Dynamic Agency and the q Theory of Investment

PETER M. DEMARZO, MICHAEL J. FISHMAN, ZHIGUO HE, and NENG WANG*

ABSTRACT

We develop an analytically tractable model integrating dynamic investment theory with dynamic optimal incentive contracting, thereby endogenizing financing constraints. Incentive contracting generates a history-dependent wedge between marginal and average q, and both vary over time as good (bad) performance relaxes (tightens) financing constraints. Financial slack, not cash flow, is the appropriate proxy for financing constraints. Investment decreases with idiosyncratic risk, and is positively correlated with past profits, past investment, and managerial compensation even with time-invariant investment opportunities. Optimal contracting involves deferred compensation, possible termination, and compensation that depends on exogenous observable persistent profitability shocks, effectively paying managers for luck.

THE EFFICIENCY OF CORPORATE investment decisions can be compromised by frictions in external financing. One important source of financial market frictions involves agency problems. Firms do not have access to as much capital as they might like, or at low enough cost, because outside investors are wary of managers' incentives to act in their own private interest. In this paper, we examine the implications of agency problems for the dynamics of firms' investment decisions and firm value.

We start with a standard dynamic model of corporate investment, the q theory of investment (see Hayashi (1982)). In the absence of fixed investment costs and no financial market frictions, the firm optimally chooses investment to equate the marginal value of capital with the marginal cost of capital (including adjustment costs). With a homogeneous production technology, the marginal value of capital, that is, marginal q, equals the average value of capital, that

is, average q.[1] This result motivates the widespread use of average q (which is relatively easy to measure) as an empirical proxy for marginal q (which is relatively difficult to measure). To this model, we introduce an agency problem. Following DeMarzo and Sannikov (2006), an agent (firm management) must be continually provided with the incentive to choose the appropriate action. The agency model matches a standard principal-agent setting in which the agent's action is unobserved costly effort, and this effort affects the mean rate of production. Alternatively, we can interpret the agency problem as one in which the agent can divert output for his private benefit. The presence of the agency problem will limit the firm's investment. Our model endogenizes the costs of external financing.

The optimal contract between investors and the agent minimizes the cost of the agency problem and has implications for the dynamics of investment and firm value. For instance, incentive contracting creates a wedge between average and marginal q that varies with firm performance. Consequently, the measurement error inherent in using average q as a proxy for marginal q will vary both over time for a given firm and across firms. The continuous-time formulation allows for a relatively simple characterization of this relation between marginal and average q. Among the predictions of the analysis, investment is positively correlated with profits, past investment, managerial compensation, and financial slack even with time-invariant investment opportunities. Despite risk-neutral managers and investors, investment decreases with firm-specific risk. More broadly, our theory suggests that financial slack, not cash flow, is the important predictor of investment after controlling for average q, thus challenging the empirical validity of using cash flow as a proxy for financial constraints as is common in the investment/cash flow sensitivity literature.[2] Optimal incentive contracting involves deferred compensation; possible termination; and compensation that depends on observable persistent profitability shocks that are beyond managerial control, effectively paying managers for luck.

The optimal incentive contract specifies, as a function of the history of the firm's profits, (i) the agent's compensation, (ii) the level of investment in the firm, and (iii) whether the contract is terminated. Termination could involve the replacement of the agent or the liquidation of the firm. Going forward, we use the terms termination and liquidation interchangeably. Through the contract, the firm's profit history determines the agent's current discounted

---

[1] Lucas and Prescott (1971) analyze dynamic investment decisions with convex adjustment costs, though they do not explicitly link their results to marginal or average q. Abel and Eberly (1994) extend Hayashi (1982) to a stochastic environment and a more general specification of adjustment costs.

[2] Fazzari, Hubbard, and Petersen (1988) (FHP) are the first to use the sensitivity of investment to cash flow (controlling for q) as a measure of a firm's financial constraints. Their logic is that the more financially constrained is a firm, the more investment will be dictated by current cash flow. A large literature follows the FHP approach. Kaplan and Zingales (1997) (KZ) provide an important critique on FHP and successors from both a theoretical (using a static model) and an empirical perspective. Much research on financial constraints has followed since the FHP–KZ debate.

expected payoff, which we refer to as the agent's "continuation payoff," W, and current investment, which in turn determines the current capital stock, K. These two state variables, W and K, completely summarize the contract-relevant history of the firm. Moreover, because of the size-homogeneity of our model, the analysis simplifies further and the agent's continuation payoff per unit of capital, $w = W/K$, becomes sufficient for the contract-relevant history of the firm.[3]

Because of the agency problem, investment is below the first-best level. The degree of underinvestment depends on the firm's realized past profitability, or equivalently, through the contract, the agent's continuation payoff (per unit of capital), w. In particular, investment is increasing in w. To understand this linkage, note that, in a dynamic agency setting, the agent is rewarded for high profits, and penalized for low profits, in order to provide incentives. As a result, the agent's continuation payoff, w, is increasing with past profitability. In turn, a higher continuation payoff for the agent relaxes the agent's incentive compatibility constraints since the agent now has a greater stake in the firm (in the extreme, if the agent owned the entire firm there would be no agency problem). Finally, relaxing the incentive compatibility constraints raises the value of investing in more capital.

In the analysis here, the gain from relaxing the incentive compatibility constraints comes by reducing the probability, within any given amount of time, of termination. If profits are low, the agent's continuation payoff w falls (for incentive reasons) and if w hits a lower threshold, the contract is terminated. We assume termination entails costs associated with hiring a new manager or liquidating assets, and show that even if these costs appear small they can have a large impact on the optimal contract and investment.

We also show that in an optimal contract the agent's payoff depends on persistent shocks to the firm's profitability even if these shocks are observable, contractible, and beyond the agent's control. When an exogenous shock increases the firm's profitability, the contract gives the agent a higher continuation payoff. The intuition is that the marginal cost of compensating the agent is lower when profitability is high because relaxing the agency problem is more valuable when profitability is high. This result may help to explain the empirical importance of absolute, rather than relative, performance measures for executive compensation. This result also implies that a profitability increase has both a direct effect on investment, as higher profitability makes investment more profitable, and an indirect effect, since with higher profitability it is optimal to offer the agent a higher continuation payoff that, as discussed earlier, leads to further investment.

As in DeMarzo and Fishman (2007a,b) and DeMarzo and Sannikov (2006), we show that the state variable, w, which represents the agent's continuation

---

[3] We solve for the optimal contract using a recursive dynamic programming approach. Early contributions that developed recursive formulations of the contracting problem include Green (1987), Spear and Srivastava (1987), Phelan and Townsend (1991), and Atkeson (1991), among others. Ljungqvist and Sargent (2004) provide in-depth coverage of these models in discrete-time settings.

payoff, can also be interpreted as a measure of the firm's financial slack. More precisely, w is proportional to the size of the current cash flow shock that the firm can sustain without liquidating, and so can be interpreted as a measure of the firm's liquid reserves and available credit. The firm accumulates reserves when profits are high, and depletes its reserves when profits are low. Thus, our model predicts an increasing relation between the firm's financial slack and the level of investment.

The agency perspective leads to important departures from standard q theory. First, we demonstrate that both average q and marginal q are increasing with the agent's continuation payoff, w, and therefore with the firm's financial slack and past profitability. This effect is driven by the nature of optimal contracts, as opposed to changes in the firm's investment opportunities. Second, we show that, despite the homogeneity of the firm's production technology (including agency costs), average q and marginal q are no longer equal. Marginal q is below average q because an increase in the firm's capital stock reduces the firm's financial slack (the agent's continuation payoff) per unit of capital, w, and thus tightens the incentive compatibility constraints and raises agency costs. The wedge between marginal and average q is largest for firms with intermediate profit histories. Very profitable firms have sufficient financial slack that agency costs are small, whereas firms with very poor profits are more likely to be liquidated (in which case average and marginal q coincide). These results imply that, in the presence of agency concerns, standard linear models of investment on average q are misspecified, and variables such as managerial compensation, financial slack, past profitability, and past investment will be useful predictors of current investment.

Related analyses of agency, dynamic contracting, and investment include Albuquerque and Hopenhayn (2004), Quadrini (2004), Clementi and Hopenhayn (2006), DeMarzo and Fishman (2007a), and Biais et al. (2010). Philippon and Sannikov (2007) analyze the optimal exercise of a growth option in a dynamic agency environment. Rampini and Viswanathan (2010, 2011) develop dynamic models of investment and capital structure with collateral constraints due to limited enforcement and explore leverage choices, the lease versus buy decision, and risk management.[4] We go beyond these analyses by providing a closer link to the theoretical and empirical investment literature. Specifically, we explore the dynamic relation between firm value, marginal q, average q, investment, and financial slack.

With discrete-time models, Lorenzoni and Walentin (2007) and Schmid (2008) also analyze the implications of agency problems for the q theory of investment. The key methodological difference is that we use the continuous-time recursive contracting methodology developed in DeMarzo and Sannikov (2006) to derive the optimal contract. This allows for a relatively simple closed-form characterization of the investment Euler equation, optimal investment dynamics, and

---

[4] In addition, our analysis owes much to the recent dynamic contracting literature, for example, Gromb (1999), Biais et al. (2007), DeMarzo and Fishman (2007b), Tchistyi (2005), Sannikov (2007), He (2009), and Piskorski and Tchistyi (2010), as well as the earlier optimal contracting literature, for example, Diamond (1984) and Bolton and Scharfstein (1990).

compensation policies. Another modeling difference is that, in Lorenzoni and Walentin (2007) and Schmid (2008), the agent must be given incentives not to default and abscond with the assets, and whether he complies is observable. This implies that, in equilibrium, the agent is never terminated. By contrast, in our analysis, whether the agent takes appropriate actions is unobservable and consequently termination does occur in equilibrium.

A growing literature in finance and macroeconomics incorporates exogenous financing frictions in the form of transaction costs of raising funds. See, for example, Kaplan and Zingales (1997), Gilchrist and Himmelberg (1998), Gomes (2001), Hennessy and Whited (2007), and Bolton, Chen, and Wang (2011), among others. This literature motivates exogenously specified financing costs with arguments based on agency problems and/or information asymmetries. In our analysis, the financing frictions stem from agency problems and are endogenously derived.

We proceed as follows. In Section I, we specify our continuous-time model of investment in the presence of agency costs. In Section II, we solve for the optimal contract using dynamic programming. Section III analyzes the implications of this optimal contract for investment and firm value. Section IV provides an implementation of the optimal contract using standard securities and explores the link between financial slack and investment. In Section V, we consider the impact of observable persistent profitability shocks on investment, firm value, and the agent's compensation. Section VI concludes. All proofs appear in the Appendix.

I. The Model

We formulate an optimal dynamic investment problem for a firm facing an agency problem. First, we present the firm's production technology. Second, we introduce the agency problem between investors and the agent. Finally, we formulate the optimal contracting problem.

A. *Firm's Production Technology*

Our model is based on a neoclassical investment setting. The firm employs capital to produce output, whose price is normalized to one (Section V considers stochastic profitability shocks). Let K and I denote the level of capital stock and gross investment rate, respectively. As is standard in capital accumulation models, the firm's capital stock K evolves according to

$$dK_t = (I_t - \delta K_t)dt, \quad t \geq 0, \tag{1}$$

where $\delta \geq 0$ is the rate of depreciation.

Investment entails adjustment costs. Following the neoclassical investment with adjustment costs literature, we assume that the adjustment cost $G(I, K)$ satisfies $G(0, K) = 0$, is smooth and convex in investment I, and is homogeneous of degree one in I and the capital stock K. Given the homogeneity of the

adjustment costs, we can write

$$I + G(I, K) \equiv c(i)K, \tag{2}$$

where the convex function c represents the total cost per unit of capital required for the firm to grow at rate $i = I/K$ (before depreciation).

We assume that the incremental gross output over time interval dt is proportional to the capital stock, and so can be represented as $K_t dA_t$, where A is the cumulative productivity process.[5] We model the instantaneous productivity dA_t in the next subsection, where we introduce the agency problem. Given the firm's linear production technology, after accounting for investment and adjustment costs we can write the dynamics of the firm's cumulative (gross of agent compensation) cash flow process Y_t for $t \geq 0$ as follows:

$$dY_t = K_t(dA_t - c(i_t)\,dt), \tag{3}$$

where $K_t dA_t$ is the incremental gross output and $K_t c(i_t)dt$ is the total cost of investment.

The contract with the agent can be terminated at any time, in which case investors recover a value lK_t, where $l \geq 0$ is a constant. We assume that termination is inefficient and generates deadweight losses. We can interpret termination as the liquidation of the firm; alternatively, in Section II, we show how l can be endogenously determined to correspond to the value that shareholders can obtain by replacing the incumbent management (see DeMarzo and Fishman (2007b) for additional interpretations). Since the firm could always liquidate by disinvesting, it is natural to specify $l \geq c'(-\infty)$.

B. The Agency Problem

We now introduce an agency conflict induced by the separation of ownership and control. The firm's investors hire an agent to operate the firm. In contrast to the neoclassical model in which the productivity process A is exogenously specified, the productivity process in our model is affected by the agent's unobservable action. Specifically, the agent's action $a_t \in [0, 1]$ determines the expected rate of output per unit of capital, so that

$$dA_t = a_t \mu dt + \sigma dZ_t, \quad t \geq 0, \tag{4}$$

where $Z = \{Z_t, \mathcal{F}_t; 0 \leq t < \infty\}$ is a standard Brownian motion on a complete probability space, and $\sigma > 0$ is the constant volatility of the cumulative

---

[5] We can interpret this linear production function as a reduced form for a setting with constant returns to scale involving other factors of production. For instance, suppose the firm has a Cobb–Douglas production function with capital and labor and both productivity z_t and labor wage ω_t shocks are i.i.d. random variables. For a given amount of capital, and with fully and instantaneously adjustable labor N, it is optimal for the firm to solve the following static problem: $\max_N \mathbb{E}(z_t K_t^\alpha N^{1-\alpha} - \omega_t N)$. This yields optimal labor demand N^* proportional to capital. Using the optimal N^*, we obtain the realized revenue net of labor cost $K_t f(z_t, \omega_t)$. The productivity shock dA_t corresponds to $f(z_t, \omega_t)$.

productivity process *A*. The agent controls the drift, but not the volatility of the process *A*. Note that the firm can incur operating losses. While these losses can accrue at an unbounded rate given the Brownian motion, we will show that the optimal contract with agency bounds cumulative losses of the firm by optimally invoking termination.

When the agent takes the action a_t, he enjoys private benefits at the rate $\lambda(1 - a_t)\mu dt$ per unit of the capital stock, where $0 \leq \lambda \leq 1$. The action can be interpreted as an effort choice; due to the linearity of private benefits, our framework is also equivalent to the binary effort setup in which the agent can *shirk*, $a = 0$, or *work*, $a = 1$. Alternatively, we can interpret $1 - a_t$ as the fraction of cash flow that the agent diverts for his private benefit, with λ equal to the agent's net consumption per dollar diverted. In either case, λ represents the severity of the agency problem and, as we show later, captures the minimum level of incentives required to motivate the agent.

Investors have unlimited wealth and are risk-neutral with discount rate $r > 0$. The agent is also risk-neutral, but with a higher discount rate $\gamma > r$. That is, we make the common assumption that the agent is impatient relative to investors. This impatience could be preference based or could arise indirectly because the agent has other attractive investment opportunities. The impatience assumption avoids the scenario in which investors indefinitely postpone payments to the agent. The agent has no initial wealth and has limited liability, so investors cannot pay negative wages to the agent. If the contract is terminated, the agent's reservation value, which is associated with his next-best employment opportunity, is normalized to zero.

C. Formulating the Optimal Contracting Problem

We assume that the firm's capital stock, K_t, and its (cumulative) cash flow, Y_t, are observable and contractible. Therefore, investment I_t and productivity A_t are also contractible.[6] To maximize firm value, investors offer a contract that specifies the firm's investment policy I_t, the agent's cumulative compensation U_t, and a termination time τ, all of which depend on the history of the agent's performance, which is given by the productivity process A_t.[7] The agent's limited liability requires the compensation process U_t to be nondecreasing. We let $\Phi = (I, U, \tau)$ represent the contract and leave further regularity conditions on Φ to Appendix A.

---

[6] Based on the growth of the firm's capital stock, the firm's investment process can be deduced from (1), and hence the firm's productivity process A_t can be deduced from (3) using I_t and Y_t.

