Dynamic Multitasking and Managerial Investment Incentives

Online Appendix

February 16, 2021

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

A 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) subject to (7) and (9), 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 $w_{CI}$ – accordingly. The solution to this benchmark case is summarized as follows:

**Proposition A.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 $\alpha_t = \lambda$, $\beta^g_i = \beta^b_i = 0$ and $I_t = I_{FB}$, $\forall t$. When $w_{t-} \in [0, w_{CI})$, $dU_t = 0$; when $w_{t-} \geq w_{CI}$ payments $dU_t$ cause $w_{t-}$ to reflect at $w_{CI}$. 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

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

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

**Proof of Proposition A.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). Hence, the incentive constraint for truthful reporting (7) binds, i.e., $\alpha = \lambda$, since it is costly to provide incentives (formally the instantaneous volatility of $w$ increases in $\alpha$). A similar argument then directly 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 $\beta^g_{CI} = \beta^b_{CI} = 0$. Formally, the sensitivities $\beta^j_{CI}, j \in \{g, b\}$ are determined from $f'_{CI}(w + \beta^g_{CI}) = f'_{CI}(w) = f'_{CI}(w + \beta^b_{CI})$ and the result again follows from the strict concavity of $f_{CI}$.\footnote{Note that here we use 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).}

Finally, interior solutions for $I$ are then given by

$$\Delta + [f_{CI}(w + \beta^g_{CI}) - f_{CI}(w + \beta^b_{CI})] - (\beta^g_{CI} - \beta^b_{CI}) f'_{CI}(w) = \frac{1}{\nu p' (I)},$$  \hspace{1cm} (A.2)

which together with $\beta^g_{CI} = \beta^b_{CI} = 0$ simplifies to $I = I_{FB}$ as characterized in (20). Intuitively, since 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, the proposed optimal investment policy is equal to first-best. Verification is then standard (cf., Hoffmann and Pfeil 2010) and therefore omitted. **Q.E.D.**
B 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 \( \mu^1 < \mu^2 < ... < \mu^M \). As in our baseline model the industry is subject to (rare) exogenous technology shocks governed by a Poisson process \( \mathbf{N} \) with intensity \( \nu \), indicating the availability of a new technology. Investment \( I \in [0, \bar{I}] \) determines the probability distribution over future profitability in the event of a technology shock, now according to

\[
p_i(I_t) := \Pr(\mu_{t+} = \mu^i|I_t) \text{ for } i = 1, ..., M. \]

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 technology shock with \( dN^i_t = 1 \) 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 A.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 now impose the following assumptions on the investment technology which are straightforward generalizations of the assumptions made in the main text to the case with \( M > 2 \).

**Assumption B.1.** The investment technology has the following properties

1. \( p'_i(I)/p_i(I) > p'_j(I)/p_j(I) \) for \( i > j \) and all \( I \), i.e., the monotone likelihood ratio

For expositional clarity we assume that it is profitable to run the firm in each state, which requires that \( \mu^1 \) is sufficiently large.

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 \( \frac{d}{dI} p_i(I) \neq 0 \).

In a slight abuse of notation \( p'_i(I) \) and \( p''_i(I) \) for \( I = 0 \) and \( I = \bar{I} \) refer to the respective right-/left-hand side derivatives.
property (MLRP) holds.

(ii) \( p_M(0) = 0 \) and \( p_1(\bar{I}) = 0 \), i.e., the best (worst) possible outcome cannot occur if \( I = 0 \) (\( I = \bar{I} \)).

(iii) \( p_i'(I)p_i''(I) < 0 \) for all \( i, I \), i.e., the first-order approach (FOA) applies.

