

The “Wall Street Walk” as a Form of Shareholder Activism

Anat R. Admati and Paul Pfleiderer*

Graduate School of Business
Stanford University

October 2005

This version April 2006

We examine whether a large shareholder can alleviate the conflict of interest between shareholders and managers through his ability to sell his shares on the basis of private information. We show that large shareholder exit often has a disciplinary impact, but that (i) the effectiveness of this mechanism can be quite different depending on whether the agency problem involves a desirable or an undesirable action from shareholders’ perspective; (ii) additional private information may increase or decrease the large shareholder’s effectiveness; and (iii) in some cases the presence of the large shareholder may exacerbate the agency problem.

Journal of Economic Literature classification: D80, D82, G23, G34, K22

Keywords: shareholder activism, monitoring, exit, manager-shareholders conflict, agency, mutual funds, corporate governance.

* We would like to thank Bob Hall, Ken Judd, Gustavo Manso, Masako Ueda, seminar participants in the University of Wisconsin, Madison, and especially Jeff Zweibel, for helpful comments. Correspondence should be sent to Prof. Anat Admati, Graduate School of Business, Stanford University, Stanford, CA 94305-5015, e-mail admati@stanford.edu.

1. Introduction

The role of active large shareholders in improving corporate performance has been discussed extensively in the last two decades. Although institutional investors such as pension funds and mutual funds hold a substantial and increasing fraction of shares in public companies in the U.S., these large shareholders typically play a limited role in overt forms of shareholder activism such as takeovers, proxy fights, strategic voting, shareholders' proposals, etc. One possible reason for this is that many forms of shareholder activism are costly to active shareholders, and these shareholders may only realize a relatively small fraction of the benefits. In other words, we have the classic "free rider" problem. In addition to this, legal barriers, agency problems affecting the incentives of the large shareholder, and the fact that many large shareholders, particularly mutual funds, are committed through their charters not to invest resources to monitor their portfolio firms, have also worked to limit activism.¹ If a large shareholder is aware that a firm's management does not act in the best interest of shareholders, it may be rational for the shareholder to follow the so-called "Wall Street Rule" or "Wall Street Walk," which leads the shareholder to sell his shares (i.e., "vote with his feet") rather than attempt to be active.

Since the Wall Street Walk seems to be an alternative to activism, it appears to be inconsistent with it. This has led some (see, e.g., Bhide (1993) and Coffee (1993)) to argue that market liquidity that allows potentially active shareholders to exit and therefore not engage in monitoring and governance activities impairs corporate governance. What seems to have not been widely recognized is the possibility that the Wall Street Walk itself can be a form of shareholder activism. An exception is Palmiter (2002, p. 1437-8), who suggests that large shareholders may be able to affect managerial decisions through the "threat (actual or implied) of selling their holdings and driving down the price of the targeted company." If managers' compensation is tied to share prices, and if the exit of a large shareholder has a negative price impact, then the presence of a large shareholder who is potentially able to trade on private information may help discipline management and improve corporate governance. This form of monitoring is consistent with behind-the-scenes negotiation with management or 'jawboning' activities, which are often considered an alternative to more costly control mechanisms and which seem to be both common and often successful in affecting managerial decisions.

Note that it is not clear a priori whether and how the ability of a large shareholder to exit might work to alleviate agency problems between managers and shareholders. First, since exit by a large shareholder is assumed to drive the price of the targeted firm down, it would appear that the large shareholder loses when he carries out a threat to exit. This leads one to question

¹ There are many papers in the law and economics literature on the role of large shareholders in corporate governance and on shareholder activism. With each of them taking a somewhat unique point of view on the subject, Bainbridge (2005), Bebchuk (2005), Black (1990), Black and Coffee (1994), Gillan and Starks (1998), Grundfest (1993) Macey (1997), Palmiter (2002), Roe (1994) and Romano (1993, 2001) discuss how shareholder activism has been practiced, describe some of the barriers that have limited its use and effectiveness, and suggest ways that large shareholders can become more involved in corporate governance.

whether the threat is credible. Second, even if the threat is credible, it is not clear that the threat will always work to benefit other existing shareholders. Third, it is not clear whether the threat of exit continues to work when in addition to the price impact created by the large shareholder's exit, the large shareholder incurs other costs of carrying out the threat such as transactions costs. Finally, it is not clear that exit by the large shareholder benefits the other shareholders if the large shareholder's exit means the loss of future monitoring benefits that would accrue if the large shareholder remained a shareholder of the firm.