[7] As we will discuss further in Section V, the firm's access to capital is implicitly determined given the investment, compensation, and liquidation policies. Note also that, given A and the investment policy, the variables K and Y are redundant and so we do not need to contract on them directly. In principle, the contract could also allow for randomized payoffs as well as investment and termination decisions. But, as we will verify later, the optimal contract with commitment does not entail randomization. The optimal contract without commitment (that is, the optimal renegotiation-proof contract) may rely on randomization; see Appendix C.

Given the contract Φ, the agent chooses an action process $\{a_t \in [0,1] : 0 \le t < \tau\}$ to solve

$$W(\Phi) = \max_{\{a_t \in [0,1]: 0 \le t < \tau\}} \mathbb{E}^a \left[\int_0^\tau e^{-\gamma t}(dU_t + \lambda(1-a_t)\mu K_t dt) \right], \quad (5)$$

where $\mathbb{E}^a(\cdot)$ is the expectation operator under the probability measure that is induced by the action process. The agent's objective function includes the present discounted value of compensation (the first term in (5)) and the potential private benefits from taking action $a_t < 1$ (the second term in (5)).

We focus on the case in which it is optimal for investors to implement the efficient action $a_t = 1$ all the time and provide a sufficient condition for the optimality of implementing this action in Appendix A. Henceforth, the expectation operator $\mathbb{E}(\cdot)$ is under the measure induced by $\{a_t = 1 : 0 \le t < \tau\}$, unless otherwise stated. We call a contract Φ *incentive compatible* if it implements the efficient action.

At the time the contract is initiated, the firm has K_0 in capital. Given an initial payoff of W_0 for the agent, the investors' optimization problem is

$$P(K_0, W_0) = \max_\Phi \mathbb{E} \left[\int_0^\tau e^{-rt} dY_t + e^{-r\tau} l K_\tau - \int_0^\tau e^{-rt} dU_t \right]$$
$$\text{s.t. } \Phi \text{ is incentive compatible and } W(\Phi) = W_0. \quad (6)$$

The investors' objective is to maximize the expected present value of the firm's gross cash flow plus termination value less the agent's compensation. The agent's expected payoff, W_0, will be determined by the relative bargaining power of the agent and investors when the contract is initiated. For example, if investors have all the bargaining power, then $W_0 = \arg\max_{W \ge 0} P(K_0, W)$, whereas if the agent has all the bargaining power, then $W_0 = \max\{W : P(K_0, W) \ge 0\}$. More generally, by varying W_0 we can determine the entire feasible contract curve.

II. Model Solution

We begin by determining optimal investment in the standard neoclassical setting without an agency problem. We then characterize the optimal contract with agency concerns.

A. A Neoclassical Benchmark

With no agency conflicts—corresponding to $\lambda = 0$, in which case there is no benefit from shirking, and/or $\sigma = 0$, in which case there is no noise to hide the agent's action—our model specializes to the neoclassical setting of Hayashi (1982), a widely used benchmark in the investment literature. Given the stationarity of the economic environment and the homogeneity of the production technology, there is an optimal investment-capital ratio that maximizes the present value of the firm's cash flows. Because of the homogeneity assumption,

we can equivalently maximize the present value of the cash flows *per unit of capital*. In other words, we have the Hayashi (1982) result that the marginal value of capital (marginal q) equals the average value of capital (average or Tobin's q), both of which are given by

$$q^{FB} = \max_i \ \frac{\mu - c(i)}{r + \delta - i}. \tag{7}$$

That is, a unit of capital is worth the perpetuity value of its expected free cash flow (expected output less investment and adjustment costs) given the firm's net growth rate $i - \delta$. To ensure that the first-best value of the firm is well defined, we impose the parameter restriction

$$\mu < c(r + \delta). \tag{8}$$

Inequality (8) implies that the firm cannot profitably grow faster than the discount rate. We also assume throughout the paper that the firm is sufficiently productive that termination/liquidation is not efficient, that is, $q^{FB} > l$.

From the first-order condition for (7), first-best investment is characterized by

$$c'(i^{FB}) = q^{FB} = \frac{\mu - c(i^{FB})}{r + \delta - i^{FB}}. \tag{9}$$

Because adjustment costs are convex, (9) implies that first-best investment is increasing with q. Adjustment costs create a wedge between the value of installed capital and newly purchased capital, in that $q^{FB} \neq 1$ in general. Intuitively, when the firm is sufficiently productive that investment has positive net present value (NPV), that is $\mu > (r + \delta)c'(0)$, investment is positive and $q^{FB} > 1$. In the special case of quadratic adjustment costs,

$$c(i) = i + \tfrac{1}{2}\theta i^2, \tag{10}$$

we have the explicit solution

$$q^{FB} = 1 + \theta i^{FB} \ \text{ and } \ i^{FB} = r + \delta - \sqrt{(r + \delta)^2 - 2\frac{\mu - (r + \delta)}{\theta}}.$$

Note that q^{FB} represents the value of the firm's cash flows (per unit of capital) prior to compensating the agent. If investors promise the agent a payoff W in present value, then absent an agency problem the agent's relative impatience ($\gamma > r$) implies that it is optimal to pay the agent W in cash immediately. Thus, the investors' payoff is given by

$$P^{FB}(K, W) = q^{FB}K - W.$$

Equivalently, we can express the agent's and investors' payoff on a per unit of capital basis, as $w = W/K$ and

$$p^{FB}(w) = P^{FB}(K, W)/K = q^{FB} - w.$$

In the neoclassical setting, the time-invariance of the firm's technology implies that the first-best investment is constant over time, and independent of the firm's history or the volatility of its cash flows. As we will explore next, agency concerns significantly alter these conclusions.

B. The Optimal Contract with Agency

We now solve for the optimal contract when there is an agency problem, that is, when $\lambda\sigma > 0$. Recall that the contract specifies the firm's investment policy I, payments to the agent U, and a termination date τ all as functions of the firm's profit history. The contract must be incentive compatible (that is, induce the agent to choose $a_t = 1$ for all t) and maximize investors' value function $P(K, W)$. Here, we outline the intuition for the derivation of the optimal contract, leaving formal details to Appendix A.

Given an incentive-compatible contract Φ and the history up to time t, the discounted expected value of the agent's future compensation is given by

$$W_t(\Phi) \equiv \mathbb{E}_t\left[\int_t^\tau e^{-\gamma(s-t)}dU_s\right]. \tag{11}$$

We call W_t the agent's continuation payoff as of date t.

The agent's incremental compensation at date t is composed of a cash payment dU_t and a change in the value of his promised future payments, captured by dW_t. To compensate for the agent's time preference, this incremental compensation must equal $\gamma W_t dt$ on average. Thus,

$$\mathbb{E}_t(dW_t + dU_t) = \gamma W_t dt. \tag{12}$$

While (12) reflects the agent's average compensation, to maintain incentive compatibility his compensation must be sufficiently sensitive to the firm's incremental output $K_t dA_t$. Adjusting output by its mean and using the martingale representation theorem (details are provided in Appendix A), we can express this sensitivity for any incentive-compatible contract as follows:[8]

$$dW_t + dU_t = \gamma W_t dt + \beta_t K_t(dA_t - \mu dt) = \gamma W_t dt + \beta_t K_t \sigma dZ_t. \tag{13}$$

To understand the determinants of the incentive coefficient β_t, suppose the agent deviates and chooses $a_t < 1$. The instantaneous cost to the agent is the expected reduction of his compensation, given by $\beta_t(1 - a_t)\mu K_t dt$, and the instantaneous private benefit is $\lambda(1 - a_t)\mu K_t dt$. Thus, to induce the agent to choose $a_t = 1$, incentive compatibility is equivalent to

$$\beta_t \geq \lambda \quad \text{for all } t.$$

Intuitively, incentive compatibility requires that the agent have sufficient exposure to the firm's realized output, as otherwise it would be profitable for the

---

[8] Intuitively, the linear form of the contract's sensitivity can be understood in terms of a binomial tree, where any function admits a state-by-state linear representation.

agent to reduce output and consume private benefits. We will further show that this incentive compatibility constraint binds. That is, the agent will face the minimum exposure that provides the incentive to choose the appropriate action ($a_t = 1$). This result follows because there is a cost to having the agent bear risk. Unlucky realizations of the productivity shocks dZ_t can reduce the agent's continuation payoff to zero and, given the agent's limited liability ($W_t \geq 0$), require termination of the contract, which is costly to investors. An optimal contract will therefore set the agent's sensitivity to $\beta_t = \lambda$ to reduce the cost of liquidation while maintaining incentive compatibility. Intuitively, incentive provision is necessary, but costly due to the reliance on the threat of ex post inefficient liquidation. Hence, the optimal contract requires the minimal necessary level of incentive provision.

Whatever the history of the firm up to date t, the only relevant state variables going forward are the firm's capital stock K_t and the agent's continuation payoff W_t. Therefore, the payoff to investors in an optimal contract after such a history is given by the value function $P(K_t, W_t)$, which we can solve for using dynamic programming techniques. As in the earlier analysis of the first-best setting, we use the scale invariance of the firm's technology to write $P(K, W) = p(w)K$ and reduce the problem to one with a single state variable $w = W/K$.

We begin with a number of key properties of the value function $p(w)$. Clearly, the value function cannot exceed the first-best, so $p(w) \leq p^{FB}(w)$. Also, as noted earlier, to deliver a payoff to the agent equal to his outside opportunity (normalized to zero), we must terminate the contract immediately as otherwise the agent could consume private benefits. Therefore,

$$p(0) = l. \tag{14}$$

Next, because investors can always compensate the agent with cash, it will cost investors at most \$1 to increase w by \$1. Therefore, $p'(w) \geq -1$, which implies that the total value of the firm, $p(w) + w$, is weakly increasing with w. In fact, when w is low, firm value will strictly increase with w. Intuitively, a higher w—which amounts to a higher level of deferred compensation for the agent—reduces the probability of termination (within any given amount of time). This benefit declines as w increases and the probability of termination becomes small, suggesting that $p(w)$ is concave, a property we will assume for now and verify shortly.

Because there is a benefit of deferring the agent's compensation, the optimal contract will set cash compensation $du_t = dU_t/K_t$ to zero when w_t is small, so that (from (13)) w_t will rise as quickly as possible. However, because the agent has a higher discount rate than investors, $\gamma > r$, there is a cost of deferring the agent's compensation. This trade-off implies that there is a compensation level $\overline{w}$ such that it is optimal to pay the agent with cash if $w_t > \overline{w}$ and to defer compensation otherwise. Thus, we can set

$$du_t = \max\{w_t - \overline{w}, 0\}, \tag{15}$$

which implies that for $w_t > \overline{w}$, $p(w_t) = p(\overline{w}) - (w_t - \overline{w})$, and the compensation level $\overline{w}$ is the smallest agent continuation payoff with

$$p'(\overline{w}) = -1. \tag{16}$$

When $w_t \in [0, \overline{w}]$, the agent's compensation is deferred ($du_t = 0$). The evolution of $w = W/K$ follows directly from the evolutions of W (see (13)) and K (see (1)), noting that $dU_t = 0$ and $\beta_t = \lambda$,

$$dw_t = (\gamma - (i_t - \delta))w_t dt + \lambda(dA_t - \mu dt) = (\gamma - (i_t - \delta))w_t dt + \lambda\sigma dZ_t. \tag{17}$$

Equation (17) implies the following dynamics for the optimal contract. Based on the agent's and investors' relative bargaining power, the contract is initiated with some promised payoff per unit of capital, w_0, for the agent. This promise grows on average at rate γ less the net growth rate ($i_t - \delta$) of the firm. When the firm experiences a positive productivity shock, the promised payoff increases until it reaches the level $\overline{w}$, at which point the agent receives cash compensation. When the firm has a negative productivity shock, the promised payoff declines, and the contract is terminated when w_t falls to zero.

Having determined the dynamics of the agent's payoff, we can now use the Hamilton–Jacobi–Bellman (HJB) equation to characterize $p(w)$ for $w \in [0, \overline{w}]$

$$rp(w) = \sup_i(\mu - c(i)) + (i - \delta)p(w) + (\gamma - (i - \delta))wp'(w) + \tfrac{1}{2}\lambda^2\sigma^2 p''(w). \tag{18}$$

Intuitively, the right side is given by the sum of instantaneous expected cash flows (the first term in brackets), plus the expected change in the value of the firm due to capital accumulation (the second term), and the expected change in the value of the firm due to the drift and volatility (using Ito's lemma) of the agent's continuation payoff w (the remaining terms). Investment i is chosen to maximize investors' total expected cash flow plus "capital gains," which, given risk neutrality, must equal the expected return $rp(w)$.

Using the HJB equation (18), we have that the optimal investment-capital ratio $i(w)$ satisfies the following Euler equation:

$$c'(i(w)) = p(w) - wp'(w). \tag{19}$$

The above equation states that the marginal cost of investing equals the marginal value of investing from the investors' perspective. The marginal value of investing equals the current per unit value of the firm to investors, $p(w)$, plus the marginal effect of decreasing the agent's per unit payoff w as the firm grows.

Equations (18) and (19) jointly determine a second-order ordinary differential equation (ODE) for $p(w)$ in the region $w_t \in [0, \overline{w}]$. We also have the condition (14) for the liquidation boundary as well as the "smooth pasting" condition (16) for the endogenous payout boundary $\overline{w}$. To complete our characterization, we need a third condition to determine the optimal level of $\overline{w}$. The condition for

optimality is given by the "super contact" condition[9]

$$p''(\overline{w}) = 0. \tag{20}$$

We can provide some economic intuition for the super contact condition (20) by noting that, using (18) and (16), (20) is equivalent to

$$p(\overline{w}) + \overline{w} = \max_i \frac{\mu - c(i) - (\gamma - r)\overline{w}}{r + \delta - i}. \tag{21}$$

Equation (21) can be interpreted as a "steady-state" valuation constraint. The left side is total firm value at $\overline{w}$ whereas the right side is the perpetuity value of the firm's cash flows given the cost of maintaining the agent's continuation payoff at $\overline{w}$ (since $\gamma > r$, there is a cost to deferring the agent's compensation). Because $\overline{w}$ is a reflecting boundary, the value attained at this point should match this steady-state level as though we remained at $\overline{w}$ forever. If the value were below this level, it would be optimal to defer the agent's cash compensation and allow his continuation payoff to increase, that is, it would be optimal to increase $\overline{w}$ until (21) is satisfied; at that point the benefit of deferring compensation further is balanced by the cost due to the agent's impatience.

We now summarize our main results on the optimal contract in the following proposition.[10]

PROPOSITION 1: *The investors' value function $P(K, W)$ is proportional to capital stock K, in that $P(K, W) = p(w)K$, where $p(w)$ is the investors' scaled value function. For $w_t \in [0, \overline{w}]$, $p(w)$ is strictly concave and uniquely solves the ODE (18) with boundary conditions (14), (16), and (20). For $w > \overline{w}$, $p(w) = p(\overline{w}) - (w - \overline{w})$. The agent's scaled continuation payoff w evolves according to (17), for $w_t \in [0, \overline{w}]$. Cash payments $du_t = dU_t/K_t$ reflect w_t back to $\overline{w}$, and the contract is terminated at the first time τ such that $w_\tau = 0$. Optimal investment is given by $I_t = i(w_t)K_t$, where $i(w)$ is defined in (19).*

The termination value l could be exogenous, for example, the capital's salvage value in liquidation. Alternatively, l could be endogenous. For example, suppose termination involves firing and replacing the agent with a new (identical) agent. Then the investors' termination payoff equals the value obtained from hiring a new agent at an optimal initial continuation payoff w_0. That is,

$$l = \max_{w_0} (1 - \kappa)p(w_0), \tag{22}$$

where $\kappa \in [0, 1)$ reflects a cost of lost productivity if the agent is replaced.

---

[9] The super contact condition essentially requires that the second derivatives match at the boundary (see Dixit (1993)).

[10] We provide necessary technical conditions and present a formal verification argument for the optimal policy in Appendix A.
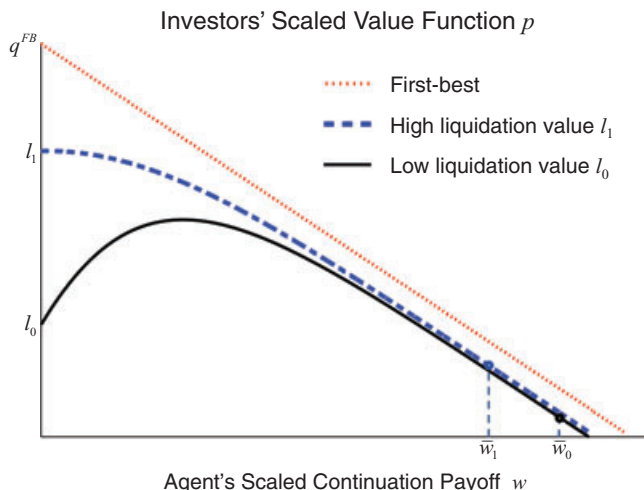
Figure 1. Investors' scaled value function $p(w)$ as a function of the agent's scaled continuation payoff w. We illustrate scenarios in which the liquidation value is high (l_1) and low (l_0).

