Somebody Stop Me: The Asset Pricing Implications of Principal-Agent Conflicts*

Juan Carlos Zelaya-Mejía†

February 2, 2023

Abstract

I show that, in a DSGE setting with heterogeneous shareholders, the stock market risk premium and volatility decrease as monitoring increases. Monitoring arises because inside shareholders have an incentive to extract private benefits from firm output, whereas outside shareholders have the incentive to limit this extraction. Monitoring varies positively with the share of outside shareholder ownership, such that monitoring represents a source of cross-sectional and time-series variation in equilibrium asset pricing moments. I present empirical evidence supporting these theoretical results.

*For comments and advice, I would like to thank my advisor Paul Ehling, Patrick Konermann, and Adam Winegar. Any remaining mistakes and opportunities for improvement are my own.

†BI Norwegian Business School

1 Introduction

What are the effects on equilibrium asset pricing moments from costly monitoring that mitigates the principal-agent problem? To answer this question, I build on the seminal paper by Albuquerque and Wang (2008). The authors develop one of the first dynamic stochastic general equilibrium asset pricing models to study the principal-agent problem in an asset pricing setting, focusing on the implications of imperfect investor protection. Monitoring is a predictable strategic response to the principal-agent problem and is therefore the natural next step in the study of the importance of principal-agent conflicts to asset pricing. Furthermore, monitoring is not expendable even in the presence of investor protection. While it is intuitive that investor protection imperfectly substitutes monitoring, it is also the case that investor protection *complements* monitoring. For example, hostile takeovers act as a market disciplinary mechanism (Manne, 1965)—i.e., a monitoring mechanism—which increases (decreases) with more (less) investor protection in the form of legislation that favors hostile takeovers.

Similar to Albuquerque and Wang (2008), my setting is a continuous time, production based general equilibrium asset pricing model with investment-specific technology shocks. The economy has one representative firm and two heterogenous agents: an inside shareholder who controls investment, dividends, and his extraction of private benefits, and an outside shareholder who restricts the insider's private benefit extraction by monitoring him. Both the insider and the outsider maximize their expected lifetime utility of consumption. I interchangeably refer to the inside shareholder as the insider (he) and to the outside shareholder as the outsider (she).

Monitoring reduces the insider's private benefit extraction by increasing its cost, but it also represents a pecuniary cost to the outsider. Furthermore, monitoring decreases in the insider's ownership and in investor protection. There are two reasons why monitoring decreases in the insider's ownership. First, the greater the insider's ownership, the more aligned his incentives are to those of the outsider. Second, when there are only two shareholders, the outsider's additional dividend cash flow from monitoring—i.e., her marginal pecuniary benefit from monitoring—decreases in the insider's ownership.

An important result of the model is that investment decreases in monitoring. Since the insider grows the firm through overinvestment when he can easily extract private benefits, it follows that monitoring improves the efficiency of capital allocation.[1] In addition, the corrective effect of monitoring on investment drives the asset pricing effects of monitoring. The effects are namely that Tobin's q increases in monitoring, and that the interest rate, the risk premium, volatility, and the dividend yield all decrease in monitoring.

As in Albuquerque and Wang (2008), investor protection restricts private benefit extraction by increasing its cost. Moreover, Tobin's q increases in investor protection, while the interest rate, the risk premium, volatility, and the dividend yield all decrease in investor protection.

The effects of the insider's ownership on investment and asset pricing moments, on the other hand, are non-monotonic. This is due to the non-monotonic effects of ownership on private benefit extraction and their transmission unto (over)investment. The non-monotonic effects of ownership on private benefit extraction arise because of monitoring. If the insider's ownership is too low, the outsider *overmonitors*, and the insider *rebels* against repressive controls by using increases in his ownership share to extract *more*—not less—private benefits. But when the insider's ownership is high enough, monitoring is 'relaxed' and the insider's private benefit extraction decreases in his share of ownership.

[1] As in Albuquerque and Wang (2008), I use the term overinvestment to refer to an investment-capital ratio that exceeds the ratio one would observe in a setting with perfect investor protection—and thus no private benefit extraction. An example of an empirical study that relates monitoring to improved capital allocation is the paper by Brav, Jiang, and Kim (2015). The authors find increases in labor productivity and in the productivity of divested plants following hedge fund activist engagements.

My empirical evidence is consistent with my model predictions. For instance, after the incidence of monitoring (the extensive margin), investment declines by up to 0.10 standard deviations, q increases by up to 0.19 standard deviations, expected returns decline by 0.15 standard deviations, and volatilities decline by up to 0.07 standard deviations.[2] The effect of monitoring on returns represents an annualized decline of approximately 11% in sample and is significant at the 1% level. Effects on investment, q, and volatilities persist for 1 year. Statistical significance obtains at the 1% level for investment and q, and at the 10% level for volatilities that or contemporaneous or up to two quarters ahead with respect to monitoring incidents.

I use Fact Set data on activism campaigns to construct empirical proxies for monitoring. My empirical measures of monitoring are (a) at the extensive margin, whether an activism campaign starts on a given firm-period, and (b) at the intensive margin, the number of campaign or monitoring days on a firm-period. To illustrate the definition of monitoring days, if a firm has two ongoing activism campaigns on a given quarter, and both campaigns last 60 calendar days, there are 120 monitoring days for that firm on that quarter.

Activism campaigns constitute a reasonable basis for empirical proxies for monitoring because they are attempts by non-managing shareholders to change the behavior of the firm's management for value-creating purposes.[3] For instance, common objectives of activism campaigns include direct engagement with management or with the Board of Directors, excluding management; voting against management, including through proxy fights; seeking increased board representation; and pursuing acquisi-

[2] I standardize variables with firm-quarter variation (investment, q, quarter-end GARCH(1,1) volatility estimates, and quarter-end dividend yields) using their cross-sectional means and standard deviations. Given that I study these variables in a setting that estimates cross-sectional variation, my approach facilitates an appropriate interpretation of the economic significance of monitoring in my regressions (Liu and Winegar, 2022).

[3] The extant literature defines both voice and exit as possible channels for monitoring, where activism campaigns fall under the former category. Monitoring by exit consists in selling a firm's shares because of discontent with the firm's performance and/or governance (Appel, Gormley, and Keim, 2019; McCahery, Sautner, and Starks, 2016).

tion by a third party (Brav, Jiang, Kim, and Partnoy, 2008; Klein and Zur, 2009; Gantchev, 2013; Bebchuck, Brav, and Jiang, and Keusch, 2020).[4]

My empirical proxy for investor protection is the Takeover Index by Cain, McKeon, and Solomon (2017). The index estimates the probability that a firm is acquired in a hostile takeover as a result of changes in state-level anti-takeover legislation. Recalling the intuition from Manne (1965), one may interpret the Takeover Index as a plausibly exogenous measure of investor protection—a representation of factors in the legal environment that improve firm governance.

This paper contributes to a strand of literature that studies the effects principal-agent conflicts on asset prices. Previous works in this literature focus on the importance of the investor protection channel. For instance, the model by Albuquerque and Wang (2008) links lower investment protection to higher investment, and further links higher investment to higher volatility and risk premia. Basak, Chabakauri, and Yavuz (2019) consider a similar model in which they endogenize the controlling shareholder's level of ownership. They propose that high investor protection may in fact be associated with higher volatility, conditional on the controlling shareholder having sufficiently high ownership.[5] My contribution to this literature is to introduce and analyze the monitoring channel.

A second contribution that this paper makes is to provide a novel definition of overmonitoring and study its effects. Seminal papers that propose the idea of excessive monitoring include Burkart, Gromb, and Panunzi (1997), as well as Pagano and Röell (1998). Burkart, Gromb, and Panunzi (1997) posit that ownership structure is an instrument to solve the trade-off between monitoring and managerial initiative. On the other hand, Pagano and Röell (1998) view ownership as a commitment device to

[4]To exemplify the cost of an aggressive activism campaign, Gantchev (2013) finds that the average proxy fight costs USD 10.71 million.

[5]Basak, Chabakauri, and Yavuz (2019) argue that when investor protection constrains private benefit extraction at a wide range of ownership levels, this induces higher ownership concentration by the controlling shareholder as a way to relax the constraint. This, in turn, increases the controlling shareholder's leverage, which increases stock return volatility.

limit agency costs, and they entertain the possibility that monitoring may be excessive only from the manager's point of view. This paper views overmonitoring from a general equilibrium perspective, where monitoring is excessive in that it distorts the incentives that ownership in the firm bestows on the inside shareholder.

Yet a third contribution of this paper is to provide novel empirical evidence on the asset pricing effects of monitoring—proxied here by activism campaigns. Recent, related papers on activism literature include Albuquerque, Fos, and Schroth (2022) and Chabakauri, Fos, and Jiang (2022). Albuquerque, Fos, and Schroth (2022) estimate the components of returns to activism using maximum likelihood estimation of a structural model. The authors define activism events as 13D filings and passive investments as those associated with 13G filings.[6] In this paper, neither monitoring nor the decision *not* to monitor necessarily entail a filing decision.[7] Chabakauri, Fos, and Jiang (2022) provide theory and evidence on insiders' strategic choice of firm ownership in anticipation of activism events. Specifically, insiders acquire more shares as a form of corporate defense. Although this paper treats insider ownership as exogenous, my conclusions about overmonitoring could inform an alternative model for insiders' accumulation of firm ownership ahead of activism events. Recall that overmonitoring and perverse incentives from inside ownership occur when the insider's ownership is too low. Thus, insiders might increase their ownership to prevent overmonitoring, consistent with their expected utility optimization.

The rest of the paper proceeds as follows: section II details the model, section III presents an analysis of equilibrium in the model, section IV discusses the asset pricing and investment implications of the equilibrium, section V presents empirical evidence

[6]The SEC mandates that investors who acquire an ownership stake above 5% in a company file either a Schedule 13D or a Schedule 13G. Investors who intend to engage in activism must file a Schedule 13D; otherwise, they can file a Schedule 13G.

[7]5,217 (52.879%) of activism campaigns in the Fact Set data set do not issue a 13D filing. One reason for observing campaigns that do not involve a 13D filing could be that activists engage larger firms and are thus not able to obtain ownership at the 5% threshold associated with 13D filings. However, firms with lower market capitalization are more common targets for activism (Brav, Jiang, Kim, and Partnoy, 2008; Becht, Franks, Grant, and Wagner, 2017).

for the model's predictions, and section VI concludes. The Appendix collects proofs and presents a welfare analysis.

2 The Model

I use a continuous time economy with agents who have infinite horizons. There is one representative firm—all firms are identical and subject to the same shocks—as well as a representative inside shareholder and a representative outside shareholder (alternatively, the insider and the outsider).

2.1 Setup

Firms only use equity financing and have a constant returns to scale production technology, $hK(t)$. Furthermore, shocks to production are investment-specific, so that the volatility of capital accumulation grows in proportion with investment. This is both analytically convenient and empirically supported (Greenwood, Hercowitz, and Krussell, 1997, 2000). Thus, capital accumulation takes place as follows:

$$dK(t) = (I(t) - \delta K(t))dt + \epsilon I(t)dZ(t), \tag{1}$$

where $I(t)$ is investment and $dZ(t)$ is a Brownian shock.

Since capital is an Itô process, so are dividends and prices:

$$dD(t) = \mu_D(t)D(t)dt + \sigma_D(t)D(t)dZ(t) \tag{2}$$

$$dP(t) = \mu_P(t)P(t)dt + \sigma_P(t)P(t)dZ(t) \tag{3}$$

Here, μ_D and μ_P are equilibrium drift processes, and σ_D and σ_P are equilibrium volatility processes. The risk-free asset is in zero net supply, and both the insider and the outsider may trade in it.

2.2 The Insider

The insider maximizes his lifetime utility of consumption,

$$E\left[\int_0^\infty e^{-\rho t}u(C_1(t))dt\right] \tag{4}$$

Let the insider have constant relative risk aversion:

$$u(C_1) = \frac{1}{1-\gamma}(C_1^{1-\gamma}-1),\ \gamma > 0 \tag{5}$$

Furthermore, the insider owns a fixed share of the firm, α. Because of consumption maximizing incentives, he extracts a fraction, $\chi(t)$, of the firm's production, $hK(t)$, as a private benefit. The insider also chooses dividends, $D(t)$, so that he trades off private benefits and dividends as income sources.[8] Given production and the insider's choice of private benefit extraction and dividends, investment is:

$$I(t) = hK(t)(1-\chi(t)) - D(t) \tag{6}$$

Private benefit extraction is costly because it is inefficient, such that inefficiency increases with the level of private benefit extraction. To illustrate, private benefit extraction may consist in the purchase of multiple vehicles or homes, so that the insider is never able to simultaneously draw utility from the entirety of these assets (he is not capable of being in more than one vehicle or more than one home simultaneously), and yet is fully responsible for the insurance, tax, and maintenance liabilities corresponding to his new assets. Thus, the cost of private benefit extraction is quadratic:

[8] There is no asymmetry of information, precluding concerns about the difference between formal and real authority to extract private benefits or to monitor (Aghion and Tirole, 1997).

$$\phi(\chi(t), hK(t)) = \frac{\eta}{2(\bar{\theta}-\theta)(1+\varsigma)}\chi(t)^2 hK(t)\ , \quad \chi(t) \in (0,1) \tag{7}$$

This expression for costs resembles that in Albuquerque and Wang (2008), with an important modification: the introduction of *monitoring*, θ, as a mechanism that moderates private benefit extraction. The level of monitoring is equal to the outsider's private cost of monitoring. Furthermore, monitoring constrains private benefit extraction by increasing its cost as monitoring increases towards the maximum possible level of monitoring, $\bar{\theta}$.[9] In the limit, as θ approaches $\bar{\theta}$, the cost of private benefit extraction reaches infinity, so that the insider abstains from extracting any private benefits. The same is true if investor protection, η, approaches infinity.

I also add the term $(1+\varsigma)$ to the insider's cost of private benefit extraction. ς is a reservation utility term that limits the proportion of cash flows that the outsider is willing use in monitoring costs. For this reason, the insider's cost of private benefit extraction is inversely proportional to ς.

The outsider has an incentive to monitor because her consumption increases in dividends and her utility is concave. Furthermore, the outsider's marginal benefit of monitoring (additional dividend cash flows) increases in her share of ownership.[10] The dividend-capital ratio dividend cash flows in the model. Later, I explain how the dividend-capital ratio increases in monitoring through reductions in both private benefit extraction and the investment-capital ratio.

Because of monitoring, the insider's ownership, α share can either involve *perverse incentives* or serve as an incentive alignment mechanism. This differs from the traditional view in the literature that ownership is an incentive alignment mechanism

[9] $\bar{\theta} \equiv \lim_{\alpha\to 0}\theta$; it is the maximum level of monitoring given the level of investor protection, η.

[10] In a setting with multiple shareholders, duplicate monitoring is useless (Pagano and Röell, 1998). Thus, only the majority outside shareholder—who has the highest marginal benefit from dividends—incurs monitoring costs. Specifically, the (majority) outside shareholder monitors the insider until her marginal benefit equals the marginal cost of monitoring. The remaining outside shareholder(s) "free-ride."

(Jensen and Meckling, 1976). As I explain in Section 3, increasing α when it is too low will incentivize the insider to extract *more*—not less—private benefits. But if insider ownership, α is high enough, increasing α aligns the insider's incentives. In particular, when the insider's ownership approaches 1, he begins to extract private benefits from his own claim to output (Jensen and Meckling, 1976).

The following expression summarizes my assumptions about private benefit extraction in general form:

$$\chi \equiv g(\theta(t),\ \alpha)\ , \quad \chi_{\theta}(\theta(t),\ \alpha) < 0 \ \ \chi_{\theta\theta}(\theta(t),\ \alpha) > 0$$

Private benefit extraction is a function of monitoring, θ, by outside shareholders, and of the insider's ownership. Additionally, private benefit extraction decreases in monitoring, but at an increasing rate, such that the marginal effectiveness of monitoring decreases at higher monitoring levels (no one can stop the insider from taking office pens home).

The insider's cash flow equation at time t is:

$$M_1(t) = \alpha(t)D(t) + \chi(t)hK(t) - \phi(\chi(t), hK(t)) \tag{8}$$

Assume that the insider can invest in the risk-free asset but is unable to trade the risky asset. Then, the insider's liquid wealth, equals his risk-free holdings: $W_1(t) = B_1(t)$, so his wealth evolves according to:

$$dW_1(t) = r(t)W_1(t) + M_1(t) - C_1(t) \tag{9}$$

Where $W_1(0) = 0$ and $r(t)$ is the equilibrium risk-free rate process.