In this paper we examine all of these issues within a model where a firm's manager has incentives that are not aligned with shareholders. We assume that a large shareholder has private information about the manager's actions and/or about the consequences of these actions to the value of the firm, and he can sell his shares (exit) based on this information. In our model all agents are rational and prices perfectly reflect all public information, including the large shareholder's trading decisions. Our model combines two common elements present in trading models and in models of executive compensation, namely that (i) large shareholders have incentives to collect information and use it for trading and (ii) explicit and implicit managerial compensation contracts often lead managers to be sensitive to market prices of their firm. The resulting disciplinary impact of large shareholder exit has not been explicitly modeled to the best of our knowledge, but it potentially explains some of the observed interaction between large shareholders and managers.

We will consider two distinct types of agency problems. In one case, the manager can take an action that is undesirable from shareholders' perspective, but which produces a private benefit to the manager. In the other case, the action is desirable from shareholders' perspective, but it is privately costly to the manager.² The manager obviously knows whether he takes the action, and we assume that he also knows the impact of the action, if taken, on the value of the firm. The manager's decision and its impact will be observed publicly at the final date of our model, but in the short run investors have less than complete information about either the manager's decision or the consequences of the action, if taken, or both. The large shareholder in our model observes some information privately before other investors and may be able to sell his shares on the basis of this information. For each of the agency problems and for a number of different information structures, we examine whether and to what extent the presence of the privately-informed large shareholder reduces the agency costs associated with the action.

While the two types of agency problems described above (one with the "bad" action and one with the "good" action from shareholders' perspective) may seem to be mirror images of one another, it turns out that they can lead to dramatically different results with respect to the disciplinary impact of the large shareholder. For example, we find that in models where the action is "bad," the presence of a large shareholder generally increases the value of the firm

² We assume for most of the paper that investors are aware of which of these agency problems is present. (Section 8 discusses "model uncertainty.") Mature firms with large cash reserves (Free Cash Flow) may be more prone to a "bad action" type of agency problem. A "good action" problem may arise when the manager is reluctant to take on risky projects that are beneficial to shareholders but which entail high cost to the manager and carry the risk of failure and job loss.

by reducing agency costs. However, in models where the action is “good,” the exit behavior of the large shareholder can actually exacerbate the agency problem and thus be harmful to shareholders. At the same time, we show that in some settings there is more scope for the large shareholder to have a positive effect when the action is “good” than when it is “bad.” This happens specifically if the large shareholder observes privately only whether the manager has acted not in accordance with shareholders’ interest but does not have additional private information about the exact consequences of the manager’s actions. The results are different in the two models because the price impact the large shareholder has when he exits depends in part on the difference in the firm value when the manager acts according to shareholders’ preferences and when he does not. This difference can have different properties depending on whether shareholders prefer that the action in question is taken or that it is not taken.

In examining the situation in which exit by the large shareholder entails costs, we show that if the cost is not too high, and as long as the large shareholder retains an information advantage over other investors when he trades, the possibility of exit may still have a disciplinary impact. In fact, the large shareholder may be more effective in disciplining the manager if his exit in and of itself reduces the value of the firm (for example, due to the loss of potential benefits brought about by future monitoring activities), because the event in which he exits inflicts a larger punishment on the manager for not acting in shareholders’ interests. If exit is so costly that the large shareholder never exits, however, then of course this mechanism fails to produce any impact.