Part (i) of Assumption B.1 (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., conditional on a technology shock, investment increases the chances of high future profitability according to \( \frac{d}{dI} F(\cdot|I) < 0 \), thus, shifting probability mass from low to high levels of profitability. Part (ii) of Assumption B.1 ensures that this FOSD shift is sufficiently strong in the sense that the realization of extremely bad (good) outcomes becomes very unlikely if investment is sufficiently high (low). A reasonable chance at realizing the highest possible outcome, thus, requires some minimal investment, which constitutes a limit to good luck, while high investment protects the firm from the worst possible outcome, thereby limiting bad luck. While parts (i) and (ii) of Assumption B.1 are not necessary for the subsequent characterization of the optimal contract, they will play a role in establishing the sign of investment distortions, in particular they are sufficient (but not necessary) to show that investment is distorted upwards for highly profitable and distorted downwards for less profitable investment technologies. We comment more on their relevance after establishing this result. Finally, as is common in moral hazard problems with continuous actions (see e.g., Holmström 1979), we impose a technical condition to ensure validity of the first-order approach, such that we can replace the incentive constraint for investment at each \( t \) by the respective first-order condition. Part (iii) of Assumption B.1 provides a sufficient (but not necessary) condition. We note that Assumption B.1 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 of the main text. Hence, the contracting problem is to find an incentive compatible truth-telling contract \( (S^*, U, \tau) \), 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.\textsuperscript{5}

**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 evolves as

$$dw_t^- = \gamma dw_t^- dt - dU_t + \alpha_t \left( d\hat{Y}_t - (\mu_t - I_t^*) dt \right) + \sum_{i=1}^{M} \beta^i_t \left[ dN^i_t - \nu p_i(I_t^*) dt \right].$$

Truth-telling and following the prescribed investment $I_t^* \in (0, \bar{I})$ is incentive compatible if and only if $\alpha_t \geq \lambda$ and\textsuperscript{6}

$$\nu \sum_{i=1}^{M} p'_i(I_t^*) \beta^i_t - \alpha_t = 0, \forall t. \quad (B.2)$$

Further, limited liability requires, as before, that $\beta^i_t \geq -w_t$ 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 $\mu^i, i \in \{1, ..., M\}$ and agent’s continuation value $w$ by $f^i(w)$.

\textsuperscript{5}Formal proofs of all steps leading to the subsequent results are available from the authors upon request.

\textsuperscript{6}Condition (B.2) 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 (iii) of Assumption B.1, also the agent’s second-order condition is satisfied

$$\nu \sum_{i=1}^{N} p''_i(I_t^*) \beta^i_t \leq 0. \quad (B.1)$$
As the agency problem for given \( w \) is independent of \( \mu^i \), 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 \( \mu^1 \), 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 out cash compensation. For \( w \in [0, \overline{w}] \) the principal’s problem can then be written as

\[
(r + \nu) f(w) = \max_{\beta \geq -w, t} \left\{ \mu^1 - I + \nu \sum_{i=1}^{M} p_i(I) \Delta^i + \frac{1}{2} \sigma^2 \lambda^2 f''(w) \right. \\
+ \left[ \gamma w - \nu \sum_{i=1}^{M} p_i(I) \beta^i \right] f'(w) \\
\left. + \nu \sum_{i=1}^{M} p_i(I) f(w + \beta^i) \right\} \tag{B.3}
\]

\[ \text{s.t. } (B.2) \]

where we already substituted the optimally binding truth-telling constraint \( \alpha = \lambda \). The optimal contract is then characterized as the solution to (B.3) 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 \( (\beta^i) \) depending on the investment outcome to satisfy incentive constraint (B.2). Taking first-order conditions in (B.3) and denoting by \( \eta(w) \geq 0 \) the Lagrange multiplier on (B.2), interior optimal values of \( \beta^i(w) \) solve

\[
\frac{f'(w) - f'(w + \beta^i(w))}{p_i'(I)/p_i(I)} = \eta(w), \tag{B.4}
\]

for \( I = I(w) \), which is exactly condition (18) in the main text. Hence, the need to provide incentives (see (B.2)) then implies that the shadow costs of delegated investment \( \eta(w) \) are strictly positive as long as signals are not perfectly informative – i.e., the likelihood ratio \( p_i'(I)/p_i(I) \) is bounded for all \( i \) – and rewarding or punishing the agent is costly, which for
$w < \bar{w}$ is reflected in the concavity of the value function. 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.1 part (i)) the agent’s compensation is, thus, increasing in the investment outcome (it is negative for $\mu^i$ small and positive for $\mu^i$ large).

