimal level properly accounting for the Markov switching structure. In order to obtain a tractable analysis with non-contractible investment, we impose the following standard assumptions on the investment technology

Assumption B.2. *The investment technology has the following properties*
i) Monotone likelihood ratio property (MLRP): $p_i'(I)/p_i(I) > p_j'(I)/p_j(I)$ *for* $i > j$*.*
ii) Validity of first-order approach (FOA): $p_i'(I)p_i''(I) < 0$ *for all* i*.*

The MLRP ensures that the realization of higher profitability is more indicative of the manager having invested a lot. It further implies a first-order stochastic dominance (FOSD) ranking on the cdf $F(\mu^i|I) = \sum_{j=1}^{i} p_j(I)$, i.e., investment increases the chances of high future profitability according to $\frac{d}{dI}F(\cdot|I) < 0$. As is common in moral hazard problems with continuous actions (see e.g., Holmstrom 1979) we further assume that the first-order approach is valid, such that we can replace the incentive constraint for investment at each t by the respective first-order condition. The second part of Assumption B.2 provides a sufficient (but not necessary) condition. We note that Assumption B.2 is clearly satisfied for the investment technology in our baseline model with binary states. Apart from the richer investment technology all other aspects of the model are as described in Section 2. Hence, the contracting problem is to find an incentive compatible truth-telling contract $(\mathbf{S}^*, \mathbf{U}, \tau)$, maximizing the principal‘s expected profit f_0 from (3) for given initial profitability $\mu_0 \in \{\mu^1, ..., \mu^M\}$, subject to incentive compatibility (4) and limited liability, while delivering expected payoff w_0 as defined in (2) to the agent. Accordingly the following derivation is a straightforward extension of the analysis in Section 3 of the main text and we will therefore be brief.[36]

Continuation Payoff and Local Incentive Compatibility Again, the contract can be written in terms of the agent's continuation payoff w_t as defined in (5) as the single state variable. In particular, analogous to Lemmas 1 and 2 the agent‘s continuation payoff

[36]Formal proofs of all subsequent results are available from the authors upon request.

evolves as

$$dw_{t^-} = \gamma dw_{t^-}dt - dU_t + \alpha_t \left(d\hat{Y}_t - (\mu_t - I_t^*)\, dt \right) + \sum_{i=1}^{M} \beta_t^i \left[dN_t^i - \nu p_i(I_t^*)dt \right],$$

and truth-telling as well as following the prescribed investment $I_t^* \in \left(0, \bar{I}\right)$ is incentive compatible if and only if $\alpha_t \geq \lambda$ and[37]

$$\nu \sum_{i=1}^{M} p_i'(I_t^*)\beta_t^i - \lambda = 0. \tag{B.4}$$

Further, limited liability requires, as before, that $\beta_t^i \geq -w_{t^-}$ for all i.

Optimal Contract As in the baseline model in the main text, the optimal contract can be derived using the dynamic programming approach where we denote the principal's value function for given profitability μ^i, $i \in \{1, ..., M\}$ and agent's continuation value w by $f^i(w)$. As the agency problem for given w is independent of μ^i, it follows from the same arguments as in the main text that the optimal contract is independent of current profitability. Hence, we can conveniently characterize the optimal contract using the principal's value function in the lowest profitability state μ^1, which, in analogy to the notation in the main text, is denoted by $f(w) := f^1(w)$. Then, extending Lemma 3 to more than two profitability states, we obtain that $f^i(w) = f(w) + \Delta^i$ for all $i \in \{2, ..., M\}$ with $\Delta^i := \frac{\mu^i - \mu^1}{r+\nu}$.