The notion that stock prices may play a role in monitoring managerial performance is certainly not new to this paper and was discussed in Holmstrom and Tirole (1993). Their model and the focus of their analysis differ from ours in several ways. In particular, Holmstrom and Tirole focus on how the ownership structure of the firm affects the value of market monitoring through its effect on liquidity and on the profits speculators realize in trading on information. In our model, by contrast, we focus on the disciplining impact of a large shareholder’s threat of exit and show that the effectiveness of this threat depends critically on the nature of the agency problem and the information structure.

The possibility that exit by a large shareholder can have positive impact on the firm is also discussed in Gopalan (2005). In this paper a sale by the informed large shareholder can serve to encourage another bidder to acquire information and implement improvements through a takeover mechanism. By contrast, our paper focuses on managerial incentives, and the impact of the large shareholder in our model is due to the effect of his trades on managerial compensation.

There is an extensive theoretical literature on shareholder activism and on the role of large shareholders in corporate governance.³ The models in this literature typically assume that the large shareholder can take a costly action, often called “monitoring,” to affect the value of the

³ A partial list includes Admati, Pfleiderer and Zechner (1994), Aghion, Bolton and Tirole (2004), Attari, Banerjee, and Noe (2006), Bolton and Von Thadden (1998), Burkart, Gromb and Panunzi (1997), Faure-Grimaud and Gromb (2004), Hartzell and Starks (2004), Huddart (1993), Kahn and Winton (1998), Katz and Owen (2001), Maug (1998), Mello and Repullo (2004), Noe (2002), Tirole (2001), and Yung (2005).

firm. Monitoring is associated with the ownership of shares, and inconsistent with exit. While the possibility that the large shareholder trades is considered in a number of these papers, the focus is generally on the incentives of the large shareholder to engage in monitoring and/or on the ownership structures that arise endogenously. In our model, by contrast, the ability of the large shareholder to exit is itself the technology by which he may affect managerial decisions. While in most of our analysis we take the information structure as given and do not incorporate the cost of acquiring information, our results have immediate implications for the case where information acquisition by the large shareholder is endogenous. This is discussed in the concluding remarks.

Empirical studies of the role and impact of large shareholders have documented various facts that are consistent with our model. In particular, Carleton, Nelson and Weisbach (1998) present evidence that large shareholders can affect firms' values through private negotiations. This is consistent with our model, because discipline through exit requires that the manager knows that the large shareholder is informed. Parrino, Sias, and Starks (2003) provide evidence to support both the notion that large shareholders are better informed than other investors, as our model assumes, and the fact that they sometimes use their private information to "vote with their feet" prior to CEO turnover. They suggest that the price impact of these trade may affect corporate decisions. Pound and Zeckhauser (1990) find that the presence of large shareholders is associated with higher expected earnings growth rates in industries where monitoring seems more likely because the information structure is more open. Sias, Starks and Titman (2001) also suggest that large shareholder trading has price impact that is likely due to superior information, consistent with our model.

The paper is organized as follows. We introduce our basic model and the two agency problems in Section 2. Section 3 analyzes the case where the large shareholder's private information includes only the manager's action and no investor observes the effect of the action on the value of the firm until the final period. In Section 4 we examine what happens when the large shareholder privately observes both the manager's decision and its impact on the firm's value. Section 5 considers the case where the manager's action is observed by all investors and the large shareholder's private information includes only the action's implications. For completeness, we analyze in Section 6 the case in which the impact of the action on the firm value is publicly known, and the private information of the large shareholder only concerns whether the action is taken or not. In Section 7 we consider situations where exit by the large shareholder is costly to the large shareholder beyond its negative price impact because of transactions costs. We also consider the possibility that exit results in an additional cost that is borne by all shareholders, specifically a cost that is associated with loss of future monitoring benefits provided by the large shareholder. Section 8 discusses briefly two extensions where investors have additional uncertainty, and Section 9 offers concluding remarks.

