Large Shareholder Activism, Risk Sharing, and Financial Market Equilibrium

Anat R. Admati and Paul Pfleiderer

Stanford University

Josef Zechner

University of Vienna

We develop a model in which a large investor has access to a costly monitoring technology affecting securities' expected payoffs. Allocations of shares are determined through trading among risk-averse investors. Despite the free-rider problem associated with monitoring, risk-sharing considerations lead to equilibria in which monitoring takes place. Under certain conditions the equilibrium allocation is Pareto efficient and all agents hold the market portfolio of risky assets independent of the specific monitoring technology. Otherwise distortions in risk sharing may occur, and monitoring activities that reduce the expected payoff on the market portfolio may be undertaken.

I. Introduction

Following the recommendations of classic portfolio theory, many individual investors, especially in the United States, hold widely diversi-

We would like to thank Alex Benos, Philippe Henrotte, David Hirshleifer, Charlie Jacklin, Charlie Kahn, Alan Kraus, Maureen McNichols, Mark Wolfson, two anonymous referees, the editor (José Scheinkman), and especially Jeff Zwiebel for helpful discussions and comments. We have also benefited from the comments of seminar participants at Stanford University, Tel Aviv University, University of British Columbia, University of California at Los Angeles, University of Chicago, University of Utah, Simon Fraser University, University of California at Berkeley, Insead, London Business School, Stockholm School of Economics, and the Norwegian School of Management. The first two authors acknowledge support from the National Science Foundation (grant SBR-9308238). The third author acknowledges support from the Social Sciences and Humanities Research Council of Canada (grant 410-90-1690).

[*Journal of Political Economy*, 1994, vol. 102, no. 6]

fied portfolios. It has been argued, however, that the push toward diversification has hurt U.S. firms inasmuch as investors who have tiny positions in many companies have little incentive to monitor the performance of any of these companies and to exert pressure on management when improvements can be made. Porter (1992), for example, argues that the patterns of ownership and the structure of capital markets in the United States have put the United States at a competitive disadvantage relative to Germany and Japan, where banks hold large positions in firms and tend to monitor those firms closely.[1] In this paper we look at the potentially conflicting goals of achieving a high rate of monitoring, which is promoted by concentrated ownership, and realizing risk-sharing gains, which usually requires a more diffuse pattern of ownership. In particular, we ask how incentives to monitor are determined when risk-averse investors can trade freely in the market and cannot make prior commitments to monitor firms at any particular level of intensity. In contrast to most financial models of asset pricing in which the payoffs of risky securities are taken to be independent of the allocation of shares, in our analysis the ownership structure affects the payoffs of firms since it affects the amount of monitoring that occurs.

While the typical individual investor has a minuscule stake in each of the firms he or she holds, many of the institutional funds through which investors achieve diversification, by virtue of their large sizes, have significant positions in many companies. In fact, most firms in the United States have large shareholders, consisting of institutional investors such as pension funds as well as private investors such as initial owners.[2] Large shareholders, and particularly institutional investors, have become increasingly active in recent years.[3] The reasons

[1] In Germany, the three largest banks control 36 percent of the voting shares in the 100 largest public companies. In Japan, a small number of banks and insurers own 30–40 percent of many public companies. For a discussion of these data and additional data and references, see Black (1992*b*).

[2] Of 456 of the Fortune 500 firms surveyed in the *1981 CDE Stock Ownership Directory: Fortune 500*, 354 (78 percent) have at least one shareholder who owns at least 5 percent of the shares; in all but 15 firms, the largest shareholder owns at least 3 percent of the firm. As is well known, the size of institutional investors has steadily increased over the past 40 years. According to Brancato (1991), institutional investors held 53.2 percent of the total market value of equity of the largest 100 companies in the United States in 1989. By 1994 the market capitalization of a single institutional investor, CalPERs (the California employees' pension fund), exceeded $80 billion.

[3] For a description of monitoring activities, see "Why Dale Hanson Won't Go Away," *Institutional Investor* (April 1990), pp. 79–84. Some of the largest pension funds in the United States have recently adopted the strategy dubbed "just vote no," which involves simply voting against management's proposals, e.g., for board members (see "Negative Votes, Positive Results," *San Jose Mercury News* [May 17, 1992]). Empirical evidence that large institutional investors are in fact better informed than the average shareholder and that they are more likely to vote against proposals with negative effects on

for this change may be a combination of the following. First, as discussed above, the size of many institutional investors has grown, and this has made monitoring more beneficial to them. Second, for various reasons takeovers have become more difficult and less common, and this has decreased the degree to which improvements can be brought about through this mechanism. Third, recent changes in proxy rules have made communication and coordination among shareholders easier, thereby lowering the effective cost of monitoring ("New Proxy Rules Embolden Shareholders," *N.Y. Times* [May 31, 1993]).

Monitoring by large shareholders typically involves identifying companies whose actions are in conflict with shareholders' interests and attempting to bring about change through negotiation with management, proxy fights, and involvement in the choice of board members. Shareholders clearly incur private costs when they are active. Even in order to "just vote no," as advocated by Grundfest (1993), information must be gathered to determine which management proposals should be challenged. Many institutional investors are also concerned that they might incur some legal liability if they take on active roles. Nevertheless, the recent trend toward shareholder activism suggests that at least some investors still find it in their best interest to engage in monitoring.

As suggested above, there is a trade-off associated with different ownership structures. While concentrated ownership encourages monitoring activities by large investors, it leads to a potential loss in risk-sharing benefits that are realized when ownership is diffuse. Since the structure of corporate ownership is determined in the marketplace, we must ask whether the market properly accounts for the social costs and benefits of various ownership structures, thereby producing the optimal structure. Other than some disclosure requirements, there are no significant impediments to a large investor's acquisition of a large stake in a firm either by the purchase of shares in the open market or by a tender offer. Conversely, when ownership is concentrated, it is relatively easy to attain a more diffuse ownership structure by selling off shares. If the market fails to produce the optimal structure, the reason must be that some agents do not fully realize the benefits or costs they produce by the size of their ownership stakes and the actions they take.

A potential source of market failure, and one that is central to this

---

stock prices is documented, e.g., by Jarrell and Poulsen (1987) and Brickley, Lease, and Smith (1988). For a discussion of the various forms of shareholder activism, the regulatory environment related to institutional shareholder activism, and some empirical findings related to this issue, see Black (1992*a*, 1992*b*).

paper, is the free-rider problem associated with shareholder activism or monitoring. This problem arises because small and passive shareholders realize the benefits of monitoring done by large shareholders but they incur none of the costs. The following simplified version of our model illustrates some of the issues that arise in this context. Suppose that there is a large, risk-averse investor, called L, who has the resources to hold a significant fraction of the shares of an existing firm. The shares not held by L are held by a fringe of small, risk-averse investors. Suppose that L can choose to invest resources in order to increase the expected payoffs of the shares to all shareholders, but small investors have no ability to do this. In this setting the first-best allocation for risk-sharing purposes is one in which L holds a fraction of the shares consistent with his relative tolerance for bearing risk. On the other hand, the first-best level of monitoring is obtained if L holds the entire firm, but this is clearly inefficient from a risk-sharing perspective.

Now suppose that trade is done via a Walrasian mechanism whereby agents submit demand functions and a market-clearing price is established. Suppose also that L cannot commit to a particular monitoring level before trading; rather, L always monitors at a level that is optimal given his equilibrium allocation of shares.[4] If L changes his holdings through trade and this leads to a different level of monitoring, then the price at which trade takes place will reflect the new monitoring level. This means that L can capture the gains of monitoring only on the shares he owns initially, but not on any newly acquired shares, for which he must pay a price that reflects the eventual monitoring. In effect, monitoring acts like a type of transaction tax on L. One of our main results is that if allocations are determined through a Walrasian mechanism in which L can commit to one round of trading, then the transaction tax described above will typically lead to less trade away from the endowments, resulting in an efficiency loss in terms of risk sharing. When there are multiple risky securities, we show that L will typically not hold the market portfolio and, in fact, may engage in monitoring activities that, although they increase the expected return on his portfolio, reduce the expected return of the market portfolio. Obviously, the use of such activities is socially suboptimal.

We obtain dramatically different results when it is not possible for

[4] For most of the paper we shall maintain the assumption that the large investor cannot commit to a particular monitoring level prior to trading. Section VII examines how our results would change if such a commitment were possible. That section also includes a discussion of mechanisms that might be available to certain large investors to make this commitment and why we believe that the no-commitment assumption is more appropriate.

L to commit to a last round of trade. The most striking result in this context obtains when L's effectiveness as a monitor does not depend on his holdings. In this case we show that the unique equilibrium allocation achieves first-best risk sharing; it is the same as the one obtained in a perfectly competitive version of the model when monitoring is not available and involves all agents' holding the market portfolio of risky assets in proportion to their relative risk tolerance.[5] This implies that the allocation is independent of the exact form of the monitoring technology. Of course, the actual level of monitoring is chosen by L optimally given this allocation, and it does depend on the costs and benefits of monitoring for the given technology. While the level of monitoring is not socially optimal (since L does not hold the entire market), there is no loss in terms of risk sharing in this model. Moreover, no activity that reduces the expected return on the market portfolio is undertaken.

In order to understand better the implications of the Walrasian trading mechanism on the equilibrium outcome, we examine another trading environment in which L can capture a fixed fraction (possibly all) of the net surplus generated by his trading and subsequent monitoring. For this model we show that L generally holds more of the risky securities and monitors at a higher level than he does in a Walrasian trading mechanism. However, the equilibrium allocation is still not efficient in terms of risk sharing, because of the inability to commit to a specific monitoring level. For example, in order to extract the surplus generated by an increased monitoring level, L's holdings must increase so as to induce the increased monitoring level, and this typically leads L to hold too much of the risky securities. If L was able to commit to a specific monitoring level prior to trading, then the overall first-best outcome would obtain; that is, L would hold the appropriate fraction of the market portfolio and monitor as though he were the sole owner of all firms.

We also discuss briefly how L's size, as measured by his risk tolerance, might be determined endogenously. The idea is that while a larger size saves in transaction costs, it leads to more monitoring, which is costly to the investors in the fund. This trade-off can lead to an interior equilibrium size for the fund, with a positive fraction of the agents choosing to invest on their own. This is a socially inefficient outcome since agents do not fully utilize potential savings in transactions costs and engage in suboptimal monitoring. The source of this result is the free-rider problem associated with monitoring.

[5] This result is obtained under a condition concerning the concavity of L's objective function. If this condition does not hold, then the type of equilibrium we define in this model does not exist.