The optimal compensation policy is then again characterized by a threshold $\overline{w}$ solving $f'(\overline{w}) = -1$. Compensation is deferred until $\overline{w}$ is reached, where the contract starts paying

[37] Condition (B.4) requires I_t^* to solve the first-order condition of the agent's maximization problem over investment I_t given the contract. It is then easy to see that under the optimal contract we must have $p_i'(I_t^*)\beta^i \geq 0$, i.e., if the probability of investment outcome i is increasing (decreasing) in investment at $I_t = I_t^*$ the agent should be rewarded (punished) if i realizes. Hence, using part (ii) of Assumption B.2, also the agent's second-order condition is satisfied

$$\nu \sum_{i=1}^{N} p_i''(I_t^*)\beta_t^i \leq 0. \tag{B.3}$$

out cash compensation. For $w \in [0, \overline{w}]$ the principal's problem can then be written as

$$(r+\nu) f(w) = \max_{\beta^i \geq -w, I} \left\{ \begin{array}{l} \mu^1 - I + \nu \sum_{i=1}^M p_i(I)\Delta^i + \frac{1}{2}\sigma^2\lambda^2 f''(w) \\ + \left[\gamma w - \nu \sum_{i=1}^M p_i(I)\beta^i\right] f'(w) \\ + \nu \sum_{i=1}^M p_i(I) f(w+\beta^i) \end{array} \right\} \quad s.t.\ (B.4) \tag{B.5}$$

where we already substituted the optimally binding truth-telling constraint $\alpha = \lambda$. The optimal contract is then characterized as the solution to (B.5) with the relevant boundary conditions as given by (13) to (15) in the main text.

Optimal Investment In order to provide incentives for investment the contract has to specify some reward/punishment (β^i) depending on the investment outcome to satisfy incentive constraint (B.4). Taking first-order conditions in (B.5) and denoting by $\eta(w) \geq 0$ the Lagrange multiplier on (B.4), interior optimal values of $\beta^i(w)$ solve

$$\frac{f'(w) - f'(w+\beta^i(w))}{p_i'(I)/p_i(I)} = \eta(w), \tag{B.6}$$

for $I = I(w)$, which is exactly condition (18) in the main text. Hence, as long as $w < \overline{w}$ and the likelihood ratio $p_i'(I)/p_i(I)$ is bounded for all i, the shadow costs of delegated investment $\eta(w)$ are strictly positive. As the principal is risk averse with respect to variation in the agent's compensation, $f''(w) < 0$, it is then optimal to reward the agent ($\beta^i(w) > 0$) for all i with strictly positive likelihood ratio and punish ($\beta^i(w) < 0$) for all i with strictly negative likelihood ratio, where the size of the respective reward and punishment is increasing in the informativeness of the respective performance signal as measured by the absolute value of the likelihood ratio. From MLRP (see Assumption B.2 part (i)) the agent's compensation is, thus, increasing in the investment outcome (it is negative for μ^i small and positive for μ^i large).

Next, consider the optimal interior investment policy, which solves the following first-order condition:

$$\nu \sum_{i=1}^M p_i'(I)\Delta^i - 1 = \frac{\partial}{\partial I}\Phi(w, I(w)), \tag{B.7}$$

where the marginal agency costs of delegated investment are given by

$$\begin{aligned}\frac{\partial}{\partial I}\Phi(w,I) &:= -\eta\nu\sum_{i=1}^{M} p_i''(I)\beta^i(w) \\ &+\nu\sum_{i\in\left\{j=1,\ldots,M:\, p_j'(I)>0\right\}} p_i'(I)\left[f(w)+\beta^i(w)f'(w)-f(w+\beta^i(w))\right] \\ &+\nu\sum_{i\in\left\{j=1,\ldots,M:\, p_j'(I)\leq 0\right\}} p_i'(I)\left[f(w)+\beta^i(w)f'(w)-f(w+\beta^i(w))\right].\end{aligned} \tag{B.8}$$

When $\frac{\partial}{\partial I}\Phi(w,I(w))=0$, first-best investment obtains. From $\eta(w)\geq 0$ and (B.3) the first term in (B.8) is unambiguously positive. Intuitively, as long as there is a relevant incentive problem, more investment requires more costly incentives. By concavity of the value function the second term is also positive while the third term is negative. Increasing investment increases the probability of the states for which costly rewards are paid, but it reduces the probability of those states for which the optimal compensation policy requires costly punishment. The following result shows that the third term dominates when the more informative investment outcomes are those requiring punishment. Then $\frac{\partial}{\partial I}\Phi(w,I(w))<0$ and the optimal investment level is distorted upwards relative to first-best.