Summing up, the insider maximizes his lifetime utility of consumption by choosing his level of private benefit extraction, the firm's dividends, and his consumption level at each period, subject to capital accumulation (1), the constraint for investment (6), his cash flow equation (8), his wealth dynamics (9), and a transversality condition that rules out hoarding of capital infinitely into the future (Appendix). The insider takes $r(t)$ as given.

2.3 The Outsider

The outside shareholder solves a consumption-asset allocation problem whereby she maximizes her lifetime utility. Let the outsider have the same form of preferences as the insider, and the same relative risk aversion. The insider's period t cash flow is:

$$M_2(t) = (1-\alpha)(1-[\theta/\bar{\theta}-\varsigma])D(t) \tag{10}$$

Given the outsider's cash flow equation, that the outsider has concave preferences, and that the insider trades off private benefits against dividends (6), the outsider decreases private benefit extraction through monitoring. Monitoring varies positively with the outsider's share of ownership, $1-\alpha$, because her marginal benefit (additional dividend cash flows) increases in her ownership. Another way to see this is that decreasing the outsider's share of ownership lowers the amount of monitoring because it necessarily raises the insider's ownership, α, which realigns the insider's incentives.[11]

To the extent that both monitoring, θ, and investor protection, η, increase the insider's cost of private benefit extraction, monitoring and investor protection function as substitutes. Thus, monitoring, θ, decreases in investor protection, η.

However, investor protection also complements monitoring. For example, hostile takeovers are a market disciplinary mechanism (Manne, 1965) that increases with investor protection, in the form of legislation that favors hostile takeovers. While such legislation makes it more costly for firms to defend themselves against hostile

[11] The intuition that increasing the insider's ownership, α, realigns the insider's incentives is valid only if α is high enough, as explained in Section 3.

takeovers, firms would never incur the costs of defense if hostile takeovers were not a credible threat. That is, investor protection does not perfectly substitute monitoring. In the model, the outsider incurs non-zero monitoring costs, $\theta/\bar{\theta}$, even under perfect investor protection, $\eta \to \infty$ (Section 3).

The general form for monitoring is thus:

$$\theta \equiv f(1-\alpha;\eta)\ , \quad \theta_{1-\alpha} > 0\ , \theta_\eta < 0$$

Which implies:

$$\frac{\partial \theta}{\partial \alpha} < 0$$

Without loss of generality, let θ have the following functional form:

$$\theta = 2\left(\frac{1-\alpha}{\eta}\right)^{1/2} \tag{11}$$

In Section 3, I show that when the insider's (outsider's) ownership is too low (high), the outsider *over*-monitors and consequently induces *perverse ownership incentives.* Specifically, when the outsider constrains private benefit extraction too much, increasing the insider's ownership share, α, leads to higher—not lower—private benefit extraction.

Letting $\pi(t)$ be the fraction of the outsider's liquid wealth that is invested in equity, and denoting the risk premium as $\lambda(t)$, such that $\lambda \equiv \mu_P(t) + Dt)/P(t) - r(t)$, the outside shareholder's wealth evolves according to:

$$\begin{aligned} dW_2(t) =& (r(t)W_2(t) + \pi(t)W_2(t)\lambda(t) - C_2(t) - (1-\alpha)[(\theta/\bar{\theta}) - \varsigma]D(t))dt + \\ & \sigma_P(t)\pi(t)W_2(t)dZ(t), \end{aligned} \tag{12}$$

where $W_2(0) = 0$.

To summarize, the outside shareholder chooses $\pi(t)$ and $C_2(t)$ each period to maximize her lifetime utility, subject to her cash-flow equation (10), her wealth dynamics (12), and a transversality condition (Appendix). The outsider takes dividends, firm value, and $r(t)$ as given.

2.4 Equilibrium

The properties that define the economy's equilibrium are as in Albuquerque and Wang (2008).

Definition 1 *An equilibrium has the following properties:*

1. $\{C_1(t),\ \chi(t),\ I(t),\ D(t):\ t \geq 0\}$ solve the controlling shareholder's problem for a given interest rate process $\{r(t):\ t \geq 0\}$
2. $\{C_2(t),\ \pi(t):\ t \geq 0\}$ solve each outside shareholder's problem given an interest rate process $\{r(t):\ t \geq 0\}$ and given stock price and dividend stochastic processes $\{P(t),\ D(t):\ t \geq 0\}$.
3. The risk-free asset market clears, such that $W_1(t) + (1 - \pi(t))W_2(t) = 0,\ \ \forall t$
4. The stock market clears, such that $1 - \alpha = \pi(t)W_2(t)/P(t),\ \ \forall t$
5. The consumption goods market clears, such that $C_1(t)+C_2(t)+I(t) = hK(t) - \phi(\chi(t), hK(t)),\ \ \forall t$

Let $\phi(\chi', hK)$ be the cost of private benefit extraction without any monitoring. In this economy, the total dead-weight loss is equal to the costs of private benefit extraction and monitoring, $\phi(\chi, hK) + (1 - \alpha)[(\theta/\bar{\theta}) - \varsigma]D(t)$, less reductions in the cost of private benefit extraction because of monitoring, $\phi(\chi', hK) - \phi(\chi, hK)$.

I follow Albuquerque and Wang (2008) in conjecturing and verifying a no-trade equilibrium. Thus, the following theorem characterizes the equilibrium:

Theorem 1 *Given assumptions 1-6 in the Appendix, there exists an equilibrium. The properties of this equilibrium are as follows. Outside shareholders invest all their liquid wealth in equity ($\pi(t) = 1$) and do not hold the risk-free asset. The outsider's consumption is equal to her entitled dividends net of monitoring costs:*

$$C_2(t) = (1-\alpha)(1-[(\theta/\bar{\theta})-\varsigma])D(t) \tag{13}$$

Insiders do not hold the risk-free asset either ($B_1(t) = 0$), and they divert a constant fraction of output:

$$\chi(t) = X \equiv (1-\alpha)\frac{(\bar{\theta}-\theta)(1+\varsigma)}{\eta} \tag{14}$$

The insider's consumption and the firm's investment and dividends are proportional to the firm's capital stock K(t), such that $C_1(t)/K(t) = M(t)/K(t) = m$, $I(t)/K(t) = i$, $D(t)/K(t) = d$. In equilibrium:

$$m = \alpha[h(1+(\bar{\theta}-\theta)(1+\varsigma)\psi) - i] \tag{15}$$

$$i = \left[\gamma(1+(1+(\bar{\theta}-\theta)(1+\varsigma)\psi)h\epsilon^2) - \sqrt{\Delta}\right]\left[\epsilon^2\gamma(\gamma+1)\right]^{-1}, \ s.t. \tag{16}$$

$$\Delta \equiv \gamma^2(1+(1+(\bar{\theta}-\theta)(1+\varsigma)\psi)h\epsilon^2)^2 - 2\epsilon^2\gamma(\gamma+1)\cdot ...$$
$$[(1+(\bar{\theta}-\theta)(1+\varsigma)\psi)h - (1-\gamma)\delta - \rho]$$

$$d = (1-X)h - i \tag{17}$$

such that $\psi = (1-\alpha)^2/(2\alpha\eta)$. In equilibrium, the processes for dividends (2), capital accumulation (1), and the stock price (3) follow geometric Brownian motions with drift and volatility coefficients as given below:

$$\mu_D = \mu_K = \mu_P = i - \delta \tag{18}$$

$$\sigma_D = \sigma_K = \sigma_P = i\epsilon \tag{19}$$

The equilibrium firm value is $P(t) = qK(t)$, where q is Tobin's q, which is equal to:

$$q = (1 - [\theta/\bar{\theta} - \varsigma])\left(1 + h(\bar{\theta} - \theta)(1 + \varsigma)\frac{1 - \alpha^2}{2\eta d\alpha}\right)^{-1}\left(\frac{1}{1 - \epsilon^2 i\gamma}\right) \tag{20}$$

The interest rate is:

$$r = \rho + \gamma(i - \delta) - \frac{\epsilon^2 i^2}{2}\gamma(\gamma + 1) \tag{21}$$

The parameter ψ partially summarizes agency costs in that it decreases in α (increasing the insider's ownership realigns incentives, given low enough monitoring) and investor protection. Additionally, as in Albuquerque and Wang (2008), q may be larger than 1 because of the $(1 - \epsilon^2 i\gamma)^{-1}$ term. Monitoring determines q through monitoring intensity, $(\theta/\bar{\theta})$, monitoring slack $(\bar{\theta} - \theta)$, and the effect of monitoring on investment and dividends. Investor protection also determines q directly and through its effect on investment and dividends.

3 Analysis of Equilibrium

3.1 Maximal Monitoring

Raising the outsider's ownership so that she reaches the maximal monitoring level, $\bar{\theta}$, increases the insider's cost of private benefit extraction prohibitively. He therefore abstains from extracting any private benefits.

Recall that only the outsider trades the risky asset. Maximal monitoring uses up nearly all of the outsider's dividend payout (save for the reservation utility term ς). Thus, under maximal monitoring, q is proportional to, and larger than ς. Specifically,

$$q^*(\theta) \equiv \lim_{\theta \to \bar{\theta}} q = \varsigma\left(1 - \epsilon^2 \gamma i^*(\theta)\right)^{-1}, \tag{22}$$

where $i^*(\theta)$ is the limit of i under maximal monitoring, $\lim_{\theta\to\bar{\theta}} i$. $i^*(\theta)$ is a constant equal to i after dropping all terms containing monitoring slack, $(\bar{\theta}-\theta)$, and agency costs, ψ. Dropping these terms has the effect of diminishing investment, which one can understand from the fact that investment, i, increases in agency costs, ψ (Section 4).

Maximal monitoring occurs when there is arbitrarily low insider ownsership, $\alpha \to 0$. Because of monitoring, q responds non-monotonically to changes in α. As I explain in Section 4, q can increase in insider ownership, α when there is overmonitoring (high α), but can decrease in α if there is relaxed monitoring (low α). This differs from the result in Albuquerque and Wang (2008), where q is strictly increasing in insider ownership, α.

Next, I show that q is higher under perfect investor protection than under maximal monitoring.

3.2 Perfect Investor Protection

Perfect investor protection—the limit as η approaches positive infinity—also fully constrains the insider's private benefit extraction. The outsider's monitoring cost, $\theta/\bar{\theta}$, is $\sqrt{1-\alpha} < 1$, i.e., it is lower than when she reaches maximal monitoring through higher ownership in the firm. Both of these results follow from L'Hôpital's rule. Appealing again to the intuition from the market for hostile takeovers, the reason why the outsider incurs non-zero monitoring costs even under perfect investor protection is that investor protection facilitates—but does not perfectly substitute—monitoring.

Under perfect investor protecion, q obtains the form:

$$q^*(\eta) = (1 - [(1-\alpha)^{1/2} - \varsigma])\frac{1}{1-\epsilon^2\gamma i^*(\eta)}, \tag{23}$$

where $i^*(\eta)$ is the limit of i under perfect investor protection, $\lim_{\eta\to\infty} i$. $i^*(\eta)$ is constant and equivalent to $i^*(\theta)$. Under perfect investor protection, q is larger than 1 if and only if $\epsilon^2 i^*(\eta)\gamma > (1-\alpha)^{1/2} - \varsigma$.

Notice that $q^*(\eta)$ and $q^*(\theta)$ differ from each other only in their numerator. It is thus easy to see that $q^*(\eta) > q^*(\theta)$; that is, q is higher under perfect investor protection than under maximal monitoring.

3.3 The Insider's Optimization Problem

Given that the insider holds zero risk-free bonds and cannot trade in the risky asset, $C_1(t) = M(t)$. Thus, the insider's optimization problem consists in trading off the costs and benefits of private benefits against dividends, and has the form:

$$\rho J_1(K) = \max_{D,\chi} \; u(M) + (I - \delta K)J_1'(K) + \frac{\epsilon^2}{2} I^2 J_1''(K) \tag{24}$$

subject to 6 and 8. That is, the insider solves a Hamilton-Jacobi-Bellman equation with one state variable (capital), where J_1 is the corresponding value function.[12]

The first order conditions with respect to D and χ are:

$$M^{-\gamma}\alpha - \epsilon^2 I J_1''(K) = J_1'(K) \tag{25}$$

$$M^{-\gamma}\left[hK - \frac{\chi h K \eta}{(\bar{\theta} - \theta)(1+\varsigma)}\right] - \epsilon^2 I J_1''(K) hK = J_1'(K) hK \tag{26}$$

Equation 25 is as in Albuquerque and Wang (2008). The marginal gain of dividends is an increase in consumption today (valued at $M^{-\gamma}\alpha$ and a reduction in the volatility of marginal utility (valued at $-\epsilon^2 I J_1''(K)$); and the marginal cost of dividends is a reduction in investment and future consumption (valued at $J_1'(K)$). Similarly, equation 26 shows that the marginal gain of private benefits is an increase in instantaneous

[12] The value function does not vary with respect to time because the agent has an infinite horizon and does not make decisions that are otherwise time-dependent.

consumption and a reduction in the volatility of marginal utility, and that the cost of private benefits is a reduction in investment and future consumption.

Given K, the marginal gain from private benefits is lower than in Albuquerque and Wang (2008). Monitoring decreases the value of instantaneous consumption through the *monitoring slack* term, $(\bar{\theta} - \theta)$. Monitoring also leads to smaller reductions in the volatility of marginal utility because i decreases in monitoring (Section 4).

Substituting 25 into 26 yields equilibrium private benefit extraction, $X \equiv \chi^* = (1-\alpha)(\bar{\theta} - \theta)(1+\varsigma)/\eta$. The expression says that private benefit extraction decreases in both monitoring and investor protection, as should be the case. Yet, for monitoring to have the intended effect on private benefit extraction, monitoring slack must be less than or equal to 1. Assumption 6 (Appendix) ensures that this is the case.

Furthermore, monitoring has to be sufficiently relaxed for ownership to realign the insider's incentives. Recall the traditional view that ownership is an incentive alignment mechanism (Jensen and Meckling, 1976). I show when this holds true and when monitoring introduces *perverse incentives* by taking partial derivatives of X with respect to α:

$$X_\alpha = \left[-\frac{\bar{\theta} - \theta}{\eta} + \frac{\theta}{2\eta} \right] (1+\varsigma), \quad s.t. \tag{27}$$

$$X_\alpha \lesseqgtr 0 \ \Leftrightarrow \theta \lesseqgtr \frac{2}{3}\bar{\theta} \tag{28}$$

$$X_{\alpha\alpha} = -3\frac{(1+\varsigma)}{\theta\eta^2} < 0 \tag{29}$$

By equations 28 and 29, private benefit extraction is concave in the insider's ownership. Denote $(2/3)\bar{\theta}$, the right hand side of equation 28, the *overmonitoring* threshold. With relaxed monitoring, i.e., when monitoring lies below the *overmonitoring* threshold, insider ownership, α, functions as an incentive realignment mechanism. That is, private benefit extraction, X, decreases in α. But with *overmonitoring*, i.e., when monitoring surpasses the *overmonitoring* threshold, perverse ownership incentives come into play: the inside shareholder uses increased ownership to *rebel* against repressive controls by extracting *more* (rather than less) private benefits.

Given the functional form for monitoring, over-monitoring and perverse incentives take place when the insider's ownership, α, is too low. Thus, the role of ownership structure is to solve the trade-off between outside control and managerial self-control.