III. Model Implications and Analysis

Having characterized the solution of the optimal contract, we first study some additional properties of $p(w)$ and then analyze the model's predictions for average q, marginal q, and investment.

A. Investors' Scaled Value Function

Using the optimal contract in Section II, we plot investors' scaled value function $p(w)$ in Figure 1 for two different termination values. The gap between $p(w)$ and the first-best value function reflects the loss due to agency conflicts. From Figure 1, we see that this loss is higher when the agent's payoff w is lower or when the termination value l is lower. Also, when the termination value is lower, the cash compensation boundary $\overline{w}$ is higher as it is optimal to defer compensation longer in order to reduce the probability of costly termination.

The concavity of $p(w)$ reveals investors' induced aversion to fluctuations in the agent's payoff. Intuitively, a mean-preserving spread in w is costly because it increases the risk of termination. Thus, although investors are risk-neutral, they behave in a risk-averse manner toward idiosyncratic risk due to the agency friction. This property fundamentally differentiates our agency model from the neoclassical Hayashi (1982) result where volatility has no effect on investment and firm value. The dependence of investment and firm value on idiosyncratic volatility in our model arises from investors' inability to distinguish the agent's actions from noise.

While $p(w)$ is concave, it need not be monotonic in w, as shown in Figure 1. The intuition is as follows. Two effects drive the shape of $p(w)$. First, as in the

first-best neoclassical benchmark of Section II.A, the higher the agent's claim w, the lower the investors' value $p(w)$, holding the total surplus fixed. This is just a wealth transfer effect. Second, increasing w allows the contract to provide incentives to the agent with a lower risk of termination. This "incentive alignment effect" creates wealth, raising the total surplus available for distribution to the agent and investors. As can be seen from the figure, the wealth transfer effect dominates when w is large, but the incentive alignment effect can dominate when w is low and termination is sufficiently costly.

If the liquidation value is sufficiently low that the value function p is non-monotonic, then, while termination is used to provide incentives ex ante, it is inefficient ex post. Inefficient termination provides room for renegotiation, since both parties will have incentives to renegotiate to a Pareto-improving allocation. Thus, the optimal contract depicted in Figure 1 is not renegotiation-proof with liquidation value l_0, whereas the contract is renegotiation-proof with liquidation value l_1. In Appendix C, we show that the main qualitative implications of our model are unchanged when contracts are constrained to be renegotiation-proof. Intuitively, renegotiation weakens incentives and has the same effect as increasing the value of the agent's outside option (which reduces investors' payoff).

Alternatively, if the agent can be fired and costlessly replaced, so that the liquidation value is endogenously determined as in (22) with $\kappa = 0$, then $p'(0) = 0$ and the optimal contract will be renegotiation-proof. We can also interpret the case with l_1 in Figure 1 in this way.

B. Average and Marginal q

Now we use the properties of $p(w)$ to derive implications for q. Total firm value, including the claim held by the agent, is $P(K, W) + W$. Therefore, average q, defined as the ratio between firm value and capital stock, is denoted by q_a and given by

$$q_a(w) = \frac{P(K, W) + W}{K} = p(w) + w. \tag{23}$$

This definition of average q is consistent with the definition of q in the first-best benchmark (Hayashi (1982)). Marginal q measures the incremental impact of a unit of capital on firm value. We denote marginal q as q_m and calculate it as

$$q_m(w) = \frac{\partial(P(K, W) + W)}{\partial K} = P_K(K, W) = p(w) - wp'(w). \tag{24}$$

While average q is often used in empirical studies due to the simplicity of its measurement, marginal q determines the firm's investment via the standard Euler equations (see (19)).

One of the most important and well-known results in Hayashi (1982) is that marginal q equals average q when the firm's production and investment technologies exhibit homogeneity as shown in our neoclassical benchmark case.
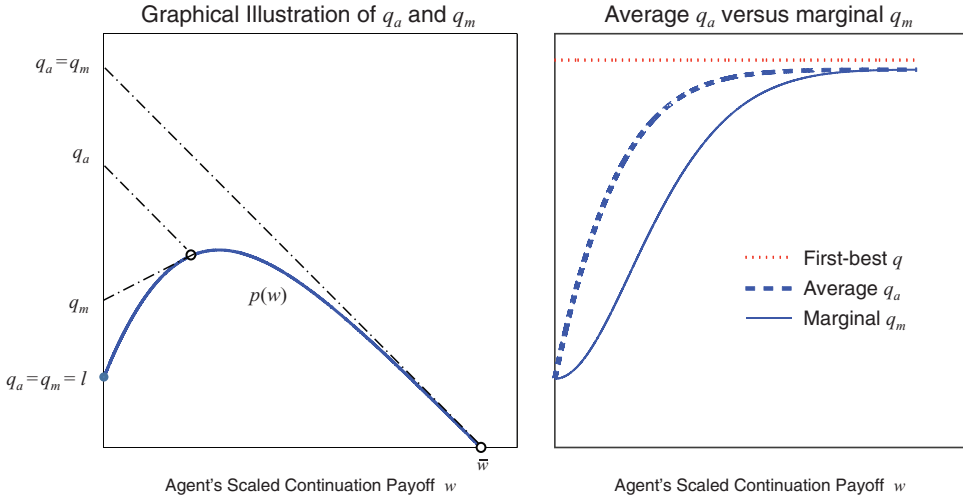
Figure 2. Average q_a **and marginal** q_m. The left panel shows a geometrical illustration of the determination of q_a and q_m. The right panel plots q_a and q_m with the first-best q^{FB}.

This result motivates the use of average q (which is relatively easy to measure) as a proxy for marginal q (which is harder to measure) in empirical investment studies. While our model also features these homogeneity properties, agency costs cause the marginal value of capital, q_m, to differ from the average value of the capital stock, q_a. In particular, comparing (7), (23), and (24) and using the fact that $p'(w) \geq -1$, we have the following inequality:

$$q^{FB} > q_a(w) \geq q_m(w). \tag{25}$$

The first inequality follows by comparing (21) and the calculation of q^{FB} in (7). Average q is above marginal q because, for a given level of W, an increase in capital stock K lowers the agent's scaled continuation payoff w, which lowers the agent's effective claim on the firm and hence induces a more severe agency problem. The wedge between average and marginal q is nonmonotone in w. See Figure 2. Average and marginal q are equal when $w = 0$ and the contract is terminated. Then $q_a > q_m$ for $w > 0$ until the cash payment region is reached, $w = \overline{w}$. At that point, the incentive benefits of w are outweighed by the agent's impatience, so that $p'(\overline{w}) = -1$ and again $q_a = q_m$. The implication for empirical investment studies is that the measurement error inherent in using average q as a proxy for marginal q varies over time for a given firm and varies across firms depending on firms' performance (which drives w). For our agency model, the relation between average and marginal q is given by equations (23) and (24).

Both average q and marginal q are functions of the agent's scaled continuation payoff w. Because $p'(w) \geq -1$, average q is increasing in w (reflecting the incentive alignment effect noted earlier). In addition, the concavity of $p(w)$ implies that marginal q is also increasing in w. In Figure 2, we plot q_a (the

vertical intercept of the line originating at $p(w)$ that has slope -1), q_m (the vertical intercept of the line tangent at $p(w)$), and the first-best average (also marginal) q^{FB}.

It is well understood that marginal and average q are forward-looking measures that capture future investment opportunities. In the presence of agency costs, it is also the case that both marginal and average q are positively related to the firm's profit history. Recall that the value of the agent's claim w evolves according to (17), and so is increasing with the past profits of the firm, and that both marginal and average q increase with w for incentive reasons. Unlike the neoclassical setting in which q is independent of the firm's history, in our setting both marginal and average q are history-dependent.

C. Investment and q

We now turn to the model's predictions for investment. First, note that the investment-capital ratio $i(w)$ in our agency model depends on w. Specifically, the first-order condition for optimal investment (19) can be written in terms of marginal q,

$$c'(i(w)) = q_m(w) = p(w) - wp'(w). \qquad (26)$$

The convexity of the investment cost function c and the monotonicity of q_m imply that investment increases with w,

$$i'(w) = \frac{q_m'(w)}{c''(i(w))} = -\frac{wp''(w)}{c''(i(w))} \geq 0, \qquad (27)$$

where the inequality is strict except at termination ($w = 0$) and the cash payout boundary ($p''(\overline{w}) = 0$).

Intuitively, when w is low, inefficient termination becomes more likely. Hence, investors optimally invest less. In the limiting case in which termination is immediate ($w = 0$), the marginal benefit of investing is just the liquidation value l per unit of capital. Thus, the lower bound on the firm's investment is given by $c'(i(0)) = l$. Assuming $c'(0) > l$, the firm will disinvest near termination.

Now consider the other limiting case in which w reaches the cash payout boundary $\overline{w}$. Because $q_m(w) < q^{FB}$ from (25), we have $i(\overline{w}) < i^{FB}$. Thus, even at this upper boundary, there is underinvestment—the strict relative impatience of the agent, that is, $\gamma > r$, creates a wedge between our solution and first-best investment. In the limit, when γ is sufficiently close to r, the difference between $i(\overline{w})$ and i^{FB} disappears. That is, the degree of underinvestment at the payout boundary depends on the agent's relative impatience.

To summarize, in addition to costly termination as a form of underinvestment, the investment-capital ratio is lower than the first-best level, that is, $i(w) < i^{FB}$ always. Thus, our model features underinvestment at all times. Figure 3 shows investors' value function and the investment-capital ratio for two different volatility levels. The positive relation between investment and the agent's continuation payoff w implies that investment is positively related to
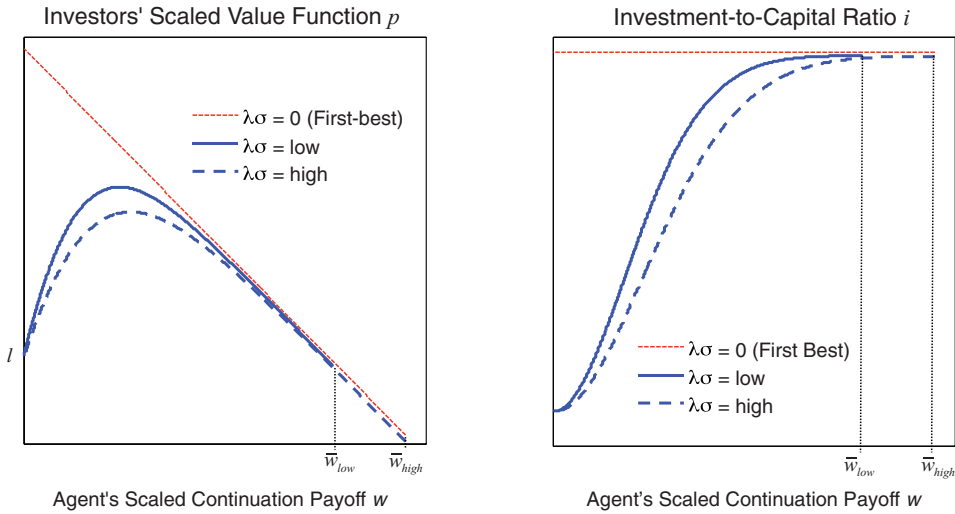
Figure 3. The effect of volatility, σ, and the severity of the agency problem, λ, on investors' scaled value function $p(w)$, and the investment-to-capital ratio $i(w)$.

past performance. Moreover, given the persistence of w, investment is positively serially correlated. By contrast, in the first-best scenario, investment is insensitive to past performance.

Figure 3 also shows that the value of the firm and the rate of investment are lower with a higher level of idiosyncratic volatility, σ. With higher volatility, firm profits are less informative regarding the agent's effort, and incentive provision becomes more costly. This effect reduces the value of the firm and the return on investment.[11] The same comparative statics would result from an increase in the rate λ at which the agent accrues private benefits (exacerbating the agency problem). In fact, from Proposition 1, firm value and the level of investment depend only on the product of λ and σ—the extent of the agency problem is determined by both firm volatility and the agent's required exposure to it.

Note also that the cash payout boundary $\overline{w}$ increases with the severity of the agency problem. As $\lambda\sigma$ increases, so does the volatility of the agent's continuation payoff w. To reduce the risk of inefficient termination, it is optimal to allow for a higher level of deferred compensation.

D. A Numerical Example

We now provide some suggestive analysis on the quantitative importance of agency. For guidance for our numerical example, we rely on the findings of Eberly, Rebelo, and Vincent (2009), who provide empirical evidence in support

---

[11] Panousi and Papanikolaou (2012) present evidence that investment is lower for firms with higher idiosyncratic risk.

Table I
The Impact of Agency Friction

The parameters are $r = 4.6\%$, $\sigma = 26\%$, $\mu = 20\%$, $\gamma = 5\%$, $\delta = 12.5\%$, and $\theta = 2$. Baseline case (III): $\lambda = 0.2$ and $l = 0.97$.

	I	II	III	IV	V	VI
Parameters						
Agency parameter, λ	0.05	0.1	0.2	0.4	0.6	0.2
Liquidation value, l	0.97	0.97	0.97	0.97	0.97	1.16
Model Outputs						
Agent payout boundary, $\overline{w}$	0.18	0.25	0.43	0.74	1.00	0.39
Average q (= marginal q) at payout boundary, $p(\overline{w}) + \overline{w}$	1.28	1.27	1.25	1.23	1.21	1.26
Maximum investor continuation payoff, $p(w_0)$	1.19	1.16	1.07	1.00	0.97	1.16
Initial agent continuation payoff, w_0	0.07	0.08	0.11	0.09	0	0
Model Predictions (%)						
Reduction in investment, $i^{FB} - i(w)$	3.02	4.05	6.18	8.41	10.0	6.17
Volatility of investment, $\lambda \sigma i'(w)$	1.09	1.22	1.49	1.63	1.65	1.45
Reduction in value, $1 - q_a(w)/q^{FB}$	3.39	4.26	6.37	9.02	11.9	6.36
Agent's share of value, $w/q_a(w)$	8.96	12.3	21.9	35.7	46.2	17.4

of Hayashi (1982). Following their work, we set the annual interest rate to $r = 4.6\%$, expected productivity to $\mu = 20\%$, and the agent's discount rate to $\gamma = 5\%$. For the full sample of large firms in Compustat from 1981 to 2003, Eberly, Rebelo, and Vincent (2009) document that the average q is 1.3 and the investment-capital ratio is 15%. Equating the first-best market-to-book ratio q^{FB} and the first-best investment-capital ratio i^{FB} to these sample averages, we set $\delta = 12.5\%$ and use quadratic adjustment costs with $\theta = 2$ and for our model (in line with estimates in Eberly, Rebelo, and Vincent (2009)).[12] We set volatility to $\sigma = 26\%$. Finally, we set the agency parameter to $\lambda = 0.2$ and the liquidation value to be $l = 0.97$ for the baseline case.

Given these baseline parameters, we have the following outputs from our model (see Table I). The maximal level of deferred compensation for the agent equals $\overline{w} = 0.43$. If the present value of the agent's future compensation exceeds this level, then it is optimal to pay the agent the difference in cash immediately. The corresponding maximal value for the firm is $q_a(\overline{w}) = 1.25$. This value is below the first-best, $q^{FB} = 1.3$, owing to the agent's relative impatience. The maximal value attainable by investors is even lower, $p(w_0) = 1.07$, due to the need to compensate the agent to provide incentives. The agent's expected compensation that maximizes the investor's value is $w_0 = 0.11$.

We simulate our model monthly, generating a sample path that lasts 20 years or until liquidation. Each simulation starts with $w_0 = \arg\max p(w)$, with the interpretation that investors own the firm and hire an agent using a contract that maximizes investors' value. We repeat the simulation 5,000 times. In Table I, we report the average data for the sample paths.