2. The General Model

There are three periods in our model. In period 0 the manager, whom we denote by M ,

decides whether or not to take a particular action. An agency problem arises because M and the shareholders of the firm have conflicting preferences with respect to this action. We will analyze two distinct models using the same notation. In one model, which we refer to as *Model B*, the action available to M is “bad” in the sense that it is undesirable from shareholders’ perspective, but the action produces a private benefit to M . In another model we analyze, *Model G*, the action is “good” for shareholders in that it increases the value of the firm, but it requires M to incur a private cost. We denote the value of the firm if M does not take the action by ν . If M takes the action in Model B, then the value of the firm decreases by $\tilde{\delta} \geq 0$ and becomes $\nu - \tilde{\delta}$, while M obtains a private benefit of $\beta > 0$. If the action is taken in Model G, then the value of the firm increases by $\tilde{\delta} \geq 0$ and becomes $\nu + \tilde{\delta}$, while M incurs a private cost of $\beta > 0$. These two models may seem like mirror images of one another but, as we will see, in our setting they can produce dramatically different results.

Our basic assumptions regarding uncertainty and the information structure are as follows. The status-quo value of the firm ν is fixed and common knowledge. This is without loss of generality in the sense that ν may be random but it is assumed that no agent has private information about it. (In fact, the value of ν will not play any significant role in our analysis.) In period zero, investors assess that $\tilde{\delta}$ has a non-trivial continuous distribution $f(\cdot)$ with support on $[0, \bar{\delta}]$, where $\bar{\delta}$ is positive and possibly infinite. M observes the realization of $\tilde{\delta}$ before making the decision whether to take the action or not. We further assume that the private cost or benefit β is fixed and known to all investors. (The possibility that β is random is discussed in Section 8.)

Since M makes his decision regarding the action after observing $\tilde{\delta}$, his strategy can be described by a function $a(\tilde{\delta})$, where $a(\tilde{\delta}) = 1$ denotes the event in which M takes the action and $a(\tilde{\delta}) = 0$ denotes the event in which he does not take it. In most of our analysis, investors will make inferences regarding the expected change in the firm value, which is given by $a(\tilde{\delta})\tilde{\delta}$, given the information they have. To simplify the notation we will use the short-hand \tilde{a} to represent M ’s action instead of $a(\tilde{\delta})$. We assume that in the final period, investors do observe the realizations of both \tilde{a} and $\tilde{\delta}$. The value of the firm is therefore $\nu - \tilde{a}\tilde{\delta}$ in Model B and $\nu + \tilde{a}\tilde{\delta}$ in Model G.

We assume that the firm is owned by many small and passive investors as well as by a large shareholder, whom we denote by L . (For expositional clarity we will use female pronouns to refer to L .) The exact ownership structure will not matter to our results, since valuation will be done under risk neutrality. We assume that L observes some private information regarding \tilde{a} and/or $\tilde{\delta}$ in period 1, and that she may be in a position to sell her shares on the basis of this information. Because it generally reflects her private information, L ’s trading decision will have an impact on the firm’s price in period 1. This in turn has the potential to affect M ’s decision if M cares about the market price of the firm in period 1.

The manager’s compensation is assumed to be linear in the realized market price of the firm in periods 1 and 2, P_1 and P_2 . Specifically, we assume that M ’s compensation is equal to $\omega_1 P_1 + \omega_2 P_2$, where ω_1 and ω_2 are non-negative coefficients representing the dependence of the compensation on the firm’s short-term (“Period 1”) and long-term (“Period 2”) price performance

respectively.⁴ We assume for most of our analysis that ω_1 and ω_2 are positive, but we will also consider the limit cases where ω_2 vanishes. If $\omega_1 = 0$, then L will not be able to affect M 's decision through her trade. The potential impact of L on M 's decision comes about through the impact of her trading decisions on P_1 . We assume that the prices, P_1 and P_2 , are set by risk-neutral, competitive market makers and therefore reflect all of the information publicly available. This means that P_2 equals $\nu - \tilde{a}\tilde{\delta}$ in Model B and $\nu + \tilde{a}\tilde{\delta}$ in Model G . In Period 1, P_1 reflects the information contained in L 's trading decision as is described in more detail below.

If M does not take the action, then his utility is simply his compensation, $\omega_1 P_1 + \omega_2 P_2$. If he takes the action in Model B , M 's utility is equal to the sum of his compensation and the private benefit β . Similarly, if M takes the action in Model G , then his utility is equal to the compensation minus the private cost β . We assume that M chooses whether to take the action or not to maximize his expected utility for every realization of $\tilde{\delta}$.