Figure 1: Private benefit extraction as a function of excess monitoring, $\hat{\theta}$, and insider ownership, α. The region to the right of the hump in the middle of the surface depicts *overmonitoring*: increasing insider ownership, α, raises private benefit extraction. The region to the left of the hump portrays relaxed monitoring: increasing α reduces private benefit extraction. Parameter values are as given in Table 1. The expressions for private benefit extraction and excess monitoring are:

$$X \equiv \chi^* = (1-\alpha)\frac{(\bar{\theta}-\theta)(1+\varsigma)}{\eta}$$

$$\hat{\theta} = \theta - (2/3)\bar{\theta}$$

$$\hat{\theta} = 2\left(\frac{(1-\alpha)}{\eta}\right)^{1/2}$$

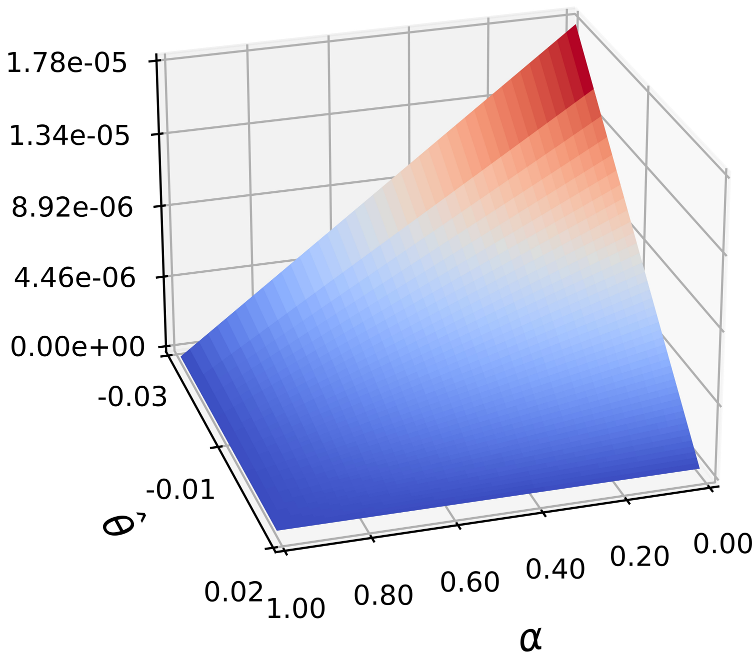


Figure 1 plots the private benefit extraction surface (Table 1 lists the parameter values I use to generate plots of model results). The bottom axes correspond to insider ownership, α, and excess monitoring, $\hat{\theta}$, where excess monitoring is the monitoring level, θ, net of the *overmonitoring* threshold, $(2/3)\bar{\theta}$. The private benefit extraction surface has the shape of a half-pyramid with a curved edge (alternatively, the surface resembles a large nose). *Overmonitoring* occurs to the right of the hump in the private benefit extraction surface. Recall that θ decreases in insider ownership, α; thus, excess monitoring, $\bar{\theta}$ also decreases in α. Starting with insider ownership, α, close to zero, increasing α traces a path *uphill* the private benefit extraction surface. Here is where the insider *rebels* because of *perverse incentives.* But to the left of the hump in the surface, there is relaxed monitoring (sufficently high insider ownership, α), so increasing α further traces a path *downhill* the private benefit extraction surface. This is where insider ownership, α, acts as an incentive alignment mechanism.

Table 1: Parameter values for plots of model results. I use the same subjective discount rate as Hansen and Singleton (1982), and ς is arbitrarily small. Otherwise, I use the parameter values employed and obtained in the calibration of the model by Albuquerque and Wang (2008) to US data.

Parameter	Variable	Value
Relative risk aversion	γ	2.0000
Subjective discount rate	ρ	0.0100
Investor protection	η	2325
Annual depreciation rate	δ	0.0800
Volatility of new investment	ϵ	0.3970
Capital productivity	h	0.1187
Outsider's reservation utility	ς	1.000e-07

3.4 The Outsider's Optimization Problem

I follow Albuquerque and Wang (2008) as well in conjecturing and verifying a constant interest rate and risk premium. Given constant α, monitoring is constant as well. This means that, as in Albuquerque and Wang (2008), the outside shareholder solves a Merton (1969) consumption-portfolio choice problem, such that:

$$\pi(t) = \frac{\lambda}{\gamma \sigma_P^2} \tag{30}$$

In addition, the no-trade equilibrium and $\pi = 1$ conditions yield:

$$\lambda = \gamma\sigma_P^2 = \gamma\epsilon^2 i^2 \tag{31}$$

The expression for the risk premium shows that a decreasing (increasing) relationship between the investment capital ratio, i, and monitoring (agency costs) implies a decreasing (increasing) relationship between the risk premium, λ, and monitoring (agency costs). I describe this in greater detail in Section 4. Additionally, note that if uncertainty, ϵ^2, or risk aversion, γ, are zero, the risk premium is zero.

3.5 Proportional Payoffs and No-Trade Equilibrium

While the outside shareholder's sole source of payoffs are the firm's dividends, the inside shareholder also obtains payoffs from net private benefits. Moreover, the insider's total payoff per share is proportional to the dividends that the outsider receives. This makes it so that the MRS for both agents is identical; therefore, both agents have the same risk attitudes and find it optimal not to trade with each other. This shows that the agents disagree only about the optimal investment level.

The equation below expresses the insider's payoff per share:

$$\frac{m}{\alpha}K = (d + ((\bar{\theta} - \theta)(1 + \varsigma)\psi + X)h)K \tag{32}$$

Hence, for each unit of dividends paid to the outsider, the insider gets $1 + h(\bar{\theta} - \theta)(1 + \varsigma)(1 - \alpha^2)/(2\alpha\eta d)$ units. This proportion is lower than if there were no monitoring.

As in Albuquerque and Wang (2008), the proportionality between the insider's payoffs and the outsider's payoffs yields identical growth rates of dividends and net payoffs to the inside shareholder between dates and states. Given $C_1(t) = M_1(t)$:

$$e^{-\rho(s-t)}\frac{U'(C_1(s))}{U'(C_1(t))} = e^{-\rho(s-t)}\left(\frac{M(s)}{M(t)}\right)^{-\gamma} = e^{-\rho(s-t)}\left(\frac{D(s)}{D(t)}\right)^{-\gamma} \tag{33}$$

And the outsider's MRS is the same:

$$e^{-\rho(s-t)}\frac{U'(C_2(s))}{U'(C_2(t))} = e^{-\rho(s-t)}\left(\frac{D(s)}{D(t)}\right)^{-\gamma} \tag{34}$$

Note that the outside shareholder also consumes her entire cash flow each period, but the terms in 10 in front of the dividend divide out, as they are constant. Note also that the identity between the MRS for the two agents exploits the fact that the two agents share the same level of relative risk aversion, γ. Finally, as Albuquerque and Wang (2008) point out, the linearity of the insider's net private benefits implies that the economy follows stochastic growth along a balanced path, so that it is ideal to scale variables scaled by capital, K.

4 Equilibrium Investment and Asset Pricing Implications

Monitoring and investor protection have same-sign impacts on investment, firm value, the interest rate, the risk premium, volatility, and the dividend yield. This is despite the fact that monitoring decreases in investor protection.

The insider's ownership, α, has non-monotonic effects on investment and asset pricing moments. The sign of the effect of insider ownership switches when insider ownership goes from being too low—such that the outsider *overmonitors* and the insider has *perverse incentives*—to being sufficiently high.

The results on the direct effect of monitoring parallel those on investor protection in Albuquerque and Wang (2008), whereas the results concerning the insider's ownership enrich their model by accounting for the insider's ownership as a driver of both the insider's private benefit extraction and the outsider's efforts to limit the insider's opportunistic behavior.

4.1 Real Investment

Proposition 1 *The equilibrium investment-capital ratio i decreases in investor protection, η, and monitoring, θ. Thus, $di/d\eta$ and $di/d\theta$ are less than zero.*

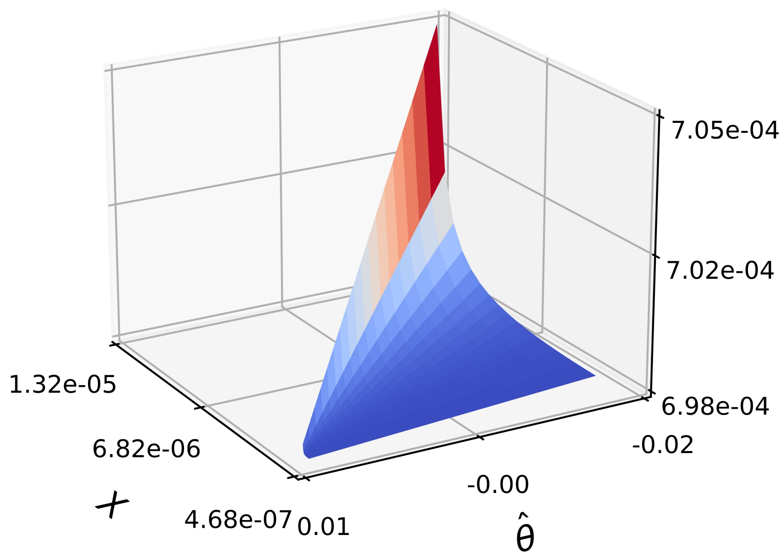
Keeping investor protection constant, lower monitoring levels allow the insider to extract more private benefits. Having a larger firm becomes more valuable to the insider when he can extract a larger share of the firm's production as private benefits. Consequently, investment increases as monitoring decreases. The same logic shows that investment increases as investor protection decreases. As such, relaxed monitoring (investor protection) leads to over-investment relative to a maximal monitoring (perfect investor protection) benchmark.

Nevertheless, the volatility of the insider's marginal utility increases in investment (equation 26). This constrains the extent to which investment responds to weaker monitoring or investor protection.

The insider's ownership, α, has non-monotonic effects on investment. When insider ownership, α, is too low, the insider rebels against *overmonitoring* by using marginal increases in α to extract more private benefits; therefore, investment increases in α. But when α is high enough for monitoring to be relaxed, marginal increases in α realign the insider's incentives, so investment decreases in α.

Figure 2 illustrates the effects of ownership on investment by plotting investment against private benefit extraction, X, and excess monitoring, $\hat{\theta}$. The surface has the shape of the half-section of a cone. The region with excess monitoring, $\hat{\theta}$, larger than zero depicts *overmonitoring* and *perverse incentives*—private benefit extraction that increases in insider ownership, α. Consistent with the insider's *perverse incentives*, investment increases in α in this region. The region with excess monitoring, $\hat{\theta}$, less than zero portrays relaxed monitoring and incentive realignment—private benefit extraction that decreases in α. Consistent with realigment of the insider's incentives, investment also decreases in α.

Figure 2: Investment as a function of private benefit extraction, X, and excess monitoring, $\hat{\theta}$. In the region with overmonitoring and perverse incentives ($\hat{\theta} > 0$), investment decreases in insider ownership, α. In the region with relaxed monitoring ($\hat{\theta} < 0$) investment decreases in α. Parameter values are as given in Table 1.



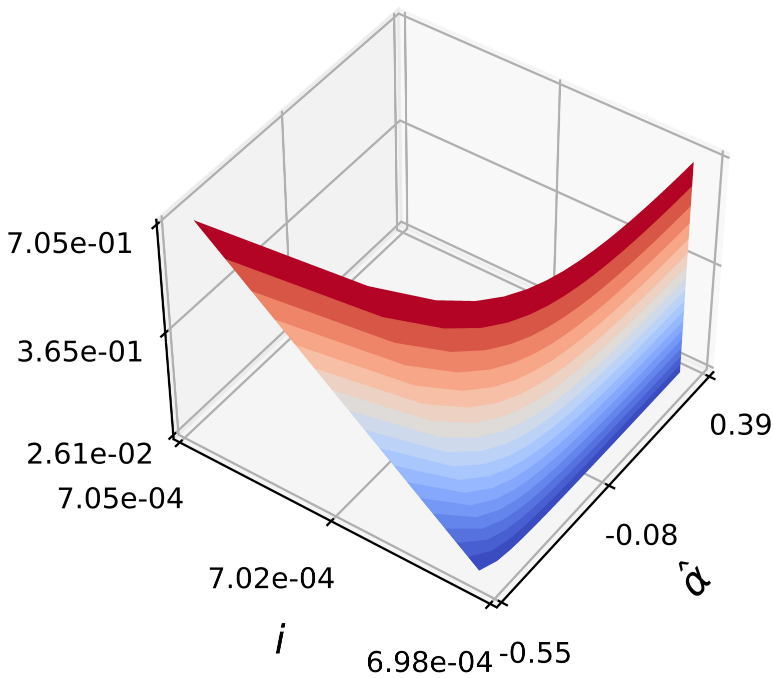
4.2 Tobin's q and Insider's Shadow Tobin's q

Proposition 2 *Tobin's q increases with investor protection, η, so $dq/d\eta$ is greater than zero. Likewise, q increases in monitoring, θ, so $dq/d\theta$ is also greater than zero.*

Monitoring and investor protection increase the value of the firm, measured by Tobin's q, by constraining the insider's extraction of private benefits. The empirical evidence in this paper and in the extant literature supports the model's prediction that firm value increases in monitoring and investor protection (cf. Brav et al., 2008; Klein and Zur, 2009; Denes et al., 2017; Cain et al., 2017).

The effect of the insider's ownership, α, on Tobin's q is non-monotonic due to the non-monotonic effects of α on private benefit extraction and their transmission unto investment. Figure 3 depicts this by plotting q against investment, i, and insider ownership in excess of the level of insider ownership required to preempt perverse ownership incentives, $\hat{\alpha}$. The shape of the surface resembles a trapezoid with one of its edges pivoted along the third dimension. Recall that investment, i, increases with insider ownership, α, when perverse ownership incentives are at play. Thus, in the region with perverse ownership incentives ($\hat{\alpha} < 0$), increasing insider ownership, α, raises q. But in the region with incentive realignment ($\hat{\alpha} > 0$), increasing α decreases q.

Figure 3: Tobin's q as a function of investment, i, and insider ownership in excess of the level of insider ownership required to preempt perverse ownership incentives, $\hat{\alpha}$. In the region with perverse ownership incentives ($\hat{\alpha} < 0$), increasing insider ownership increases q. In the region with incentive realignment ($\hat{\alpha} > 0$), increasing insider ownership decreases q. Parameter values are as given in Table 1.



This differs from the result in Albuquerque and Wang (2008), where $dq/d\alpha$ is strictly larger than zero. The intuition for why q increases in α in the region with perverse incentives is that raising the insider's ownership reduces *overmonitoring* and brings the insider closer to the incentive alignment region. Similarly, q decreases in α in the incentive alignment region because lowering the outsider's ownership reduces monitoring, θ, which is value increasing.

The insider's shadow valuation and shadow Tobin's q both have the same form as in Albuquerque and Wang (2008):

$$\hat{P}(t) = \frac{1}{\alpha} E_t \left[\int_t^\infty e^{-\rho(s-t)} M(s) \frac{M(s)^{-\gamma}}{M(t)^{-\gamma}} ds \right] = \frac{1}{1 - \epsilon^2 i \gamma} K(t) \tag{35}$$

$$\hat{q} = \frac{1}{1 - \epsilon^2 i \gamma} \tag{36}$$

However, the insider's shadow valuation and shadow Tobin's q are lower with monitoring than without it. This follows from Proposition 1, whereby $di/d\theta < 0$.

With maximal monitoring ($\alpha \to 0$), the insider's valuation explodes. This is because the insider's cash flow per share (15) also explodes. Moreover, although the insider extracts no private benefits because his cost for doing so is prohibitively high (7), he is subject to a perverse ownership incentive for benefit extraction (28).

The insider's private valuation is larger than q under perfect investor protection: $\hat{q} > q = q^*(\eta)$. This is also true for the insider's valuation under perfect investor protection, $\hat{q}^*(\eta)$—unlike in Albuquerque and Wang (2008), in which $\hat{q}^*(\eta) > q^*(\eta)$ obtains. The first inequality is due to (a) the fact that investment decreases in investor protection (Proposition 1) and (b) the outsider's monitoring costs.[13] The second inequality, $\hat{q}^*(\eta) > q^*(\eta)$, is solely due to the outsider's monitoring costs.

With imperfect investor protection, $\hat{q} > q^*(\eta) > q$, where the second inequality follows from Proposition 2. This result that resembles the one in Albuquerque and Wang (2008); however, applying Proposition 1, one can see that $\hat{q}, q^*, q$ are all lower than in Albuquerque and Wang (2008) because of monitoring.

4.3 Risk-Free Rate

As usual, the interest rate (21) consists of a subjective discount rate (ρ), a consumption growth term ($\gamma(i - \delta)$), and a precautionary savings term ($-\epsilon^2 i^2 \gamma(\gamma + 1)/2$).

[13] Earlier, I use intuition from the market for hostile takeovers to explain why the outsider monitors under perfect investor protection: investor protection facilitates—but does not perfectly substitute—monitoring.

Intuitively, ρ is the measure at which agents discount cash flows in a risk-neutral world ($\gamma = 0$). Furthermore, if agents observe high investment, they rationally expect a higher consumption level tomorrow, and thus require a higher interest rate to convince them to save towards increased future consumption. Agents also require a higher interest rate if they are unwilling to substitute consumption between periods (low $1/\gamma$, high γ). Lastly, higher investment also negatively impacts the interest rate by raising the volatility of capital accumulation and output, $\epsilon^2 i^2$, thereby increasing agents' willingness to save and the supply of capital. Proposition 3 below summarizes the effect of investor protection and monitoring on the interest rate.