[12] We are not attempting a full calibration exercise. Rather, we use the first-best benchmark as a proxy to calculate our parameter values and obtain suggestive results regarding the potential impact of agency.
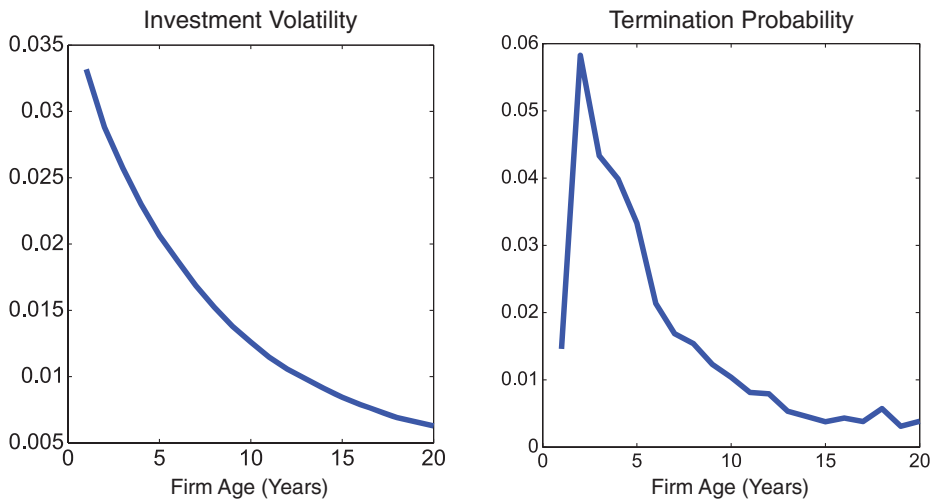
Figure 4. Investment volatility and termination probability conditional on firm age from simulation. Baseline parameters are $r = 4.6\%$, $\mu = 20\%$, $\gamma = 5\%$, $\delta = 12.5\%$, $\theta = 2$, $\sigma = 26\%$, $\lambda = 0.2$, and $l = 0.97$.

As Table I shows, for our baseline parameters, agency costs reduce the average investment rate 6.18% below the first-best level, $i^{FB} = 15\%$. Also, unlike the first-best in which the investment rate is constant, with agency costs, investment volatility is 1.49%. As Figure 4 shows, investment volatility decreases with firm age. The intuition is that older firms have survived to be old because of good performance, and good performance relaxes the agency problem (by raising w) and so investment is closer to first-best and less volatile. Figure 4 also shows the annual probability of termination conditional on firm age. The termination likelihood is increasing in age for the younger firms but decreasing in age for older firms. The intuition regarding the younger firms is that the agent begins with some surplus and it will take time for bad performance to erode this surplus and cause a termination. For the older firms, the figure shows a survivorship bias. The longer the firm has survived, the higher the likelihood that performance has been good over the firm's life, and consequently the lower the likelihood of termination. As shown in Table I, the effect of these investment distortions and the possibility of termination lead to an average reduction in firm value of 6.37%. Finally, we report the agent's average share of total firm value to be 21.9%.

Cases I to V of Table I also illustrate comparative statics as we change the magnitude of the agency problem, given by λ. As the agency problem becomes more severe, the agent becomes exposed to greater risk to provide incentives. This greater risk exposure increases the risk of termination; as a result, the payout boundary $\overline{w}$ increases to allow for a larger potential buffer of deferred compensation. Not surprisingly, the total value of the firm and the maximum value to investors also decline with the severity of the agency problem. The

agent's initial surplus, however, changes nonmonotonically with λ—there is a value to providing the agent with higher surplus to avoid early termination, but the total value-added of operating the firm is also declining. Indeed, for Case V with $\lambda = 0.6$, the maximal value of the operating firm to its investors is equal to its liquidation value, and so the agent is given no initial surplus (if λ were any higher it would be optimal for investors to shut down the firm immediately). The final rows in the table demonstrate that with a more severe agency problem, investment volatility and the average deviation from first-best increase, as does the agent's share of firm value.

Case VI considers an increase in the liquidation value relative to Case III, the base case. Naturally, with a higher liquidation value, firm value rises but the payout boundary and initial agent surplus fall as less buffer against termination is needed. Deviations from the first-best are reduced due to the decline in the cost of incentive provision via termination and the deferral of compensation. The agent's average share also falls, which is not surprising given the decline in the initial and maximal continuation values.

Cases V and VI also have an alternative interpretation. If the manager can be fired and costlessly replaced, then the value of the firm to investors upon termination, $l = p(0)$, should equal the maximal value attainable with a new manager, $p(w_0)$. Thus, Cases V and VI show the effect of agency when there is no "direct" cost of terminating the manager. The effect on investment and firm value is still significant because of the possible deferral of compensation and the fact that future managers capture rents if the initial manager is fired.[13]

IV. Implementing the Optimal Contract

In Section III, we characterized the optimal contract in terms of an optimal mechanism. In this section, we consider implications of the optimal mechanism for the firm's financial slack, and explore the link between financial slack and investment.

Recall that the dynamics of the optimal contract are determined by the evolution of the agent's continuation payoff w_t. Because termination occurs when $w_t = 0$, we can interpret w_t as a measure of the firm's "distance" to termination. Indeed, from (17), the largest short-run shock dA_t to the firm's cash flows that can occur without termination is given by w_t/λ. This suggests that we can interpret $m_t = w_t/\lambda$ as the firm's available financial slack, that is, the largest short-run loss the firm can sustain before the agent is terminated and a change of control occurs.

We can formalize this idea in a variety of ways. Financial slack may correspond to the firm's cash reserves (as in Biais et al. (2007)), a line of credit

---

[13] Because future managers start with an initial surplus of zero, the rent each one captures is infinitesimal. But, because replacing the manager is costless, there will be a replacement "frenzy" whenever the termination boundary is hit, with an infinite number of replacements over a very short period. This case captures a limiting case of a more realistic scenario in which there is some small cost of replacing the manager.

(as in DeMarzo and Sannikov (2006) and DeMarzo and Fishman (2007b)), or a combination of the firm's cash and available credit. Payments to investors may correspond to payouts on debt, equity, or other securities. Rather than attempt to describe all possibilities, we'll describe one simple way to implement the optimal contract, and then discuss which of its features are robust.

Specifically, suppose the firm meets its short-term financing needs by maintaining a cash reserve. (Recall that the firm will potentially generate operating losses, and so needs access to cash or credit to operate.) Let M_t denote the level of cash reserves at date t. These reserves earn interest rate r, and once they are exhausted, the firm cannot operate and the contract is terminated.

The firm is financed with equity. Equity holders require a minimum payout rate of[14]

$$dD_t = [K_t(\mu - c(i_t)) - (\gamma - r)M_t]dt.$$

The first component of the dividend, $K_t(\mu - c(i_t))$, corresponds to the firm's expected free cash flow. The second component, $(\gamma - r)M_t$, adjusts for the relative impatience of the agent, and is negligible when $\gamma \approx r$. If the agent fails to meet this minimal payout rate, the contract is terminated. Other than this constraint, the agent has discretion to choose an effort level a_t, the firm's investment-capital ratio i_t, as well as additional payout X_t in excess of the minimum payout rate described earlier. The agent is compensated by receiving a fraction λ of any "special" dividends X_t.

Under this implementation, the firm's cash reserves will grow according to

$$dM_t = rM_tdt + dY_t - dD_t - dX_t. \tag{28}$$

The value of the firm's equity is given by

$$S_t = \mathbb{E}_t\left[\int_t^\tau e^{-r(s-t)}(dD_t + (1-\lambda)\,dX_t) + e^{-r(\tau-t)}lK_t\right], \tag{29}$$

where τ is the first stochastic (hitting) time such that $M_t = 0$. The expected payoff to the agent is given by

$$W_t = \mathbb{E}_t\left[\int_t^\tau e^{-\gamma(s-t)}\lambda\,dX_t\right]. \tag{30}$$

The following proposition establishes that the above specification implements the optimal contract.

PROPOSITION 2: *Suppose the firm has initial cash reserves M_0 and can operate as long as $M_t \geq 0$ and it maintains the minimum payout rate dD_t. Then it is optimal for the agent to choose effort $a_t = 1$ and to choose the investment-capital ratio i_t given in Proposition 1. The agent accumulates cash reserves M_t until $m_t \equiv M_t/K_t = \overline{w}/\lambda$, and pays out all cash in excess of this amount. Given this policy,*

---

[14] If $dD_t < 0$, we interpret this as the maximum rate that the firm can raise new capital, for example, by issuing equity. We can show, however, that if $\lambda = 1$, then $dD_t > 0$.

the agent's payoff is $W_t = \lambda M_t$, which coincides with the continuation payoff of Proposition 1. Finally, the firm's stock price satisfies $S_t = (p(\lambda m_t) + m_t)K_t$.

In this implementation, regular dividends are relatively "smooth" and approximately correspond to the firm's *expected* rate of free cash flow. The cash flow fluctuations induced by the firm's productivity shocks are absorbed by the firm's cash reserves until the maximal level of reserves is achieved or the firm runs out of cash. Also, because the above financial policy implements the optimal contract, there is no ex ante change to the policy (such as issuance of alternative securities) that will make shareholders better off.[15]

The above implementation is not unique. For example, the minimum dividend payouts could be equivalently implemented as required coupon payments on long-term debt or preferred stock. (Such an implementation may be more natural if termination is interpreted as liquidating the firm, as opposed to firing the manager.) Also, rather than solely use cash reserves, the firm may maintain its financial slack M_t through a combination of cash and available credit, and the contract will terminate once these are exhausted.[16] Again, financial slack M_t will be proportional to the agent's continuation payoff W_t in the optimal contract. In fact, because W_t is a measure of the firm's "distance to termination" in the optimal contract, its relation to the firm's financial slack is a robust feature of *any* implementation.

In our implementation, financial slack per unit of capital is given by $m = w/\lambda$. Intuitively, firms with less severe agency problems, that is, lower λ, have more financial slack to avoid liquidation. We can also reinterpret some of our earlier results in terms of this measure of financial slack:

- Financial slack is positively related to past performance.
- Average q (corresponding to enterprise value plus agent rents) and marginal q increase with financial slack.[17]
- Investment increases with financial slack.
- Expected agent cash compensation (over any time interval) increases with financial slack.

---

[15] If the optimal contract is not renegotiation-proof, then, not surprisingly, there may be ex post improvements available to shareholders. See Appendix C for further discussion.

[16] The implementation developed here using cash reserves is similar to that in Biais et al. (2007). However, we could also follow DeMarzo and Fishman (2007b) and DeMarzo and Sannikov (2006) by providing the firm financial slack with a credit line, which is also common in practice. With a credit line, the implementation involves a lower minimum dividend payout rate; it would be reduced by the interest due on the credit line and the risk-free interest rate on the firm's unused credit. This reduction reflects that (i) the firm would now be paying interest on the drawn balance, and (ii) the firm would not be earning interest on the cash balance.

[17] Commonly used empirical measures of q do not include all future managerial compensation. And while under our definition q increases with financial slack, as Figure 2 shows, the value to investors $p(w)$ need not be monotone and is eventually decreasing. Thus, common empirical measures of q, for example, $p(w)$, might show a (false) negative correlation between q and financial slack.

- The maximal level of financial slack is higher for firms with more volatile cash flows and firms with lower liquidation values.[18]

The investment literature often focuses on the positive relation between firms' cash flow and investment; see, for example, Fazzari, Hubbard, and Peterson (1988), Hubbard (1998), and Stein (2003). Although our results are concistent with this effect, our analysis suggests that financial slack (a stock, rather than flow, measure) has a more direct effect on investment. It is also worth noting that our dynamic agency model does not yield a sharp prediction on the sensitivity of di/dm with respect to financial slack. That is, as shown by Kaplan and Zingales (1997) (in a model with exogenous financing costs), it is difficult to sign d^2i/dm^2 without imposing strong restrictions on the cost of investment. Their result can be understood in the context of our model from (27), where it is clear that $i'(w)$, and therefore di/dm, depends on the convexity of the investment cost function c''. Thus, d^2i/dm^2 will depend on the third-order derivatives of the cost function.

V. Persistent Profitability Shocks

The only shocks in our model thus far are the firm's idiosyncratic temporary productivity shocks. Although these shocks have no effect in the neoclassical setting, they obscure the agent's actions to create an agency problem. The optimal incentive contract then implies that these temporary idiosyncratic shocks have a persistent impact on the firm's investment, growth, and value.

In this section, we extend the model to allow for persistent observable shocks to the firm's profitability. These shocks differ in two important ways from the firm's temporary productivity shocks. First, these profitability shocks are observable and can be contracted on. Second, because these profitability shocks are persistent, they will affect the firm's optimal rate of investment even in the neoclassical setting.

Our goal is to explore the interaction of public persistent shocks with the agency problem, as well as the consequences for investment, financial slack, and managerial compensation. As a benchmark, if the profitability shocks were purely transitory, they would have no effect on the firm's investment or the agent's compensation, with or without an agency problem. Investors would simply absorb the shocks, insulating the firm and the agent. As we will show, however, if the profitability shocks are persistent they will affect both the optimal level of investment and the agent's compensation, with the latter effect having an additional feedback on the firm's investment.

We extend our model in Section V.A. Next, we analyze the interaction effects in Section V.B. Finally, in Section V.C we examine the relation between investment and financial slack controlling for average q.

---

[18] Note, however, that, while the optimal contract (in terms of payoffs and net cash flows) only depends on $\lambda\sigma$, in this implementation the maximal level of financial slack is given by $\overline{m} = \overline{w}/\lambda$. So, although $\overline{w}$ increases with λ, the maximal level of financial slack $\overline{m}$ tends to decrease with the level of private benefits (λ).

A. The Model and Solution Method

We extend the basic model by introducing a stochastic profitability variable, π_t. To keep the analysis simple, we model π_t as a two-state Markov regime-switching process.[19] Specifically, $\pi_t \in \{\pi^L, \pi^H\}$ with $0 < \pi^L < \pi^H$. Let ξ^n be the transition intensity out of state $n = L$ or H to the other state. Thus, for example, given a current state L at date t, the state changes to H with probability $\xi^L dt$ over the time interval $(t, t + dt)$. The state π_t is observable to investors and the agent and is contractible. The firm's operating profit is given by the following modification of (3):

$$dY_t = K_t(\pi_t dA_t - c(i_t)dt). \qquad (31)$$

One interpretation of π is the output price, but more generally, it can correspond to any observable factor affecting the firm's profitability.

Let $P(K, W, \pi)$ denote investors' value function when capital stock is K, the agent's continuation payoff is W, and the state is π. Again, using the scale invariance of the firm's technology, we conjecture that for $n = L$ or H and $w = W/K$, we can write the value function as

$$P(K, W, \pi^n) = K p_n(w), \qquad (32)$$

where $p_n(w)$ represents investors' scaled value per unit of capital in state n.

To determine the dynamics of the agent's scaled continuation payoff w, we must first consider how the agent's payoff is affected if the state changes. Suppose the state changes from π^L to π^H. How should the agent's scaled continuation payoff w respond to this exogenous shock? In designing the optimal contract, investors optimally adjust the agent's continuation payoff to minimize agency costs. When the state π_t switches from π^L to π^H, the firm becomes more profitable. In Figure 5, this is captured by the expansion of investors' value function, that is, $p_H(w) \geq p_L(w)$ for any w. Because firm value is higher, the benefit of avoiding termination/liquidation is also higher, and we will show that this decreases the marginal cost of increasing the agent's payoff, that is, $p'_H(w) \geq p'_L(w)$ for any w. This observation suggests that it is optimal to increase the agent's continuation payoff w, and thus the firm's financial slack, when profitability improves in order to reduce agency costs.

To formalize this effect, let $\psi_{nm}(w)$ denote the endogenous adjustment of w conditional on a jump from state π^n to the alternative state π^m, so that the agent's scaled continuation payoff changes from w just prior to the jump to $w + \psi_{nm}(w)$ immediately after. The optimal adjustment should equate the marginal cost of compensating the agent before and after the jump. Given that investors have to deliver an additional dollar of compensation to the agent, what is their marginal cost of doing so in each state? The marginal cost is

---

[19] Piskorski and Tchistyi (2010) consider a model of mortgage design in which they use a Markov switching process to describe interest rates. They were the first to incorporate such a process in a continuous-time contracting model, and our analysis follows their approach.
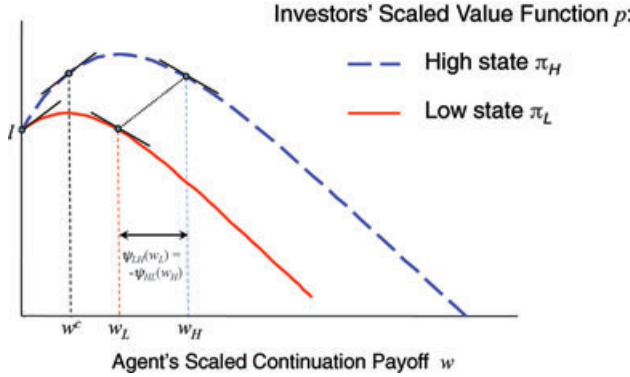
Figure 5. Determination of the agent's continuation payoff (compensation) adjustment $\psi_{HL}(w)$ **or** $\psi_{LH}(w)$ **given exogenous profitability changes.** As shown, the optimal compensation policy implies that in the interior region, $p'_L(w_L) = p'_H(w_L + \psi_{LH}(w_L)) = p'_H(w_H) = p'_L(w_H + \psi_{HL}(w_H))$.

captured by the marginal impact of w on investors' value function, that is, $p'_n(w)$. Therefore, the compensation adjustment ψ_{nm} is given by

$$p'_n(w) = p'_m(w + \psi_{nm}(w)), \tag{33}$$

which is feasible as long as $p'_n(w) \leq p'_m(0)$. If $p'_n(w) > p'_m(0)$, the agent's payoff jumps to zero ($\psi_{nm}(w) = -w$) and the contract terminates in order to minimize the difference in the marginal cost of compensation. See Figure 5; there, if the state is high and $w < w^c$, where w^c is determined by

$$p'_H(w^c) = p'_L(0),$$

a jump to the low state triggers termination.