Proposition B.3. *Fix $w\in(0,\bar{w})$, take first-best investment to be interior $I^{FB}\in\left(0,\bar{I}\right)$, and consider the most informative performance signal $k=\arg\max_j\left|p_j'(I)/p_j(I)\right|$. Then, as long as the limited liability constraint is slack, optimal investment is distorted upwards, $I(w)>I^{FB}$, if the likelihood ratio $p_k'(I)/p_k(I)$ is sufficiently negative at I^{FB}, while it is distorted downwards, $I(w)<I^{FB}$, if the respective likelihood ratio is sufficiently positive at I^{FB}.*

Proof of Proposition B.3. We assume for simplicity that the optimal investment $I(w)$ solving (B.7) is unique. Note first that from (B.6) $\eta\to 0$ as $\left|p_k'(I)/p_k(I)\right|\to\infty$ such that optimal investment approaches first-best in the limit. Now, as $\eta\to 0$ the first term in (B.8) goes to zero. Further, (B.6) implies that $\beta^j\to 0$ for all $j\neq k$ and it follows from $p_k'(I)/p_k(I)<0$ that also the second term in (B.8) goes to zero, while the third term stays strictly positive. The result then follows from continuity and the first-order condition for optimal investment in (B.7). The argument showing the underinvestment result is analogous (cf., also the proof of Proposition 2). **Q.E.D.**

Proposition B.3 clearly illustrates that the sign of the investment distortion is driven by the relative informativeness of the available performance signals as captured by the respective likelihood ratios.[38] In particular, if low (high) realizations of μ^i allow for relatively cost-effective incentive provision, optimal incentive pay focusses on these states which are, hence, severely punished (highly rewarded). As a consequence the realization of these states induces a high variation in the agent's continuation value which is costly to the principal, who therefore distorts investment upwards (downwards) in order to make these costly states less likely.

We close this discussion by noting that the above analysis as well as the main results extend straightforwardly to the case where the principal observes additional informative signals about investment expenditures other than realized profitability. As illustrated above, what matters for investment distortions as well as optimal incentives for delegated investment is only the informativeness of the respective performance signals. To build intuition consider the special case where the principal observes some signal s informative about investment expenditures (e.g., an accounting measure) which is a sufficient statistic for realized profitability μ, i.e., given the signal s, profitability μ does not contain any further information about investment. Formally μ is a garbling of s that might, e.g., be confounded by additional factors affecting profitability which are beyond the agent's control.[39] Then, following Holmstrom's (1979) sufficient statistics result the agent's incentive compensation should only condition on s and not on realized profitability such that investment distortions as characterized in Proposition B.3 depend only on the likelihood ratio distribution of the sufficient statistic.

B.3 Implementation of the Optimal Contract

In this Appendix we illustrate one particular implementation of the optimal contract as characterized in Section 3 of the main text. The implementation follows DeMarzo et al. (2012) and is based on cash reserves as a measure of financial slack, equity and a portfolio of derivative securities or insurance contracts.

In particular, the firm is equity financed and uses cash reserves to cover its short-term liquidity needs. Let M_t denote the level of cash reserves, earning interest r. Cash reserves grow if cash flows are positive and they are used to cover operating losses, which corresponds to negative cash flows. Further, equity holders' require a minimum dividend, given by

$$dD_t = [\mu_t - I_t - (\gamma - r) M_t] \, dt, \tag{B.9}$$

[38] Note that Proposition B.3 does not require MLRP (Assumption B.2 (i)) to hold. The following interpretation still uses MLRP to identify the states μ^i with a negative (positive) likelihood ratio as low (high) states.