Proposition 3 *The interest rate decreases in investor protection, η, and monitoring, θ, if and only if $1 > \epsilon^2(\gamma + 1)i$.*

As noted above, investment raises the interest rate through the economic growth term ($\gamma(i - \delta)$) and lowers the interest rate through the precautionary savings term ($-\epsilon^2 i^2 \gamma(\gamma + 1)/2$). The growth term dominates when $1 > \epsilon^2(\gamma + 1)i$; otherwise, the precautionary savings term dominates. Albuquerque and Wang (2008) provide evidence that in a zero monitoring economy, the above inequality holds, such that the growth term in the expression for the interest rate dominates, and the interest rate decreases in investor protection, η.

Thus, momentarily assume that $1 > \epsilon^2(\gamma+1)i$, such that the interest rate increases in investment. Additionally, recall that investment responds non-monotonically to insider ownership, α (because of the non-monotonic effects of α on private benefit extraction). Therefore, the interest rate responds non-monotonically to insider ownership. When α is too low (high enough), investment increases (decreases) in α, so the interest rate increases (decreases) in α. This result differs from that in Albuquerque and Wang (2008), in which the interest rate is strictly decreasing in α. The discrepancy is due to the presence of monitoring.

4.4 Volatility, Risk Premium, and Expected Return

Proposition 4 *Return volatility, σ_P; the risk premium, λ; and expected returns decrease in monitoring, θ, and investor protection, η .*

First, recall that investment decreases in investor protection and monitoring (Proposition 1). Second, note that investment raises the volatility of capital accumulation and output, which in turn raises the risk premium, $\lambda = \gamma\sigma_P^2 = \gamma\epsilon^2 i^2$. Therefore, reducing monitoring or investor protection increases volatility and the equity risk premium. The expected return on equity is equal to the interest rate, r, plus the equity risk premium, λ. Both of these moments decrease in monitoring and investor protection, so the expected return on equity decreases in monitoring and investor protection.

As before, the effects of insider ownership, α, on return volatility, the risk premium, and expected returns are non-monotonic. The reason is that these moments increase in investment, and α has non-monotonic effects on private benefit extraction, which transmit unto investment. This differs from the result in Albuquerque and Wang (2008), where return volatility σ_P, the risk premium λ, and expected returns all decrease in the insider's ownership, α.

4.5 Dividend Yield

Proposition 5 *The dividend yield,*

$$y = \rho + (\gamma - 1)\left(i - \delta - \frac{\gamma}{2}\epsilon^2 i^2\right), \tag{37}$$

decreases (increases) in monitoring, θ, and investor protection, η, when $\gamma > 1$ ($\gamma < 1$).

The sign of the dividend yield's sensitivity to the investment-capital ratio, i, determines how the dividend yield responds to monitoring and investor protection. Note that the expression for dy/di is $(\gamma - 1)(1 - \gamma\epsilon^2 i)$, and that the proof for Proposition 1 (Appendix) establishes that $1 > \gamma\epsilon^2 i$. Thus, whether γ is larger than or less than 1 determines whether the dividend yield increases or decreases in investment, which in turn determines whether the dividend yield decreases or increases in monitoring and investor protection.

Assume momentarily that γ is larger than 1, such that the dividend yield increases in investment.[14] Since investment decreases in monitoring, the dividend yield decreases in monitoring as well. Thus, monitoring increases firm value proportionally more than it does dividends. Similarly, since investment decreases in investor protection, so does the dividend yield.

Once again, the effects of insider ownership, α, are non-monotonic and feed in through private benefit extraction and investment.

5 Empirical Evidence

5.1 Summary of Results

Consistent with my model predictions, my firm-level regression evidence shows that investment, expected returns, and GARCH(1,1) volatility estimates decrease in monitoring, while q increases in monitoring. Results for dividend yields are statistically negligible. Next, I describe the data I use and my results in detail.

5.2 Data

I take monthly stock price and return data from CRSP; use quarterly firm fundamentals from Compustat; use the Fama French data provided in WRDS; and take activism campaign data from Fact Set. The sample only includes standard codes for common stock (codes 10 and 11) and excludes industries that have been historically regulated:

[14]The signs discussed in this analysis switch if γ is less than 1.

banking and financials (Standard Industrial Classification (SIC) codes 6000–6999), railroads (SIC codes 4000–4099), airlines (SIC codes 4500–4599), utilities (SIC codes 4900–4999), public administration and non-classifiables (SIC codes 9000-9999). The study period is from January 1996 to December 2016.[15]

To avoid underweighting monitored firms, I curtail the sample by matching monitored firms to non-monitored firms with the same SIC code which are also in the same Fama-French SMB and HML portfolios. If a monitored firm obtains no matches using 4-digit SIC codes and Fama-French decile portfolios, I use 3-digit SIC codes and Fama-French quintile portfolios, as in Brav, Jiang, Partnoy, and Thomas (2008).

Moreover, I divide the sample into two to formulate separate analyses for dependent variables that have variation at the firm-quarter level and those that have variation at the firm-month level. I employ cross-sectional analysis in the quarterly sample and a difference-in-differences analaysis in the monthly sample. I explain the distinct methodologies in greater detail in subsection 5.3. The quarterly sample contains 6,276 unique PERMNOs and 4,816 campaign events, while the monthly sample contains 7,677 PERMNOs and 5,375 campaign events.

Table 2 defines the variables I use for empirical analysis. My proxy variables for monitoring rely on activism campaign data from Fact Set, which includes non-hedge fund activist campaigns.[16] The proxy variables are (a) at the extensive margin, whether an activism campaign starts on a given firm-period ($\mathbb{1}(MON_{i,t}^{s,p})$), and (b) at the intensive margin, the number of campaign or monitoring days on a given firm-period ($N(MON_{i,t}^{d,p})$). For these proxies, a firm-period is a firm-quarter or a firm-month, depending on the level of variation under scrutiny. To illustrate the case with firm-quarter variation, $N(MON_{i,t}^{d,q})$, if there are two ongoing campaigns

[15]Data for the Takeover Index—my proxy for investor protection—ends on 2016.

[16]3,955 (40.087%) of campaigns in the Fact Set data set are hedge fund campaigns. For studies of shareholder activism that include non-hedge fund activism campaigns, see Klein and Zur (2006) and Greenwood and Schor (2009). For studies that only use hedge fund activism campaigns, see Brav, Jiang, Partnoy, and Thomas (2008); Gantchev (2013); Becht, Frank, Grant, and Wagner (2017); and Bebchuck, Brav, Jiang, and Keusch (2020).

on a given quarter, and each of them lasts 60 days, the number of monitoring days is 120.[17] To the extent that time expended on a campaign measures the effort cost to the activist(s), $N(MON_{i,t}^{d,p})$ represents the empirical counterpart of the outsider's private cost of monitoring, θ, in the model.

Table 2: Variable names and descriptions. The prefix f indicates the lead operator for a given number of months, l indicates lags, z indicates standardization, and z^c indicates standardization using cross-sectional means and standard deviations. Variables without an f or l prefix are contemporaneous.

Variable	Description
$\mathbb{1}(MON_{i,t}^{s,q})$	Indicator for campaign starts in a firm-quarter
$\mathbb{1}(MON_{i,t}^{s,m})$	Indicator for campaign starts in a firm-month
$N(MON_{i,t}^{d,q})$	Number of monitoring days in a firm-quarter
$N(MON_{i,t}^{d,m})$	Number of monitoring days in a firm-month
$TOIND_{i,t}$	Cain, McKeon, and Solomon (2017) Takeover Index: probability of becoming the target of a hostile takeover
$i_{i,t}$	Investment, as in Baker, Stein, and Wurgler (2003)
$Q_{i,t}$	Tobin's Q, as in Duchin, Ozbas, and Sensoy (2010)
$h_{i,t}^{MKT}$	Quarter-end GARCH(1,1) volatility estimates for quarterly cumulative returns under the market model
$Y_{i,t}$	Quarter-end dividend yield
$ME_{i,t}$	Market equity, in thousands of USD
$z(\epsilon_{i,t}^{MKT})$	Excess returns residualized with respect to the market factor and standardized
$z(R_{i,t}^{e})$	Standardized excess returns
$z(R_{i,t}^{ex,MKT})$	Ex-dividend returns net of the CRSP value weighted index and standardized
$z(R_{i,t}^{MKT})$	Cum-dividend returns net of the CRSP value weighted index and standardized
$z(R_{i,t}^{ex})$	Standardized ex-dividend returns
$z(R_{i,t})$	Standardized cum-dividend returns

I base my monitoring proxies on activism campaigns because these campaigns are attempts by non-managing shareholders to change the behavior of the firms' managers. For example, common objectives of activism campaigns include direct engagement with management or with the Board of Directors, excluding management; voting against management, including through proxy fights; seeking increased board representation; and pursuing acquisition by a third party (Brav, Jiang, Kim, and Partnoy, 2008; Klein and Zur, 2009; Gantchev, 2013; Bebchuck, Brav, and Jiang, and

[17]The only difference between $N(MON_{i,t}^{d,m})$ and $N(MON_{i,t}^{d,q})$ is the length of the period over which I measure monitoring days.

Keusch, 2020).[18]

The proxy for investor protection is the Cain, McKeon, and Solomon (2017) Takeover Index (henceforth, $TOIND_{i,t}$). $TOIND_{i,t}$ estimates the probability that a firm is acquired in a hostile takeover using state-level changes in anti-takeover laws as predictor variables. Theory propounds that hostile takeovers act as a market disciplinary mechanism—that is, as a channel for improved firm governance (Manne, 1965). Thus, $TOIND_{i,t}$ is a plausibly exogenous measure of investor protection—a set of legal environment factors that improve firm governance.[19]

To measure investment, i, I follow Baker, Stein, and Wurgler (2003) and use the quarterly change in total assets, $\Delta A_t/(0.5 \times A_t - 0.5 \times l3(A_t))$, normalizing changes as in Davis and Haltiwanger (1992). The measure of Tobin's q follows Duchin, Ozbas, and Sensoy (2010), with the exception that market equity data comes from CRSP, rather than Compustat.

To study effects on expected returns, I residualize monthly stock-level excess returns with respect to the market factor and standardize. This yields the variable $z(\epsilon_{i,t}^{MKT})$. The reason for residualizing is that the impact of monitoring on returns should be firm-specific and thus orthogonal to the market factor.

For volatility, I use quarter-end, stock-level GARCH(1,1) estimates, $h_{i,t}^{MKT}$. The measure of returns is quarter-end excess returns compounded at quarterly frequency; and I employ the market model for the conditional mean. For both expected returns and volatilities, I omit the regression intercept from the conditional mean.

Dividend yields, $Y_{i,t}$, are rolling 12-month sums of ordinary cash dividends divided

[18] A campaign may have multiple objectives, so different alternatives need not be mutually exclusive. A campaign might start with a simple statement of the activist's demands, but evolve into a more costly attempt to gain board representation, which can ultimately turn into a proxy fight (Gantchev, 2013). Empirical evidence shows that a firm's management is more likely to acquiesce to activists' demands if a campaign is likely to be successful even under management's resistance, or if the campaign represents high reputation costs to the firm's management (Bebchuck, Brav, Jiang, and Keusch, 2020).

[19] Cain, McKeon, and Solomon (2017) report $TOIND_{i,t}$ by GVKEY (firm) and fiscal year.

by the contemporaneous stock price at quarter-end.[20] Finally, $ME_{i,t}$ is market equity at the firm level.[21]

Figure 4 displays the evolution of equal-weighted monitoring and investor protection throughout the study period. Both monitoring and investor protection have risen over time, highlighting the increasing relevance of both monitoring and investor protection.[22]

Table 3: Summary statistics. Variables are as defined in Table 2. SD stands for standard deviation; p(25), p(50), and p(75) stand for the 25th, 50th, and 75th percentiles, respectively; and N stands for the number of non-missing observations.

Panel A: Quarterly Data						
	Mean	SD	p(25)	p(50)	p(75)	N
$\text{l3}(i_{i,t})$	-0.003	0.152	-0.045	-0.003	0.032	24,888
$q_{i,t}$	9.689	0.327	9.646	9.790	9.852	24,800
$\epsilon_{i,t}^{MKT}$	2.735	38.729	-15.640	-2.170	12.170	3,399
$h_{i,t}^{MKT}$	1,443.993	3,911.805	286.528	643.438	1,261.080	13,517
$Y_{i,t}$	0.031	0.098	0.008	0.015	0.028	4,992
$\mathbb{1}(MON_{i,t}^{s})$	0.088	0.283	0.000	0.000	0.000	21,959
$\text{N}(MON_{i,t}^{d})$	10.274	29.135	0.000	0.000	0.000	25,166
$l12(TOIND_{i,t})$	0.128	0.077	0.075	0.111	0.161	20,432
$ME_{i,t}$	3,568.300	20911.275	40.154	132.127	468.257	25,054

Panel B: Monthly Data						
	Mean	SD	p(25)	p(50)	p(75)	N
$\epsilon_{i,t}^{MKT}$	0.793	22.141	-8.152	-0.771	6.691	38,396
$R_{i,t}^{ex}$	1.964	23.494	-7.837	0.000	8.561	38,396
$R_{i,t}^{ex,MKT}$	1.258	22.536	-7.874	-0.543	7.144	38,396
$R_{i,t}^{MKT}$	1.146	22.531	-7.999	-0.650	7.023	38,396
$R_{i,t}^{ex}$	0.020	0.235	-0.078	0.000	0.086	38,396
$R_{i,t}$	0.021	0.235	-0.077	0.001	0.087	38,396
$\mathbb{1}(MON_{i,t}^{s,m})$	0.015	0.120	0.000	0.000	0.000	51,245
$\text{N}(MON_{i,t}^{c,m})$	0.063	0.283	0.000	0.000	0.000	56,365
$l12(TOIND_{i,t})$	0.128	0.081	0.068	0.109	0.161	46,226
$ME_{i,t}$	3,798.761	20538.325	45.532	164.824	707.033	38,401

Table 3 lists summary statistics. In quarterly data, about 9% of observations

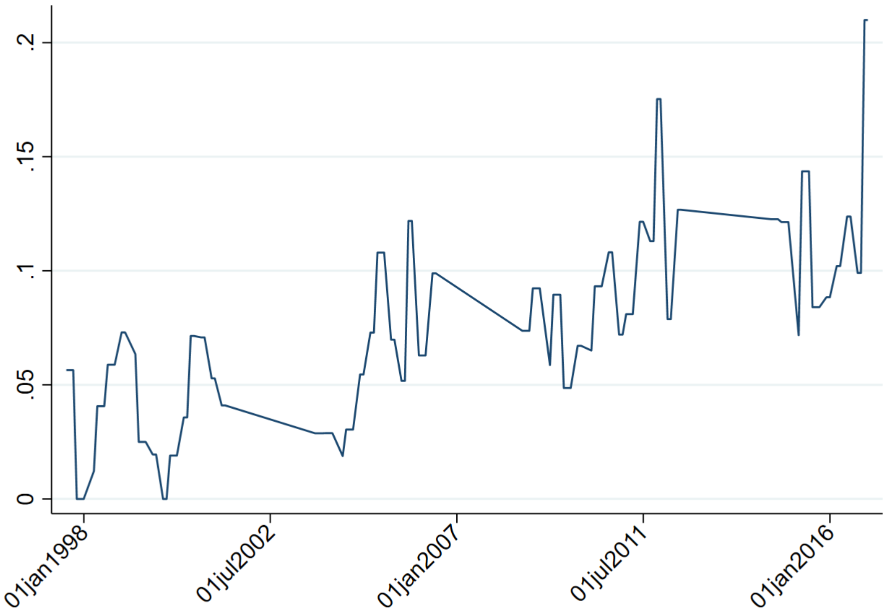
[20]Ordinary dividends exclude M&A cash flows, buybacks, and new issues. Following Pettenuzzo, Sabbatucci, and Timmermann (2020), I aggregate ordinary dividends at the PERMNO-date level; then I compute rolling 12-month sums.

[21]Some firms have multiple stock issuances. In these cases, I keep the stock that has higher market equity. Additionally, I keep only primary link codes (P, C) from the WRDS CRSP-Compustat merged database; however, I do use non-primary codes (J) to compute total firm market equity.

[22]There has also been a rise in the share of activism events that result in settlements, which involve a "standstill" from activists' part and concessions by management in favor of activists' objectives (Bebchuck, Brav, Jiang, and Keusch, 2020).

Figure 4: Quarterly equal-weighted monitoring and investor protection for the sample. I measure monitoring as (Panel A) the number of campaign starts, $N(MON_{i,t}^{s,q})$; and (Panel B) the number of campaign days($N(MON_{i,t}^{d,q})$). Investor protection (Panel C) is proxied by the Cain, McKeon, and Solomon (2017) Takeover Index, $TOIND_{i,t}$, which measures the probability that a firm undergoes a hostile takeover as a target.