The above discussion leads to the following dynamics for the agent's continuation value w. As before, cash compensation is deferred up to a threshold $\overline{w}^n$, but now the threshold depends on the state. Letting N_t denote the cumulative number of regime changes up to time t, the dynamics for the agent's scaled continuation payoff with state π^n and $w_t \in [0, \overline{w}^n]$ are given by

$$dw_t = (\gamma - (i_t - \delta))w_t dt + \lambda(dA_t - \mu dt) + \psi_{nm}(w_t)(dN_t - \xi^n dt). \tag{34}$$

Also as before, the diffusion martingale term $\lambda(dA_t - \mu dt)$ describes the agent's binding incentive constraint, implied by the concavity of investors' scaled value functions in both regimes (see Appendix D).[20] The jump martingale term $\psi_{nm}(w_t)(dN_t - \xi^n dt)$ has a zero drift, and this guarantees that the agent's continuation payoff W grows at γ on average, taking into account the net capital

---

[20] The incentive provision $\lambda K_t(dA_t - \mu dt)$ does not scale with the profitability variable π_t. This is consistent with the interpretation of the agent's action as effort. In contrast, the incentive provision would scale with π_t under the interpretation that the agent's action involves cash diversion. Otherwise, the qualitative conclusions of the model would remain unchanged.

accumulation rate, $i_t - \delta$, along the equilibrium path. In Appendix D, we provide a formal characterization of the optimal contract, and derive the following key properties:

PROPOSITION 3: *With state π^n, the agent's continuation payoff evolves according to* (34) *until $w_t = 0$ and the contract is terminated or until $w_t = \overline{w}^n$ and the agent receives cash. The investors' value functions p_L, p_H are concave, with $p_L < p_H$ for $w > 0$ and $p'_L < p'_H$ for $w \leq \overline{w}^H$. Thus, the compensation adjustment ψ_{LH} is positive and ψ_{HL} is negative. Moreover, if the state is π^H and w_t is low enough such that $p'_H(w_t) \geq p'_L(0)$, then $\psi_{HL}(w_t) = -w_t$ and the contract is immediately terminated if the state jumps to π^L.*

B. Model Implications

Here, we discuss a number of implications of our model for the impact of profitability shocks on the agent's compensation, the firm's financial slack, and the optimal level of investment.

B.1. Agent Compensation

An important implication of our results is that the agent's compensation will be affected by persistent shocks to the firm's profitability, even when these shocks are publicly observable and unrelated to the agency problem. Specifically, the agent is rewarded when the state improves, $\psi_{LH} > 0$, and is penalized—and possibly immediately terminated—when the state worsens, $\psi_{HL} < 0$. This result is in contrast with conventional wisdom that optimal contracts should insulate managers from exogenous shocks and compensate them based solely on relative performance measures. Rather, managerial compensation will *optimally* be sensitive to the *absolute* performance of the firm.[21]

The intuition for this result is that an increase in the firm's profitability makes it efficient to reduce the likelihood of termination by increasing the level of the agent's compensation. Thus, the optimal contract shifts the agent's compensation from low states to high states. More generally, in a dynamic agency context, the optimal contract smoothes the *marginal cost of compensation*, increasing the agent's rents (and thus aligning incentives) in states in which the incentive problem is more costly.[22]

---

[21] Bertrand and Mullainathan (2001) present evidence that CEO compensation depends on exogenous shocks and Jenter and Kanaan (2010) present evidence that the firing of CEOs depends on exogenous shocks.

[22] Although dynamic models more naturally allow for changes in profitability and changes in the agent's compensation, a similar result can be derived in a static model as a comparative static result. For instance, consider a static model in which higher agent compensation relaxes the agency problem. And suppose the agent and investors sign a contract before the realization of a profitability shock. Then the optimal contract will raise the agent's compensation in high-profitability states and lower the agent's compensation in low-profitability states. Of course, a static model cannot generate *changes* in compensation that result from *changes* in profitability as only one profitability level and one compensation level will be observed.

We assumed here that the liquidation/termination value of the firm is independent of the current state, thus making termination relatively more costly in the high state. If termination corresponds to firing and replacing the agent, then, although $l_H > l_L$ (firm value upon replacement is higher when the state is high), the qualitative results discussed earlier remain unchanged. Indeed, as long as it is costly to replace the agent ($\kappa > 0$ in (22)), then $w^c > 0$ and the agent may be fired and replaced if the state worsens.

However, if l_H and l_L are specified in some alternative manner, it is possible that termination would be sufficiently less costly when the state was high so that it would be optimal to reduce the agent's compensation (and thereby increase the risk of termination) when the state improved. But while these specific results could change depending on such assumptions, the more important qualitative result, namely, that the agent's compensation is affected by persistent observable shocks, would remain.

B.2. Hedging and Financial Slack

We can also interpret these results in the context of the firm's financial slack. Because $\psi_{LH} > 0$, it is optimal to increase the firm's available slack when the state improves and decrease slack when the state worsens. This sensitivity could be implemented, for example, by having the firm hold a financial derivative that pays off when the state improves. A similar adjustment to financial slack might be implemented with convertible debt—when the state improves and firm value increases, bondholders convert their securities and financial slack increases. Whatever the specific implementation, it is optimal for the firm to *increase* the sensitivity of its cash position to the state. Notably, it is *not* optimal for the firm to hedge a change in the state. Rather, the firm's hedging policy should smooth the changes in the marginal value of financial slack.[23]

B.3. Profitability, Financial Slack, and Investment

Investment depends on the firm's profitability. See Figure 6; in the left panel, the solid line depicts the change in the investment-capital ratio when profitability improves, $i_H(w + \psi_{LH}(w)) - i_L(w)$. We can decompose this change into two components. There is a direct effect that investment is more profitable in the high state, and so $i_H(w) - i_L(w) > 0$. The dashed line depicts this direct effect. There is also an indirect effect, $i_H(w + \psi_{LH}(w)) - i_H(w) > 0$, arising from the optimal discrete adjustment of financial slack when profitability changes. Because the agent's continuation payoff w, and thus the firm's financial slack, will optimally increase with an increase in profitability, ($\psi_{LH}(w) > 0$), the agency problem is reduced and this also makes investment more profitable. So the direct effect understates the impact of a profitability shock on the

---

[23] This motive for hedging is related to that of Froot, Scharfstein, and Stein (1993), who suggest that firms should hedge to fund their investment needs when external capital is costly. Here, firms smooth the agency costs that underlie their cost of capital.
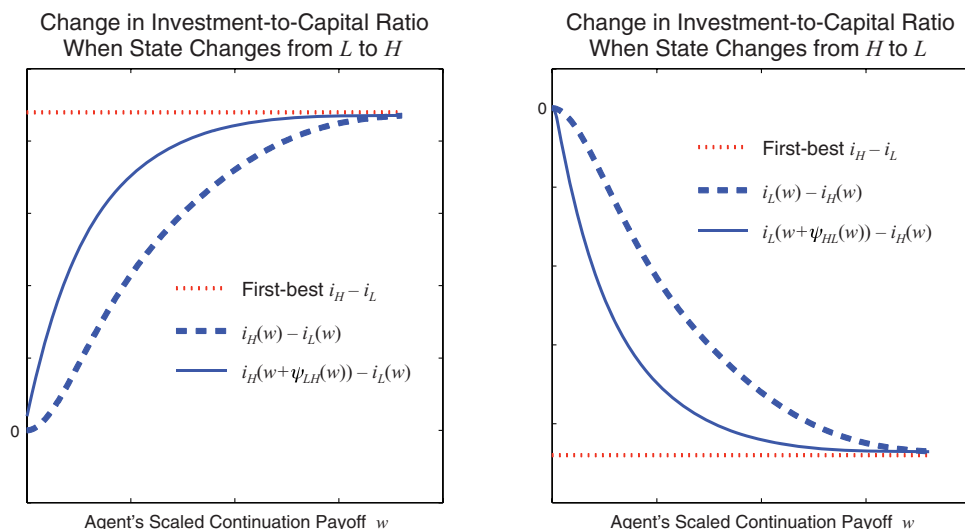
Figure 6. Profitability shocks and investment.

investment-capital ratio. The difference between the solid and dashed lines measures this indirect effect. As illustrated in Figure 6, the indirect effect vanishes when financial slack is high.

These results have implications for investment regressions (motivated either implicitly or explicitly by investment functions). With stochastic profitability shocks, it is important to jointly analyze the firm's investment and financial slack decisions. Treating financial slack (e.g., cash plus available credit) as predetermined in an investment regression misses the indirect effect. This would lead one to underpredict the investment response to a profitability shock by ignoring the optimal contemporaneous adjustment of financial slack.

Also note that the investment response to a profitability shock, depicted by the solid line in Figure 6, is smaller than the investment response in the first-best case, depicted by the horizontal line. Although agency costs induce underinvestment in both regimes, they also dampen the impact of profitability shocks on investment.

C. Investment, Financial Slack, and Average q

In Section IV, we noted the positive relation between investment and financial slack that results from the dynamics of the agency problem. More interesting empirically, however, is the relation between investment and financial slack that survives after controlling for average (or Tobin's) q. We could investigate this effect in our basic model by considering heterogeneity across firms in other model parameters, such as the firm's average profitability. With the stochastic profitability model developed here, however, we can make an even stronger point by considering the relation between investment, financial slack, and
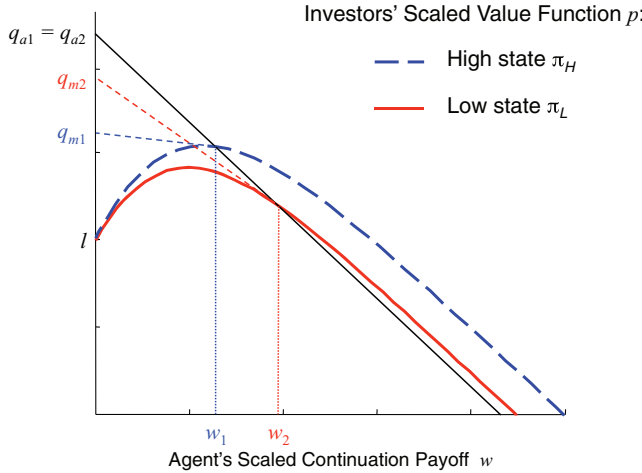
Figure 7. Comparison of marginal q, holding average q fixed.

average q for a single firm. Specifically, for a given firm, and a given average q, we will show that investment is increasing with financial slack.

Figure 7 depicts the investor value functions p_L and p_H for each state. Consider two situations that lead to the same average q: the high state with low financial slack w_1, and the low state with high financial slack w_2. These points can be seen to have the same average $q = p(w) + w$, as they lie along a line with slope -1. Thus, the figure shows that financial slack and profitability are substitutes in the determination of average q.

The figure also shows that, despite average q being equal, marginal q differs for these two points. Marginal $q = p(w) - wp'(w)$ can be determined as the intercept with the vertical p-axis of the line tangent to the value function from each point. As shown in the Figure 7, the situation with the higher w_2 (and low state) has a higher marginal q_2. Because investment is increasing in marginal q, even after controlling for average q, higher financial slack leads to a higher investment rate.

In Figure 8, we illustrate the sensitivity of investment to financial slack for the whole range of average q in our example. For each level of average q, we compute the difference in financial slack (Δw) and investment (Δi) across the two states, and plot the ratio $\Delta i/\Delta w$. The figure shows that this sensitivity is positive, that is, more financial slack is associated with higher investment, holding average q fixed. The figure also shows that the relation between investment and slack is nonmonotone so that the sensitivity of investment to financial slack is not necessarily higher for more financially constrained firms.

VI. Concluding Remarks

By synthesizing an agency-based model of financial frictions and neoclassical investment theory, our model generates a number of predictions about financing and investment dynamics. Optimal contracting implies that agency
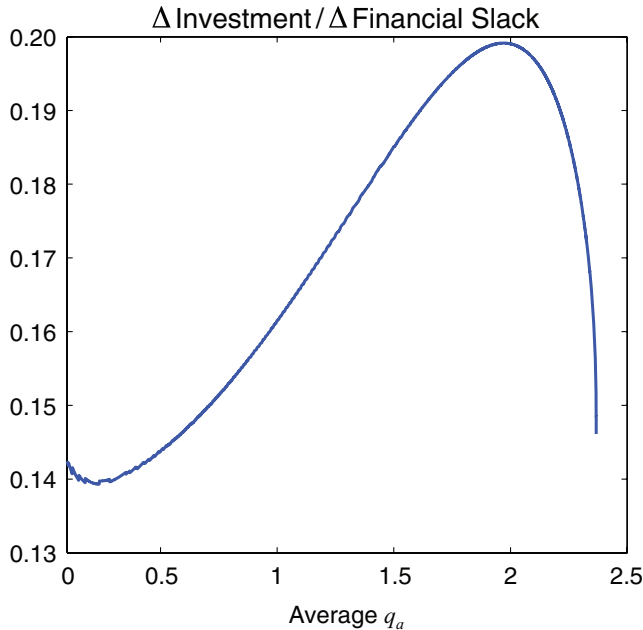
Figure 8. The ratio of the investment difference to the financial slack difference, holding average q fixed.

costs introduce a history-dependent wedge between marginal q and average q. Consequently, the measurement error inherent in using average q as a proxy for marginal q will vary over time with the firm's performance. Because the agent is rewarded for past success by holding a larger future stake in the firm (that is, a higher continuation payoff), agency costs fall and thus the return to investment rises when the firm performs well. Hence, investment is positively correlated with past profitability, past investment, managerial compensation, and financial slack even with time-invariant investment opportunities. Also, even with risk-neutrality, investment decreases with firm-specific risk because such risk hides the agent's performance and exacerbates the agency problem.

To illustrate the effect of profitability shocks on firm value and investment dynamics in the presence of agency conflicts, we extend our model to allow the firm's profitability to vary stochastically over time. Here, we show that investment increases with financial slack after controlling for average q. More broadly, our theory suggests that financial slack, not cash flow, is the important predictor of investment after controlling for average q, thus challenging the empirical validity of using cash flow as a proxy for financial constraints as is common in the investment/cash flow sensitivity literature. We also show that the agent's compensation will depend not only on the firm's hidden transitory shocks, but also on observable persistent profitability shocks even though these shocks are beyond the agent's control. This result may help to explain the empirical relevance of absolute performance evaluation ("paying for luck"). In ongoing research, we continue to explore the sensitivity of

managerial compensation to external shocks in a range of dynamic agency contexts.

The analysis of profitability shocks illustrates the importance of treating financial slack as endogenous and subject to change in the event of a profitability shock. In the setting we consider, it is optimal for financial slack to increase (decrease) in response to a positive (negative) profitability shock. Hence, in estimating the investment response to such a shock, treating financial slack as predetermined would miss the response of investment to the optimal adjustment of financial slack.

Our analysis also highlights the importance of including the present value of future managerial compensation when estimating q. Commonly used empirical measures of q do not explicitly adjust for managerial compensation, but instead are based on the market value of the firm's assets, estimated from the market value of equity and the book value of debt and other liabilities. To the extent that managers' future compensation is already reflected on the balance sheet (such as, for example, equity resulting from prior stock or option grants, or cash held in reserve to pay the manager), this approach is consistent with the theory. However, if a large part of managerial compensation arises from future salary and bonuses, option grants, etc., then typical empirical measures of q may be somewhat misspecified, and likely lie between $p(w)$, the value of the firm to its investors, and $p(w) + w$, the value to both investors and managers.