There are a number of papers related to ours. However, we are not aware of any paper that focuses on monitoring in a portfolio setting with multiple risky securities and risk-averse investors and considers a Walrasian-type trading mechanism for determining the allocation of shares. Shleifer and Vishny (1986) explore the role of a large shareholder in improving a firm's expected performance through a takeover mechanism. All investors are assumed to be risk neutral, and in most of the analysis, the large investor's holdings are given exogenously. Shleifer and Vishny discuss briefly the large investor's choice of ownership before the monitoring decision is made, but they limit their analysis to the case in which the large investor can commit to adjusting his position only once. They show that if the large investor starts with zero initial holdings, then he has no incentive to acquire shares and become an active shareholder. As we show here, this result does not hold with risk aversion.[6]

In a closely related paper, Huddart (1992) discusses a model in which a large shareholder can monitor the effort taken by the manager through the observation of a costly signal. His model involves a particular monitoring technology in which monitoring affects both the expected payoff and the variance of the payoff. By contrast, we allow for multiple risky securities and employ a general reduced-form model of the monitoring technology, but we assume that monitoring affects only the expected payoffs of risky securities.[7] Huddart (1991) models a manager who owns a fraction of the firm and can affect the firm's payoff by expending effort. The paper discusses the endogenous determination of the manager's holdings and concludes that if trading is anonymous, the manager's equilibrium holdings are zero, leading to a diminished joint surplus. These two papers assume, as we do, constant absolute risk aversion and normal distributions. However, our model is more explicit in describing the trading mechanisms by which allocations are determined.

Our analysis focuses on the notion that large shareholders provide a *public good* and that they incur a *private cost* in doing so. Ownership of a significant fraction of a firm, however, often confers various private benefits as well, and they have been analyzed and measured in

[6] In another related paper, which focuses on takeovers, Kyle and Vila (1991) analyze a model with a large investor who first buys shares in a market and then possibly makes a tender offer to take over the firm and improve its payoffs. In their model, in contrast to ours, all agents are risk-neutral, and the takeover premium is exogenously specified. The presence of liquidity or "noise" traders allows successful takeovers to occur in some cases even when the raider does not initially hold a substantial fraction of the firm.

[7] This is consistent with much of the informal descriptions of shareholder activism. In particular, large shareholders rarely get directly involved with production decisions that would affect the riskiness of the payoffs.

some previous papers. For example, Barclay and Holderness (1989) attempt to infer the private benefits of control from the difference between block prices and the corresponding postannouncement exchange price of a share.[8] Zwiebel (1990) analyzes equilibrium ownership structures in a model in which the private benefits of control are divisible.

A paper related to our discussion in Section IV is DeMarzo and Bizer (1993), which analyzes various economic settings in which it is not possible to commit to a last round of trade. For an agency model similar to ours, DeMarzo and Bizer characterize the sequential equilibrium outcomes of their trading model.[9] An important difference between their paper and our analysis is that they assume a finite set of possible prices at which trade can occur.

The rest of the paper is organized as follows. Section II describes the basic model. In Section III we analyze the case of one round of trading. Section IV discusses equilibrium when a commitment to a last round of trade is not possible. In Section V we analyze the case in which L is able to extract at least some of the surplus generated by his trade and monitoring. Section VI illustrates by an example how L's size can be determined endogenously. In Section VII we summarize our results and contrast them with the case in which commitment to a specific monitoring level can be made. Concluding remarks are offered in Section VIII.

II. The Model

We model a financial market with N risky securities that represent equity shares of N firms. Without loss of generality, each risky security's total supply is normalized to one, so that the supply vector for all securities is the vector of ones, denoted by e. In addition, there is a riskless security, which is in perfectly elastic supply and whose gross return is normalized to one.

We assume that there is one large investor who has access to some

[8] Although in most cases the block price is higher than the exchange price, suggesting a premium that may be interpreted as measuring private benefits of control, these authors find that in about 20 percent of the cases there is actually a discount associated with the block price. Such discounts suggest that there is a private cost (as in our analysis) as well as a private benefit to being a large investor. Indeed, Barclay and Holderness (1989, p. 392) state that "more substantial discounts are often associated with firms in severe financial distress at the time of trade, suggesting that the private costs of block ownership are likely to increase during times of financial difficulty. In such cases, for example, blockholders are likely to spend considerable time monitoring management, and they face an increased threat of litigation brought by disgruntled minority shareholders."

[9] This equilibrium coincides with our globally stable allocation when the latter exists.

monitoring activities by which he can affect the expected payoffs of the firms' securities. We refer to this large investor as L throughout the paper. There is also a continuum of small investors who behave perfectly competitively.

Time is divided into three periods. Initially, L has an endowment of the shares of N firms given by the vector ω. We assume that the elements of ω are nonnegative and less than or equal to one. In period 1, shares can be traded in a securities market and investors establish their equilibrium allocations. The details of the trading process will be discussed in subsequent sections. In period 2, L chooses a vector of monitoring levels m, as discussed below. Finally, in period 3, the firms' payoffs are realized and consumption takes place.

Note that we assume that the monitoring decision is made in period 2 given the allocation of shares that investors hold at the start of that period. Investor L cannot commit to any specific monitoring levels prior to period 2.[10] Monitoring takes place from period 2 to period 3, and there is no further trading during this time.

We make the following specific assumptions about investor preferences, securities returns, and monitoring technologies.

ASSUMPTION 1. *Constant absolute risk aversion.*—All investors (including L) maximize the expected utility of wealth resulting from their portfolios' final payoffs. Their utility functions exhibit constant absolute risk aversion. The risk tolerance of L is ρ.[11] The aggregate trading behavior of the small investors is modeled as that of a representative, price-taking investor with a risk tolerance of τ.

ASSUMPTION 2. *Continuous and convex monitoring technologies.*—Investor L has access to K monitoring technologies, where $K \leq N$ without loss of generality. The level of monitoring is a K vector m, which is assumed without loss of generality to be a nonnegative vector in R^K. The total cost of monitoring at level m if L's holdings are given by α is $C(m, \alpha)$, where (i) $C(0, \alpha) = 0$ for all α, (ii) $C(\cdot, \cdot)$ is twice continuously differentiable with respect to m, (iii) $\partial C(m, \alpha)/\partial m_k > 0$ for all k, (iv) $\partial C(m, \alpha)/\partial \alpha_i \leq 0$ for all i, and (v) $\partial C(m, \alpha)/\partial \alpha_i = 0$ if $\alpha_i \geq 1$.[12] The vector of expected payoffs given that L monitors at level m and holds position α is given by $\mu(m, \alpha)$. Unless otherwise noted, $\mu(\cdot, \cdot)$ is continuous and differentiable, $\mu(m, \alpha) = \mu(m, 0)$ for all m if

[10] See Sec. VII for a discussion of the case in which commitment to a specific monitoring level is possible.

[11] The risk tolerance coefficient is the inverse of the Arrow-Pratt coefficient of absolute risk aversion, which is assumed constant.

[12] Condition iv implies that if L increases his holdings in the stock of firm i, his cost of monitoring at a particular level m does not increase and may decrease. Condition v implies that if L holds all of a firm's shares, monitoring costs cannot be further reduced by acquiring a larger position made possible by other investors shorting firm i's stock.

$\alpha \leq 0$, and $\partial\mu(m, \alpha)/\partial\alpha_i = 0$ if $\alpha_i \geq 1$. Moreover, for every $\alpha \geq 0$, $\alpha'\mu(m, \alpha) - C(m, \alpha)$ is a globally concave function of m. Thus the optimal monitoring level for a given allocation α is a well-defined function $m(\alpha)$. Finally, $\partial y'\mu(m, \alpha)/\partial\alpha_i \geq 0$ at $y = \alpha$ and $m = m(\alpha)$.[13]

ASSUMPTION 3. *Normal distributions.*—Securities payoffs are jointly normally distributed, and the variance-covariance matrix of payoffs, denoted by V, is positive definite and not affected by the monitoring activities of L.

We have assumed in assumption 2 that, when the monitoring level and L's holdings in other securities are held constant, his cost of monitoring is weakly decreasing and the expected payoff of his portfolio is weakly increasing as his holding of any security increases. This is motivated by the notion that a larger shareholder might be a more effective monitor, since he may have more influence on management through voting or other means. In the subsequent analysis we shall derive some results for the special case of our model in which the effectiveness and cost of monitoring are independent of L's holdings. We refer to such technologies as *allocation-neutral* as defined below.

DEFINITION. A monitoring technology is *allocation-neutral* iff, for all m, for all α, and for all i and j,

$$\frac{\partial C(m, \alpha)}{\partial \alpha_i} = 0$$

and

$$\frac{\partial \mu_j(m, \alpha)}{\partial \alpha_i} = 0.$$

Table 1 summarizes the notation that will be used throughout the paper and gives the section in which each variable is introduced.

III. Equilibrium with One Round of Trading

In this section we analyze the model under the assumption that trading takes place only once (or that it is possible for L to commit to trade only once). We show that the availability of monitoring generally leads to a loss in risk sharing because it introduces a "transaction tax" that makes trading more costly for L.

We model trading as a Walrasian process in which L is strategic; that is, he takes into account the effect of his trade on the price. If L

[13] Assumption 2 implies that when the large investor holds portfolio α and does the optimal monitoring for that portfolio, a small increase in his holding of firm i does not decrease the expected payoff (net of monitoring costs) of his portfolio. This assumption holds, in particular, if $\partial\mu_i(m, \alpha)/\partial\alpha_j$ is nonnegative for $i = j$ and zero for $i \neq j$.

TABLE 1

SUMMARY OF NOTATION

L	The large investor (Sec. II)
ρ	The risk tolerance of L (Sec. II)
τ	The (aggregate) risk tolerance of the small investors (Sec. II)
N	The number of risky securities (Sec. II)
e	The N vector of ones (Sec. II)
ω	The initial endowment of shares held by L (Sec. II)
α	L's holdings of the N risky securities (generic term) (Sec. II)
K	The number of monitoring technologies available to L (Sec. II)
m	L's monitoring level (generic term) (Sec. II)
$C(m, \alpha)$	The cost to L of monitoring at level m when his holdings are α (Sec. II)
$\mu(m, \alpha)$	The expected payoffs vector when L monitors at level m and holds α (Sec. II)
V	The variance-covariance matrix of the N firms' payoffs (Sec. II)
$m(\alpha)$	The optimal level of monitoring for L when his allocation is α (Sec. III)
$P(\alpha)$	The Walrasian price vector when L holds α (and monitors accordingly) (Sec. III)
$\Psi(\alpha)$	The certainty equivalent obtained by L given the allocation α (net of monitoring costs) (Sec. III)
α^*	L's competitive equilibrium allocation; $\alpha^* = \rho e/(\rho + \tau)$ (Sec. III)
α_P	L's equilibrium allocation if he is passive and there is one trading round (Sec. III)
α_A	L's equilibrium allocation if he is active and there is one trading round (Sec. III)
α_G	A globally stable allocation (Sec. IV)
α_S	The surplus-maximizing allocation (Sec. V)
$\Upsilon(\alpha)$	The (total) certainty equivalent obtained by the small investors when L holds α and they hold $e - \alpha$ (Sec. V)
κ	The transactions costs associated with trading in the securities market (Sec. VI)
ϕ	The fraction of investors choosing to invest in a large fund (Sec. VI)

NOTE.—The section number refers to the section in which the variable first appears.

holds the vector α of the risky securities after trading, then since he cannot commit in advance to specific monitoring levels, he will choose the monitoring levels $m(\alpha)$ to maximize his net benefits from monitoring

$$m(\alpha) = \underset{m}{\text{argmax}} \, [\alpha' \mu(m, \alpha) - C(m, \alpha)]. \quad (1)$$

Note that L enjoys the benefits of monitoring only as they affect the payoff of his portfolio, but he pays the full cost. Obviously, the first-best level of monitoring is achieved at $m(e)$, where e is the vector of ones, that is, if monitoring is chosen as though L is the sole owner of all the securities.