We assume that L may be subject to a liquidity shock in period 1. Specifically, there is probability $0 < \theta < 1$ that, independent of her private information, L will need to sell her entire stake in period 1.⁵ While the value of θ is common knowledge, only L knows her actual motives for trading when she trades. We generally assume that $\theta > 0$, but our model is also well defined in the limit case where $\theta = 0$, i.e., when L is never subject to a liquidity shock. As will become clear, most of our results will apply to this case as well, and we will often use it for illustration. The main complication for the $\theta = 0$ case is that additional equilibria can arise that are not the limit of any equilibrium for the case $\theta > 0$ as θ vanishes.⁶ The equilibria we analyze for the $\theta > 0$ cases always have a well-defined limit as θ goes to zero, and are indeed equilibria of the model in which $\theta = 0$.⁷ If she is not subject to a liquidity shock, L chooses whether or not to

⁴ We take the form of M 's compensation as exogenous here. Presumably, M 's compensation balances risk sharing and various agency considerations. If the agency problem considered in this paper was the only relevant problem in the contracting environment between M and shareholders, then it would be reasonable to endogenize the dependence of the compensation on prices. However, we believe that in reality M 's compensation is designed to solve a more complex problem. We therefore limit ourselves to asking what disciplinary impact L can have given this compensation function. It might be interesting to examine the robustness of our results to other compensation functions or to attempt to endogenize the compensation within a more general agency framework.

⁵ It simplifies our analysis that there is only one quantity that L sells when she is subject to a liquidity shock. However, our results would not change if the liquidity shock entailed L selling less than her entire stake. Our model can also be analyzed under the assumption that L receives a shock that forces her to buy shares or a shock that might involve either buying or selling. We have not examined systematically all the variations of our model under these alternative trading assumptions, since our focus is on the disciplinary impact of the Wall Street Walk, i.e., the possibility of exit by the large shareholder. For those versions of the model that we have analyzed, however, which are those discussed in Section 3, we have found that having a positive probability of a liquidity buy trade does not change the qualitative nature of our results.

⁶ For example, there may be equilibria in which L never sells and therefore never has effect on M 's decisions through her trading, or equilibria where L sells only some of her shares. Eliminating these equilibria would require restrictions on out-of-equilibrium beliefs.

⁷ Thus, whenever we discuss L 's equilibrium strategies for the case $\theta = 0$, we will assume

sell her shares if the expected value of the firm given all her information is smaller than P_1 , the price at which she would exit, where P_1 incorporates the information communicated by the sale.

We will analyze the Bayesian-Nash equilibria of Model B and Model G under various assumptions concerning what L and other investors observe in period 1. In such equilibria M , using his information, makes the optimal decision regarding his action, taking L 's trading strategy as given. Similarly, L , based on her information, determines whether to sell her shares in the event that she is not subject to a liquidity shock.⁸ Both M and L take as given the fact that P_1 will reflect the conditional expectation of $\tilde{a}\tilde{\delta}$ based on the information available to investors, including L 's trading decision, in period 1.

For most of our analysis we will make the following tie-breaking assumptions: (i) if M is indifferent between taking the action and not taking the action then he takes the action; (ii) if L is indifferent between selling her shares and not selling, she sells her shares. Generally, when we use these assumptions, they will not change the set of equilibria, because $\tilde{\delta}$ has a continuous distribution and indifference will hold for at most one realization of $\tilde{\delta}$. We will not employ these assumptions in cases where they would change the set of equilibria non-trivially, e.g., in the limit case where $\omega_2 = 0$ or in the model where $\tilde{\delta}$ is publicly observable in period 1 (which is analyzed in Section 6).