[39] Hence, profitability might be far more volatile than the sufficient statistic.

which is paid out of cash reserves M_t. This minimum dividend comprises of the expected free cash flow $\mu_t - I_t$ minus an adjustment factor that reflects the discounting difference $\gamma - r$. If the firm fails to meet the minimum payout rate (B.9), or the cash holdings are exhausted, the manager is laid off, which is critical in providing incentives for the manager not to divert cash flows. Incentives for investment can then be provided by creating exposure of the firm's financial slack M_t to the investment outcome, which we formalize through a portfolio of derivative securities contingent on the investment outcome.[40] Derivatives are fairly priced given investors' beliefs on the level of investment I_t. Concretely, we stipulate that holding a state-price security that pays one unit in case of an investment success (failure) incurs flow costs of $P^g = \nu p(I_t)$ $(P^b = \nu(1 - p(I_t)))$, so that the instantaneous net payoff from holding such a security is given by $dS_t^j = dN_t^j - P_t^j dt$, for $j \in \{g, b\}$. We denote the number of securities of type $j \in \{g, b\}$ held by the firm at time t by n_t^j, and require the firm to hold at any time a portfolio of size

$$n_t^j = \frac{\beta^j(\lambda M_t)}{\lambda}, j \in \{g, b\}, \tag{B.10}$$

of the respective security.[41] Else, the manager is replaced. Other than that, the manager is free to choose investment and to distribute cash in form a special dividend X at any time. The manager receives compensation in form of fraction λ of this special dividend. Overall, the firm's cash reserves, thus, follow

$$dM_t = rM_t dt + d\hat{Y}_t + n_t^g dS_t^g + n_t^b dS_t^b - dD_t - dX_t. \tag{B.11}$$

When M_t hits zero for the first time, the firm can no longer pay the minimum dividend dD and, thus, goes into restructuring. In this process the incumbent manager is fired and equity holders realize a payoff corresponding to L_τ, where τ denotes the first time at which M_t falls to zero. The value of the firm's equity claim is then given by

$$P(M_t, \mu_t) = E_t\left[\int_t^\tau e^{-r(s-t)}(dD_s + (1-\lambda)dX_s) + e^{-r(\tau-t)}L_\tau\right], \tag{B.12}$$

[40]Note that we stipulate that the investment outcome is a verifiable event. In our interpretation of investment into absorptive capacity (see footnote 6) an investment success could, e.g., be a patent granted to the firm and an investment failure a patent granted to a competitor. In this case, successes and failures would be reflected in the firm's stock price, implying that the required exposure could be created by derivatives based on this underlying. For an alternative implementation in a setting with only downside risk see Biais et al. (2010). There the firm is requested to maintain an insurance contract against accident costs which, for incentive reasons, entails only partial coverage and a down-sizing covenant.

[41]Note from $\beta^b < 0$ (see (19)), the firm holds a short position in security b, i.e., $n_t^b < 0$.

and we have the following result:

Proposition B.4. *Suppose the firm has initial cash reserves M_0 and can operate as long as $M_t \geq 0$. When the manager is fired unless he maintains the minimum payout rate dD_t and holds a derivative security portfolio n_t^g, n_t^b, it is optimal for him to refrain from diverting funds and to implement the optimal investment profile as characterized in Proposition 1. The firm accumulates cash M_t until $M_t = \overline{w}/\lambda$, and pays out cash in excess of this amount. Given this policy, the manager's payoff is $w_t = \lambda M_t$, which coincides with the continuation value of Proposition 1, and the equity value satisfies $P(M_t, \mu_t) = f(\lambda M_t, \mu_t) + M_t$.*

Proof of Proposition B.4. Under the proposed implementation, cash reserves evolve according to

$$\begin{aligned} dM_t &= \gamma M_t dt + \left(d\widehat{Y}_t - (\mu_t - I_t)\,dt\right) + \frac{\beta^g(\lambda M_t)}{\lambda}\left(dN_t^g - \nu p(I_t)\,dt\right) \\ &\quad + \frac{\beta^b(\lambda M_t)}{\lambda}\left(dN_t^b - \nu(1-p(I_t))\,dt\right) - dX_t. \end{aligned}$$

Now, define $w_t = \lambda M_t$ to get

$$\begin{aligned} dw_t &= \lambda dM_t = \gamma w_t dt + \lambda\left(d\widehat{Y}_t - (\mu_t - I_t)\,dt\right) + \beta^g(w_t)\left(dN_t^g - \nu p(I_t)\,dt\right) \\ &\quad + \beta^b(w_t)\left(dN_t^b - \nu(1-p(I_t))\,dt\right) - \lambda dX_t. \end{aligned}$$

Letting $dU_t = \lambda dX_t$, incentive compatibility under the proposed implementation then follows from incentive compatibility of the optimal contract characterized in Proposition 1 and the agent's value is given by w_t. Note further, that the agent is indifferent as to when to issue the special dividend.