We assume that small investors have "rational expectations" in their assessment of the expected payoffs of the risky securities, effectively anticipating the monitoring level chosen by L.[14] If small investors conjecture that L's final holdings are given by α, then their aggregate demand function for the securities is given by

$$\tau V^{-1}[\mu(m(\alpha), \alpha) - P], \tag{2}$$

where P is the price vector of the securities, V is the variance-covariance matrix of the payoff, and τ is the aggregate risk tolerance level of the small investors. Market clearing implies that small investors must be holding $e - \alpha$ in equilibrium, that is, all the shares not held by L. Thus the market-clearing price vector is given by

$$P(\alpha) = \mu(m(\alpha), \alpha) - \frac{1}{\tau} V(e - \alpha). \tag{3}$$

The first term represents the correct assessment that the expected payoff of the security is $\mu(m(\alpha), \alpha)$. The last term represents the adjustment for risk. Note that both terms are affected by L's holdings α.

For the analysis in the rest of the paper it will be convenient to define

$$\Psi(\alpha) \equiv \alpha'(\mu(m(\alpha), \alpha)) - C(m(\alpha), \alpha) - \frac{\alpha' V \alpha}{2\rho}. \tag{4}$$

This represents the certainty equivalent obtained by L (net of monitoring costs) when he holds α.

Now denote by $\alpha_A(\omega)$ the equilibrium holdings of the risky securities by L if his initial endowment is ω and monitoring is possible (so L is "active"). Then

$$\alpha_A(\omega) \in \underset{\alpha}{\operatorname{argmax}}\, \Psi(\alpha) - (\alpha - \omega)' P(\alpha). \tag{5}$$

Note that L is not a price taker: he behaves strategically in that he takes into account the effect of his holdings on the equilibrium price.

There are two benchmark allocations that will play a role in our discussion. First, define $\alpha_P(\omega)$ to be L's equilibrium holdings of the risky assets if his endowment is ω and if monitoring is not possible, that is, if L is "passive." Clearly, $\alpha_P(\omega)$ satisfies (5) for each ω when m

[14] Note that if small investors know the model's parameters, they will be able to infer L's holdings from the equilibrium price. In our model, there is a one-to-one function relating the price and L's holdings α. In a more complicated model there might be multiple large investors, or agents may not know certain parameters. Then unless α is directly observable (or demands can be made contingent on its value), small investors will not be able to infer the expected payoff vector precisely on the basis of the price. In this case small investors will assess the expected payoffs with some statistical error. While we have chosen the simplest setting, we conjecture that the qualitative nature of our results will not change if this complication is included in the model.

is constrained to be zero. Another benchmark portfolio is the competitive equilibrium allocation, which we denote by α^*. In this normal-exponential model with two agents, it is easy to show that

$$\alpha^* = \frac{\rho}{\rho + \tau} e. \qquad (6)$$

Note that α^* is independent of the initial endowment ω and that it is proportional to e, the vector of ones, which is the supply. That is, in the competitive equilibrium, each investor holds securities in proportion to their market weights. Moreover, in this model it is well known that α^* is the unique Pareto-efficient allocation, where optimal risk sharing is achieved. Note also that both $\alpha_P(\omega)$ and α^* are independent of the monitoring technology. The difference between them is that $\alpha_P(\omega)$ is the allocation that emerges when L is strategic, and α^* is the equilibrium allocation when L is perfectly competitive.

A. *The Case of One Risky Security*

Suppose that there is one risky security and (without loss of generality) L has one monitoring activity available. We can assume, essentially without loss of generality, that the expected payoff of the risky asset is given by $\mu(m, \alpha) = \mu_0 + m$.[15] The following result characterizes the equilibrium allocation of L.

PROPOSITION 1. Assume that the monitoring technology is allocation-neutral. Then the availability of monitoring generally leads to less trading away from the endowment and to less efficient risk sharing. Specifically, (i) $\alpha_A(\omega) < \alpha_P(\omega)$ if $\omega < \alpha^*$, (ii) $\alpha_A(\omega) > \alpha_P(\omega)$ if $\omega > \alpha^*$, and (iii) $\alpha_A(\alpha^*) = \alpha_P(\alpha^*) = \alpha^*$. If the technology is not allocation-neutral, then parts i and iii do not always hold; that is, it is possible that $\omega \leq \alpha^*$ but $\alpha_A(\omega) > \alpha_P(\omega)$.

To gain some intuition for this result, note that monitoring gives rise to something like a transactions tax. When L buys shares, he increases his monitoring but does not receive the benefits of this higher level of monitoring on the newly acquired shares. The reason is that the price he pays for the shares already reflects the increase in the expected payoffs brought about by the higher monitoring level. Conversely, when L sells shares, he reduces his monitoring level. But these cost savings are partially eroded by the loss realized on the

[15] This follows because any effect the holdings α have on the expected payoffs can be captured in the cost function by the appropriate transformation. Specifically, let $y = \mu(m, \alpha)$. Assume that, for every α, $\mu(m, \alpha)$ is an invertible function of m. (This is satisfied if $\mu(m, \alpha)$ is strictly increasing in m.) Then we can write $m = \mu^{-1}(y, \alpha)$. We now relabel the monitoring choice to be y. By construction, $\mu(\mu^{-1}(y, \alpha), \alpha) = y$ for all α, and the cost function is now given by $\hat{C}(y, \alpha) = C(\mu^{-1}(y, \alpha), \alpha)$.

shares sold since the price at which they are sold reflects the lower expected payoffs associated with a reduced level of monitoring. Therefore, L captures the benefit of monitoring only on his endowment ω.

Note that when monitoring is not possible, L always trades in the direction of α^*, since this improves risk sharing. (Because he is strategic, however, L holds back and does not trade all the way to α^*.) When monitoring is possible and the monitoring technology is allocation-neutral, the transactions tax discussed above leads L to trade less in the direction of α^*. That is, he buys fewer shares when his endowment is below α^* and sells fewer shares when his endowment is above α^*. If the monitoring technology is not allocation-neutral, however, an incentive arises for L to increase his holdings since this makes monitoring cheaper at the margin. His holdings may then be larger than α^* even if his endowment is lower than α^*.

An interesting limit case of our model is one in which investors are risk-neutral (i.e., both τ and ρ go to infinity). In this case, trading cannot provide any risk-sharing gains. If the monitoring technology is allocation-neutral, then it can be shown that the unique equilibrium allocation is equal to the initial endowment; that is, no trade takes place. This is a manifestation of the transactions tax described above and is similar to the Grossman and Hart (1980) takeover paradox. If the technology is not allocation-neutral, then trade may take place with L increasing his stake in the firm so as to take advantage of the economies of scale in monitoring.

To illustrate proposition 1 as well as a number of other results that follow, we shall consider the following example.

Example: Quadratic cost function.— Let $\rho = \tau = V = 1$, and suppose that the cost of monitoring is given by

$$C(m, \alpha) = \frac{\gamma m^2}{2}, \quad \text{for } \gamma > 0, \tag{7}$$

and $\mu(m, \alpha) = m$. This monitoring technology is allocation-neutral, since for any fixed monitoring level m, the effectiveness and cost of monitoring are independent of α. For this example the optimal level of monitoring given holdings of α is $m(\alpha) = \alpha/\gamma$. The equilibrium allocation is the α that maximizes

$$\frac{\alpha^2}{2\gamma} - \frac{\alpha^2}{2} - (\alpha - \omega)\left[\frac{\alpha}{\gamma} - (1 - \alpha)\right], \tag{8}$$

which is given by

$$\alpha_A(\omega) = \frac{\omega(1 + \gamma) + \gamma}{1 + 3\gamma}. \tag{9}$$

Note that as the marginal cost of monitoring, measured by γ, increases without bound, $\alpha_A(\omega)$ converges to $(1 + \omega)/3$, which is exactly the equilibrium allocation $\alpha_P(\omega)$ when L is passive. This is intuitive, since when monitoring becomes prohibitively expensive, L chooses a very small monitoring level, which vanishes in the limit and is therefore irrelevant for the allocation.

B. The Case of Multiple Risky Securities

In this subsection we briefly examine, without attempting a full analysis of the general case, some issues that arise when there are many risky securities.[16] In particular, we examine (i) whether L holds the securities in proportion to their supply, that is, whether he holds the market portfolio of risky securities; (ii) how the holdings of L relate to the relative cost of monitoring individual firms; and (iii) whether L might employ a costly activity that increases the expected payoff of one security while at the same time decreasing the expected payoff of another security by a larger amount (thereby reducing the expected payoff on the market portfolio).

Suppose that there are N securities and assume first that L has access to N monitoring activities such that each activity affects the expected payoffs of exactly one of the securities (but not others). Specifically, we assume that, for $i = 1, 2, \ldots, N$, $\mu_i(m) = m_i$ and $C(m, \alpha) = \Sigma_{i=1}^{N} (\gamma_i m_i^2/2)$. Thus the monitoring technology is separable across assets in both costs and benefits. Let the initial endowment of L be $\omega = 0$, and define Q to be a diagonal matrix with $Q_{ii} = (2\gamma_i)^{-1}$ and $Q_{ij} = 0$ for $i \neq j$. Recall that e is the vector of ones (the market supply vector). The following proposition states that only under quite special conditions will L purchase assets according to their market weights.

PROPOSITION 2. In the example above, unless $V^{-1}Qe$ is proportional to e, L does not hold the market portfolio; that is, his equilibrium holdings are not proportional to e.