Panel A: Equal-weighted average $N(MON_{i,t}^{s,q})$



Panel B: Equal-weighted average $N(MON_{i,t}^{d,q})$

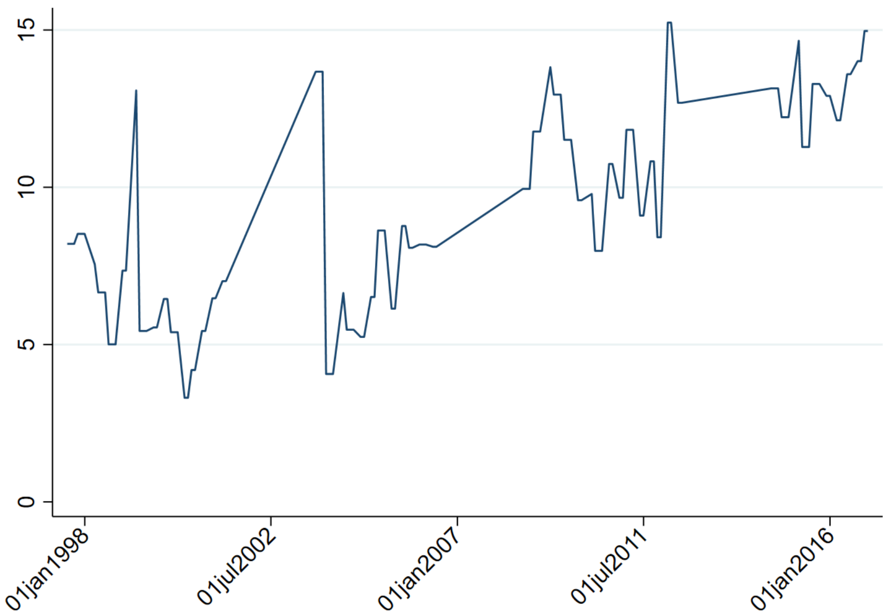
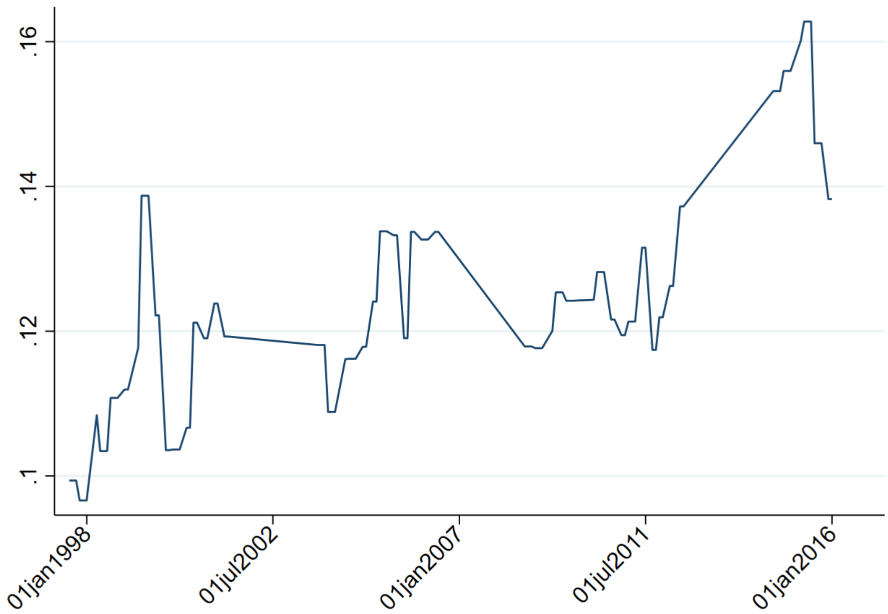


Figure 4, cont.

Panel C: Equal-weighted average $TOIND_{i,t}$



correspond to campaign starts ($\mathbb{1}(MON_{i,t}^{s,q})$), the average number of monitoring days ($N(MON_{i,t}^{d,q})$) is 10.274, and the average probability that a firm is the target of a hostile takeover ($TOIND_{i,t}$) is about 13%. In monthly data, about 2% of observations correspond to campaign starts ($\mathbb{1}(MON_{i,t}^{s,m})$), the average number of monitoring days is 0.067 ($N(MON_{i,t}^{d,q})$), and the average probability that a firm is the target of a hostile takeover is also about 13% ($TOIND_{i,t}$).[23] In both samples, firms concentrate at market equity, $ME_{i,t}$, below USD 5 million.[24]

5.3 Empirical Strategy

I study the firm-level impact of monitoring on investment, Tobin's q, expected returns, excess return volatilities, and dividend yields. To that end, I use two regression

[23]The means of monitoring variables decline because I exploit the window that follows monitoring events in the monthly sample, unlike in the quarterly sample (see subsection 5.3).

[24]Small firm size is consistent with the finding that the likelihood of becoming a target for activism decreases in market equity (Brav, Jiang, Kim, and Partnoy, 2008; Becht, Franks, Grant, and Wagner, 2017). $ME_{i,t}$ has skewness of 12.240 in the quarterly sample and 11.642 in the monthly sample.

models. The first model exploits firm-quarter level variation, while the second model exploits firm-month level variation. In all regressions, I cluster standard errors by PERMNO.[25]

$$y_{i,t} = \alpha + \beta f(MON^{q}_{i,t}) + \gamma' x_{i,t} + \tau_t + \nu_{i,t} \tag{38}$$

Equation 38 states the first regression model, which analyzes the impact of monitoring on investment, q, excess return volatilities, and dividend yields (all of which vary at the firm-quarter level). I take leads of each dependent variable at 3, 6, 9, and 12 months. I also measure contemporaneous effects on all variables except for investment, since investment should not respond contemporaneously to monitoring.

Regressions corresponding to Equation 38 use cross-sectionally standardized variables (with the exception of the monitoring dummy). That is, I subtract the cross-sectional mean and divide by cross-sectional standard deviations. Given that I study cross-sectional variation through the model in Equation 38, my chosen standardization approach facilitates an appropriate interpretation of the economic significance of monitoring (Liu and Winegar, 2022).

$f(MON^{q}_{i,t})$ is (a) in regressions studying the extensive margin of monitoring, an indicator the start of a monitoring event, $\mathbb{1}(MON^{s,q}_{i,t})$, or (b) in regressions studying the intensive margin of monitoring, the count of monitoring days, $z^c(N(MON^{d,q}_{i,t}))$. $x_{i,t}$ is a vector containing the control variables $z^c(l12(TOIND_{i,t}))$ and $z^c(l3(ME_{i,t}))$, and τ_t are time fixed effects.

$$\begin{aligned} y_{i,a,t} = \alpha &+ \beta_1 f(MON^{m}_{i,a,c}) POST_{a,t} + \beta_2 f(MON^{m}_{i,a,c}) + \beta_3 POST_{a,t} + \gamma' x_{i,a,t} \\ &+ \delta_i + \delta_a + \tau_t + \nu_{i,a,t} \end{aligned} \tag{39}$$

The second regression model, given in 39, analyzes the impact of monitoring on expected returns, which vary at the firm-month level. In addition to the main proxy for expected returns, $z(\epsilon^{MKT}_{i,a,t})$, I employ other return measures as dependent variables:

[25] I construct the sample such that PERMNOs and GVKEYs match one-to-one.

excess returns, abnormal returns (net of returns on the CRSP value-weighted index), and total returns. I include cum- and ex-dividend variants for the latter two measures.

In regressions conforming to Equation 39, I standardize all variables (save for the monitoring dummy) in the traditional way, since I contemplate both time-series and cross-sectional variation. The new subscript, a, indexes a group of affiliated or matching firms, consisting of a monitored firm and its matched non-monitored firms.

$f(MON^{m}_{i,a,c})$ is analogous to $f(MON^{q}_{i,t})$, but captures variation in monitoring at the firm-month level—both on the extensive margin ($\mathbb{1}(MON^{s,m}_{i,a,c})$) and on the intensive margin ($z(N(MON^{d,m}_{i,a,c}))$). Note the time subscript c, which indicates that monitoring variables are set to the time of the start of the most recent campaign for the monitored firm in a given group, a.

The variable $POST_{a,t}$ is a dummy for whether firms in a matched group, a, are within 2 years past the start of a campaign for the monitored firm in the group. The window for $POST_{a,t}$ renews when there is an ongoing campaign if the monitored firm in the group becomes the target of a new campaign. Furthermore, note that the window for $POST_{a,t}$ is 1 year longer than the window for q. The reason for the mismatch in the windows is that the effects that the model predicts are at odds with the dynamics of my empirical proxies. The model says that, in the same instant, q increases in monitoring, while expected returns decrease in monitoring. But empirically, q and realized returns move in the same direction on a given period. An extended window of analysis for expected returns, relative to the window for q, allows for price appreciation in the first part of the window and decreased expected returns in the latter part of the window.

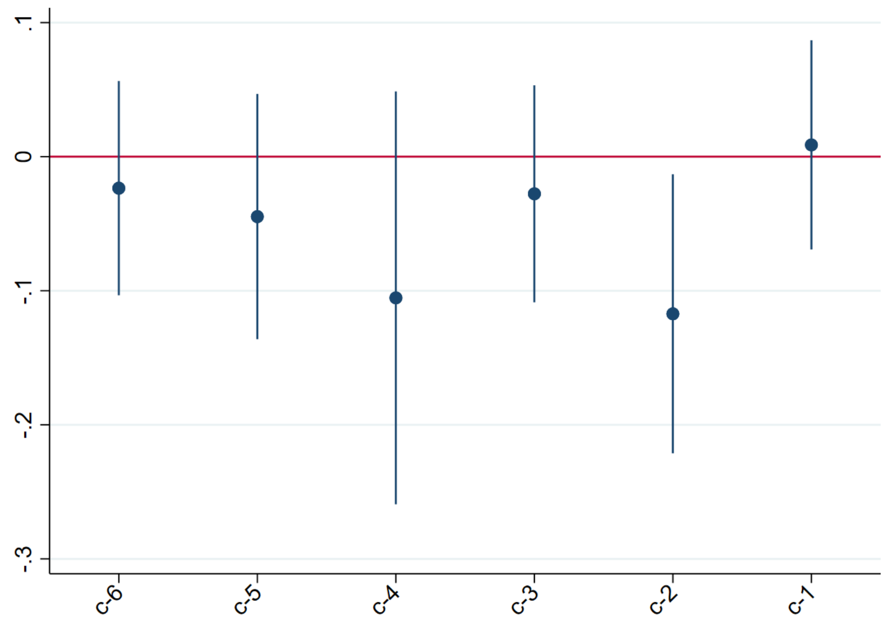
Similar to before, $x_{i,a,t}$ is a vector containing containing the controls $z(l12(TOIND_{i,a,t}))$ and $z(l3(ME_{i,a,t}))$, δ_i are firm fixed effects, δ_a are matched group fixed effects, and τ_t are time fixed effects. I only employ firm fixed effects in extensive margin regressions, to compare monitored firms against non-monitored firms. Firm fixed effects, δ_i, sub-

Figure 6: Parallel trends test: extensive margin (Panel A) and intensive margin (Panel B) of monitoring. The plot displays interaction term coefficients, β_k, from a regression of the form:

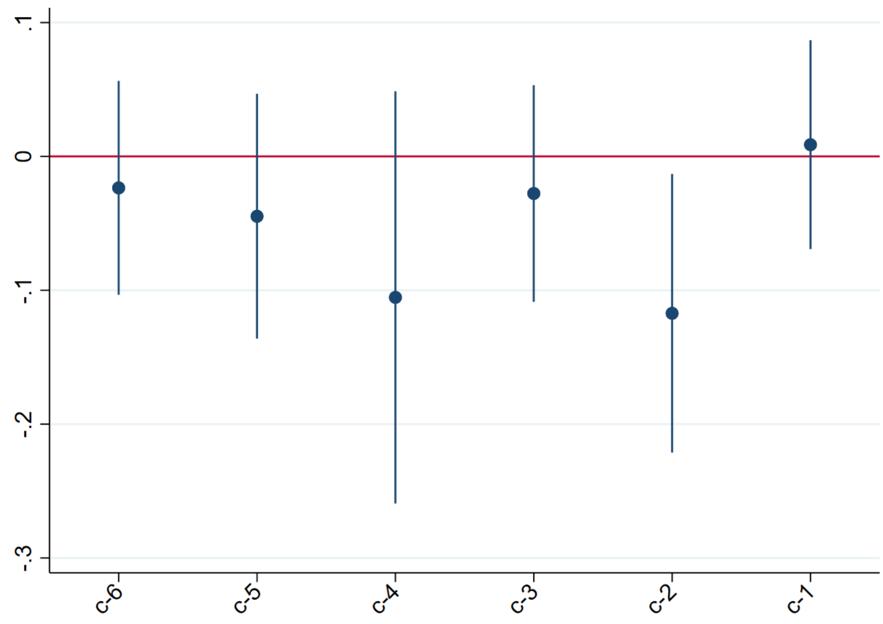
$$y_{i,a,t} = \alpha + \Sigma_{k=-6}^{-1} \beta_k f(MON_{i,a,c}^{m}) \mathbb{1}(c+k)_{a,t} + \beta f(MON_{i,a,c}^{m}) + \gamma' x_{i,a,t} + \delta_i + \delta_a + \tau_t + \nu_{i,a,t}.$$

The dependent variable is standardized, residualized excess returns, $z(\epsilon_{i,a,t}^{MKT})$. At the extensive margin, $f(MON_{i,a,c}^{m}) = \mathbb{1}(MON_{i,a,t}^{s,m})$; and at the intensive margin, $f(MON_{i,a,c}^{m}) = z(N(MON_{i,a,c}^{d,m}))$. $\mathbb{1}(c+k)_{a,t}$ is an indicator for whether firms in a matched group, a, are k periods away from the start of a monitoring event in the group. Bands depict coefficients' 95% confidence intervals.

Panel A: Extensive Margin



Panel B: Intensive Margin



sume the monitoring dummy, $\mathbb{1}(MON^{s,m}_{i,a,t})$, in extensive margin regressions. Matched group effects, δ_a are only in place in intensive margin regressions, to compare firms that are monitored *more* against those that are monitored *less*, including those that are not monitored at all.

Given the setup, in extensive margin regressions, the coefficient β_1 captures the impact on $y_{i,a,t}$ after the incidence of monitoring at time $c < t$. In intensive margin regressions, β_1 is the impact on $y_{i,a,t}$ after a 1 standard deviation increase in monitoring days (relative to peer firms) at time c. Figures 6 shows that the data satisfies the parallel trends assumption; thus, β_1 is indeed identified in the terms that I have outlined.[26]

5.4 Results

Table 4 presents estimates for the extensive margin effects of monitoring on investment, q, volatility, and dividend yields. The indicator for campaign starts, $\mathbb{1}(MON^{s;q}_{i,t})$, is the main explanatory variable. Consistent with the model, investment (Panel A) and excess return volatilities (Panel C) decline with monitoring, whereas Tobin's q (Panel B) increases with monitoring. Moreover, the incidence of monitoring suscitates a decline in investment of between 0.07 and 0.10 standard deviations. The decline persists from 1 quarter to 1 year after the monitoring event, and is statistically significant at the 1% level. Monitoring also provokes a decline in volatilities of between 0.06 and 0.07 standard deviations. This effect persists for 1 year, with significance at the 10% level from the incidence of monitoring until 2 quarters ahead from the event. For Tobin's q, a monitoring incident leads to an increase of between 0.15 and 0.19 standard deviations. The effect persists for 1 year and is significant at the 1% level.

[26] I omit the treatment period from the plots for parallel trends tests for presentation purposes. It is widely documented in the activism literature that there is a positive effect on cumulative returns in a [-20,20] or [-10,+30] day window around the onset of a campaign (cf. Brav, Jiang, Kim, and Partnoy, 2008; Collin-Dufresne and Fos, 2015).

Table 4: Extensive margin effects of monitoring on investment, q, volatilities, and dividend yields. Variables are as defined in Table 2. The main independent variable is $\mathbb{1}(MON_{i,t}^{s;q})$, and the control variables are $z(TOIND_{i,t})$ and $z(l3(ME_{i,t}))$. Standard errors are clustered at the PERMNO level, and all regressions include time fixed effects. β is the coefficient on $\mathbb{1}(MON_{i,t}^{s;q})$, and t-statistics are stated in parentheses. One, two, and three stars (*) indicate significance at the 10%, 5%, and 1% level, respectively.