In order to understand the subsequent analysis of investment distortions, it is now instructive to compare how the distribution of the (marginal) costs of providing incentives $\left| f'(w) - f'(w + \beta^i(w)) \right|$ across investment outcomes $i = 1, \ldots, M$ changes with the implemented level of investment. To do so, note that from (B.4) the optimal incentive scheme for investment equalizes the expected (marginal) costs of providing incentives through rewards and punishment (akin to (19) for the two state case):

$$\sum_{i \in \{j = 1, \ldots, M: p'_j(I) > 0\}} p_i(I) \left[ f'(w) - f'(w + \beta^i(w)) \right] = \sum_{i \in \{j = 1, \ldots, M: p'_j(I) \leq 0\}} p_i(I) \left[ f'(w + \beta^i(w)) - f'(w) \right].$$

(B.5)

Now, as investment increases, the implied FOSD shift in the outcome distribution (which follows from Assumption B.1 part (i)) directly implies that, under the optimal compensation policy, the costs of providing incentives upon realization of a low outcome triggering punishment (right-hand side) have to increase faster on average than the costs upon realization of a high outcome warranting a reward (left-hand side).\footnote{Formally, we have}

$$\sum_{i \in \{j = 1, \ldots, M: p'_j(I) \leq 0\}} p_i(I) \frac{\partial \left[ f'(w + \beta^i(w)) - f'(w) \right]}{\partial I} > \sum_{i \in \{j = 1, \ldots, M: p'_j(I) > 0\}} p_i(I) \frac{\partial \left[ f'(w) - f'(w + \beta^i(w)) \right]}{\partial I}.$$
providing incentives are optimally shifted from high towards low investment outcomes, which – while becoming less and less likely – are, thus, more and more costly upon realization. This insight will be crucial in understanding distortions in optimal investment to which we turn next.

From (B.3) the optimal interior investment policy solves the following first-order condition:

\[ \nu \sum_{i=1}^{M} p_i'(I) \Delta^i - 1 = \frac{\partial}{\partial I} \Phi(w, I(w)), \]  

(B.6)

where the marginal agency costs of delegated investment are given by

\[ \frac{\partial}{\partial I} \Phi(w, I) := -\eta \nu \sum_{i=1}^{M} p''_i(I) \beta^i(w) \]

+ \nu \sum_{i=\{j=1,..,M: p'_j(I) > 0\}} p'_i(I) \int_{0}^{\beta^i} [f'(w) - f'(w + x)] dx \]

(B.7)

+ \nu \sum_{i=\{j=1,..,M: p'_j(I) \leq 0\}} p'_i(I) \int_{\beta^i}^{0} [f'(w + x) - f'(w)] dx.

When \( \frac{\partial}{\partial I} \Phi(w, I(w)) = 0 \), first-best investment obtains. From \( \eta(w) \geq 0 \) and (B.1) the first term in (B.7) 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

\[ p_i(I) \int_{0}^{\beta^i} |f'(w + x) - f'(w)| dx. \]

While optimal compensation \( \beta^i \) is chosen to equalize expected marginal compensation costs (see B.5), optimal investment reflects the change in expected total compensation costs, arising from the investment-dependent FOSD shift in the outcome distribution (see B.7).
costly punishment. Now, recall from the discussion of the optimal compensation policy above that, as the implemented investment level increases, costs of providing incentives are more and more concentrated in low investment outcomes triggering punishment. Thus, the third term becomes increasingly important for more profitable investment technologies and eventually may even dominate such that \( \frac{\partial}{\partial I} \Phi(w, I(w)) < 0 \). In this case the optimal investment level is distorted upwards relative to first-best, while we have underinvestment else. The following Proposition formalizes these results and provides a straightforward extension of Proposition 2 to the case of an arbitrary number of states.

**Proposition B.2.** Fix \( w \in (0, \bar{w}) \) and suppose that Assumption B.1 holds. Then, as long as the limited liability constraint is slack, optimal investment is distorted

(i) downwards, \( I(w) < I_{FB} \), for sufficiently low values of \( I_{FB} > 0 \),

(ii) upwards, \( I(w) > I_{FB} \), for sufficiently high values of \( I_{FB} < \bar{I} \).