To see how L's portfolio depends on the relative costs, suppose that the payoffs of the two firms are independently and identically distributed but that the marginal cost of monitoring firm 1 is lower than that of firm 2 (i.e., $\gamma_1 < \gamma_2$). Then it can be shown that L will hold a *lower* fraction of firm 1 than of firm 2; that is, he holds relatively more of those securities for which the marginal cost of monitor-

[16] For the general multiasset case it should be noted that part iii of proposition 1 remains valid. That is, if the monitoring technology is allocation-neutral and if the initial holdings of L are equal to α^*, then there is no trade. Parts i and iii of the proposition do not have analogous statements in the multiasset case.

ing is *higher*. This seemingly counterintuitive result is explained by noting that in this example a higher marginal monitoring cost effectively allows L to commit to a lower monitoring level, lowering his total monitoring costs. The equilibrium holdings depend on the *total* cost of the monitoring that L will incur once his portfolio position has been taken, which is equal to $\alpha_i^2/2\gamma_i$ for firm i if the marginal cost is γ_i. This is decreasing in γ_i. In this case the free-rider problem is most severe when the marginal cost of monitoring is low, leading to a relatively larger risk-sharing efficiency loss.

In the example above, each monitoring activity is "firm-specific" and affects the expected payoff of exactly one security, without any effect on the expected payoff of others. In general, of course, monitoring may affect simultaneously the expected payoffs of multiple securities. For example, it may change the share of the industry profits enjoyed by one firm relative to others in the industry. In fact, some monitoring activities might be socially "wasteful" in the sense that, at a cost, they increase the expected payoff of one firm while at the same time decreasing by as much as or even more than the expected payoff of other firms. If L does not hold the market portfolio in equilibrium, the possibility arises that he might employ such activities.

To see that this is indeed possible, suppose that there are two securities and two monitoring activities available. The monitoring activities are p and r for "*p*ositive" and value *r*educing," respectively. The payoffs of the two securities are independent of each other, each having a variance of one, and $\rho = \tau = 1$. The expected payoffs of the securities are given by

$$\mu(m_p, m_r) = \begin{pmatrix} 1 & -12 \\ 0 & 11 \end{pmatrix} \begin{pmatrix} m_p \\ m_r \end{pmatrix} \tag{10}$$

for nonnegative m_p and m_r. Note that the second monitoring activity leads to an increase in the expected payoff of security 2 while at the same time decreasing the expected payoff of security 1 by a greater amount. Assume that the cost of monitoring is given by

$$C(m, \alpha) = \frac{m_p^2}{2} + \frac{m_r^2}{2}. \tag{11}$$

Clearly, in the first-best allocation, where L holds the market portfolio of risky assets, the value-reducing activity will not be employed. In this example, however, it can be shown that if L begins with a zero endowment in both securities, his equilibrium holdings of the risky securities are given by $\alpha_A = (32/116, 35/116)$, and the optimal monitoring levels are $m(\alpha_A) = (32/116, 1/116)$. Note that in equilibrium, L holds more of security 2 than security 1 and uses the value-reducing activity (at a level of 1/116).

IV. Equilibrium with Repeated Trading Rounds

In the last section we analyzed a situation in which there is exactly one round of trading prior to the realization of the payoffs. This essentially requires either that it is feasible to trade only once or that L can make a commitment to trade only once. However, if it is possible to trade repeatedly before the actual monitoring decision has to be made, then L may be unable to commit to a final round of trade. In this section we consider a trading environment in which after any round of trade further trading is possible. The inability to commit to a last round of trade generally erodes L's strategic advantage. Under some conditions we show that the resulting equilibrium allocation is the same as the competitive equilibrium allocation α^* independent of the specific monitoring technology. Thus optimal risk sharing is achieved in these cases. The monitoring level chosen is still socially suboptimal (being equal to $m(\alpha^*)$ and not $m(e)$). In other cases the model produces an unstable outcome, and the type of equilibrium we define does not exist.

To gain some intuition, suppose first that monitoring is not possible, and $\omega \neq \alpha^*$. With one round of trade the equilibrium allocation is $\alpha_P(\omega) = \rho(1 + \omega)/(2\rho + \tau)$. Now suppose that investors have traded to this allocation, and imagine that a new round of trade becomes possible. Then further trading will take place. Of course, if agents anticipate that further rounds of trading are forthcoming, this would alter their demands in the first round, and $\alpha_P(\omega)$ will not be an equilibrium.[17] Similar considerations arise if monitoring is possible: the equilibrium allocations found in the last section do not survive repeated trading opportunities. We shall seek allocations that do endure even when additional trading rounds are possible. We call such allocations globally stable.

DEFINITION. An allocation α_G is globally stable iff (i)

$$\alpha_G \in \underset{\alpha}{\operatorname{argmax}} \, [\Psi(\alpha) - \Psi(\alpha_G) - (\alpha - \alpha_G)' P(\alpha_G)]$$

[17] The situation is similar to that of a durable-good monopolist who is unable to commit not to reduce the price of the good when consumers are patient. The Coase conjecture (see, e.g., Tirole 1988) states that the monopolist would in fact lose all the monopoly power and that the equilibrium would be identical to that obtained in a perfectly competitive world. For a discussion of this and additional references, see DeMarzo and Bizer (1993).

and (ii) for every $\omega \in [0, 1]$, such that $\omega \neq \alpha_G$,

$$\Psi(\alpha_G) - \Psi(\omega) - (\alpha_G - \omega)' P(\alpha_G) > 0.$$

Recall that $P(\alpha)$ is the market-clearing price that obtains when L holds α and monitors accordingly. By construction, small investors are content to hold the rest of the supply of the risky securities $e - \alpha$ if trade takes place at a price $P(\alpha)$ and if this indeed is the final allocation. A globally stable allocation has two desirable properties. First, when small investors conjecture that the *final* holdings of L will be α_G and therefore are willing to trade only at $P(\alpha_G)$, L has no incentive to trade away from α_G. Thus, when one starts from a globally stable allocation α_G, the conjecture that α_G is indeed L's final allocation is justified. Second, when one starts at any initial allocation $\omega \neq \alpha_G$, if small investors conjecture that L's final allocation will be α_G, then trade to the allocation α_G, at prices $P(\alpha_G)$, is desirable for L. Given the first property, the conjecture that this is indeed the last round of trade will be justified in equilibrium. Thus, if α_G is globally stable, the final allocation will be α_G no matter what the final allocation is.[18]

To illustrate the definition above, consider again the quadratic–cost function example, with $N = \rho = \tau = V = 1$, $\mu(m, \alpha) = m$, and $C(m, \alpha) = \gamma m^2/2$. Here the optimal monitoring level is given by $m(\alpha) = \alpha/\gamma$, and the certainty equivalent obtained by L when he holds α is

$$\Psi(\alpha) = \alpha^2\left(\frac{1}{2\gamma} - \frac{1}{2}\right). \tag{12}$$

If $\gamma > 1$, then it can be shown that $\alpha_G = \alpha^* = 0.5$ is a globally stable allocation. To see this, observe that when traders conjecture that L's final allocation is $\alpha^* = 0.5$, then they are willing to trade shares of the risky asset only at the price

$$P(0.5) = m(0.5) - \frac{1}{\tau} V(1 - 0.5) = \frac{1}{2\gamma} - \frac{1}{2}. \tag{13}$$

If L starts at an allocation equal to 0.5 and can trade only at the price $P(0.5)$, he will have no incentive to trade, since $\Psi(\alpha) - (\alpha - 0.5)P(0.5)$ is maximized at $\alpha = 0.5$ and this is the unique optimizer.

[18] The concept we define is a static concept that does not consider explicitly the equilibrium of a corresponding trading game among investors. The trading game implicit in our discussion is one in which Walrasian trading takes place in a sequence of markets. There is no discounting, and agents' payoff is determined by the final (or limit) allocation to which trade settles. If the allocations achieved through trading do not converge, then the payoff is determined as though no trade takes place. This is similar to the game analyzed in DeMarzo and Bizer (1993).

Now consider the case $\gamma < 1$. We shall show that in this case there does not exist any globally stable allocation. This can be verified by showing that there is no allocation α_o that satisfies

$$\alpha_o \in \underset{\alpha}{\operatorname{argmax}} [\Psi(\alpha) - \Psi(\alpha_o) - (\alpha - \alpha_o)P(\alpha_o)]$$
$$= \underset{\alpha}{\operatorname{argmax}} \left\{(\alpha^2 - \alpha_o^2)\left(\frac{1}{2\gamma} - \frac{1}{2}\right) - (\alpha - \alpha_o)\left[\frac{\alpha_o}{\gamma} - (1 - \alpha_o)\right]\right\}. \quad (14)$$

The convexity of $\Psi(\cdot)$ implies that the expression in (14) has a *minimum* at $\alpha = 0.5$. Indeed, when $\alpha_o < 0.5$, L is better off if he trades to any allocation $\alpha > 0.5$ at a price $P(\alpha_o)$. Similarly, when $\alpha_o > 0.5$, L is better off if he trades to any allocation $\alpha < 0.5$ at a price $P(\alpha_o)$. If $\alpha_o = 0.5$, L is better off if he trades to any allocation different from 0.5 at the price $P(0.5)$. As we shall show below, it is true more generally that the convexity of $\Psi(\cdot)$ plays an important role in determining whether a globally stable allocation exists.[19]

To illustrate the case $\gamma < 1$ in the example above, where no globally stable allocation exists, suppose that $\omega = 0$ and $\gamma = ½$. If trade takes place exactly once, then the resulting allocation would be $\alpha = 0.2$, since this maximizes $\Psi(\alpha) - \alpha P(\alpha) = (\alpha^2/2) - \alpha(3\alpha - 1)$. However, once L has acquired 0.2 shares, if one more trading round becomes available, then further trade to $\alpha = 0.32$ will take place, since this maximizes $\Psi(\alpha) - (\alpha - 0.2)P(\alpha) = (\alpha^2/2) - (\alpha - 0.2)(3\alpha - 1)$. It is clear, however, that if investors realize that once L moves to 0.2 in the first round there will be another round of trade to 0.32, then they would not trade at $P(0.2)$ in the initial round (since the final payoffs are determined by the final allocation). It is interesting to note that the limit of the sequence of allocations given by {0, 0.2, 0.32, 0.392, . . .} is 0.5. In fact 0.5 is the limit from any starting point. (If α_m is the allocation on the mth round, then $\alpha_{m+1} = [1 + 3\alpha_m]/5$, which always converges to 0.5.) One might think that it should be possible for L to trade from whatever initial allocation he has to $\alpha = 0.5$ at the price $P(0.5)$. However, it is easy to show that $\Psi(0.5) - \Psi(\alpha) - (0.5 - \alpha)P(0.5) < 0$ for all $\alpha \neq 0.5$, so L never wants to move to 0.5 at the price $P(0.5)$. (Indeed, $\alpha = 0.5$ is the *worst* allocation for L if he must trade at $P(0.5)$.)