Because our model is based on a constant returns to scale investment technology (as in Hayashi (1982)), it is not well suited to address questions relating firm size and growth. Indeed, if we control for past performance or financial slack, size does not matter in our framework. If we fail to control for past performance/slack, then larger firms—because they are more likely to have had good recent performance in order to become large—will tend to grow faster than worse performing firms that are thus smaller. That is, as noted earlier, investment is positively serially correlated. If we were to incorporate decreasing returns to scale into our setting, we would find the highest growth for small firms with good recent performance (and high recent growth), and the slowest growth for large firms with poor recent performance (and low recent growth).

Appendix A: Proof of Proposition 1

We impose the usual regularity condition on the payment policy

$$\mathbb{E} \left(\int_0^\tau e^{-\gamma s} dU_s \right)^2 < \infty. \tag{A1}$$

We further require that

$$\mathbb{E} \left[\int_0^T (e^{-rt} K_t)^2 dt \right] < \infty \text{ for all } T > 0, \tag{A2}$$

and

$$\lim_{T \to \infty} \mathbb{E}(e^{-rT} K_T) = 0. \tag{A3}$$

Both regularity conditions place certain restrictions on the investment policies. Since the project is terminated at τ, throughout we take the convention that $H_T \mathbf{1}_{\{T > \tau\}} = H_\tau$ for any stochastic process H.

Throughout the proof, we follow the adjustment cost $c(i)$ as the commonly used quadratic form.

LEMMA 1: *For any contract* $\Phi = \{U, \tau\}$, *there exists a progressively measurable process* $\{\beta_t : 0 \le t < \tau\}$ *such that the agent's continuation value* W_t *evolves according to*

$$dW_t = \gamma W_t dt - dU_t + \beta_t K_t (dA_t - \mu dt) \tag{A4}$$

under $a_t = 1$ *all the time. The contract* Φ *is incentive compatible if and only if* $\beta_t \ge \lambda$ *for* $t \in (0, \tau)$.

In the optimal contract, $\beta_t = \lambda$, so the above dynamics in (A4) are the dynamics in (13). The proof is similar to DeMarzo and Sannikov (2006); we omit the proof here.

Now we verify that the contract and the associated investment policy derived from the HJB equation are indeed optimal. The evolution of $w = \frac{W}{K}$ follows easily from the evolutions of W and K. The key ODE (18), under the quadratic adjustment cost and the associated optimal investment policy, is

$$(r + \delta)p(w) = \mu + \frac{(p(w) - wp'(w) - 1)^2}{2\theta} + p'(w)(\gamma + \delta)w + \frac{\lambda^2 \sigma^2}{2} p''(w). \tag{A5}$$

LEMMA 2: *The scaled investors' value function* $p(w)$ *is concave on* $(0, \overline{w})$.

Proof: By differentiating (A5), we obtain

$$(r + \delta)p' = -\frac{(p - wp' - 1)wp''}{\theta} + (\gamma + \delta)wp'' + (\gamma + \delta)p' + \frac{\lambda^2 \sigma^2}{2} p'''. \tag{A6}$$

Evaluating (A6) at the upper boundary $\overline{w}$, and using $p'(\overline{w}) = -1$ and $p''(\overline{w}) = 0$, we find

$$\frac{\lambda^2 \sigma^2}{2} p'''(\overline{w}) = \gamma - r > 0.$$

Therefore, $p''(\overline{w} - \epsilon) < 0$. Now let $q(w) = p(w) - wp'(w)$. We have

$$(r + \delta)q(w) = \mu + \frac{(q(w) - 1)^2}{2\theta} + (\gamma - r)wp'(w) + \frac{\lambda^2 \sigma^2}{2} p''(w).$$

Suppose that there exists some $\widetilde{w} < \overline{w}$ such that $p''(\widetilde{w}) = 0$. Choose the largest $\widetilde{w}$ such that $p''(w + \epsilon) < 0$. Evaluating the above equation at $\widetilde{w}$, we have

$$(r + \delta)q(\widetilde{w}) = \mu + \frac{(q(\widetilde{w}) - 1)^2}{2\theta} + (\gamma - r)\widetilde{w}p'(\widetilde{w}).$$

Since $q(\widetilde{w}) < q^{FB}$, and $(r + \delta)q^{FB} = \mu + \frac{(q^{FB} - 1)^2}{2\theta}$, we have $p'(\widetilde{w}) < 0$. Therefore, evaluating (A6) at point $\widetilde{w}$, we obtain

$$(r + \delta)p'(\widetilde{w}) = p'(\widetilde{w})(\gamma + \delta) + \frac{\lambda^2 \sigma^2}{2} p'''(\widetilde{w}),$$

which implies that $p'''(\widetilde{w}) = \frac{2(r - \gamma)}{\lambda^2 \sigma^2} p'(\widetilde{w}) > 0$. This is inconsistent with the choice of $\widetilde{w}$ where $p''(\widetilde{w}) = 0$ but $p''(\widetilde{w} + \epsilon) < 0$. Therefore, $p(\cdot)$ is strictly concave over the whole domain $(0, \overline{w})$. Q.E.D.

Take any incentive-compatible contract $\Phi = (I, U, \tau)$. For any $t \leq \tau$, define its auxiliary gain process $\{G\}$ as

$$G_t(\Phi) = \int_0^t e^{-rs}(dY_s - dU_s) + e^{-rt}P(K_t, W_t)$$

$$= \int_0^t e^{-rs}\left(K_s dA_s - I_s ds - \frac{\theta I_s^2}{2K_s} - dU_s\right) + e^{-rt}P(K_t, W_t), \quad \text{(A7)}$$

where the agent's continuation payoff W_t evolves according to (13). Under the optimal contract Φ^*, the associated optimal continuation payoff W_t^* has a volatility $\lambda \sigma K_t$, and $\{U^*\}$ reflects W_t^* at $\overline{W}_t^* = \overline{w}K_t$.

Recall that $w_t = W_t/K_t$ and $P(K_t, W_t) = K_t p(w_t)$. Ito's lemma implies that, for $t < \tau$,

$$e^{rt}dG_t = K_t \left\{ \begin{bmatrix} -rp(w_t) + \mu - I_t/K_t - \frac{\theta}{2}(I_t/K_t)^2 + (I_t/K_t - \delta)(p(w_t) - w_t p'(w_t)) \\ + \gamma w_t p'(w_t) + \frac{\beta_t^2 \sigma^2}{2} p''(w_t) \\ + [-1 - p'(w_t)]dU_t/K_t + \sigma[1 + \beta_t p'(w_t)]dZ_t \end{bmatrix} dt \right\}.$$

Under the optimal investment policy I_t^*/K_t and optimal incentive policy $\beta_t = \lambda$, the first piece in the bracket—which is just our (18)—stays at zero always, whereas other investment and incentive policies will make this term nonpositive (here we use the concavity of p). The second piece captures the optimality of the cash payment policy. It is nonpositive since $p'(w_t) \geq -1$, but equals zero under the optimal contract.

Therefore, for the auxiliary gain process we have

$$dG_t(\Phi) = \mu_G(t)dt + e^{-rt}K_t\sigma[1 + \beta_t p'(w_t)]dZ_t,$$

where $\mu_G(t) \leq 0$. Let $\varphi_t \equiv e^{-rt}K_t\sigma[1 + \beta_t p'(w_t)]$. Conditions (A1) and (A2) imply that $\mathbb{E}[\int_0^T \varphi_t dZ_t] = 0$ for $\forall T > 0$ (note that p' is bounded). Furthermore, under

Φ investors' expected payoff is

$$\widetilde{G}(\Phi) \equiv \mathbb{E}\left[\int_0^\tau e^{-rs}dY_s - \int_0^\tau e^{-rs}dU_s + e^{-r\tau}lK_\tau\right].$$

Then, given any $t < \infty$,

$$\begin{aligned}
\widetilde{G}(\Phi) &= \mathbb{E}[G_\tau(\Phi)]\\
&= \mathbb{E}\left[G_{t\wedge\tau}(\widetilde{\Phi}) + \mathbf{1}_{t\leq\tau}\left[\int_t^\tau e^{-rs}dY_s - e^{-rs}dU_s\right] + e^{-r\tau}lK_\tau - e^{-rt}P(K_t, W_t)\right]\\
&= \mathbb{E}[G_{t\wedge\tau}(\widetilde{\Phi})] + e^{-rt}\mathbb{E}\left\{\left[\int_t^\tau e^{-r(s-t)}(dY_s - dU_s) + e^{-r(\tau-t)}lK_\tau\right.\right.\\
&\qquad\left.\left. - P(K_t, W_t)\right]\mathbf{1}_{t\leq\tau}\right\}\\
&\leq G_0 + (q^{FB} - l)\mathbb{E}[e^{-rt}K_t].
\end{aligned}$$

The first term of the inequality follows from the negative drift of $dG_t(\Phi)$ and the martingale property of $\int_0^{t\wedge\tau}\varphi_s dZ_s$. The second term in the inequality follows from

$$\mathbb{E}_t\left[\int_t^\tau e^{-r(s-t)}(dY_s - dU_s) + e^{-r(\tau-t)}lK_\tau\right] \leq q^{FB}K_t - w_t K_t,$$

which is the first-best result, and

$$q^{FB}K_t - w_t K_t - P(K_t, W_t) < (q^{FB} - l)K_t,$$

as $w + p(w)$ is increasing ($p' \geq -1$). But due to (A3), we have $\widetilde{G} \leq G_0$ for all incentive-compatible contracts. On the other hand, under the optimal contract Φ^*, investors' payoff $\widetilde{G}(\Phi^*)$ achieves G_0 because the above weak inequality holds in equality when $t \to \infty$.

Finally, we require that the agent's shirking benefit, $\phi \equiv \lambda\mu$, be sufficiently small to ensure the optimality of $a = 1$ all the time. Similar to DeMarzo and Sannikov (2006) and He (2009), there is a sufficient condition for the optimality of $a = \{\mu\}$ against $a_t = 0$ for some t. Let $\widehat{w} = \arg_w \max p(w)$. We require that

$$\frac{(p(w) - wp'(w) - 1)^2}{2\theta} \leq (r + \delta)p(w) - p'(w)[(\gamma + \delta)w - \phi] \quad \text{for all } w.$$

Since the left side is increasing in w, and the right side dominates $p(\frac{\phi}{\gamma+\delta}) - \frac{\gamma-r}{r+\delta}(p(\widehat{w}) - p(\frac{\phi}{\gamma+\delta}))$ (see the proof in DeMarzo and Sannikov (2006)), a sufficient condition is

$$\frac{(p(\overline{w}) + \overline{w} - 1)^2}{2\theta} \leq p\left(\frac{\phi}{\gamma+\delta}\right) - \frac{\gamma-r}{r+\delta}\left(p(\widehat{w}) - p\left(\frac{\phi}{\gamma+\delta}\right)\right).$$

Appendix B: Proof for Proposition 2

We need to verify that under the proposed scheme, the agent's value function given the relevant state (M, K) is $V(M, K) = \lambda M$, and his optimal policy is the one we obtained in Proposition 1. We take the "guess and verify" approach. Under the proposed implementation, the evolution of the cash reserve M is

$$
\begin{aligned}
dM_t &= rM_t dt + dY_t - dD_t - dX_t \\
&= rM_t dt + K_t dA_t - K_t c(i_t) dt - [K_t \mu dt - K_t c(i_t) dt - (\gamma - r) M_t dt] - dX_t \\
&= \gamma M_t dt + K_t (dA_t - \mu dt) - dX_t,
\end{aligned}
$$

and $dK_t = I_t dt - \delta K_t dt$. Given the agent's value function $V(M, K) = \lambda M$, we have his HJB equation as

$$
\gamma \lambda M dt = \sup_{a_t \in [0,1], dX_t \geq 0} \lambda(\gamma M_t dt + K_t(a_t \mu dt - \mu dt) - dX_t) + \lambda K_t (1 - a_t) \mu dt + \lambda dX_t,
$$

where the first term is $\mathbb{E}_t[V_M(M, K) dM]$ under the policy a_t and dX_t, the second term is the agent's private benefit by exerting a_t, and the third term is his portion of special dividend. Then

$$
\begin{aligned}
\gamma \lambda M dt &= \sup_{a_t \in [0,1], dX_t} \lambda(\gamma M dt + K(a_t \mu dt - \mu dt) - dX) + \lambda K(1 - a_t) \mu dt + \lambda dX \\
&= \gamma \lambda M dt,
\end{aligned}
$$

where the action choice a_t drops out because of the binding incentive compatibility constraint. This proves that the agent's value function and his optimal policy under the proposed implementation coincide with those in Proposition 1. In particular, as the agent is indifferent between investment policies, he will follow the optimal investment policy we derived in Proposition 1.

Now we verify that the firm's equity value satisfies $S_t = (p(\lambda m_t) + m_t) K_t$. Since

$$
\begin{aligned}
S_t &= \mathbb{E}_t \left[\int_t^\tau e^{-r(s-t)} (dD_s + (1 - \lambda) dX_s) + e^{-r(\tau-s)} l K_s \right] \\
&= \mathbb{E}_t \left[\int_t^\tau e^{-r(s-t)} (dD_s + dX_s - \lambda dX_s) + e^{-r(\tau-s)} l K_s \right]
\end{aligned}
$$

and $dD_s + dX_s = rM_s dt + dY_s - dM_s$, we have

$$
S_t = \mathbb{E}_t \left[\int_t^\tau e^{-r(s-t)} (dY_s - \lambda dX_s) + e^{-r(\tau-t)} l K_\tau + e^{-r(s-t)} [rM_s dt - dM_s] \right]. \quad \text{(B1)}
$$

Now, because $\lambda dX_s = dU_s$, according to the definition of $P(W_t, K_t)$ we know that

$$
S_t = P(W_t, K_t) + \mathbb{E}_t \left[\int_t^\tau e^{-r(s-t)} [rM_s ds - dM_s] \right].
$$

Using integration by parts,

$$
\int_t^\tau e^{-r(s-t)} dM_s = e^{-r(\tau-t)} M_\tau - M_t + \int_t^\tau e^{-r(s-t)} rM_s ds.
$$

Since $M_\tau = 0$, the second part in (B1) is M_t, and therefore $S_t = P(W_t, K_t) + M_t = (p(\lambda m_t) + m_t)K_t$.   Q.E.D.

Appendix C: Renegotiation-Proof Contract

As we have indicated in Section III.A, our contract may not be renegotiation-proof. Intuitively, whenever $p'(w) > 0$ (or, $P_W(W, K) > 0$), both parties may achieve an ex post Pareto-improving allocation by renegotiating the contract. Therefore, the value function $p(w)$ that is renegotiation-proof must be weakly decreasing in the agent's scaled continuation payoff w. Moreover, $p^{RP}(w)$ has an (endogenous) renegotiation boundary $\underline{w}^{RP}$, where the scaled investors' value function $p^{RP}(w)$ has the following boundary conditions:

$$p^{RP}(\underline{w}^{RP}) = l, \tag{C1}$$

$$p^{RP\prime}(\underline{w}^{RP}) = 0. \tag{C2}$$

Specifically, $\underline{w}^{RP}$ (rather than $w = 0$ in the baseline model) becomes the lower bound for the agent's scaled continuation payoff w during the equilibrium employment path. The scaled investors' value function $p^{RP}(w)$ solves the ODE (18) for $w \in [\underline{w}^{RP}, \overline{w}^{RP}]$, with two sets of free-boundary conditions: the boundary conditions (16) and (20) at the payout boundary $\overline{w}^{RP}$, and the boundary conditions (C1) and (C2) at the renegotiation boundary $\underline{w}^{RP}$.

The dynamics of the scaled agent's payoff w takes the following form:

$$dw_t = (\gamma + \delta - i(w_t))w_t dt + \lambda \sigma dZ_t - du_t + (d\upsilon_t - \underline{w}^{RP} dM_t), \tag{C3}$$

where the first (drift) term implies that the expected rate of change for the agent's scaled continuation payoff w is $(\gamma + \delta - i(w))$, the second (diffusion) term captures incentive provisions in the continuation-payoff region (away from the boundaries), and the third term, the nondecreasing process u, captures the reflection of the process w at the upper payment boundary $\overline{w}^{RP}$. Unlike the dynamics (17) for the agent's scaled payoff process w without renegotiation, the last term $d\upsilon_t - \underline{w}^{RP} dM_t$ in the dynamics (C3) captures the effect at the renegotiation boundary. The nondecreasing process υ reflects w at the renegotiation boundary $\underline{w}^{RP}$. The intensity of the counting process dQ is $d\upsilon_t/\underline{w}^{RP}$; once $dQ = 1$, w becomes zero, and the firm is liquidated.[24] Note that the additional term $d\upsilon_t - \underline{w}^{RP} dQ_t$ is a compensated Poisson process and hence a martingale increment.