		Panel A		
	β	(t)	R^2	N
$z^c(f3(i_{i,t}))$	-0.103***	(-3.786)	0.002	14,785
$z^c(f6(i_{i,t}))$	-0.078***	(-2.806)	0.002	14,384
$z^c(f9(i_{i,t}))$	-0.088***	(-3.303)	0.002	13,985
$z^c(f12(i_{i,t}))$	-0.074***	(-2.585)	0.002	13,609
		Panel B		
	β	(t)	R^2	N
$z^c(q_{i,t})$	0.180***	(4.916)	0.007	15,605
$z^c(f3(q_{i,t}))$	0.149***	(3.575)	0.006	12,384
$z^c(f6(q_{i,t}))$	0.185***	(5.064)	0.007	12,219
$z^c(f9(q_{i,t}))$	0.187***	(4.496)	0.008	12,366
$z^c(f12(q_{i,t}))$	0.194***	(4.853)	0.008	10,809
		Panel C		
	β	(t)	R^2	N
$z^c(h_{i,t}^{MKT})$	-0.066*	(-1.675)	0.010	8,793
$z^c(f3(h_{i,t}^{MKT}))$	-0.071*	(-1.788)	0.010	8,671
$z^c(f6(h_{i,t}^{MKT}))$	-0.068*	(-1.714)	0.010	8,514
$z^c(f9(h_{i,t}^{MKT}))$	-0.067	(-1.639)	0.010	8,342
$z^c(f12(h_{i,t}^{MKT}))$	-0.060	(-1.442)	0.010	8,175
		Panel D		
	β	(t)	R^2	N
$z^c(Y_{i,t})$	-0.028	(-0.498)	0.008	2,691
$z^c(f3(Y_{i,t}))$	-0.038	(-0.642)	0.010	2,230
$z^c(f6(Y_{i,t}))$	0.034	(0.410)	0.010	2,149
$z^c(f9(Y_{i,t}))$	0.101	(1.166)	0.010	2,265
$z^c(f12(Y_{i,t}))$	0.086	(1.073)	0.009	2,124

Table 5: Extensive margin effects of monitoring on expected returns. Variables are as defined in Table 2. The control variables are $z(TOIND_{i,a,t})$ and $z(l3(ME_{i,a,t}))$. Standard errors are clustered at the PERMNO level. T-statistics are stated in parentheses. One, two, and three stars (*) indicate significance at the 10%, 5%, and 1% level, respectively.

	$z(\epsilon^{MKT}_{i,a,t})$	$z(R^{e}_{i,a,t})$	$z(R^{ex,MKT}_{i,a,t})$	$z(R^{MKT}_{i,a,t})$	$z(R^{ex}_{i,a,t})$	$z(R_{i,a,t})$
$\mathbb{1}(MON^{s,m}_{i,a,c})POST_{a,t}$	-0.145***	-0.139***	-0.145***	-0.144***	-0.140***	-0.139***
	(-3.962)	(-3.894)	(-3.921)	(-3.894)	(-3.921)	(-3.894)
$\mathbb{1}(MON_{i,a,c})$	0.122***	0.118***	0.125***	0.123***	0.120***	0.118***
	(3.126)	(3.159)	(3.228)	(3.159)	(3.228)	(3.159)
$POST_{a,t}$	0.013	0.022	0.023	0.023	0.022	0.022
	(0.710)	(1.245)	(1.257)	(1.245)	(1.257)	(1.245)
Observations	25254	25254	25254	25254	25254	25254
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.0328	0.124	0.0393	0.0393	0.123	0.124

Results for dividend yields (Panel D) are more difficult to interpret. Yields decline contemporaneously and 1 quarter after the incidence of monitoring, but at statistically neglibible magnitudes. From 2 quarters and until 1 year after a monitoring incident, dividend yields increase, albeit without statistical significance.

Table 5 shows the extensive margin effects of monitoring on expected returns. Consistent with the model, the expected returns of monitored firms are lower than those of non-monitored peers after a monitoring event. The effect is similar in magnitude across expected return proxies—ranging from -0.139 to -0.145 standard deviations—and is statistically significant at the 1% level for all return measures. The annualized size of the effect corresponds to a decline in returns of 11% in the case of the main return proxy, $z(\epsilon^{MKT}_{i,t})$.

Table 6 shows results on the intensive margin, using monitoring days, $z(N(MON^{d,q}_{i,t}))$, as the main explanatory variable. As predicted by the model, investment (Panel A) and excess return volatilities (Panel D) decline in response to increased monitoring, while Tobin's q (Panel B) increases in response to higher monitoring. A 1 standard deviation increase in monitoring days leads to a decline in investment of between 0.02 and 0.03 standard deviations. As in the extensive margin, this effect is in place from

Table 6: Intensive margin effects of monitoring on investment, q, volatilities, and dividend yields. Variables are as defined in Table 2. The main independent variable is $N(MON_{i,t}^{d,q})$, and the control variables are $z(TOIND_{i,t})$ and $z(l3(ME_{i,t}))$. Standard errors are clustered at the PERMNO level, and all regressions include time fixed effects. β is the coefficient on $N(MON_{i,t}^{d,q})$, and t-statistics are stated in parentheses. One, two, and three stars (*) indicate significance at the 10%, 5%, and 1% level, respectively.

Panel A				
	β	(t)	R^2	N
$z^c(f3(i_{i,t}))$	-0.023***	(-3.411)	0.001	16,863
$z^c(f6(i_{i,t}))$	-0.034***	(-4.883)	0.002	16,409
$z^c(f9(i_{i,t}))$	-0.025***	(-3.818)	0.001	15,968
$z^c(f12(i_{i,t}))$	-0.027***	(-3.846)	0.002	15,555

Panel B				
	β	(t)	R^2	N
$z^c(q_{i,t})$	0.056***	(5.074)	0.010	17,820
$z^c(f3(q_{i,t}))$	0.061***	(5.376)	0.011	14,256
$z^c(f6(q_{i,t}))$	0.070***	(6.486)	0.012	14,021
$z^c(f9(q_{i,t}))$	0.071***	(6.169)	0.012	14,133
$z^c(f12(q_{i,t}))$	0.076***	(6.462)	0.012	12,390

Panel C				
	β	(t)	R^2	N
$z^c(h_{i,t}^{MKT})$	-0.027*	(-1.705)	0.012	10,354
$z^c(f3(h_{i,t}^{MKT}))$	-0.027*	(-1.669)	0.012	10,198
$z^c(f6(h_{i,t}^{MKT}))$	-0.028	(-1.633)	0.012	10,002
$z^c(f9(h_{i,t}^{MKT}))$	-0.029	(-1.567)	0.012	9,793
$z^c(f12(h_{i,t}^{MKT}))$	-0.030	(-1.582)	0.012	9,602

Panel D				
	β	(t)	R^2	N
$z^c(Y_{i,t})$	-0.017	(-0.819)	0.004	3,255
$z^c(f3(Y_{i,t}))$	0.019	(0.619)	0.006	2,717
$z^c(f6(Y_{i,t}))$	0.065*	(1.658)	0.010	2,614
$z^c(f9(Y_{i,t}))$	0.069*	(1.653)	0.010	2,745
$z^c(f12(Y_{i,t}))$	0.080*	(1.656)	0.012	2,590

Table 7: Intensive margin of effects of monitoring on expected returns. Variables are as defined in Table 2. The control variables are $z(TOIND_{i,t})$ and $z(l3(ME_{i,t}))$. Standard errors are clustered at the PERMNO level. T-statistics are stated in parentheses. One, two, and three stars (*) indicate significance at the 10%, 5%, and 1% level, respectively.

	$z(\epsilon^{MKT}_{i,a,t})$	$z(R^{e}_{i,a,t})$	$z(R^{ex,MKT}_{i,a,t})$	$z(R^{MKT}_{i,a,t})$	$z(R^{ex}_{i,a,t})$	$z(R_{i,a,t})$
$z(N(MON^{d,m}_{i,a,c}))POST_{a,t}$	-0.047***	-0.043***	-0.046***	-0.044***	-0.044***	-0.043***
	(-3.536)	(-3.156)	(-3.247)	(-3.156)	(-3.247)	(-3.156)
$z(N(MON^{d,m}_{i,a,c}))$	0.038***	0.035***	0.037***	0.036***	0.035***	0.035***
	(3.653)	(3.351)	(3.394)	(3.351)	(3.394)	(3.351)
$POST_{a,t}$	0.001	0.012	0.013	0.013	0.013	0.012
	(0.050)	(0.729)	(0.754)	(0.729)	(0.754)	(0.729)
Observations	19886	19886	19886	19886	19886	19886
Group Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.0816	0.172	0.0872	0.0874	0.172	0.172

1 quarter to 1 year after the monitoring event and is significant at the 1% level. An additional standard deviation of monitoring also leads to a decline in volatilities of about 0.03 standard deviations, which persists for 1 year and is significant at the 10% level contemporaneously and 1 quarter after the increase in monitoring. Tobin's q increases in response to more monitoring, with magnitudes ranging from about 0.6 to 0.8 standard deviations for 1 year after the uptick in monitoring days. Increases in q are statistically significant at the 1% level.

As with the analysis at the extensive margin, results for dividend yields (Panel D) are varied and thus unclear. A 1 standard deviation increase in monitoring days, $N(MON^{d,q}_{i,t})$, leads to decreased yields contemporaneously (without statistical significance) but also leads to increases in yields from 1 quarter to 1 year after the increase in monitoring. These increases attain statistical significance at the 10% level from 2 quarters to 1 year after the increase in monitoring.

Table 7 shows the intensive margin effects of monitoring on expected returns. Conforming to the model's predictions, the expected returns of firms that had a 1 standard deviation *higher* level of monitoring at the start of a campaign are *lower* than those of peers who were monitored less at the start of a campaign. Again, the effect is similar in magnitude across expected return proxies—ranging from -0.043 to -0.047 standard deviations—and is statistically significant at the 1% level for all return measures. The annualized size of the effect corresponds to a decline in returns of 3.4% in the case of the main return proxy, $z(\epsilon_{i,t}^{MKT})$. It is worth noting that in the monthly sample, a 1 standard deviation increase in monitoring days corresponds to only 0.28 additional monitoring days—meaning that small increases in monitoring intensity have sizeable effects on expected returns.

Summing up, firm-quarter and firm-month level regressions support the predictions in my model that investment, expected returns, and volatilites decrease in monitoring, whereas q increases in monitoring. Results for yields are varied and therefore unclear.

6 Conclusion

This paper explores how asset pricing moments respond to monitoring—an important feature of the equilibrium of an economy in which inside shareholders can extract private benefits. Monitoring increases firm value and reduces the interest rate, the risk premium, volatility, and the dividend yield. As long as monitoring is relaxed enough, increasing the inside shareholder's share of ownership reduces private benefit extraction by raising its marginal cost. Thus, with relaxed monitoring, insider ownership acts as an *incentive alignment* mechanism. However, *overmonitoring* breaks down the mechanism by inducing *perverse incentives*: the inside shareholder uses higher ownership levels to rebel by increasing his extraction of private benefits.

Furthermore, the non-monotonic effect of insider ownership on private benefit extraction propagates into the economy through investment. For instance, if the insider's ownership is too low, such that the outsider *overmonitors*, increasing the insider's ownership raises investment, which increases the volatility of returns and expected returns.

My empirical results support the model predictions that investment, excess returns, and the volatility of returns decrease in monitoring, and that q increases in monitoring. However, my results are inconclusive concerning the impact of monitoring on dividend yields.

To conclude, monitoring is prevalent and has substantial economic impacts. As such, the role of monitoring in asset pricing frameworks warrants further research. For example, one could study a setting with multiple outside shareholders, in which only the shareholder with the maximum share of outside ownership (who has the highest marginal benefit from dividends) monitors the insider. Additionally, one could endogenize monitoring and allow shareholders to trade.

Appendix

Next, I present the proofs for the theorem and the propositions in the body of the paper. The proofs make use of the assumptions below.

Assumption 1 $h > \rho + \delta(1-\gamma)$

Assumption 2 $1 - \alpha < \eta$

Assumption 3 $\gamma\left(1 + (1 + (\bar{\theta} - \theta)(1 + \varsigma)\psi)h\epsilon^2\right)^2 > 2\epsilon^2(\gamma+1)[(1+(\bar{\theta}-\theta)(1+\varsigma)\psi h - (1-\gamma)\delta - \rho]$

Assumption 4 $(1 - X)h > i$

Assumption 5 $\rho + (\gamma - 1)(i - \delta) - \gamma(\gamma - 1)i^2\epsilon^2/2 > 0$

Assumption 6 $(1 - (1 - \alpha)^{1/2}(1 + \varsigma)) \leq \eta^{1/2}/2$

Assumption 1 guarantees positive investment for risk-neutral firms even with perfect investor protection. Assumption 2 bounds the fraction of private benefits that the insider extracts below 1 (monitoring reduces that fraction further). Assumption 3 ensures that the investment-capital ratio is a real number, and Assumption 4 ensures that dividends are positive. Assumption 5 guarantees positive Tobin's q and dividend yield. Assumption 6 makes it so that monitoring costs reduce private benefit extraction compared to a setting with no monitoring. Assumptions 1-5 are as in AW, except that I modify Assumption 3. I also add Assumption 6.

A Proof of Theorem 1

A.1 Inside Shareholder's Value Function

First, guess and verify the controlling shareholder's value function:

$$J_1(K) = \frac{1}{1-\gamma}\left(A_1 K^{1-\gamma} - \frac{1}{\rho}\right) \tag{40}$$

$$J_1'(K) = A_1 K^{-\gamma}$$

$$J_1''(K) = -\gamma A_1 K^{-\gamma-1}$$

$$\frac{\partial p J_1(K)}{\partial D} : M^{-\gamma}\alpha + A_1 K^{-\gamma-1}\gamma\epsilon^2 I = A_1 K^{-\gamma}$$

$$\Leftrightarrow M^{-\gamma}\alpha = A_1 K^{-\gamma} - A_1 K^{-\gamma-1}\gamma\epsilon^2 I$$

$$\Leftrightarrow M^{-\gamma}\alpha = A_1 K^{-\gamma}(1 - K^{-1}\gamma\epsilon^2 I)$$

$$\Leftrightarrow \frac{M^{-\gamma}\alpha}{K^{-\gamma}} = A_1(1-\gamma\epsilon^2 i)\ , \quad i = \frac{I}{K}$$

$$\Leftrightarrow m^{-\gamma}\alpha = A_1(1-\gamma\epsilon^2 i)\ , \quad m = \frac{M}{K} \tag{41}$$

Plugging the optimal private benefit extraction into the insider's cash flow constraint (8):

$$M_1(t) = \alpha D(t) + \frac{1-\alpha}{\eta}(\bar{\theta}-\theta)(1+\varsigma)hK(t) - \frac{\eta}{2}X^2 hK$$

$$\Leftrightarrow m = \alpha d + \frac{h}{2\eta}(\bar{\theta}-\theta)(1+\varsigma)\left(1-\alpha^2\right) \tag{42}$$

Moving back to the flow of funds constraint for investment (6):

$$I(t) = hK(t) - D(t) - X(t)hK(t)$$

$$\Leftrightarrow d = h(1-X) - i\ , \tag{43}$$

Plugging (43) into (42):

$$m = \alpha\left(h(1-X) - i + \frac{h}{2\eta\alpha}(\bar{\theta}-\theta)(1+\varsigma)\left(1-\alpha^2\right)\right)$$

Collect terms scaled by productivity, h:

$$m = \alpha(h(1 + (\bar{\theta} - \theta)(1 + \varsigma)\psi) - i)\ , \qquad \underbrace{\psi = \frac{(1-\alpha)^2}{2\eta\alpha}}_{\text{agency cost parameter}} \tag{44}$$

Back to the HJB equation. Substituting in J_1', J_1'':

$$\rho\left[\frac{1}{1-\gamma}\left(A_1 K^{1-\gamma} - \frac{1}{\rho}\right)\right] = u(M) + (I - \delta K)A_1 K^{-\gamma}) - \frac{\epsilon^2}{2} I^2 \gamma A_1 K^{-\gamma-1}$$

$$\Leftrightarrow \rho\frac{A_1}{1-\gamma} - \frac{1}{(1-\gamma)K^{1-\gamma}} = \frac{m^{1-\gamma}}{1-\gamma} - \frac{1}{1-\gamma}K^{1-\gamma} + (i-\delta)A_1 - \frac{\epsilon^2}{2} i^2 \gamma A_1$$

$$\Leftrightarrow \frac{m^{1-\gamma}}{1-\gamma} - \rho\frac{A_1}{1-\gamma} + (i-\delta)A_1 - \frac{\epsilon^2}{2} i^2 \gamma A_1$$