Next, consider the valuation of the equity claim, which follows from arguments similar to those in DeMarzo et al. (2012), in particular their Proposition 2. Substituting from (B.11), (B.12) can be written as

$$\begin{aligned} P(M_t, \mu_t) &= E_t\left[\int_t^\tau e^{-r(s-t)}\left((\mu_s - I_s)\,ds - dU_s\right) + e^{-r(\tau-t)}L_\tau\right] \\ &\quad + E_t\left[\int_t^\tau e^{-r(s-t)}\left(rM_s ds - dM_s\right)\right] \\ &= f(W_t, \mu_t) + M_t = f^i(\lambda M_t) + M_t, \end{aligned}$$

where we have used integration by parts. **Q.E.D.**

Note that in the implementation given in Proposition B.4, the state-price securities are not used to hedge against investment failure. By contrast, although firm value is a concave function of financial slack, the firm's derivative position deliberately creates exposure of its financial position to the uncertain investment outcome in order to provide the appropriate incentives for investment. Let us also comment a bit more on the interpretation of the restructuring process triggered when $M_t = 0$. In this event, which could be interpreted as insolvency, the firm needs to raise cash from the capital market, which involves a fixed cost of k. With the equity value net of the required cash injection given by $f(\lambda M_t, \mu_t)$, the firm continues to operate with a new manager and initial cash reserves of $M^* = w^*/\lambda$. Hence, we have $P(0, \mu) = P(M^*, \mu) - k$, where we interpret k as the cost of raising external funds, which captures the key financing friction in our model.

B.4 Numerical Implementation

To solve numerically for the optimal contract of Proposition 1, we take the following iteration steps.

1. Solve for the principal's value function $f^{(0)}$ and the free boundary $\overline{w}^{(0)}$ without technology shocks. That is, we solve the ODE in (12) with $\nu = 0$ (thus the initial investment is $I^{(0)} = 0$ and the initial rewards and punishments are $\beta^{g^{(0)}} = \beta^{b^{(0)}} = 0$).

2. Given $f^{(0)}$, $\overline{w}^{(0)}$, and $\beta^{b^{(0)}}$, update the optimal investment scheme $I^{(1)}$ according to (16) and subject to incentive compatibility, i.e., $\beta^{g}\left(I^{(1)}\right) = \beta^{b^{(0)}} + \lambda / \left(\nu p'(I^{(1)})\right)$.

3. Given $\beta^{g^{(0)}}$, $\beta^{b^{(0)}}$, and $I^{(1)}$, update the principal's value function $f^{(1)}$ and the free boundary $\overline{w}^{(1)}$.

4. Given $f^{(1)}$, $\overline{w}^{(1)}$, and $I^{(1)}$, update the optimal rewards and punishments $\beta^{g^{(1)}}$ and $\beta^{b^{(1)}}$ according to the first order condition (19), subject to the (binding) incentive constraint (8) and the limited liability constraint (9). That is, we solve (19) for $\beta^{b^{(1)}}$, such that $\beta^{g^{(1)}} = \beta^{b^{(1)}} + \lambda / \left(\nu p'(I^{(1)})\right)$ and $\beta^{b^{(1)}} \geq w$.

5. Repeat steps 2 to 4 until the problem converges. The convergence criterion is

$$\max\left[\sup_{w}\left|I^{(i+1)} - I^{(i)}\right|, \sup_{w}\left|f^{(i+1)} - f^{(i)}\right|\right] < 10^{-5}.$$