We illustrate the contracting behavior at the renegotiation boundary through the following intuitive way. When the agent's poor performance drives w down to $\underline{w}^{RP}$, the two parties run a lottery. With a probability of $d\upsilon_t/\underline{w}^{RP}$, the firm is liquidated. If the firm is not liquidated, the agent stays at the renegotiation

---

[24] Technically speaking, the counting process has a survival probability $\Pr(Q_t = 0) = \exp(-\upsilon_t/\underline{w}^{RP})$.
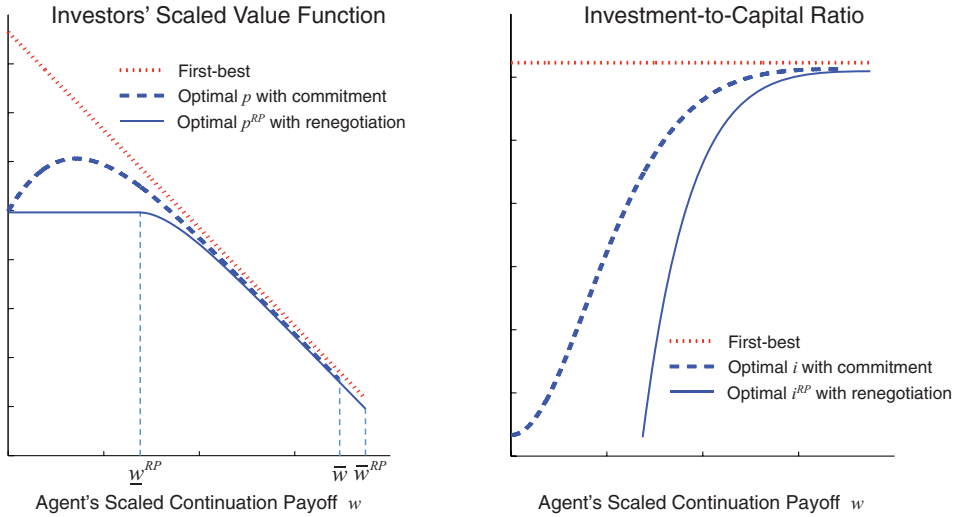
Figure C1. Renegotiation proofness. The original scaled value function $p(w)$ is not renegotiation-proof because $p'(0) > 0$. For the renegotiation-proof contract, $\underline{w}^{RP}$ is the lower bound for the agent's scaled continuation payoff w, with the following properties: $p(\underline{w}^{RP}) = p(0) = l$ and $p'(\underline{w}^{RP}) = 0$. The value function $p^{RP}(w)$ solves the ODE (18) subject to the boundary conditions (16) and (20) and the above-stated conditions at $\underline{w}^{RP}$.

boundary $\underline{w}^{RP}$. Here, the stochastic termination is to achieve the "promise-keeping" constraint so that w is indeed the scaled continuation payoff with expected growth rate $\gamma + \delta - i(w)$ as specified in Proposition 1. To see this, by running this lottery, the agent could potentially lose $(dv_t/\underline{w}^{RP}) \cdot \underline{w}^{RP} = dv_t$, which just compensates the reflection gain dv_t if the firm is not liquidated.

Since renegotiation further worsens the agency conflict, we expect not only a greater value reduction for investors, but also a stronger underinvestment distortion. Figure C1 shows this result.

Appendix D: Technical Details for Section V

D.1. Characterization of Optimal Contracting with Stochastic Profitability

Fix regime L as the current regime (similar results hold for regime H upon necessary relabelling.) Based on (31) and (34) in Section V, the following Bellman equation holds for $P(K, W, \pi^L)$:

$$
\begin{aligned}
r P(K, W, \pi^L) = \sup_{I,\,\Psi} \; & (\mu \pi^L K - I - G(I, K)) + (I - \delta K) P_K \\
& + (\gamma W - \Psi_{LH}(K, W)\xi^L) P_W + \frac{\lambda^2 \sigma^2 K^2}{2} P_{WW} \\
& + \xi^L \big(P(K, W + \Psi_{LH}(K, W), \pi^H) - P(K, W, \pi^L) \big), \quad \text{(D1)}
\end{aligned}
$$

where investment I and the compensation adjustment $\Psi_{LH}(K, W)$ (both un-scaled) are state-dependent controls.

The investment policy $I(K, W, n)$, by taking a first-order condition (FOC), is similar to the baseline case. The FOC for optimal $\Psi_{LH}(K, W)$, given that the solution takes an interior solution, yields that

$$P_W(K, W, \pi^L) = P_W(K, W + \Psi_{LH}(K, W), \pi^H). \tag{D2}$$

As discussed in the main text, the optimal contract equates the marginal cost of delivering compensation, that is, $-P_W$, across different Markov states at any time. However, in general, the solution of $\Psi_{nm}(K, W)$ might be binding (corner solution), as the agent's continuation payoff after the regime change has to be positive. Therefore, along the equilibrium path the optimal $\Psi_{nm}(K, W)$ might bind, that is, $\Psi_{nm}(K, W) + W \geq 0$ holds with equality.

The scale invariance remains: with $w = W/K$, we let $p_n(w) = P(K, W, \pi^n)/K$, $i_n(w) = I(K, W, \pi^n)/K$, $\psi_{nm}(w) = \Psi_{nm}(K, W, n)/K$, and upper payment boundary $\overline{w}_n = \overline{W}(K, \pi^n)/K$. Similar to equation (19),

$$i_n(w) = \frac{P_K(K, W, n) - 1}{\theta} = \frac{p_n(w) - wp'_n(w) - 1}{\theta}. \tag{D3}$$

Combining this result with the above analysis regarding $\psi_{nm}(w)$'s (notice that $P_W(K, W, \pi^n) = p'_n(w)$), the following proposition characterizes the ODE system $\{p_n\}$ when profitability is stochastic.

PROPOSITION 4: *For $0 \leq w \leq \overline{w}_n$ (the continuation-payoff region for regime n), the scaled investor's value function $p_n(w)$ and the optimal payment threshold $\overline{w}_n$ solve the following coupled ODEs:*

$$
\begin{aligned}
(r + \delta)p_L(w) = {} & \mu\pi^L + \frac{(p_L(w) - wp'_L(w) - 1)^2}{2\theta} \\
& + p'_L(w)[(\gamma + \delta)w - \xi^L\psi_{LH}(w)] + \frac{\lambda^2\sigma^2}{2}p''_L(w) \\
& + \xi^L(p_H(w + \psi_{LH}(w)) - p_L(w)), \quad 0 \leq w \leq \overline{w}_L,
\end{aligned}
$$

$$
\begin{aligned}
(r + \delta)p_H(w) = {} & \mu\pi^H + \frac{(p_H(w) - wp'_H(w) - 1)^2}{2\theta} \\
& + p'_H(w)[(\gamma + \delta)w - \xi^H\psi_{HL}(w)] + \frac{\lambda^2\sigma^2}{2}p''_H(w) \\
& + \xi^H(p_L(w + \psi_{HL}(w)) - p_H(w)), \quad 0 \leq w \leq \overline{w}_H
\end{aligned}
\tag{D4}
$$

subject to the boundary conditions at the upper boundary $\overline{w}_n$,

$$p'_n(\overline{w}_n) = -1, \tag{D5}$$

$$p''_n(\overline{w}_n) = 0, \tag{D6}$$

and the left boundary conditions at liquidation,

$$p_n(0) = l_n, \quad n = 1, 2.$$

The scaled endogenous compensation adjustment functions $\psi_{nm}(w)$ satisfy

$$p'_n(w) = p'_m(w + \psi_{nm}(w))$$

if $w + \psi_{nm}(w) > 0$ (interior solution); otherwise, $\psi_{nm}(w) = -w$. For $w > \overline{w}_n$ (cash payment regions), $p_n(w) = p_n(\overline{w}_n) - (w - \overline{w}_n)$.

Proof: See Appendix D2.

In solving the coupled ODEs in Proposition 4 with the optimal payment threshold $\overline{w}_n$ and the endogenous compensation adjustment functions $\psi_{nm}(w)$, we take the following numerical iteration steps:

(1) Solve for $p_L^{(0)}$ and $p_H^{(0)}$ without profitability changes (that is, setting ξ's to zero.) This essentially amounts to solving a nonlinear ODE with a free-boundary condition about $\overline{w}_n^{(0)}$.
(2) Given $p_L^{(0)}$ and $p_H^{(0)}$, obtain the compensation adjustment function $\psi_{HL}^{(0)}(w)$ based on Proposition 4.
(3) Given $\psi_{HL}^{(0)}(w)$ and $p_L^{(0)}$, calculate $p_H^{(1)}$ by solving the high-state one-dimensional nonlinear ODE (that is, the second equation) in equation (D4). The resulting solution is $p_H^{(1)}$ with the free boundary $\overline{w}_H^{(1)}$.
(4) Obtain $\psi_{LH}^{(0)}(w)$ based on $p_H^{(1)}$ and $p_L^{(0)}$, and then calculate $p_L^{(1)}$ by solving the low-state one-dimensional nonlinear ODE (that is, the first equation) in (D4). The resulting solution is $p_L^{(1)}$ with the free boundary $\overline{w}_L^{(1)}$.
(5) Given $p_L^{(1)}$ and $p_H^{(1)}$, update $\psi_{HL}^{(1)}(w)$.
(6) Repeat the procedures after Step 3 until convergence obtains. The convergence criterion is

$$\max \left[\sup_w \left(p_L^{(j+1)} - p_L^{(j)} \right), \quad \sup_w \left(p_H^{(j+1)} - p_H^{(j)} \right) \right] < 10^{-5}.$$

D.2. *Proof for Proposition 3*

LEMMA 3:   *Both p_n's are strictly concave for $0 \le w < \overline{w}_n$.*

Proof: Denote two states as n, m. By differentiating (D4), we obtain

$$(r + \delta) p'_n = -\frac{(p_n - w p'_n - 1) w p''_n}{\theta} + p''_n \cdot \left[(\gamma + \delta) w - \xi^n \psi_{nm}(w) \right]$$
$$+ p'_n (\gamma + \delta - \xi^n \psi'_{nm}(w)) + \frac{\lambda^2 \sigma^2}{2} p'''_n$$
$$+ \xi^n (p'_m(w + \psi_{nm}(w))(1 + \psi'_{nm}(w)) - p'_n).$$

Notice that, when $\psi_{nm}(w)$ takes an interior solution, $p'_m(w + \psi_{nm}(w)) = p'_n(w)$; otherwise, $\psi'_{nm}(w) = -1$. Either condition implies that

$$
(r + \delta)p'_n = -\frac{(p_n - wp'_n - 1)wp''_n}{\theta} + p''_n \cdot [(\gamma + \delta)w - \xi^n \psi_{nm}(w)] \\
+ p'_n(\gamma + \delta) + \frac{\lambda^2\sigma^2}{2}p'''_n,
$$
(D7)

which takes a similar form as in (A6). Similar to the argument in the proof of Proposition 2, we can show that $p''_n(\overline{w}_n - \epsilon) < 0$ and $p''_m(\overline{w}_m - \epsilon) < 0$.

Denote $\mu_n = \mu\pi^n$. Let $q_n(w) = p_n(w) - wp'_n(w)$, that is, the marginal q that captures the investment benefit. We have

$$
(r + \delta + \xi^n)q_n(w) = \mu_n + \frac{(q_n(w) - 1)^2}{2\theta} + \xi^m q_m(w + \psi_{nm}(w)) \\
+ (\gamma - r)wp'_n(w) + \frac{\lambda^2\sigma^2}{2}p''_n
$$

$$
(r + \delta + \xi^m)q_m(w + \psi_{nm}(w)) = \mu_m + \frac{(q_m(w + \psi_{nm}(w)) - 1)^2}{2\theta} + \xi^n q_n(w) \\
+ (\gamma - r)(w + \psi_{nm}(w))p'_m(w + \psi_{nm}(w)) \\
+ \frac{\lambda^2\sigma^2}{2}p''_m(w + \psi_{nm}(w)).
$$
(D8)

Recall that the first-best pair (q_n^{FB}, q_m^{FB}) solves the system

$$
(r + \delta + \xi^n)q_n^{FB} = \mu_n + \frac{\left(q_n^{FB} - 1\right)^2}{2\theta} + \xi^n q_m^{FB}
$$

$$
(r + \delta + \xi^m)q_m^{FB} = \mu_m + \frac{\left(q_m^{FB} - 1\right)^2}{2\theta} + \xi^m q_n^{FB}
$$

It is understood throughout the proof that, as in the Hayashi (1982) model with quadratic adjustment cost, (q_n, q_m) in equation (D8) and (q_n^{FB}, q_m^{FB}) take the smaller root in solving the quadratic equations.

Suppose that there exist some points so that p_n is convex. Pick the largest $\widetilde{w}$ such that $p''_n(\widetilde{w}) = 0$ but $p''_n(\widetilde{w} + \varepsilon) < 0$, and $p''_n(w) \leq 0$ for $w \in (\widetilde{w}, \overline{w}_n)$. If $\psi_{nm}(\widetilde{w})$ is interior, then let

$$
k = p'_n(\widetilde{w}) = p'_m(\widetilde{w} + \psi_{nm}(\widetilde{w})), \quad p''_n(\widetilde{w}) = p''_m(\widetilde{w} + \psi_{nm}(\widetilde{w}))(1 + \psi'_{nm}(\widetilde{w})) = 0.
$$

Clearly, if $p''_m(\widetilde{w} + \psi_{nm}(\widetilde{w})) = 0$, then

$$(r + \delta + \xi^n)q_n(\widetilde{w}) = \mu_n + \frac{(q_n(\widetilde{w}) - 1)^2}{2\theta} + \xi^m q_m(\widetilde{w} + \psi_{nm}(\widetilde{w}))$$
$$+ (\gamma - r)\widetilde{w}k$$

$$(r + \delta + \xi^m)q_m(\widetilde{w} + \psi_{nm}(\widetilde{w})) = \mu_m + \frac{(q_m(\widetilde{w} + \psi_{nm}(\widetilde{w})) - 1)^2}{2\theta}$$
$$+ \xi^n q_n(\widetilde{w}) + (\gamma - r)(\widetilde{w} + \psi_{nm}(\widetilde{w}))k.$$

Since a positive k in the above system will imply that $q_n > q_n^{FB}$ and $q_m > q_m^{FB}$, we must have $k < 0$. Thus, evaluating (A6) at the point $\widetilde{w}$, we obtain

$$\frac{\lambda^2 \sigma^2}{2} p'''_n(\widetilde{w}) = (r - \gamma)p'_n(\widetilde{w}) = (r - \gamma)k > 0.$$

This is inconsistent with the choice of $\widetilde{w}$ where $p''_n(\widetilde{w}) = 0$ but $p''_n(\widetilde{w} + \varepsilon) < 0$. Notice that the above argument applies to the case $p''_m(\widetilde{w} + \psi_{nm}(\widetilde{w})) > 0$.

The case of $1 + \psi'_{nm}(\widetilde{w}) = 0$ but $p''_m(w + \psi_{nm}(w)) < 0$ is ruled out easily: with $\psi_{nm}(\cdot) \in C^1$, $\psi_{nm}(\widetilde{w}) = -\psi_{mn}(\widetilde{w} + \psi_{nm}(\widetilde{w}))$ implies that $-1 = \psi'_{nm}(\widetilde{w}) = -\psi'_{mn}(\widetilde{w} + \psi_{nm}(\widetilde{w}))(1 + \psi'_{nm}(\widetilde{w})) = 0$, a contradiction.

Finally, consider the case where $\psi_{nm}(\widetilde{w})$ is binding at $-w$. Take the same approach; notice that in this case the points after regime switching are exactly zero. Therefore, the same argument applies, and $p_n(\cdot)$ is strictly concave over the whole domain $(0, \overline{w}_n)$.   Q.E.D.

Once the concavity of p_n's is established, the verification argument is similar to the baseline case in Appendix A.

Now we show the properties of the compensation adjustment functions. For simplicity, take $\xi^n = \xi^m = \xi$. We focus on $\psi_{HL}(w)$. Once $\psi_{HL}(w) < 0$ is shown, it immediately follows that $\psi_{LH}(w) > 0$.