**Proof of Proposition B.2.** The proof follows the same steps as the proof of Proposition 2. Assume for simplicity that the optimal investment \( I(w) \) solving (B.6) is unique. Clearly, by setting \( \beta^i(w) = 0 \) for all \( i \) such that the agent chooses \( I(w) = 0 \), we have \( \Phi(w, I) = 0 \), whereas \( \Phi(w, I) > 0 \) whenever interior \( I(w) > 0 \) has to be incentivized. It then is a direct implication of the continuity of \( \Phi(\cdot, I) \),\(^9\) that the marginal agency costs of investment \( \phi(w, I) \) are positive at \( I_{FB} \) whenever \( I_{FB} \) is sufficiently small and part (i) follows from (B.6). To show part (ii), observe, first, that by Assumption B.1 part (i), \( p_1(I) \) is strictly increasing as we approach \( \bar{I} \) such that together with Assumption B.1 part (ii) an investment level equal to \( \bar{I} \) can be incentivized via punishing the agent off-equilibrium for the realization of \( \mu^1 \). Hence, \( \Phi(w, \bar{I}) = 0 \) and the result follows from \( \Phi(w, I) > 0 \) for interior investment levels and continuity. Q.E.D.

Proposition B.2 follows from the same robust economic intuition as Proposition 2 in the main text (which essentially is a special case with only two investment outcomes): Recall that investment induces a first order stochastic dominance shift in the output distribution (see Assumption B.1 (i)) such that high (low) investment outcomes are unlikely under the

\(^9\)Continuity follows from Assumption B.1 (i) and (ii).
first-best investment policy when $I_{FB}$ is sufficiently low (high). From (B.5), incentives for investment are then optimally provided via high reward (severe punishment) for these unlikely and, thus, in a likelihood-ratio sense very informative outcomes. As a consequence the realization of these outcomes induces a high variation in the agent’s continuation value which is costly to the principal, who therefore distorts investment downwards (upwards) in order to make the outcomes – which are particularly costly in terms of incentive provision – less likely.

We next comment on the sufficient conditions imposed to show the results in Proposition B.2. In doing so, we focus on the overinvestment case (part (ii) in Proposition B.2) for expositional clarity, the discussion of the underinvestment case (part (i) in Proposition B.2) is symmetric. The key economic channel underlying the results in Proposition B.2 is that higher investment shifts probability mass from low to high investment outcomes, which is implied by Assumption B.1 (i). Assumption B.1 (ii) ensures that this shift is sufficiently strong at the extreme in that the worst possible outcome does no longer occur with positive probability if investment is sufficiently high, $p_1(\breve{I}) = 0$, such that the associated likelihood ratio goes to infinity in absolute terms.\footnote{Note that Assumption B.1 (ii) is rather weak for high $M$, in that it does not restrict any moment of the outcome distribution other than the probabilities of the 2 most extreme investment outcomes.} This effectively ensures that, as $I$ approaches $\breve{I}$, the informativeness of the worst possible investment outcome eventually dominates that of all favorable outcomes such that incentives are optimally provided exclusively via punishment for bad performance. Accordingly, overinvestment obtains for $I_{FB}$ sufficiently close to $\breve{I}$,\footnote{Clearly, whether the informativeness of one or multiple low outcomes (with $p'_i(I) < 0$) grows large as $I \to \breve{I}$ does not matter for the argument. We chose the condition in Assumption B.1 part (i) to illustrate that it is sufficient for one low outcome to become highly informative. Under MLRP, i.e., Assumption B.1 part (i), this is always the worst possible outcome, but for the overinvestment result to hold it could be any $i$ with $p'_i(I) < 0$.} While this limit argument is needed to obtain analytical results,\footnote{The reason is that the cost of providing incentives as captured by the curvature of the principal’s value function changes with the agent’s continuation value. This together with the fact that a closed-form solution to the principal’s value function does not exist already in much simpler settings without delegated investment, and, thus, has to be characterized implicitly based on the solution to a boundary} it is easy to show numerically that overinvestment for highly profitable investment...
technologies holds more generally. Intuitively, it is just needed that bad investment outcomes are sufficiently informative relative to good investment outcomes for high $I_{FB}$, such that incentives are mainly provided via punishment for poor performance. This is already apparent from our analysis with two investment outcomes $M = 2$ in the main text. In particular, as we established there in Example 1, overinvestment is not restricted to the limit as $p_1(I_{FB}) \to 0$, but instead holds as long as the low state is sufficiently unlikely given first-best investment, i.e., for all $I_{FB}$ with $p(I_{FB}) > 0.74$. That is, for overinvestment to occur, the probability of failure has to fall below 0.26.