It is not clear what one should generally predict about the equilibrium of this trading environment when there does not exist a globally stable allocation. Since there is no obvious allocation to trade to, a natural conjecture might be that L will be "trapped" at his endowment

[19] Unfortunately, we are unable to provide a simple interpretation of this property from the primitives of our model.

and no trade will take place. Note, however, that the endowment might be extremely inefficient for risk-sharing purposes. (For example, the endowment in the previous example is $\omega = 0$, which is clearly inefficient.) It is possible that other trading mechanisms, such as block trading or tender offers, can facilitate trading in such cases, or that this problem will lead to the development of commitment technology to assure a last trading round.[20]

Proposition 3 characterizes the set of globally stable allocations for the case of one risky security. A stronger result for the case in which the monitoring technology is allocation-neutral, which applies also in the multiasset case, is provided in proposition 4.[21]

PROPOSITION 3. Assume that $N = 1$. Then there is at most one globally stable allocation. If $\Psi(\alpha)$ is concave, then there exists a unique globally stable allocation $\alpha_G \in (0, 1)$.

Recall that we have already shown in proposition 1 that, with one round of trade, if the monitoring technology is allocation-neutral and if the initial allocation is equal to the competitive equilibrium allocation α^*, then there will be no further trading. The following result shows that in this case if $\Psi(\cdot)$ is concave, then α^* is the unique globally stable allocation. Note that this result holds also for the case of multiple securities.

PROPOSITION 4. Assume that $\Psi(\cdot)$ is strictly concave and that the monitoring technology is allocation-neutral. Then there exists a unique globally stable allocation α_G, which is equal to the competitive equilibrium allocation α^*.

Under the conditions of proposition 4, the possibility of monitoring does not distort the asset holdings away from the optimal risk-sharing allocation. No matter what the initial allocation and independent of the specific monitoring technology, the only globally stable allocation in this case is α^*. In a sense, we have obtained a separation of the ownership and monitoring decisions: when the monitoring technology is allocation-neutral, the equilibrium allocation becomes "neutral" to the technology. Note that the equilibrium monitoring level is $m(\alpha^*)$. While this is not the socially optimal level of monitoring, it is clear

[20] DeMarzo and Bizer (1993) study a related model in which there is no commitment to a last round of trade. An important assumption in their model is that there is a finite set of possible prices at which trade can take place. DeMarzo and Bizer find that for any initial allocation there exists a unique equilibrium for their trading game. If $\Psi(\cdot)$ is concave, this equilibrium corresponds to our globally stable allocation, and it does not depend on the initial endowment. If $\Psi(\cdot)$ is not concave, then in general their equilibrium depends on the initial endowment and is sensitive to the precise specification of the price grid.

[21] We conjecture that proposition 3 holds also if there is more than one asset. Proposition 4 addresses the general case of allocation-neutral technologies, but we have been unable to prove a uniqueness result for the case in which there are many assets and the technology is not allocation-neutral.

that L will never employ monitoring activities that reduce the expected payoff on the market portfolio, that is, such that the total increase in expected payoffs of some securities is more than offset by the decrease in the expected payoffs of others.

V. Surplus-Maximizing Allocations

With the Walrasian trading mechanism used in the last two sections, prices at which trade takes place always reflect the monitoring that L will do at his final allocation. As a result L captures the benefits of monitoring only on his initial endowment of shares but not on any of the shares he acquires in the market. In this section we assume that L can capture some of the gains realized by the other investors as a consequence of his trading and monitoring activities. We show that in this case more monitoring gets done, but since L is still unable to commit to a specific monitoring level prior to trade, the increased monitoring is accompanied by a distortion in risk sharing, whereby L generally bears more risk than he would in the competitive allocation.

Before providing further motivation for the trading environment we seek to model in this section, we shall define more formally the surplus generated by L's trade. Recall that $\Psi(\cdot)$, defined in equation (4), is the certainty equivalent L obtains from his position (net of the price he pays in the transaction). Thus, when L trades from his endowment ω to another allocation α, his certainty equivalent changes from $\Psi(\omega)$ to $\Psi(\alpha)$. We can define an analogous expression for the small shareholders, which we denote $\Upsilon(\alpha)$. This is the total certainty equivalent obtained by the small shareholders when L holds α (and they hold $e - \alpha$) of the risky securities. It is easy to show that

$$\Upsilon(\alpha) = (e - \alpha)' \mu(m(\alpha), \alpha) - \frac{1}{2\tau}(e - \alpha)' V(e - \alpha). \qquad (15)$$

The total net surplus obtained by all agents when L trades from ω to α is given by

$$\Psi(\alpha) + \Upsilon(\alpha) - [\Psi(\omega) + \Upsilon(\omega)]. \qquad (16)$$

While we do not model this explicitly, there are a number of ways in which L might be able to capture a fixed fraction of this total surplus. For example, he might be able to make a conditional tender offer for $\alpha - \omega$ shares. The extent to which the gains can be captured by L would then depend on the probability that any particular shareholder is pivotal in determining the success of the tender offer. As long as this probability is positive, some of the gain can be captured by L. Alternatively, L might negotiate with other large (but passive) shareholders to acquire some of their holdings. The way in which gains are split among the parties is determined by the bargaining power of each. It is easy to see, however, that the final allocation that emerges if L attempts to maximize his expected utility under these conditions does not depend on how the gains are split, since in all cases the final allocation is the one that maximizes the total surplus.[22] The following result characterizes the surplus-maximizing allocation, which we denote by α_S.

PROPOSITION 5. The surplus-maximizing allocation α_S is independent of ω and different from (in the case of one risky asset, strictly larger than) the competitive equilibrium allocation α^*.

To illustrate this result, consider again the quadratic–cost function example, in which $N = 1$, $C(m, \alpha) = \gamma m^2/2$, and $\mu(m, \alpha) = m$. In this case, α_S maximizes

$$\frac{\alpha}{\gamma} - \frac{\alpha^2}{\gamma} - \frac{\alpha^2 V}{2\rho} - \frac{(1-\alpha)^2 V}{2\tau}. \tag{17}$$

Since this is a quadratic function, it is easy to derive the solution and verify the conclusions of the proposition in this case. For example, if $\gamma = 2$ and $V = \rho = \tau = 1$, then $\alpha_S = 0.6$, which is larger than the competitive allocation $\alpha^* = 0.5$. Intuitively, the surplus-maximizing allocation for L balances the risk-bearing costs of holding a higher fraction than α^* with the gains from increased monitoring that obtain with a larger allocation, gains that L can now be compensated for. In this example, as γ goes to zero, so that monitoring becomes very cheap, the monitoring gains overwhelm the risk-sharing gains and α_S goes to one. In another extreme situation, if all agents become risk-neutral (ρ and τ go to infinity), risk-sharing considerations do not exist, α_S goes to one, and the first-best level of monitoring is chosen. This result is in sharp contrast to the corresponding result in the case of one trading round in a Walrasian market, where risk-neutral agents do not trade at all, and so when $\omega = 0$ no monitoring takes place in equilibrium.

We have seen in Section III*B* that when there is more than one

[22] To see this suppose first that L can capture all the surplus generated by his trading. For example, assume that there is one other passive shareholder and L can make a take-it-or-leave-it offer to trade the relevant block of shares with the passive shareholder. Then it is clear that if L wishes to acquire $\alpha - \omega$ shares (sell if this is negative), he must offer to pay $\Upsilon(\omega) - \Upsilon(\alpha)$, which extracts the surplus generated by his trade from the other shareholders. Obviously L maximizes his expected utility if he chooses α to maximize the expression in (16). More generally, suppose that L can capture a fixed fraction $\theta \in (0, 1]$ of the total gain generated by trading from ω to α shares, where θ is a parameter that can depend on the specific value of ω. Then the price paid for the trade, call it $P(\omega, \alpha)$, must solve $\Psi(\alpha) - \Psi(\omega) - P(\omega, \alpha) = \theta\{\Psi(\alpha) + \Upsilon(\alpha) - [\Psi(\omega) + \Upsilon(\omega)]\}$. Again, maximizing L's expected utility is equivalent to maximizing the total surplus. Note also that this allocation is clearly independent of θ and ω.

risky asset, with one trading round L will generally not hold the market portfolio in equilibrium. Moreover, in a simple example we have seen that L would tend to concentrate his holdings in securities for which the marginal monitoring cost is relatively high. It can be shown that the surplus-maximizing portfolio can also differ from the market portfolio. However, in the same parametric example, L's surplus-maximizing allocation is actually concentrated in assets whose relative marginal cost of monitoring is relatively low. The reason for this is that now L can capture some of the gains of monitoring, alleviating to some extent the free-rider problem. Nevertheless, since the surplus-maximizing allocation is different from the market portfolio, it is possible again that L utilizes monitoring activities that reduce the expected payoff of the market portfolio.[23]

The results in this section were derived under the assumption that the move from ω to α_S represents the only transaction and, in particular, that no trade occurs after this transaction is completed. Note first that, since α_S is independent of ω, L does not wish to trade again using the same surplus extraction mechanism by which he had arrived at α_S. However, if trading in a Walrasian market is possible after one arrives at the allocation α_S and before the monitoring decision is made, then the analysis changes in line with our discussion in Section IV, and obviously the surplus-maximizing allocation is not stable.

VI. Endogenous Fund Size: An Example

So far the size of L, as measured by his risk tolerance coefficient, has been taken to be exogenously determined. In this section we analyze a simple example that suggests how the size of L might be determined endogenously. Specifically, we assume that L is an institutional fund representing a group of small investors, and we allow small investors to choose between investing through this institutional fund and investing directly on their own accounts. Each small investor trades off a savings in transactions costs achieved by investing through the fund against that portion of the fund's monitoring costs that will be paid by him if he joins the fund. The equilibrium size of the fund is such that the transactions cost savings are exactly offset by the cost of monitoring. We continue to assume that the fund cannot commit to a specific monitoring level before establishing its portfolio position.

For simplicity assume that there is one risky asset. Each small investor chooses to trade either through the fund or on his own account. We assume a continuum of small investors, and we let ϕ represent

[23] Examples to illustrate all the points above can be obtained from the authors on request.

the fraction of investors investing in the fund. We also assume that the supply of the risky asset is one share per capita. Now define κ as a *fixed* transactions cost incurred by the small investor if he trades in the securities market directly. If the agent invests through the fund, he shares this cost with the other investors in the fund. We assume that κ is sufficiently small that, for any $\phi > 0$, the transactions cost *per fund investor* is essentially zero and can be ignored. Furthermore, to simplify the analysis we assume that all investors have constant absolute risk aversion with identical risk tolerance levels equal to one. In this case, a fund investing for a fraction of investors equal to ϕ can be represented by a single investor with risk tolerance ϕ, whereas the remaining fraction of $1 - \phi$ individual investors can be represented by a single competitive investor with risk tolerance $1 - \phi$.