Given a particular realization of $\tilde{\delta}$, M must decide whether to take the action. It is easy to see that in the benchmark case in which L is not present and \tilde{a} is not observed by investors until period 2, M will take the action if and only if $\tilde{\delta} \leq \beta/\omega_2$ in Model B and if and only if $\tilde{\delta} \geq \beta/\omega_2$ in model G. For most of our analysis we assume that $\beta/\omega_2 < \bar{\delta}$, i.e., that there is a positive probability that M acts in the interests of the shareholders (refraining from taking the action in Model B and taking the action in Model G).⁹ For a strategy of M that specifies whether he takes the action or not as a function of $\tilde{\delta}$, the *ex ante* expected value of the firm in Model B is $\nu - E(\tilde{\delta}\tilde{a})$. Similarly, in model G, the *ex ante* value of the firm given M 's strategy is $\nu + E(\tilde{\delta}\tilde{a})$.

Note that the best outcome from shareholders' perspective in Model B is that M never takes the action, which means that highest value of the firm in this case is ν . Since the firm value is reduced by $\tilde{\delta}$ whenever $\tilde{a} = 1$ relative to this best case, the *ex ante* expected agency cost associated with the action in Model B is $E(\tilde{\delta}\tilde{a})$. Analogously, the first best from shareholders' perspective

that they are obtained as the limit of her strategies when $\theta > 0$. For example, if L is indifferent between selling and not selling her shares for a given value of $\tilde{\delta}$ if $\theta = 0$, but she strictly prefers to sell for the same value of $\tilde{\delta}$ in the model whenever $\theta > 0$, then we will assume that L sells her shares also in the model with $\theta = 0$.

⁸ We do not need to consider trades by L that cannot arise from a liquidity shock, since such trades will generally lead to zero expected profits under reasonable assumptions about investors' beliefs. However, when investors cannot be sure whether L 's trade is based on private information or liquidity, L makes positive profits when he trades on his information. See the discussion of L 's information advantage in Section 7.

⁹ This assumption, which holds trivially when $\bar{\delta}$ is infinite, is made for ease of presentation only. If $\omega_2 > 0$ our results are easily modified when it does not hold. We will address the limit case where $\omega_2 = 0$ separately in some of our analysis.

in Model G is that M always takes the action, which increases the value of the firm from ν to $\nu + \tilde{\delta}$. Since the increase of $\tilde{\delta}$ is not realized whenever $\tilde{a} = 1$, the *ex ante* expected agency cost in Model G is equal to $E(\tilde{\delta}(1 - \tilde{a})) = E(\tilde{\delta}) - E(\tilde{\delta}\tilde{a})$. Thus, in Model B the *ex ante* expected agency cost is reduced if $E(\tilde{\delta}\tilde{a})$ is made lower, while the opposite is true in Model G. We will be interested in the impact that L 's presence has on the *ex ante* expected agency cost in the two models, which from now on we will simply refer to as agency cost. This impact is measured by the difference between the agency cost in the equilibrium where L is not present and the agency cost in the equilibrium when L is present. It will be useful to use the following terms:

Definition: Consider the impact that L 's presence has on the agency cost associated with the action.

- (i) An equilibrium is *disciplining* if L 's presence has a positive impact, i.e., the agency cost is lower when L is present than when she is not present.
- (ii) An equilibrium is *non-disciplining* if L 's presence has no impact on the agency cost.
- (iii) An equilibrium is *dysfunctional* if L 's presence has a negative impact, i.e., the agency cost is higher when L is present than when she is not present.

It is easy to see that in order for the equilibrium to be disciplining, L must have some information about \tilde{a} in period 1, and that some of her information at that point must be private. We will examine L 's impact under various information structures that satisfy this condition. We will at times also compare, for the same model specifications, L 's effectiveness in Model B vs. Model G. To distinguish the different information structures, we will use superscripts to denote the information observed by L in period 1, and subscripts to denote the information (if any) that is publicly observed by all investors in period 1. For example, Model B^a is Model B where L observes \tilde{a} in period 1 and investors do not observe either \tilde{a} or $\tilde{\delta}$ directly until period 2; Model $G_a^{\tilde{\delta}}$ is Model G where \tilde{a} is observed publicly in period 1, and, in addition, L observes $\tilde{\delta}$ privately in period 1. Note that when there is no subscript, investors do not observe either \tilde{a} or $\tilde{\delta}$ in period 1. When there is no superscript, the model is the base or benchmark model where L is not present.