This last observation points at a limit to the result that firms with high (low) returns to investment overinvest (underinvest) in intangible capital (see Proposition B.2). That is, the underlying, economically robust effect that higher investment expenditures shift probability mass from low to high investment outcomes, may not be strong enough to warrant a sufficiently asymmetric provision of investment incentives. In the following, we formalize this intuition and quantify a boundary to our result by imposing a lower bound on the probabilities of the most extreme investment outcomes which, in turn, restricts their relative informativeness. For simplicity, we assume that the problem is symmetric in the sense that $|p_i'(I)| = p'(I) \forall i \in \{1, \ldots, M\}$.  

**Example B.1.** Assume that $M = 4$ and the respective probabilities $p_i(I_i) := \Pr(\mu_{i+} = \mu^i|I_i)$ for $i = 1, \ldots, 4$ are, for $0 \leq I \leq \bar{I} = 1$, given by

\[
\begin{align*}
p_1(I) &= (a + \phi) - \phi \sqrt{I} \\
p_2(I) &= (a + \phi + x) - \phi \sqrt{I} \\
p_3(I) &= (b - \phi + x) + \phi \sqrt{I} \\
p_4(I) &= (b - \phi) + \phi \sqrt{I}
\end{align*}
\]

---

13 That is, we consider a violation of part (ii) of Assumption B.1, which captures the economically robust effect underlying our results. On the other hand, we refrain from relaxing part (i), which would most likely not yield any interesting insights, or part (iii), which is crucial to keep the problem tractable.

14 Hence, the relative informativeness across states, as captured by the respective likelihood ratios $|p'_i|/p_i$, only depends on $p_i$.  

---

11
implying that $p_1(I) = a \geq 0$ and $p_4(0) = b \geq 0$. We further set $x = 1/2 - (a + b)$, ensuring that $\sum_{i=1}^{4} p_i(I) = 1$, and $a + b < 1/2$ for strict MLRP to hold. We stipulate that $\phi = 1/6$ and all other parameters are as in the baseline specification in Example 1 in the main text. Then the following statements hold whenever the limited liability constraint is slack:

(i) If part (ii) of Assumption B.1 is satisfied, i.e., $a = b = 0$, investment is distorted upwards (downwards) when $I_{FB}$ is higher (lower) than the threshold 0.32.

(ii) Assume $b = 0$. Then, if $a > 0.18$, investment is distorted downwards for all $I_{FB} > 0$.

(iii) Assume $a = 0$. Then, if $b > 0.23$, investment is distorted upwards for all $I_{FB} < 1$.

Example B.1 nicely illustrates the robust intuition underlying our predictions on investment distortions. Investment is distorted upwards (downwards) if optimal incentives are predominately provided via costly punishment following unfavorable investment outcomes (reward following favorable investment outcomes). The choice of how to optimally provide incentives is in turn driven by the relative informativeness of favorable and unfavorable investment outcomes as captured by the respective likelihood ratios. Part (i) of Example B.1 corresponds to the benchmark case in which Assumption B.1 (ii) is satisfied and, thus, the worst (best) investment outcome will become extremely informative under sufficiently high (low) investment levels. Accordingly, if the returns to investment are sufficiently high (low), investment will be distorted above (below) first-best. Next, part (ii) of Example B.1 shows that if the probability of the worst investment outcome under the maximum possible investment level, $p_1(I)$, is greater than 0.18, overinvestment will never arise under the optimal contract, no matter how profitable the investment technology is. The reason is that with $p_1(I) > 0.18$, the relative informativeness of the lowest state (as captured by its likelihood ratio relative to that of the highest state), remains sufficiently low, such that it is optimal to provide costly incentives mainly via rewards for favorable outcomes – even at very high levels of investment expenditures. This pushes the principal towards underinvestment in order to reduce the probability of favorable states and the associated incentive costs. Finally, part (iii) of Example B.1, considers the opposite case, in which the highest possible investment outcome occurs with a strictly positive probability even
if the firm does not invest at all, i.e., \( p_M(0) > 0.23 \). In this case, the informativeness of favorable investment outcomes remains so low (relative to unfavorable outcomes), that incentives are mainly provided by costly punishment, even at low levels of investment. As a result, it is always (i.e., independent of the profitability of the investment technology) optimal to distort investment above first-best (as long as the limited liability constraint is slack).