Consider again the monitoring technology described by the quadratic–cost function example, where $C(m, \alpha) = C(m) = \gamma m^2/2$ and $\mu(m, \alpha) = m$.[24] Assume also that $\gamma > 1$. If the fund holds fraction α of the total number of shares, then its optimal monitoring level is $m(\alpha) = \alpha/\gamma$. If the fund is unable to commit to a single round of trade, then it will trade to the globally stable allocation $\alpha_G = \alpha^* = \phi$.[25] Intuitively, since all investors are assumed to have the same level of risk tolerance and since the per capita supply of the risky asset is one share, the fund takes the position $\alpha_G = \phi$ so that each investor in the fund has one share, that is, $\alpha_G/\phi = 1$. The monitoring cost paid by the fund per investor in the fund is equal to

$$\frac{1}{\phi}\left(\frac{\alpha_G^2}{2\gamma}\right) = \frac{\phi}{2\gamma}. \tag{18}$$

In equilibrium the fund size adjusts so that this cost of investing in the fund is equated with the savings in transactions costs investors realize by joining the fund, that is, $\phi/2\gamma = \kappa$. However, if $\gamma > (2\kappa)^{-1}$, then the transactions cost saving realized in joining the fund is so large that it overwhelms the monitoring cost of being in the fund even when all investors have joined the fund. The equilibrium size of the fund is therefore

$$\phi^* = \min(2\gamma\kappa, 1). \tag{19}$$

Note that the optimal fund size increases with κ, the size of individual investors' transactions costs, and with γ, the marginal cost of moni-

[24] Since we are normalizing everything on a per capita basis, the cost of the monitoring technology must be interpreted as the cost of monitoring at level m per investor in the economy. Since in general not every investor in the economy will invest in the fund, the cost per investor in the fund will be higher.

[25] Note that $\Psi(\alpha) = (\alpha^2/\gamma) - (\alpha^2 V/2\phi)$ and $P(\alpha) = (\alpha/\gamma) - [V(1 - \alpha)/(1 - \phi)]$.

toring. Both results are intuitive. First, higher transactions costs make it more advantageous for investors to invest through funds. Second, a higher marginal cost of monitoring mitigates the free-rider problem faced by the fund, and its optimal size increases. Since the socially optimal situation involves all agents' investing through the fund (i.e., $\phi^* = 1$), the free-rider problem leads to a socially inefficient outcome unless $2\gamma\kappa \geq 1$. This means that if $\kappa < 1/2\gamma$, welfare can be improved by any measure that increases κ to $1/2\gamma$. At the higher level of κ, all investors invest through the fund, there is no loss in risk sharing, and the first-best level of monitoring is achieved. Note also that since no investor trades on his own account, the higher level of the fixed transactions cost is not paid by anyone and therefore does not result in any deadweight loss.

VII. Commitment to Specific Monitoring Levels

The basic model analyzed so far assumes that L cannot commit to a specific level of monitoring before trading: he always monitors at a level that is optimal given his final holdings. As we discuss at the end of this section, we believe that it is actually difficult for large investors to make credible commitments to specific monitoring levels and, in particular, to commit to refraining from monitoring. Nevertheless, the examination of the case in which L can commit to a specific monitoring level is useful because it reveals the value of such a commitment. It also enables us to review and summarize the results of the previous sections.

The analysis below is done in the context of the quadratic–cost function example, where $N = V = \rho = \tau = 1$, $\mu(m) = m$, and $C(m, \alpha) = m^2$. Our results are summarized in figure 1. The graphs on the left of the figure show the final allocation of shares held by L as a function of L's initial endowment ω. The graphs on the right show the level of monitoring undertaken by L, again as a function of ω. Within these graphs, the lines marked N pertain to our analysis so far, where *no* commitment to a monitoring level is possible before trading, and those marked with a C apply to cases in which L can commit to some chosen monitoring level before trading. The three rows of graphs correspond to the three trading mechanisms considered in Sections III, IV, and V, respectively.

In interpreting figure 1, one should note that in this example the competitive equilibrium allocation is $\alpha^* = 0.5$, and the first-best monitoring level, that is, the level that maximizes the total net benefit from monitoring $\mu(m) - C(m, \alpha)$, is given by $m(1) = 0.5$. We make the following observations regarding the results illustrated in the figure.

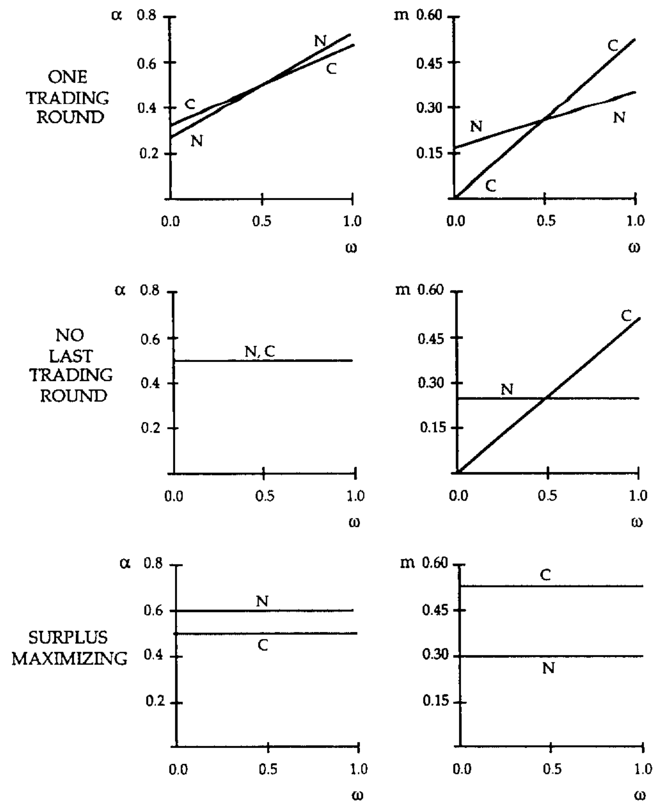

FIG. 1

i) When prices are determined in a Walrasian trading mechanism (i.e., in both the first and the second rows of the figure), the monitoring level that L would like to commit to is $m(\omega)$, that is, the optimal level for the initial endowment rather than the one that is optimal for the final allocation. In particular, if L's initial endowment is $\omega = 0$, then he commits to not monitoring at all. The reason is that the transaction price reflects the level of monitoring done at the final allocation, so that L does not benefit from an increase or a decline in the monitoring level relative to the original allocation.

ii) When there is only one round of trade (the top row of the figure), the ability to commit to a specific monitoring level improves risk sharing relative to the case of no commitment: for every $\omega \neq 0.5$, the allocation given by the line C (in the top row of the figure) is closer to 0.5 than the allocation given by the line N. In fact, because the transactions tax associated with trading disappears, the allocation is the same as the one that would obtain if L was passive and did not monitor at all. The monitoring level, however, is still $m(\omega)$ as discussed above.

iii) The globally stable allocation is unaffected by whether L can commit to a specific monitoring level. In all cases, and independent of the endowment, the allocation coincides with the competitive allocation $\alpha^* = 0.5$. In the case without commitment to a specific monitoring level, the monitoring level, which is $m(\alpha^*)$, becomes independent of the endowment, whereas with commitment it is given by $m(\omega)$, which does vary with ω.

iv) The observations above imply that when prices are determined in a Walrasian trading mechanism, the ability to commit to a specific monitoring level leads to less monitoring when $\omega < 0.5$ and to more monitoring when $\omega > 0.5$.

v) In the case of surplus maximization (the third row of the figure), the ability to commit to a specific monitoring level leads to first-best results: the final allocation is Pareto optimal and the socially optimal amount of monitoring is undertaken, because L can extract the surplus generated by his monitoring activities. By contrast, in the base case without commitment to a specific monitoring level, L overinvests in shares (holds 0.6 rather than 0.5) and underinvests in monitoring (m is equal to 0.3 rather than 0.5).

It is clear that L is generally better off if he can commit to a specific monitoring level. The extent to which the ability to commit is valuable, however, depends on L's initial endowment and on the trading mechanism. Figure 2 shows the difference in the certainty equivalent obtained by L when commitment to a specific monitoring level is possible relative to the case in which it is not. We first note that with the Walrasian trading mechanisms, if L's endowment ω is equal to α^*, then the ability to commit to a specific monitoring level has no

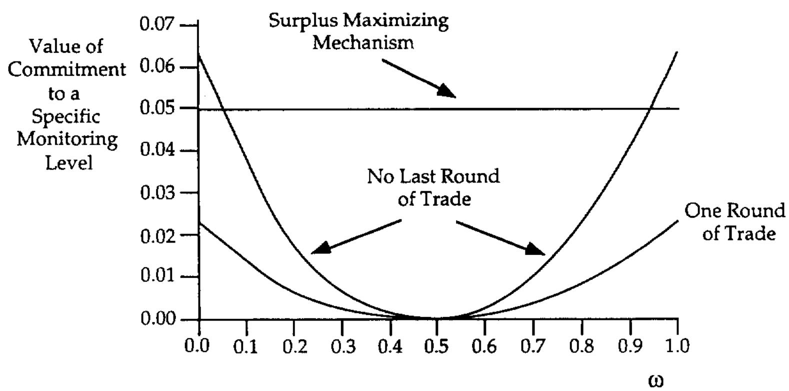

FIG. 2

value, since in this case L does not trade away from his endowment, and so $m(\alpha) = m(\alpha^*)$. The further L's endowment is from α^*, the more trading away from his endowment L does, and so the ability to commit, which leads to the elimination of the transaction tax associated with trading, is more valuable. We also see that the value of commitment to a specific monitoring level is higher when there is no last round of trade relative to the case of one trading round. The reason is that with one round of trading, L trades less away from his endowment, and so the monitoring choice is closer to the one he would like to commit to. In the case of surplus maximization, the value of commitment to a particular monitoring level is independent of ω and is always relatively large. In this case, commitment allows L to generate more surplus, as he can choose the appropriate monitoring level without compromising diversification.