3. Action-Only Monitoring

We start our analysis by assuming that L observes privately whether M has taken the action, i.e., the realization of \tilde{a} . However, neither L nor other investors are assumed to observe the realization of $\tilde{\delta}$ until period 2. The models in this section therefore are denoted by the superscript a (and no subscript).

Consider first Model B^a . It is easy to see that, since the action reduces the value of the firm, in every equilibrium of this model L sells her shares whenever she observes that M has taken the action.¹⁰ Let E_s be the expected value of $\tilde{a}\tilde{\delta}$ conditional on L selling her shares and E_{ns} be the

¹⁰ It is immediate that L must *weakly* prefer to sell if the action is taken, and we can invoke

expected value of $\tilde{a}\tilde{\delta}$ conditional on L not selling her shares. Then the price of the firm in period 1, P_1 , can take on two possible values, $\nu - E_s$ if L sells her shares, and $\nu - E_{ns}$ if she does not. Note that, when $\theta > 0$, L might be forced to sell for liquidity reasons even if M does not take the action, and this, together with the strategies of L and M , will need to be incorporated into the determination of E_s and E_{ns} .

If M takes the action for a particular δ , P_2 is equal to $\nu - \delta$. Since L exits with probability 1 when M takes the action, P_1 is equal to $\nu - E_s$. Thus, M 's expected utility is

$$\beta + \omega_1(\nu - E_s) + \omega_2(\nu - \delta). \quad (1)$$

If M does not take the action then $P_2 = \nu$, and P_1 is equal to $\nu - E_s$ with probability θ and $\nu - E_{ns}$ with probability $1 - \theta$, because L sells if and only if she is subject to a liquidity shock. Thus, M 's expected utility if he does not take the action is

$$\omega_1(\nu - \theta E_s - (1 - \theta)E_{ns}) + \omega_2\nu. \quad (2)$$

Comparing these, we conclude that M will take the action if and only if

$$\beta - (1 - \theta)\omega_1(E_s - E_{ns}) - \omega_2\delta \geq 0. \quad (3)$$

The potential impact of L 's presence comes about through the second term in the equation above, $(1 - \theta)\omega_1(E_s - E_{ns})$. This term depends on the difference between the first-period price given a sale by L and the first-period price given that L retains its shares. To the extent that exit reflects negative information, this is negative. The absolute value of this difference measures the extent to which L exerts "punishment" on M by selling her shares and driving the price down when M takes the action.

Note that, since $\omega_2 > 0$, the left-hand side of (3) is decreasing in δ . This implies that if M prefers to take the action for a given δ , then he must strictly prefer to take it for all smaller values. An equilibrium of Model B^a will therefore be characterized by a cutoff point x such that the action is taken if and only if $\tilde{\delta} \leq x$. Given such a strategy for M , and since L sells her shares if M takes the action or if she is subject to a liquidity shock, we have

$$E_s(x) = \frac{\Pr(\tilde{\delta} \leq x)E(\tilde{\delta} \mid \tilde{\delta} \leq x)}{\theta + (1 - \theta)\Pr(\tilde{\delta} \leq x)}; \quad E_{ns}(x) = 0. \quad (4)$$

the tie-breaking assumption that she sells when she is indifferent. As we will see, in the unique equilibrium of our model L in fact strictly prefers to sell her shares when the action is taken.

Note that we use the notation $E_s(x)$ to signify the dependence of E_s on the cutoff point x . Since investors can infer that $\tilde{a} = 0$ (the action was not taken) if L does not sell her shares, $E_{ns}(x) = 0$ independent of x . The calculation of $E_s(x)$ takes into account that with probability θ a sale by L is due to a liquidity shock and therefore is uninformative about $\tilde{a}\tilde{\delta}$, while with probability $1 - \theta$ a sale implies that $\tilde{a} = 1$ and therefore, given M 's strategy, that $\tilde{\delta} \leq x$. Note that for a given x , $E_s(x)$ is decreasing in θ , and in the limit case where $\theta = 1$, $E_s(x) = \Pr(\tilde{\delta} \leq x)E(\tilde{\delta} \mid \tilde{\delta} \leq x)$. Conversely, when θ vanishes, E_s approaches $E(\tilde{\delta} \mid \tilde{\delta} \leq x)$ since in this case a sale conveys perfectly that $\tilde{a} = 1$.