C 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

\[
dD_t = [\mu_t - I_t - (\gamma - r) M_t] dt, \tag{C.1}
\]

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 (C.1), 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.\(^{15}\) Derivatives are

\(^{15}\)Note that we stipulate that the investment outcome is a verifiable event. In our interpretation of investment into absorptive capacity (see footnote 10 in the main text) 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
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 \( dS^j_t = dN^j_t - P^j_t dt \), 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 \( n^j_t \), and require the firm to hold at any time a portfolio of size

\[
n^j_t = \frac{\beta^j (\lambda M_t)}{\lambda}, j \in \{g, b\}, \quad (C.2)
\]

of the respective security.\(^{16}\) 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 \( \lambda \) of this special dividend.

Overall, the firm’s cash reserves, thus, follow

\[
dM_t = rM_t dt + dY_t + n^g_t dS^g_t + n^b_t dS^b_t - dD_t - dX_t. \quad (C.3)
\]

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

\[
P(M_t, \mu_t) = E_t \left[ \int_t^\tau e^{-r(s-t)} (dD_s + (1 - \lambda) dX_s) + e^{-r(\tau-t)} L_\tau \right], \quad (C.4)
\]

and we have the following result:

**Proposition C.1.** 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 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 downsizing covenant.

\(^{16}\)Note from \( \beta^b < 0 \) (see (19)), the firm holds a short position in security \( b \), i.e., \( n^b_t < 0 \).
holds a derivative security portfolio \( n_v^g, n_v^b \), 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 the threshold \( M_t^{pb} := \bar{w}/\lambda \), 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 C.1.** Under the proposed implementation, cash reserves evolve according to

\[
\begin{align*}
    dM_t &= \gamma M_t dt + \left( d\tilde{Y}_t - (\mu_t - I_t) dt \right) + \frac{\beta^g (\lambda M_t)}{\lambda} (dN_t^g - \nu p (I_t) dt) \\
    &\quad + \frac{\beta^b (\lambda M_t)}{\lambda} (dN_t^b - \nu (1 - p (I_t)) dt) - dX_t.
\end{align*}
\]

Now, define \( w_t = \lambda M_t \) to get

\[
\begin{align*}
    dw_t &= \lambda dM_t = \gamma w_t dt + \lambda \left( d\tilde{Y}_t - (\mu_t - I_t) dt \right) + \beta^g (w_t) (dN_t^g - \nu p (I_t) dt) \\
    &\quad + \beta^b (w_t) (dN_t^b - \nu (1 - p (I_t)) dt) - \lambda dX_t.
\end{align*}
\]

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 (C.3), (C.4) can be written as

\[
\begin{align*}
P(M_t, \mu_t) &= E_t \left[ \int_t^\tau e^{-r(s-t)} (\mu_s - I_s) ds - dU_s \right] + e^{-r(\tau-t)} L_\tau \\
&\quad + E_t \left[ \int_t^\tau e^{-r(s-t)} (rM_s ds - dM_s) \right] \\
&= f(W_t, \mu_t) + M_t = f^i(\lambda M_t) + M_t,
\end{align*}
\]
where we have used integration by parts. Q.E.D.

Note that in the implementation given in Proposition C.1, 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.

D 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 $f(0)$, the replacement value $L_r(0) = \max_w f(0)(w) - k$, and the free boundary $\bar{w}(0)$ 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 $f(0)$, $L_r(0)$, $\bar{w}(0)$, and $\beta^b(0)$, update the optimal investment scheme $I(1)$ according to (16) and subject to incentive compatibility, i.e., $\beta^g(I(1)) = \beta^b(0) + \lambda / (\nu p'(I(1)))$.

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

4. Given $f(1)$, $L_r(1)$, $\bar{w}(1)$, 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 $\beta^{(1)}$, such that $\beta^{(1)} = \beta^{(1)} + \lambda / (\nu p(I^{(1)}))$ and $\beta^{(1)} \geq w$.

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

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