Note that our analysis in this section indicates that monitoring will generally be done in equilibrium even if commitment to a specific level is possible. The only exception to that is the case in which L starts with zero holdings and the trading mechanism is Walrasian. Nevertheless, we believe that the analysis of the previous sections is more appropriate and that it is generally difficult for large investors to make commitments to particular monitoring levels. The key issue concerns what mechanisms might exist by which such a commitment can be made credible. A simple way for an institutional fund to commit to a specific monitoring level would be to specify in its charter which monitoring activities management can engage in and, possibly, the extent to which resources can be spent on monitoring. However, implementation of this solution might be problematic. First, it would be difficult to determine ex post what constitutes monitoring and which management actions are in violation of the institution's charter. Second, after trading has been completed, it would be in the best interest of the institution's shareholders to remove the monitoring restriction from the charter and let management choose the monitoring level optimally. Alternatively, the compensation of the institution's manager could be designed to induce the choice of a particular monitoring level. Again, after the institution has traded to its optimal portfolio, its shareholders would find it optimal to renegotiate the compensation contract with the manager, thereby eliminating the commitment role of the initial compensation schedule.[26] Other incen-

[26] It might be argued that a free-rider problem prevents this from occurring. Suppose that if some individual investor in the fund attempts to lobby for a change in the charter or the compensation contract, he incurs a significant cost in terms of time, effort, and money. While the collective benefit to the members of the fund might be large if the charter or compensation contract is altered, that individual's benefit will typically be quite small relative to the personal cost of bringing about the change. Thus

tive problems might also make it suboptimal to use managerial compensation as a commitment device for future monitoring behavior.[27]

VIII. Concluding Remarks

In this paper we have developed a framework to analyze the effects of large shareholder activism on securities market equilibrium. A free-rider problem arises in our model because passive shareholders benefit from the large investor's monitoring activities, but they do not incur the costs associated with monitoring. Despite this problem we have shown that in a portfolio context with risk-averse investors, large shareholder activism is consistent with equilibrium even if the initial holdings of the large investor are zero. Moreover, under some conditions the equilibrium holdings of all investors in our model are independent of the monitoring technologies. Under these conditions, all investors, whether passive or active, hold the market portfolio of risky securities. In other cases the portfolio choice of the large investor will be affected by the available monitoring technology and by his initial endowment of shares.

We now make some observations and discuss directions for future research. First, our analysis has dealt only with the case of one large investor. In reality, of course, there are many institutions and other large investors who could be active. Understanding the interactions between large investors in the presence of monitoring capabilities is an interesting direction for future research. Note that when there are multiple large investors, we might observe specialization in monitoring that could be a consequence either of different investors having or developing specialized monitoring technologies (e.g., in specific industries) or of the equilibrium choice of monitoring activities by different large investors. This may lead large investors to hold portfolios different from the market portfolio. Optimal risk sharing might still be attainable if these large investors represent mutual funds and small investors can hold multiple funds such that their combined holdings sum to the market portfolio.

---

no investor in the fund will find it in his interest to take the initiative in promoting the change. Note, however, that the manager himself could propose the change coupled with an increase in his compensation that rewards him for his effort in lobbying for the change. Since this reward to the manager would be spread among all of the fund's investors, the cost per investor of rewarding the manager will generally be small relative to the benefit per investor of making the change. In this case all would vote with the manager for the change, if the only alternative is to leave the charter and compensation contract unaltered.

[27] As discussed in Sec. VIII, an effective commitment to a particular monitoring level might be possible in the context of multiple large investors who vary in their ability to use private information generated through monitoring to alter their portfolios.

The analysis in this paper does not capture any possible informational asymmetries between the active shareholder and other investors. Such asymmetries may arise as a by-product of the monitoring activity itself, which may include careful studies of the firm's performance and managerial decisions and might involve direct contact between the large shareholder and management. This raises the possibility that the large shareholder will be able to make trading profits on the basis of such information. These profits might indeed partially compensate the large shareholder for the costs involved in his monitoring activities.[28]

Our model of the monitoring technology takes as primitives a general cost function and stipulates that the expected payoffs on the securities are a general function of the monitoring level (and possibly of the holdings of the large investor). We have not derived the costs and effects of monitoring from a model of individual behavior, the production technology of the firm, or the general economic environment in which firms operate (e.g., the form of strategic interaction between firms). In a complete analysis it would be necessary to model the monitoring technology in more detail. Such a model might suggest further restrictions on the technology and produce sharper predictions.

The monitoring technology and the resulting equilibrium holdings and monitoring levels are also affected by the regulatory and institutional environment. For example, partly for regulatory reasons, the ownership structure in the United States is much more diffuse than that in Japan, Germany, and the United Kingdom. There are also various rules pertaining to proxy fights, the degree of communication and coordination between shareholders, and so forth that affect monitoring technologies. Within the framework of our model, a more detailed model of the monitoring technology that would also take into account the regulatory constraints might offer policy implications.

In our analysis the large investor is modeled as an optimizing individual. The model abstracts from whether this is a private investor or an institution. However, there may be a distinction between private investors and institutional investors such as pension funds, since in the latter there may be additional agency problems between the managers of the fund (who make the monitoring decisions) and its shareholders. Modeling this agency problem in detail and then integrating this into our framework would be interesting.

[28] Note, however, that some institutional investors, e.g., index funds, are restricted in their ability to use private information to alter their portfolios. This may explain why such investors would tend to engage in less monitoring than large investors who can use private information in their trading.

Many institutions in the United States trade very frequently rather than hold the same portfolio for a long time. This would affect their incentives to monitor. Our model is static in the sense that there is only one decision concerning monitoring and, in effect, monitoring occurs instantaneously. An extension to a more dynamic model, where monitoring takes time and portfolio turnover decisions are made frequently, might be better able to address some of these issues.

Finally, the possibility that the monitor uses "value-reducing" activities in a world of multiple securities illustrates the complex interactions between the allocation of asset holdings and corporate policies. One fundamental issue concerns what is or should be the objective function of the firm in an economy in which investors hold diversified portfolios and in which some investors might be able to affect managerial decisions in several firms. Suppose, for example, that all investors hold the market portfolio. Then firms that act in the interests of shareholders might well refrain from competing against each other in the product market. Note, however, that the shareholders are also consumers and might therefore be hurt by such behavior. Moreover, stakeholders in the firm include its employees, suppliers, and so forth. Obviously, a complex model that accounts for all these elements is beyond the scope of this paper, but our approach here might be useful in future attempts to analyze these important issues.

Appendix

Proof of Proposition 1

The optimal allocation $\alpha_A(\omega)$ solves

$$\omega m'(\alpha) - C_1(m(\alpha), \alpha)m'(\alpha) - C_2(m(\alpha), \alpha) + V\left(-\frac{\alpha}{\rho} + \frac{-2\alpha + 1 + \omega}{\tau}\right) = 0. \tag{A1}$$

By assumption, $C_2(m, \alpha) \equiv 0$. Also, $C_1(m, \alpha) > 0$ for every m and α. The first-order condition for the optimal monitoring choice implies $\alpha \equiv C_1(m(\alpha), \alpha)$. We can rewrite equation (A1) as

$$(\omega - \alpha)m'(\alpha) + V\left(-\frac{\alpha}{\rho} + \frac{-2\alpha + 1 + \omega}{\tau}\right) = 0. \tag{A2}$$

To prove part i, assume that $\omega < \alpha^*$. Note that $\alpha_P(\omega)$ solves the equation $-(\alpha/\rho) + [(-2\alpha + 1 + \omega)/\tau] = 0$. We have

$$\omega < \alpha_P(\omega) = \frac{\rho(1 + \omega)}{2\rho + \tau} < \alpha^*. \tag{A3}$$

This means that the left-hand side of (A2) is negative at $\alpha = \alpha_P(\omega)$. Now observe that

$$\frac{\partial}{\partial \alpha}\left(-\frac{\alpha}{\rho} + \frac{-2\alpha + 1 + \omega}{\tau}\right) < 0. \tag{A4}$$

It can be shown using assumption 2 that $m'(\alpha) > 0$.[29] This and (A4) imply that the first-order condition (A2) cannot hold for any $\alpha \geq \alpha_P(\omega)$. Similar arguments can be used to prove parts ii and iii of the proposition.

A case in which $C_2(m, \alpha) < 0$ and parts i and iii do not hold is given in the following example. Let $\rho = \tau = V = 1$, and $\mu(m, \alpha) = m$. Suppose that the cost of monitoring is given by

$$C(m, \alpha) = \begin{cases} \gamma m^2/2\alpha & \text{if } 0 < \alpha \leq 1 \\ \gamma m^2/2 & \text{if } \alpha > 1. \end{cases} \tag{A5}$$

For this example, if $\alpha \leq 1$, $m(\alpha) = \alpha^2/\gamma$ and the optimal final position $\alpha_A(\omega)$ is the α that maximizes

$$\frac{\alpha^3}{2\gamma} - \frac{\alpha^2}{2} - (\alpha - \omega)\left[\frac{\alpha^2}{\gamma} - (1 - \alpha)\right], \tag{A6}$$

as long as this is not larger than one; otherwise $\alpha_A(\omega) = 1$. Thus

$$\alpha_A(\omega) = \min\left\{1, \frac{2\omega}{3} - \gamma + \left[\left(\frac{2\omega}{3} - \gamma\right)^2 + \frac{2\gamma(1 + \omega)}{3}\right]^{1/2}\right\}. \tag{A7}$$

Note that if $\omega \in (1/5, 1/2]$, then $\alpha_A(\omega) > \alpha_P(\omega)$. This demonstrates that parts i and iii of the proposition may not hold when $C_2(m, \alpha) < 0$. Q.E.D.

Proof of Proposition 2

It is easy to see that the optimal monitoring level for each technology i is given by

$$m_i(\alpha) = \frac{\alpha_i}{\gamma_i}. \tag{A8}$$

Then α is chosen to maximize

$$-\alpha' Q\alpha - \frac{1}{2\rho}\alpha' V\alpha + \frac{1}{\tau}\alpha' V(e - \alpha). \tag{A9}$$

The solution to this maximization problem is given by

$$\alpha_A = \left(2Q + \frac{1}{\rho}V + \frac{2}{\tau}V\right)^{-1}\frac{Ve}{\tau}. \tag{A10}$$

The proposition now follows immediately. Q.E.D.

[29] To do this consider the original specification of the monitoring technology in which the expected payoff is $\mu(m, \alpha)$. Assuming that the technology is allocation-neutral, we have $m'(\alpha) = -\mu_1(m, \alpha)/[\alpha\mu_{11}(m, \alpha) - C_{11}(m, \alpha)]$, which is positive since $\alpha\mu(m, \alpha) - C(m, \alpha)$ is, by assumption, a globally concave function of m and $\mu(m, \alpha)$ is an increasing function of m. This shows that $m'(\alpha)$ is positive under the original specification for the monitoring technology. This remains true once we have transformed the specification for the monitoring level as in n. 6 since the transformation is monotonically increasing.