Using $m^{1-\gamma} = m \cdot m^{-\gamma}$ and plugging in m and $m^{-\gamma}\alpha$:

$$m^{1-\gamma} = \underbrace{\frac{A_1(1-\gamma\epsilon^2 i)}{\alpha}}_{m^{-\gamma}/\alpha} \cdot \underbrace{\alpha((1 + (\bar{\theta} - \theta)(1 + \varsigma)\psi)h - i)}_{m}$$

$$\Leftrightarrow m^{1-\gamma} = A_1(1 - \gamma\epsilon^2 i)[(1 + (\bar{\theta} - \theta)(1 + \varsigma)\psi)h - i] \tag{45}$$

Thus, returning to the HJB equation:

$$\frac{A_1}{1-\gamma}(1 - \gamma\epsilon^2 i)[(1 + (\bar{\theta} - \theta)(1 + \varsigma)\psi)h - i] - \rho\frac{A_1}{-\gamma} + (i-\delta)A_1 - \frac{\epsilon^2}{2} i^2 \gamma A_1$$

$$\Leftrightarrow \frac{A_1}{1-\gamma}\underbrace{(\rho - (i-\delta)(1-\gamma) + \frac{\epsilon^2}{2} i^2 \gamma(1-\gamma))}_{\text{dividend yield, } y} =$$

$$\frac{A_1}{1-\gamma}\underbrace{(1 - \gamma\epsilon^2 i)[(1 + (\bar{\theta} - \theta)(1 + \varsigma)\psi)h - i]}_{\text{dividend yield, } y}$$

$$\Leftrightarrow \frac{\epsilon^2}{2} i^2 \gamma(\gamma - 1) + \gamma\epsilon^2 i^2 - \gamma i - \gamma i \epsilon^2 (1 + (\bar{\theta} - \theta)(1 + \varsigma)\psi)h +$$

$$(1 + (\bar{\theta} - \theta)(1 + \varsigma)\psi)h - (1-\gamma)\delta - \rho + i - i = 0$$

$$\Leftrightarrow \frac{\epsilon^2}{2} i^2 \gamma(\gamma+1) - \gamma i (1 + (1 + (\bar{\theta} - \theta)(1+\varsigma)\psi) h \epsilon^2) +$$
$$(1 + (\bar{\theta} - \theta)(1+\varsigma)\psi) h - (1-\gamma)\delta - \rho = 0$$

This is a quadratic function for i:

$$i = \frac{1}{\epsilon^2 \gamma(\gamma+1)} [\gamma(1 + (1 + (\bar{\theta} - \theta)(1+\varsigma)\psi) h \epsilon^2) \pm \sqrt{\Delta}] \tag{46}$$

$s.t.$

$$\Delta = \gamma^2 (1 + (1 + (\bar{\theta} - \theta)(1+\varsigma)\psi) h \epsilon^2)^2 -$$
$$2\gamma(\gamma+1)\epsilon^2 [(1 + (\bar{\theta} - \theta)(1+\varsigma)\psi) h - (1-\gamma)\delta - \rho]$$

The assumptions require $\Delta > 0$. For $\epsilon = 0$, $i = [(1+(\bar{\theta}-\theta)(1+\varsigma)\psi)h-(1-\gamma)\delta-\rho]/\gamma$. And by continuity , for $\epsilon > 0$, i is the smaller of the two roots (Albuquerque and Wang, 2008).

Solving for A_1 by the expression for m^γ in equation (41):

$$A_1 = \frac{m^{-\gamma}\alpha}{1 - \epsilon^2 i \gamma}$$

Applying the expression for $m^{1-\gamma}$ in equation (45):

$$A_1 = \frac{m^{1-\gamma}}{y} \tag{47}$$

To finish the verification, check the transversality condition:

$$\lim_{T\to\infty} E\left[e^{-\rho T} |J_1(K(T))|)\right] = 0$$

Recall that $J_1(K) = \frac{1}{1-\gamma}\left[A_1 K^{1-\gamma} - \frac{1}{\rho}\right]$. Furthermore, only K has T on its index, so check:

$$\lim_{T\to\infty} E\left[e^{-\rho T} K(T)^{1-\gamma}\right] = 0$$

Applying Itô's lemma on $\log K(t)^{1-\gamma}$ and integrating from 0 to T,

$$K(T)^{1-\gamma} = K(0)^{1-\gamma} \exp\left\{(1-\gamma)\left[(i-\delta-\frac{1}{2}\epsilon^2 i^2)T + \epsilon i d Z(T)\right]\right\}$$

Substituting, into the transversality condition and using the expectation of a log normal random variable, one verifies:

$$\lim_{T\to\infty} e^{-\rho T} K(0)^{1-\gamma} \exp\left\{(1-\gamma)\left[i-\delta-\frac{\epsilon^2 i^2}{2}+\frac{(1-\gamma)^2}{2}\epsilon^2 i^2\right]T\right\},$$

which indeed is zero as long as $\rho > 0$ and $y > 0$, which is true by assumption.

A.2 Outside Shareholder's Value Function

Following the no-trade conjecture (no trade in the risky asset), the outside shareholder invests all his liquid wealth in the risky asset. He also optimally consumes all dividends each period, net of monitoring expenses:

$$J_2(K_0) = E\left[\int_0^\infty e^{-\rho t}\frac{1}{1-\gamma}([(1-\alpha)(1-[(\theta/\bar{\theta})-\varsigma])dK(t)]^{(1-\gamma)}-1)dt\right]$$

Grouping exponential terms:

$$\begin{aligned}=&\frac{1}{1-\gamma}E\Bigg[\int_0^\infty \Bigg([(1-\alpha)(1-[(\theta/\bar{\theta})-\varsigma])dK(0)]^{1-\gamma}\cdot \\ &\exp\left\{(1-\gamma)(i-\delta-\frac{1}{2}\epsilon^2 i^2)t+(1-\gamma)\epsilon i Z(t)-\rho t\right\}-\exp\{-\rho t\}\Bigg)dt\Bigg]\end{aligned}$$

$$\begin{aligned}=&\frac{1}{1-\gamma}\int_0^\infty \Bigg([(1-\alpha)(1-[(\theta/\bar{\theta})-\varsigma])dK(0)]^{1-\gamma}\cdot \\ &\exp\left\{[(1-\gamma)[(i-\delta-\frac{1}{2}\epsilon^2 i^2+\frac{1}{2}(1-\gamma)\epsilon^2 i^2]-\rho]t\right\}-\exp\{-\rho t\}\Bigg)dt\end{aligned}$$

Evaluating the integral and using Assumption 5,

$$=\frac{1}{1-\gamma}\left([(1-\alpha)(1-[(\theta/\bar{\theta})-\varsigma])dK(0)]^{1-\gamma}\frac{1}{\rho-(1-\gamma)[(i-\delta-\frac{1}{2}\gamma\epsilon^2 i^2]}-\frac{1}{\rho}\right)$$

Substituting for the dividend yield:

$$= \frac{1}{1-\gamma}\left([(1-\alpha)(1-[(\theta/\bar{\theta})-\varsigma])dK(0)]^{1-\gamma}\frac{1}{y}-\frac{1}{\rho}\right)$$

And summarizing,

$$J_2(K_0) = \frac{1}{1-\gamma}\left(A_2K(0)^{1-\gamma}-\frac{1}{\rho}\right), \tag{48}$$

$$A_2 = \frac{[(1-\alpha)(1-[(\theta/\bar{\theta})-\varsigma])d]^{1-\gamma}}{y}$$

Using the definition of the dividend yield, $y = \frac{D}{P}\frac{1/K}{1/K} = \frac{d}{q}$:

$$A_2 = \frac{[(1-\alpha)(1-[(\theta/\bar{\theta})-\varsigma])]^{1-\gamma}d^{1-\gamma}}{d/q} = \frac{q[(1-\alpha)(1-[(\theta/\bar{\theta})-\varsigma])]^{1-\gamma}}{d^{\gamma}}$$

The value functions of both shareholders have the same form. Thus, both transversality conditions also have the same form, and the outsider's transversality condition is satisfied.

Let us continue with the Merton consumption and portfolio choice for the outside shareholder. It is possible to substitute straightforwardly into Merton '69, equations 42 and 43:

$$C_2(t) = \frac{\rho}{\gamma} - (1-\gamma)\left[\frac{\lambda^2}{2\sigma_p^2\gamma^2}+\frac{r}{\gamma}\right](1-\alpha)qK(t)\ ,\ \ W_2(t) = (1-\alpha)qK(t) \tag{49}$$

$$\pi(t) = \frac{\lambda}{\gamma\sigma_p^2} \tag{50}$$

Recalling that (1) $\pi = 1$, (2) the equilibrium condition that P and K both follow a geometric Brownian motion with the same drift coefficient, and that (3) $K(t)$ has drift $\epsilon I(t)$, the equilibrium equity risk premium, λ, has the form

$$\lambda = \gamma\sigma_p^2 = \gamma\sigma_K^2 = \gamma\epsilon^2 i^2. \tag{51}$$

Next, obtain the SDF process from the outside shareholder's marginal utility, $\xi_t = e^{-\rho t}C_2(t)^{-\gamma}$. Using the outsider's consumption, $C_2(t) = (1-\alpha)(1-[(\theta/\bar{\theta})-\varsigma])dK(t)$, and applying Itô's lemma:

$$\frac{d\xi_2(t)}{\xi_2(t)} = -(\rho + \gamma(i-\delta) - \frac{1}{2}\gamma(\gamma+1)\epsilon^2 i^2)dt - \epsilon i dZ(t) \tag{52}$$

Thus, the interest rate is:

$$\rho + \gamma(i-\delta) - \frac{1}{2}\gamma(\gamma+1)\epsilon^2 i^2 \tag{53}$$

Given the SDF process, the firm's market value from the perspective of the outside shareholder is:

$$P(t) = \frac{1}{1-\alpha}E_t\left[\int_t^\infty \frac{\xi\,(s)}{\xi\,(t)}(1-\alpha)(1-[(\theta/\bar{\theta})-\varsigma])D(s)ds\right]$$

Solving the integral gives:

$$P(t) = d(1-[(\theta/\bar{\theta})-\varsigma])K(t)\cdot\frac{1}{\rho-(1-\gamma)\left(i-\delta-\frac{1}{2}\gamma\epsilon^2 i^2\right)}$$

Using the definitions for y, A_1, and q:

$$q = (1-[(\theta/\bar{\theta})-\varsigma])\frac{\alpha d}{m}\frac{1}{1-\epsilon i\gamma}$$

$$= (1-[(\theta/\bar{\theta})-\varsigma])\left(1+h(\bar{\theta}-\theta)(1+\varsigma)\left(\frac{1-\alpha^2}{2\eta d\alpha}\right)\right)^{-1}\left(\frac{1}{1-\epsilon^2 i\gamma}\right)$$

Q.E.D.

B Proofs of Propositions

B.1 Proof of Proposition 1

Given $f(x)$, the quadratic function for i, $f(0)$ gives $h(1+(\bar{\theta}-\theta)(1+\varsigma)\psi)-\delta(1-\gamma)-\rho$, which is larger than zero by assumption 1. Taking i as the smaller of the roots for f, $f(0)$ implies $i>0$.

Taking the total derivative of f with respect to η yields

$$\begin{aligned}\frac{df}{d\theta} = 0 = &-\frac{di}{d\eta}\left[\gamma\left(1-\epsilon^2\gamma i+\epsilon^2[h(1+(\bar{\theta}-\theta)\psi)-i]\right)\right] \\ &+\frac{d\psi}{d\eta}\left[h(\bar{\theta}-\theta)(1+\varsigma)\left(1-\epsilon^2\gamma i\right)\right] \\ &+\left(\frac{d\bar{\theta}}{d\eta}-\frac{d\theta}{d\eta}\right)\left[h(1+\varsigma)\psi\left(1-\epsilon^2\gamma i\right)\right],\end{aligned}$$

$$s.t.\quad \frac{d\psi}{d\eta}=-\frac{(1-\alpha)^2}{2\alpha\eta^2},\quad \frac{d\theta}{d\eta}=-\left(\frac{1-\alpha}{\eta^3}\right)^{1/2},\quad \frac{d\bar{\theta}}{d\eta}=-\eta^{-3/2}.$$

This expression implies that $di/d\eta<0$. To see this, first note that Assumption 5 guarantees that $f(i)<f(\epsilon^{-2}\gamma^{-1})$, such that $1-\epsilon^2\gamma i>0$. Additionally, Assumption 4 implies that $h(1+(\bar{\theta}-\theta)\psi)>i$. Thus, the first term above is negative. The second term is negative, since $\theta<\bar{\theta}$ and $d\psi/d\eta<0$. And the last term is negative because $d\bar{\theta}/d\eta$, $d\theta/d\eta<0$ and $|d\bar{\theta}/d\eta|>|d\theta/d\eta|$. Therefore, rearranging for $di/d\eta$ yields a negative term.

By the implicit function theorem:

$$\frac{\partial i}{\partial\theta}=-\frac{\partial f/\partial\theta}{\partial f/\partial i}$$

$$=\frac{1}{\gamma}\frac{h(1+\varsigma)\psi\left(\epsilon^2 i\gamma-1\right)}{1-\epsilon^2 i\gamma+\left(1+(\bar{\theta}-\theta)(1+\varsigma)\psi)h-i\right)\epsilon^2}<0$$

Q.E.D.

It is convenient now to use the implicit function again to obtain $di/d\psi$, as I use this term in subsequent proofs:

$$\frac{\partial i}{\partial \psi} = -\frac{\partial f/\partial \psi}{\partial f/\partial i}$$

$$= \frac{1}{\gamma}\frac{(\bar{\theta}-\theta)(1+\varsigma)h\left(1-\epsilon^2 i\gamma\right)}{1-\epsilon^2 i\gamma+\left(1+(\bar{\theta}-\theta)(1+\varsigma)\psi)h-i\right)\epsilon^2} > 0$$

B.2 Proof of Proposition 2

I first prove that q decreases in η and then prove that q decreases in θ. Note that $q = d/y$. Then,

$$\frac{d\log q}{d\eta} = \frac{1}{d}\left[-h\frac{dX}{d\eta} - \frac{di}{d\eta} - \frac{d}{y}\frac{dy}{d\eta}\right]$$

$$= \frac{1}{d}\left[h(1-\alpha)\frac{(\bar{\theta}-\theta)(1+\varsigma)}{2\eta^2} - \frac{di}{d\eta}\left(1+q(\gamma-1)\left[1-\gamma\epsilon^2 i\right]\right)\right]$$

$$= \frac{1}{d}\Bigg[h(1-\alpha)\frac{(\bar{\theta}-\theta)(1+\varsigma)}{2\eta^2} - \frac{di}{d\eta}\left(1+h(\bar{\theta}-\theta)(1+\varsigma)\frac{1-\alpha^2}{2\eta d\alpha}\right)^{-1}\cdot$$
$$\left((\gamma-1)(1-[(\theta/\bar{\theta})-\varsigma])+1+h(\bar{\theta}-\theta)(1+\varsigma)\frac{1-\alpha^2}{2\eta d\alpha}\right)\Bigg]$$

Where the above is larger than zero, applying $di/d\eta < 0$ (Proposition 1), and $1 > (1-[(\theta/\bar{\theta})-\varsigma])$. As for θ:

$$\frac{d\log(q)}{d\theta} = \frac{1}{d}\left[-h\frac{dX}{d\theta} - \frac{di}{d\theta} - q\frac{dy}{d\theta}\right]$$

$$= \frac{1}{d}\left[\frac{h(1-\alpha)(1+\varsigma)}{\eta} - \frac{di}{d\theta}\left(1+q(\gamma-1)\left(1-\gamma\epsilon^2 i\right)\right)\right]$$

$$= \frac{1}{d}\Bigg[\frac{h(1-\alpha)(1+\varsigma)}{\eta} - \frac{di}{d\theta}\left(1+h(\bar{\theta}-\theta)(1+\varsigma)\frac{1-\alpha^2}{2\eta d\alpha}\right)^{-1}\cdot$$
$$\left((\gamma-1)(1-[(\theta/\bar{\theta})-\varsigma])+1+h(\bar{\theta}-\theta)(1+\varsigma)\frac{1-\alpha^2}{2\eta d\alpha}\right)\Bigg]$$

Where, similar to the previous statement, the above is larger than zero by applying $di/d\theta < 0$ (Proposition 1) and $1 > (1-[(\theta/\bar{\theta})-\varsigma])$.

Q.E.D.