Because $p_H(w) > p_L(w)$ while $p_L(0) = p_H(0) = l$, we know $p'_H(0) > p'_L(0)$, and $\psi_{HL}(w) = -w$ when w is small. The claim that $\psi_{HL}(w) = -w$ for w lower than a threshold w^c such that $p'_H(w^c) = p'_L(0)$ follows from the concavity of both p's. We focus on the area where the jump functions take interior solutions.

Consider $\Delta q(w) = q_H(w) - q_L(w + \psi_{HL}(w))$, which is the difference between marginal q's across the two points before and after a jump. This difference starts at zero (both $q = l$), and must be positive in the payment boundary, as eventually $q_L = p_L(w) + w < p_H(w) + w = q_H$. Notice that

$$\Delta q'(w) = -w p''_H(w) + (w + \psi_{HL}(w))(1 + \psi'_{HL}(w))p''_L(w + \psi_{HL}(w)),$$

and the slope becomes zero on the upper boundary $\overline{w}_H$. Focus on w's that are below the payment region. When ψ takes interior solutions, $p'_H(w) = p'_L(w + \psi_{HL}(w))$ and $p''_H(w) = p''_L(w + \psi_{HL}(w))(1 + \psi'_{HL}(w))$. Therefore, $\Delta q'(w) = p''_H(w)\psi_{HL}(w)$ and $\psi_{HL}(w) = 0$ if and only if $\Delta q'(w) = 0$. Moreover, $\Delta q(w)$ is decreasing if and only if $\psi_{HL}(w)$ is positive.

The following lemma shows that $\psi_{HL}(\overline{w}_H) = \overline{w}_L - \overline{w}_H \le 0$ on the upper payment boundary; later on we show $\psi_{HL}(\overline{w}_H) < 0$ strictly.

LEMMA 4: *The upper payment boundary $\overline{w}_H \geq \overline{w}_L$ so that $\psi_{HL}(\overline{w}_H) = \overline{w}_L - \overline{w}_H \leq 0$.*

Proof: We prove by contradiction. Suppose that $\psi_{HL}(\overline{w}_H) = \overline{w}_L - \overline{w}_H > 0$. According to (D8), at the upper boundary we have $(r + \delta + \xi)\overline{q}_n = \mu_n + \frac{(\overline{q}_n - 1)^2}{2\theta} + \xi\overline{q}_m - (\gamma - r)\overline{w}_n$. Therefore,

$$(r + \delta + 2\xi)(\overline{q}_H - \overline{q}_L) = \mu_H - \mu_L + \frac{(\overline{q}_H - \overline{q}_L)(\overline{q}_H + \overline{q}_L + 2)}{2\theta} \qquad (D9)$$
$$+ (\gamma - r)(\overline{w}_L - \overline{w}_H).$$

Moreover, $\psi_{HL}(\overline{w}_H) > 0$ implies that $\Delta q(w)$ is decreasing for $w < \overline{w}_H$ around its vicinity. Take $\widehat{w}$, which is the largest w such that $\psi_{HL}(\widehat{w}) = 0$ and $\Delta q'(\widehat{w}) = 0$. (The existence of such $\widehat{w}$ follows from the fact $\psi_{HL}(0+) < 0$.) This implies that $\Delta q(\widehat{w})$ reaches its local maximum and $\Delta q(\widehat{w}) > \overline{q}_H - \overline{q}_L$, and $\psi'_{HL}(\widehat{w}) \geq 0$. However, denoting $q_L(\widehat{w}) = \widehat{q}_L$ and $q_H(\widehat{w}) = \widehat{q}_H$, (D8) implies that $(r + \delta + \xi)\widehat{q}_n = \mu_n + \frac{(\widehat{q}_n - 1)^2}{2\theta} + \xi\widehat{q}_m + (\gamma - r)\widehat{w}p'_n(\widehat{w}) + \frac{\lambda^2\sigma^2}{2}p''_n(\widehat{w})$, which says that

$$(r + \delta + 2\xi)\Delta q(\widehat{w}) = \mu_H - \mu_L + \frac{\Delta q(\widehat{w})(\widehat{q}_L + \widehat{q}_H + 2)}{2\theta}$$
$$+ \frac{\lambda^2\sigma^2}{2}[p''_H(\widehat{w}) - p''_L(\widehat{w})]. \qquad (D10)$$

Comparing to (D9), since $\widehat{q}_L + \widehat{q}_H < \overline{q}_L + \overline{q}_H$ as q is increasing with w, in order for $\Delta q(\widehat{w}) > \overline{q}_H - \overline{q}_L$ the last term in (D10) must be strictly positive, or $p''_H(\widehat{w}) > p''_L(\widehat{w})$. But recall that $p''_H(\widehat{w}) = p''_L(\widehat{w})(1 + \psi'_{HL}(\widehat{w}))$; concavity of p_n's and $\psi'_{HL}(\widehat{w}) \geq 0$ yields a contradiction. Q.E.D.

Now we proceed to show that $\psi_{HL}(w) \leq 0$ always. Since $\psi_{HL}(\overline{w}_H) \leq 0$ and $\psi_{HL}(0+) < 0$, for any w such that $\psi_{HL}(w) > 0$, we can always find two points $\widehat{w} < \widetilde{w}$ closest to w such that $\psi_{HL}(\widehat{w}) = \psi_{HL}(\widetilde{w}) = 0$, $\Delta q'(\widehat{w}) = \Delta q'(\widetilde{w}) = 0$, $\Delta q(\widehat{w}) > \Delta q(\widetilde{w})$, $\psi'_{HL}(\widehat{w}) > 0$, and $\psi'_{HL}(\widetilde{w}) < 0$. We have

$$(r + \delta + 2\xi)\Delta q(\widehat{w}) = \mu_H - \mu_L + \frac{\Delta q(\widehat{w})(\widehat{q}_L + \widehat{q}_H + 2)}{2\theta} + \frac{\lambda^2\sigma^2}{2}p''_L(\widehat{w})\psi'_{HL}(\widehat{w})$$
$$< \mu_H - \mu_L + \frac{\Delta q(\widehat{w})(\widehat{q}_L + \widehat{q}_H + 2)}{2\theta}$$

and $(r + \delta + 2\xi)\Delta q(\widetilde{w}) = \mu_H - \mu_L + \frac{\Delta q(\widetilde{w})(\widetilde{q}_L + \widetilde{q}_H + 2)}{2\theta} + \frac{\lambda^2\sigma^2}{2}p''_L(\widetilde{w})\psi'_{HL}(\widetilde{w}) > \mu_H - \mu_L + \frac{\Delta q(\widetilde{w})(\widetilde{q}_L + \widetilde{q}_H + 2)}{2\theta}$. Finally, because $\widehat{q}_L + \widehat{q}_H < \widetilde{q}_L + \widetilde{q}_H$, $\Delta q(\widehat{w}) < \Delta q(\widetilde{w})$, a contradiction.

A similar argument rules out the case that there $\exists \widehat{w}$, such that $\psi_{HL}(\widehat{w}) = \psi'_{HL}(\widehat{w}) = 0$, while $\psi_{HL}(w) \leq 0$ for all w. Suppose not. Then $\Delta q'(\widehat{w}) = 0$ and

$$(r + \delta + 2\xi)\Delta q(\widehat{w}) = \mu_H - \mu_L + \frac{\Delta q(\widehat{w})(\widehat{q}_L + \widehat{q}_H + 2)}{2\theta}.$$

For $w = \widehat{w} + \epsilon$, after neglecting some terms with an order higher than ϵ^2 (note that $\psi_{HL}(\widehat{w} + \epsilon)$ is in the order of ϵ^2)

$$(r + \delta + 2\xi)\Delta q(\widehat{w} + \epsilon) = \mu_H - \mu_L + \frac{\Delta q(\widehat{w} + \epsilon)(q_L(\widehat{w} + \epsilon) + \widehat{q}_H(\widehat{w} + \epsilon) + 2)}{2\theta}$$

$$+ (\gamma - r)\psi_{HL}(\widehat{w} + \epsilon)p_H'(\widehat{w} + \epsilon)$$

$$+ \frac{\lambda^2 \sigma^2}{2}p_L''(\widehat{w} + \epsilon)\psi_{HL}'(\widehat{w} + \epsilon).$$

The first term in the second line is in a lower order than ϵ, as $\psi_{HL}(\widehat{w}) = \psi_{HL}'(\widehat{w}) = 0$. The second term is positive. Because $q_L(\widehat{w} + \epsilon) + \widehat{q}_H(\widehat{w} + \epsilon) - (\widehat{q}_L + \widehat{q}_H)$ is in the order of ϵ, this contradicts $\Delta q'(\widehat{w}) = 0$.

Finally, we rule out the case of $\overline{w}_H = \overline{w}_L$ so $\psi_{HL}(\overline{w}) = 0$. At $\overline{w}$, we have $\psi_{HL}'(\overline{w}) = 0$, because $p_H''(w) = p_L''(w)(1 + \psi_{HL}'(\overline{w}))$ and $p_L'''(w) = p_H'''(w) = \frac{2(\gamma - r)}{\lambda^2 \sigma^2}$. Now consider the point $\overline{w} - \epsilon$:

$$(r + \delta + 2\xi)\Delta\overline{q}(\overline{w} - \epsilon) = \mu_H - \mu_L + \frac{\Delta\overline{q}(\overline{w} - \epsilon)(\overline{q}_L(\overline{w} - \epsilon) + \overline{q}_H(\overline{w} - \epsilon) + 2)}{2\theta}$$

$$+ (\gamma - r)\psi_{HL}(\overline{w} - \epsilon)p_H'(\overline{w} - \epsilon)$$

$$+ \frac{\lambda^2 \sigma^2}{2}p_L''(\overline{w} - \epsilon)\psi_{HL}'(\overline{w} - \epsilon).$$

One can show that, compared to the value at $\overline{w}$, $\overline{q}_L(\overline{w} - \epsilon) + \overline{q}_H(\overline{w} - \epsilon)$ is ϵ^2 order smaller. Again, the first term in the second line is in a lower order, and the second term is negative. Therefore, $\Delta\overline{q}(\overline{w} - \epsilon)$ should be ϵ^2 order smaller. However,

$$\Delta\overline{q}''(\overline{w}) = p_H'''(\overline{w})\psi_{HL}(\overline{w}) + p_H''(\overline{w})\psi_{HL}'(\overline{w}) = 0,$$

a contradiction. Q.E.D.

REFERENCES

Abel, Andrew B., and Janice C. Eberly, 1994, A unified model of investment under uncertainty, *American Economic Review* 84, 1369–1384.

Albuquerque, Rui, and Hugo Hopenhayn, 2004, Optimal lending contracts and firm dynamics, *Review of Economic Studies* 71, 285–315.

Atkeson, Andrew, 1991, International lending with moral hazard and risk of repudiation, *Econometrica* 59, 1069–1089.

Bertrand, Marianne, and Sendhil Mullainathan, 2001, Are CEOs rewarded for luck? The ones without principles are, *Quarterly Journal of Economics* 116, 901–932.

Biais, Bruno, Thomas Mariotti, Guillaume Plantin, and Jean-Charles Rochet, 2007, Dynamic security design: Convergence to continuous time and asset pricing implications, *Review of Economic Studies* 74, 345–390.

Biais, Bruno, Thomas Mariotti, Jean-Charles Rochet, and Stéphane Villeneuve, 2010, Large risks, limited liability, and dynamic moral hazard, *Econometrica* 78, 73–118.

Bolton, Patrick, Hui Chen, and Neng Wang, 2011, A unified theory of Tobin's q, corporate investment, financing, and risk management, *Journal of Finance* 66, 1545–1578.

Bolton, Patrick, and David Scharfstein, 1990, A theory of predation based on agency problems in financial contracting, *American Economic Review* 80, 93–106.

Clementi, Gian Luca, and Hugo A. Hopenhayn, 2006, A theory of financing constraints and firm dynamics, *Quarterly Journal of Economics* 121, 229–265.

DeMarzo, Peter, and Michael Fishman, 2007a, Agency and optimal investment dynamics, *Review of Financial Studies* 20, 151–188.

DeMarzo, Peter, and Michael Fishman, 2007b, Optimal long-term financial contracting, *Review of Financial Studies* 20, 2079–2128.

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

Diamond, Douglas W., 1984, Financial intermediation and delegated monitoring, *Review of Economic Studies* 51, 393–414.

Dixit, Avinash, 1993, *The Art of Smooth Pasting*, Vol. 55 in Fundamentals of Pure and Applied Economics (Harwood Academic Publishers, Routledge, NY).

Eberly, Janice C., Sergio Rebelo, and Nicolas Vincent, 2009, Investment and value: A neoclassical benchmark, Working paper, Northwestern University.

Fazzari, Steven, Glenn Hubbard, and Bruce Petersen, 1988, Financing constraints and corporate investment, *Brookings Papers on Economic Activity* 1, 141–195.

Froot, Kenneth A., David Scharfstein, and Jeremy Stein, 1993, Risk management: Coordinating corporate investment and financing policies, *Journal of Finance* 48, 1629–1658.

Gilchrist, Simon, and Charles Himmelberg, 1998, Investment, fundamentals, and finance, in Ben S. Bernanke and Julio J. Rotemberg, eds.: *NBER Macroeconomics Annual* (MIT Press, Cambridge, MA).

Gomes, Joao F., 2001, Financing investment, *American Economic Review* 91, 1263–1285.

Green, Edward J., 1987, Lending and the smoothing of uninsurable income, in Edward Prescott and Nancy Wallace, eds.: *Contractual Arrangements for Intertemporal Trade* (University of Minnesota Press, Minneapolis, MN).

Gromb, Denis, 1999, Renegotiation in debt contracts, Working paper, MIT.

Hayashi, Fumio, 1982, Tobin's marginal q and average q: A neoclassical interpretation, *Econometrica* 50, 213–224.

He, Zhiguo, 2009, Optimal executive compensation when firm size follows geometric Brownian motion, *Review of Financial Studies* 22, 859–892.

Hennessy, Christopher A., and Toni M. Whited, 2007, How costly is external financing? Evidence from a structural estimation, *Journal of Finance* 62, 1705–1745.

Hubbard, Glenn, 1998, Capital markets imperfections and investment, *Journal of Economic Literature* 36, 193–225.

Jenter, Dirk, and Fadi Kanaan, 2011, CEO turnover and relative performance evaluation, *Journal of Finance* (forthcoming).

Kaplan, Steven, and Luigi Zingales, 1997, Do investment-cash flow sensitivities provide useful measures of financing constraints? *Quarterly Journal of Economics* 112, 169–215.

Ljungqvist, Lars, and Thomas J. Sargent, 2004, *Recursive Macroeconomic Theory* (MIT Press, Cambridge, MA).

Lorenzoni, Guido, and Karl Walentin, 2007, Financial frictions, investment and Tobin's q, Working paper, MIT.

Lucas, Robert E. Jr., and Edward C. Prescott, 1971, Investment under uncertainty, *Econometrica* 39, 659–681.

Panousi, Vasia, and Dimitris Papanikolaou, 2012, Investment, idiosyncratic risk, and ownership, *Journal of Finance* 67, 1113–1148.

Phelan, Christopher, and Robert Townsend, 1991, Computing multi-period, information constrained optima, *Review of Economic Studies* 58, 853–881.

Philippon, Thomas, and Yuliy Sannikov, 2007, Real options in a dynamic agency model, with applications to financial development, IPOs, and business risk, Working paper, NYU Stern School.

Piskorski, Tomasz, and Alexei Tchistyi, 2010, Optimal mortgage design, *Review of Financial Studies* 23, 3098–3140.

Quadrini, Vincenzo, 2004, Investment and liquidation in renegotiation-proof contracts with moral hazard, *Journal of Monetary Economics* 51, 713–751.

Rampini, Adriano, and S. Viswanathan, 2010, Collateral, risk management, and the distribution of debt capacity, *Journal of Finance* 65, 2293–2322.

Rampini, Adriano, and S. Viswanathan, 2011, Collateral and capital structure, Working paper, Duke University.

Sannikov, Yuliy, 2007, Agency problems, screening, and increasing credit lines, Working paper, Princeton University.

Schmid, Lukas, 2008, A quantitative dynamic agency model of financing constraints, Working paper, Duke University.

Spear, Stephen, and Sanjay Srivastava, 1987, On repeated moral hazard with discounting, *Review of Economic Studies* 54, 599–617.

Stein, Jeremy, 2003, Agency, investment and corporate investment, in George Constantinides, Milton Harris, and René Stulz, eds.: *Handbook of the Economics of Finance* (Elsevier, North Holland).

Tchistyi, Alexei, 2005, Security design with correlated hidden cash flows: The optimality of performance pricing, Working paper, University of California, Berkeley.