Proof of Proposition 3

Assume that $\alpha_1 \neq \alpha_2$ and that both α_1 and α_2 are globally stable. Then

$$\Psi(\alpha_1) - \Psi(\alpha_2) - (\alpha_1 - \alpha_2)P(\alpha_2) \leq 0 \tag{A11}$$

and

$$\Psi(\alpha_2) - \Psi(\alpha_1) - (\alpha_2 - \alpha_1)P(\alpha_1) \leq 0. \tag{A12}$$

Together these expressions imply that

$$(\alpha_1 - \alpha_2)[P(\alpha_1) - P(\alpha_2)] = (\alpha_1 - \alpha_2)[\mu(m(\alpha_1), \alpha_1) - \mu(m(\alpha_2), \alpha_2)] + \frac{V(\alpha_1 - \alpha_2)^2}{\tau} \leq 0 \tag{A13}$$

or, since $V > 0$,

$$(\alpha_1 - \alpha_2)[\mu(m(\alpha_1), \alpha_1) - \mu(m(\alpha_2), \alpha_2)] < 0. \tag{A14}$$

But $\mu(m(\alpha), \alpha)$ is increasing in α, so (A14) is not possible. Thus there is at most one globally stable allocation.

To prove the second part we need to show that $\Psi'(0) - P(0) > 0$ and $\Psi'(1) - P(1) < 0$. Since by assumption $\Psi'(\alpha)$ is decreasing and $P(\alpha)$ is increasing,

$$\begin{aligned}\Psi'(\alpha) - P(\alpha) &= \alpha m'(\alpha) + m(\alpha) - C_1(m(\alpha), \alpha)m'(\alpha) - C_2(m(\alpha), \alpha) \\ &\quad - \frac{\alpha V}{\rho} - m(\alpha) + \frac{(1-\alpha)V}{\tau} \\ &= -C_2(m(\alpha), \alpha) - \frac{\alpha V}{\rho} + \frac{(1-\alpha)V}{\tau}.\end{aligned} \tag{A15}$$

The second equality follows from the fact that $\alpha = C_1(m(\alpha), \alpha)$ for all α because of the optimality of $m(\alpha)$. Using (A15), we see that $\Psi'(0) - P(0) = -C_2(m(0), 0) + (V/\tau)$, which is positive since $C_2(m(0), 0) \leq 0$, and $\Psi'(1) - P(1) = -C_2(m(1), 1) - (V/\rho)$, which is negative since $C_2(m(1), 1) = 0$. Uniqueness follows from strict concavity. Q.E.D.

Proof of Proposition 4

If α_G is a globally stable allocation, then

$$0 \in \underset{\delta}{\operatorname{argmax}} \left\{ (\alpha_G + \delta)' \mu(m(\alpha_G + \delta)) - C(m(\alpha_G + \delta)) - \frac{1}{2\rho}(\alpha_G + \delta)' V(\alpha_G + \delta) - \delta' \left[\mu(m(\alpha_G)) - \frac{1}{\tau} V(e - \alpha_G) \right] \right\}. \tag{A16}$$

(Note that we have written both $\mu(\cdot, \cdot)$ and $C(\cdot, \cdot)$ as functions of m alone since they do not depend on α.) Differentiating the maximand with respect to δ and evaluating the expression at $\delta = 0$, we obtain[30]

$$\alpha_G'\left[\frac{\partial\mu(m(\alpha_G))}{\partial m'}\right]\frac{\partial m(\alpha_G)}{\partial\alpha'} - \frac{\partial C(m(\alpha_G))}{\partial m'}\frac{\partial m(\alpha_G)}{\partial\alpha'} - \frac{1}{\rho}\alpha_G' V + \frac{1}{\tau}(e - \alpha_G)' V. \tag{A17}$$

Since by definition $m(\alpha)$ maximizes $\alpha'\mu(m) - C(m)$, it follows that

$$\alpha_G'\left[\frac{\partial\mu(m(\alpha_G))}{\partial m'}\right] - \frac{\partial C(m(\alpha_G))}{\partial m'} = 0. \tag{A18}$$

This means that (A17) can be simplified to

$$-\frac{1}{\rho}\alpha_G' V + \frac{1}{\tau}(e - \alpha_G)' V. \tag{A19}$$

This vanishes at $\alpha_G = \alpha^* = \rho e/(\rho + \tau)$. The concavity of $\Psi(\cdot)$ guarantees that the second-order conditions for maximization hold. Q.E.D.

Proof of Proposition 5

The optimal allocation α_S maximizes

$$\begin{aligned}\Psi(\alpha) + \Upsilon(\alpha) &= \alpha\mu(m(\alpha),\alpha) - C(m(\alpha),\alpha) - \frac{1}{2\rho}\alpha' V\alpha \\ &\quad + (e-\alpha)'\mu(m(\alpha),\alpha) - \frac{1}{2\tau}(e-\alpha)' V(e-\alpha) \\ &= e'\mu(m(\alpha),\alpha) - C(m(\alpha),\alpha) - \frac{1}{2\rho}\alpha' V\alpha \\ &\quad - \frac{1}{2\tau}(e-\alpha)' V(e-\alpha).\end{aligned} \tag{A20}$$

It follows that α_S solves

$$\begin{aligned}&e'\left[\frac{\partial\mu(m(\alpha_S),\alpha_S)}{\partial m'}\right]\left[\frac{\partial m(\alpha_S)}{\partial\alpha'}\right] - \frac{\partial C(m(\alpha_S),\alpha_S)}{\partial m'}\left[\frac{\partial m(\alpha_S)}{\partial\alpha'}\right] \\ &+ e'\left[\frac{\partial\mu(m(\alpha_S),\alpha_S)}{\partial\alpha'}\right] - \frac{\partial C(m(\alpha_S),\alpha_S)}{\partial\alpha'} - \frac{1}{\rho}\alpha_S' V + \frac{1}{\tau}(e-\alpha_S)' V = 0.\end{aligned} \tag{A21}$$

To show that $\alpha_S \neq \alpha^*$, we need to show only that

$$\begin{aligned}&e'\left[\frac{\partial\mu(m(\alpha^*),\alpha^*)}{\partial m'}\right]\left[\frac{\partial m(\alpha^*)}{\partial\alpha'}\right] - \frac{\partial C(m(\alpha^*),\alpha^*)}{\partial m'}\left[\frac{\partial m(\alpha^*)}{\partial\alpha'}\right] \\ &+ e'\left[\frac{\partial\mu(m(\alpha^*),\alpha^*)}{\partial\alpha'}\right] - \left[\frac{\partial C(m(\alpha^*),\alpha^*)}{\partial\alpha'}\right]\end{aligned} \tag{A22}$$

is not equal to zero. (Recall that $-(V\alpha^*/\rho) + [V(e - \alpha^*)/\tau] = 0$.) Since

[30] In writing an expression for the derivative, we use the following convention: if $x(y)$ is an M vector that is a function of the N vector y, then $\partial x(y)/\partial y'$ is an $M \times N$ matrix in which element (i, j) is the derivative of $x_i(y)$ with respect to y_j.

$$(\alpha^*)'\left[\frac{\partial\mu(m(\alpha^*),\alpha^*)}{\partial m'}\right] = \frac{\partial C(m(\alpha^*),\alpha^*)}{\partial m'}, \tag{A23}$$

the first two terms of (A22) are equal to

$$(e - \alpha^*)'\left[\frac{\partial\mu(m(\alpha^*),\alpha^*)}{\partial m'}\right]\left[\frac{\partial m(\alpha^*)}{\partial \alpha'}\right], \tag{A24}$$

which is a positive vector. Since α^* is proportional to e, it follows by assumption 2 that the third term is a nonnegative vector. Since, by assumption 2, $\partial C(m(\alpha^*), \alpha^*)/\partial\alpha'$ is a nonpositive vector, it follows that the entire expression is not equal to zero. Obviously with one asset we can conclude that $\alpha_S > \alpha^*$ since (A22) is positive and $-(V\alpha/\rho) + [V(1 - \alpha)/\tau]$ is decreasing in α. Q.E.D.

References

Barclay, Michael J., and Holderness, Clifford G. "Private Benefits from Control of Public Corporations." *J. Financial Econ.* 25 (December 1989): 371–95.

Black, Bernard S. "Agents Watching Agents: The Promise of Institutional Investor Voice." *UCLA Law Rev.* 39 (April 1992): 811–93. (*a*)

———. "The Value of Institutional Investor Monitoring: The Empirical Evidence." *UCLA Law Rev.* 39 (April 1992): 895–939. (*b*)

Brancato, Carolyn K. "The Pivotal Role of Institutional Investors in Capital Markets." In *Institutional Investing: Challenges and Responsibilities of the 21st Century,* edited by Arnold W. Sametz and James L. Bicksler. Homewood, Ill.: Business One Irwin, 1991.

Brickley, James A.; Lease, Ronald C.; and Smith, Clifford W., Jr. "Ownership Structure and Voting on Antitakeover Amendments." *J. Financial Econ.* 20 (January/March 1988): 267–91.

DeMarzo, Peter, and Bizer, David. "Sequential Trade." Working paper. Evanston, Ill.: Northwestern Univ., Kellogg School Management, 1993.

Grossman, Sanford J., and Hart, Oliver D. "Takeover Bids, the Free-Rider Problem, and the Theory of the Corporation." *Bell J. Econ.* 11 (Spring 1980): 42–64.

Grundfest, Joseph A. "Just Vote No: Minimalist Strategies for Dealing with Barbarians inside the Gate." *Stanford Law Rev.* 45 (April 1993): 857–937.

Huddart, Steven. "Ownership and Information as Substitutes." Research Paper no. 1138. Stanford, Calif.: Stanford Univ., Grad. School Bus., 1991.

———. "The Effect of a Large Shareholder on Corporate Value." Working paper. Stanford, Calif.: Stanford Univ., Grad. School Bus., 1992.

Jarrell, Gregg A., and Poulsen, Annette B. "Shark Repellents and Stock Prices: The Effects of Antitakeover Amendments since 1980." *J. Financial Econ.* 19 (September 1987): 127–68.

Kyle, Albert S., and Vila, Jean-Luc. "Noise Trading and Takeovers." *Rand J. Econ.* 22 (Spring 1991): 54–71.

Porter, Michael E. *Capital Choices: Changing the Way America Invests in Industry.* Washington: Council on Competitiveness and Harvard Bus. School, 1992.

Shleifer, Andrei, and Vishny, Robert W. "Large Shareholders and Corporate Control." *J.P.E.* 94, no. 3, pt. 1 (June 1986): 461–88.

Tirole, Jean. *The Theory of Industrial Organization.* Cambridge, Mass.: MIT Press, 1988.

Zwiebel, Jeffrey. "Block Investment and Partial Benefits of Corporate Control." Working paper. Cambridge: Massachusetts Inst. Tech., 1990.