B.3 Proof of Proposition 3

Recall that $r = \rho + \gamma(i - \delta) - (1/2)\epsilon^2 i^2 \gamma(\gamma + 1)$. Then:

$$\frac{dr}{d\psi} = \frac{\gamma di}{d\psi}\left(1 - (\gamma + 1)\varepsilon^2 i\right)$$

$$\frac{dr}{d\theta} = \frac{\gamma di}{d\theta}\left(1 - (\gamma + 1)\varepsilon^2 i\right)$$

And by applying $di/d\psi > 0$ and $di/d\theta < 0$ (Proposition 1), $1 > \epsilon^2(\gamma + 1)i \implies dr/d\psi > 0,\ dr/d\theta < 0$.

Q.E.D.

B.4 Proof of Proposition 4

Recall that $\sigma_P^2 = \epsilon^2 i^2$ and $\lambda = \gamma\sigma_P^2$. Recall as well that $d\psi/d\eta,\ d\psi/d\alpha < 0$, and $di/d\psi > 0$. Thus, σ_P and λ decrease in η. However, it is not necessarily the case that σ_P and λ decrease in α.

Recall that $d\theta/d\eta > 0,\ d\theta/d\alpha < 0$, and $di/d\theta < 0$. This also implies that σ_P and λ decrease in η, and—contrary to the implication from the derivatives with respect to ψ—that σ_P and λ increase in α.

Whether volatility and the risk premium decrease or increase in α depends on which of the transmission channels for α—ψ or θ—dominates. When α is too low such that the outsider *overmonitors* and the insider *rebels*, investment increases in α, and so do σ_P and λ. The opposite is true when α is high enough.

Turning to total equity returns $(r + \lambda)$:

$$\frac{d\gamma\epsilon^2 i^2 + r}{d\psi} = \frac{di}{d\psi}\gamma(\epsilon^2 i + 1 - \epsilon^2 i\gamma)$$

$$\frac{d\gamma\epsilon^2 i^2 + r}{d\theta} = \frac{di}{d\theta}\gamma(\epsilon^2 i + 1 - \epsilon^2 i\gamma)$$

By Proposition 1, $di/d\psi > 0$ and $di/d\theta < 0$; and by Assumption 5, $1 > \epsilon^2 i\gamma$. Thus, total returns decrease in η through both the ψ and θ channels.

However, as with volatility and the risk premium, whether total returns increase or decrease in α depends on which of ψ or θ is the more dominant transmission channel. Because of the non-monotonic effects of α on private benefit extraction, their spillover unto investment, and the sensitivity of total returns to investment, total returns increase in α when α is too low and perverse incentives are at play. But total returns decrease in α when the insider's ownership is sufficiently high.

Q.E.D.

B.5 Proof of Proposition 5

The expression for the dividend yield is $y = \rho + (\gamma - 1)(i - \delta - (\gamma/2)\epsilon^2 i^2)$. Hence,

$$\frac{dy}{d\psi} = \frac{di}{d\psi}(\gamma - 1)(1 - \gamma\epsilon^2 i) \gtrless 0 \Leftrightarrow \gamma \gtrless 1$$

$$\frac{dy}{d\theta} = \frac{di}{d\theta}(\gamma - 1)(1 - \gamma\epsilon^2 i) \gtrless 0 \Leftrightarrow \gamma \lessgtr 1$$

Therefore, the dividend yield decreases in investor protection, η, when γ is greater than 1 and increases in investor protection when γ is less than 1. Additionally, when γ is greater than 1, the dividend yield increases (decreases) in α if the θ (ψ) channel dominates. The θ channel dominates when α is too low, whereas the ψ channel dominates when α is high enough. The dividend yield's derivatives change sign when γ is less than 1.

Q.E.D.

C Maximal Monitoring and Perfect Investor Protection

In this section, I prove the form of q under maximal monitoring and perfect investor protection.

C.1 Maximal Monitoring

To attain maximal monitoring ($\theta \to \bar{\theta}$), one must arbitrarily decrease α. Now notice that

$$\lim_{\alpha \to 0} \psi = -(1/\eta),$$

and that

$$\lim_{\theta \to \bar{\theta}} \Delta = \gamma^2(1+h\epsilon^2)^2 - 2\epsilon^2\gamma(\gamma+1)[h-(1-\gamma)\delta-\rho].$$

Therefore,

$$\lim_{\theta \to \bar{\theta}} i = \left[\gamma\left(1+h\epsilon^2\right) - \left[\gamma^2\left(1+h\epsilon^2\right)^2 - 2\epsilon^2\gamma(\gamma+1)[h-(1-\gamma)\delta-\rho]\right]^{1/2}\right] \cdot \left[\epsilon^2\gamma(\gamma+1)\right]^{-1},$$

Which is a constant, $i^*(\theta)$. It then follows that

$$\lim_{\theta \to \bar{\theta}} q = \varsigma[1-\epsilon^2\gamma i^*(\theta)]^{-1}$$

Q.E.D.

C.2 Perfect Investor Protection

Recall that $\eta \to \infty$ characterizes perfect investor protection. Now define $i^*(\eta) \equiv \lim_{\eta \to \infty} i$, and then note that $i^*(\eta) = i^*(\theta)$. Turning to monitoring terms, $\lim_{\eta \to \infty} \bar{\theta} = 0$ and $\lim_{\eta \to \infty} \theta = 0$. Additionally, $\lim_{\eta \to \infty}(\theta/\bar{\theta}) = (1-\alpha)^{1/2}$.

The form for $q^*(\eta) \equiv \lim_{\eta \to \infty} q$ is then immediate.

Q.E.D.

D Welfare Analysis

This section of the Appendix considers how the outsider's utility cost and the insider's utility gain respond to monitoring. I follow Albuquerque and Wang (2008) by first defining the outsider's utility cost and the insider's utility gain, subsequently verifying that these two elements decrease in investor protection, and lastly proving that they decrease in monitoring.

As in Albuquerque and Wang (2008), the outsider's utility cost is the fraction of capital, $1 - \zeta_2$, that she would be willing to give up to move from a state of imperfect investor protection to perfect investor protection. Similarly, the insider's utility gain is the fraction of capital, $\zeta_1 - 1$ that he would require to transition from imperfect investor protection to perfect investor protection. Both the outsider's utility cost and the insider's utility cost take the same form as in Albuquerque and Wang (2008).

Let J_2^* define the outsider's value function under perfect investor protection. Then, the expression $J_2^*(\zeta_2 K(0)) = J_2(K(0))$ identifies ζ_2. Recall that

$$J_2(K(0)) = \frac{1}{1-\gamma}\left(\frac{[(1-\alpha)(1-[(\theta/\bar{\theta})-\varsigma])d]^{1-\gamma}}{y}K(0)^{1-\gamma} - \frac{1}{\rho}\right)$$

and define

$$J_2^*(K(0)) = \frac{1}{1-\gamma}\left(\frac{\left[(1-\alpha)(1-[(\theta/\bar{\theta})-\varsigma])d^*\right]^{1-\gamma}}{y^*}K(0)^{1-\gamma} - \frac{1}{\rho}\right),$$

such that d^*, y^* are the dividend-capital ratio and dividend yield under perfect investor protection. Then,

$$\zeta_2 = \frac{d}{d^*}\left(\frac{y^*}{y}\right)^{1/(1-\gamma)}.$$

Similarly, defining J_1^* as the insider's value function under perfect investor protection, $J_1^*(\zeta_1 K(0)) = J_1(K(0))$ identifies the outsider's utility gain. Moreover, define

$$J_1^*\left(K_0\right) = \frac{1}{1-\gamma}\left(\frac{m^{*1-\gamma}}{y^*}K(0)^{1-\gamma} - \frac{1}{\rho}\right)$$

where m^* is the outsider's cash flow per unit of capital under perfect investor protection. Then,

$$\zeta_1 = \frac{m}{m^*}\left(\frac{y^*}{y}\right)^{1/(1-\gamma)}$$

From the expressions for ζ_2:

$$\frac{d\log\zeta_2}{d\eta} = \frac{d\log d}{d\eta} - \frac{1}{1-\gamma}\frac{d\log y}{d\eta}$$

$$= \frac{1}{d}\frac{d\left(\left(1-X\right)h - i\right)}{d\eta} + \frac{1}{y}\frac{di}{d\eta}\left(1-\epsilon^2 i\gamma\right)$$

$$= \frac{1}{d}\frac{h(1-\alpha)}{2\eta^2}(\bar{\theta}-\theta)(1+\varsigma)\Big) - \frac{1}{d}\frac{di}{d\eta}\left(\frac{\hat{q}-q}{\hat{q}}\right) > 0$$

where the inequality follows from $\hat{q} - q > 0$.

As for ζ_1:

$$\frac{d\log m}{dn} - \frac{1}{1-\gamma}\frac{d\log y}{d\eta}$$

$$= \frac{\alpha h\psi}{2\eta}(\bar{\theta}-\theta)(1+\varsigma) - 2\frac{2\alpha h\psi}{2\eta}(\bar{\theta}-\theta)(1+\varsigma) - \frac{\alpha}{m}\frac{di}{d\eta} + \frac{1-\gamma\epsilon^2 i}{y}\frac{di}{d\eta}$$

$$= -\frac{\alpha h}{m}\left(\frac{\psi(\bar{\theta}-\theta)(1+\varsigma)}{2\eta}\right) < 0$$

Thus, the outsider's utility cost and the insider's utility gain decrease in investor protection. Next, I show that the outsider's utility cost and the insider's utility gain both decrease in monitoring.

Starting with the outsider's utility cost:

$$\frac{d\log\zeta_2}{d\theta} = \frac{1}{d}\frac{d((1-X)h-i)}{d\theta} - \frac{1}{1-\gamma}\frac{d\log y}{d\theta}$$

$$= \frac{1}{d}\left(-h\frac{dX}{d\theta} - \frac{di}{d\theta}\right) + \frac{1}{y}(1-\gamma\epsilon^2 i)\frac{di}{d\theta}$$

$$= \frac{1}{d}\left(h\frac{(1-\alpha)(1+\varsigma)}{\eta}\right) - \frac{1}{d}\frac{di}{d\theta}\left(\frac{\hat{q}-q}{\hat{q}}\right) > 0,$$

where the inequality follows from Proposition 1. Moving to the insider's utility cost:

$$\frac{d\log\zeta_1}{d\theta} = \frac{1}{m}\frac{dm}{d\theta} - \frac{1}{1-\gamma}\frac{1}{y}\frac{dy}{d\theta}$$

$$= \frac{-\alpha}{m}\left(h(1+\varsigma)\psi + \frac{di}{d\theta}\right) + \frac{1}{y}(1-\gamma\epsilon^2 i)\frac{di}{d\theta}$$

$$= -\frac{\alpha}{m}h(1+\varsigma)\psi < 0$$

Q.E.D.

References

Aghion, P., & Tirole, J. (1997). Formal and real authority in organizations. *Journal of political economy*, *105*(1), 1–29.

Albuquerque, R., Brandão-Marques, L., Ferreira, M. A., & Matos, P. (2019). International corporate governance spillovers: Evidence from cross-border mergers and acquisitions. *The Review of Financial Studies*, *32*(2), 738–770.

Albuquerque, R., Fos, V., & Schroth, E. (2022). Value creation in shareholder activism. *Journal of Financial Economics*, *145*(2), 153–178.

Appel, I. R., Gormley, T. A., & Keim, D. B. (2019). Standing on the shoulders of giants: The effect of passive investors on activism. *The Review of Financial Studies*, *32*(7), 2720–2774.

Backus, M., Conlon, C., & Sinkinson, M. (2021). Common ownership in america: 1980–2017. *American Economic Journal: Microeconomics*, *13*(3), 273–308.

Basak, S., Chabakauri, G., & Yavuz, M. D. (2019). Investor protection and asset prices. *The Review of Financial Studies*, *32*(12), 4905–4946.

Bebchuk, L. A., Brav, A., Jiang, W., & Keusch, T. (2020). Dancing with activists. *Journal of Financial Economics*, *137*(1), 1–41.

Becht, M., Franks, J., Grant, J., & Wagner, H. F. (2017). Returns to hedge fund activism: An international study. *The Review of Financial Studies*, *30*(9), 2933–2971.

Becht, M., Franks, J., Mayer, C., & Rossi, S. (2009). Returns to shareholder activism: Evidence from a clinical study of the hermes uk focus fund. *The Review of Financial Studies*, *22*(8), 3093–3129.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307–327.

Brav, A., Jiang, W., & Kim, H. (2015). The real effects of hedge fund activism: Productivity, asset allocation, and labor outcomes. *The Review of Financial Studies*, *28*(10), 2723–2769.

Brav, A., Jiang, W., Partnoy, F., & Thomas, R. (2008). Hedge fund activism, corporate governance, and firm performance. *The Journal of Finance*, *63*(4), 1729–1775.

Burkart, M., Gromb, D., & Panunzi, F. (1997). Large shareholders, monitoring, and the value of the firm. *The quarterly journal of economics*, *112*(3), 693–728.

Cain, M. D., McKeon, S. B., & Solomon, S. D. (2017). Do takeover laws matter? evidence from five decades of hostile takeovers. *Journal of Financial Economics*, *124*(3), 464–485. https://doi.org/https://doi.org/10.1016/j.jfineco.2017.04.003

Chabakauri, G., Fos, V., & Jiang, W. (2022). Trading ahead of barbarians' arrival at the gate: Insider trading on non-inside information. *Columbia Business School Research Paper Forthcoming*.

Clifford, C. P. (2008). Value creation or destruction? hedge funds as shareholder activists. *Journal of Corporate Finance*, *14*(4), 323–336.

Collin-Dufresne, P., & Fos, V. (2015). Do prices reveal the presence of informed trading? *The Journal of Finance*, *70*(4), 1555–1582.

Davis, S. J., & Haltiwanger, J. (1992). Gross job creation, gross job destruction, and employment reallocation. *The Quarterly Journal of Economics*, *107*(3), 819–863.

Denes, M. R., Karpoff, J. M., & McWilliams, V. B. (2017). Thirty years of shareholder activism: A survey of empirical research. *Journal of Corporate Finance*, *44*, 405–424.

Dimson, E., Karakaş, O., & Li, X. (2015). Active ownership. *The Review of Financial Studies*, *28*(12), 3225–3268.

Edmans, A., & Holderness, C. G. (2017). Blockholders: A survey of theory and evidence. *The handbook of the economics of corporate governance*, *1*, 541–636.

Gantchev, N. (2013). The costs of shareholder activism: Evidence from a sequential decision model. *Journal of Financial Economics*, *107*(3), 610–631.

Gogineni, S., Linn, S. C., & Yadav, P. K. (2020). Vertical and horizontal agency problems in private firms: Ownership structure and operating performance. *Journal of Financial and Quantitative Analysis (Forthcoming)*.

Greenwood, J., Hercowitz, Z., & Krusell, P. (1997). Long-run implications of investment-specific technological change. *The American economic review*, 342–362.

Greenwood, J., Hercowitz, Z., & Krusell, P. (2000). The role of investment-specific technological change in the business cycle. *European Economic Review*, *44*(1), 91–115.

Greenwood, R., & Schor, M. (2009). Investor activism and takeovers. *Journal of Financial Economics*, *92*(3), 362–375.

Hansen, L. P., & Singleton, K. J. (1982). Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica: Journal of the Econometric Society*, 1269–1286.

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

Karpoff, J. M., Malatesta, P. H., & Walkling, R. A. (1996). Corporate governance and shareholder initiatives: Empirical evidence. *Journal of financial economics*, *42*(3), 365–395.

Klein, A., & Zur, E. (2009). Entrepreneurial shareholder activism: Hedge funds and other private investors. *The Journal of Finance*, *64*(1), 187–229.

Liu, Z., & Winegar, A. (2022). Economic magnitudes within and across finance. *SSRN*.

Manne, H. G. (1965). Mergers and the market for corporate control. *Journal of Political economy*, *73*(2), 110–120.

McCahery, J. A., Sautner, Z., & Starks, L. T. (2016). Behind the scenes: The corporate governance preferences of institutional investors. *The Journal of Finance*, *71*(6), 2905–2932.

Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The review of Economics and Statistics*, 247–257.

Nissim, D. (2019). Ebitda, ebita, or ebit? *Columbia Business School Research Paper*, (17-71).

Pagano, M., & Röell, A. (1998). The choice of stock ownership structure: Agency costs, monitoring, and the decision to go public. *The Quarterly Journal of Economics*, *113*(1), 187–225.

Pettenuzzo, D., Sabbatucci, R., & Timmermann, A. (2020). Cash flow news and stock price dynamics. *The Journal of Finance*, *75*(4), 2221–2270.