Now consider M 's decision whether or not to take the action. Since $\tilde{\delta}$ has a continuous distribution, it is easy to see that in any disciplining equilibrium, M must be indifferent between taking and not taking the action at the equilibrium cutoff $\tilde{\delta} = x_B$.¹¹ Thus, x_B must satisfy

$$\beta - (1 - \theta)\omega_1(E_s(x_B) - E_{ns}(x_B)) - \omega_2x_B = 0. \quad (5)$$

Let us now turn to Model G^a. Since the action increases the value of the firm, in every equilibrium L sells her shares if the action is *not* taken (or if she is subject to a liquidity shock), and retains her shares if the action is taken. Let E_s and E_{ns} be the conditional expectations of $\tilde{a}\tilde{\delta}$ given sale and no sale by L respectively. If M takes the action for $\tilde{\delta} = \delta$, then $P_2 = \nu + \delta$, and, since L sells if and only if she is subject to a liquidity shock, P_1 is equal to $\nu + E_s$ with probability θ and $\nu + E_{ns}$ with probability $1 - \theta$. Thus, M 's expected utility if he takes the action is

$$-\beta + \omega_1(\theta(\nu + E_s) + (1 - \theta)(\nu + E_{ns})) + \omega_2(\nu + \delta). \quad (6)$$

If M does not take the action, then $P_1 = \nu + E_s$, since L sells with probability 1, and $P_2 = \nu$. Thus, M 's expected utility if he does not take the action is

$$\omega_1(\nu + E_s) + \omega_2\nu \quad (7)$$

It follows that M prefers to take the action if and only if

$$-\beta + (1 - \theta)\omega_1(E_{ns} - E_s) + \omega_2\delta \geq 0. \quad (8)$$

¹¹ Our tie-breaking assumption is that M takes the action for $\tilde{\delta} = x_B$, but of course with a continuous distribution this does not matter to the calculation of $E_s(x_B)$ and $E_{ns}(x_B)$ and therefore does not affect the equilibrium.

The left-hand side of (8) is increasing in δ , so an equilibrium for this model involves a cutoff point x such that M takes the action if and only if $\tilde{\delta} \geq x$. Since L sells her shares if she is subject to a liquidity shock or if M does not take the action, we have for Model G^a , again using the notation $E_s(x)$ and $E_{ns}(x)$ to signify the dependence of prices on M 's cutoff point,

$$E_s(x) = \frac{\theta \Pr(\tilde{\delta} \geq x) E(\tilde{\delta} | \tilde{\delta} \geq x)}{\theta + (1 - \theta) \Pr(\tilde{\delta} < x)}; \quad E_{ns}(x) = E(\tilde{\delta} | \tilde{\delta} \geq x). \quad (9)$$

In Model G^a , no sale by L communicates to investors that the action was definitely taken ($\tilde{a} = 1$), and thus that $\tilde{\delta} \geq x$. A sale by L communicates that either the action was not taken (and therefore $\tilde{a} = 0$ and $\tilde{\delta} < x$), or that L was subject to a liquidity shock, which is uninformative about \tilde{a} and $\tilde{\delta}$. Fixing x , $E_s(x)$ is increasing in θ . In the limit when $\theta = 1$, a sale is uninformative about $\tilde{a}\tilde{\delta}$ and thus $E_s(x) = \Pr(\tilde{\delta} \geq x) E(\tilde{\delta} | \tilde{\delta} \geq x)$. As θ vanishes, however, a sale implies that the action was definitely not taken, and thus in the limit $E_s(x) = 0$.

An equilibrium for Model G^a will be characterized by a cutoff x_G such that M is indifferent between taking and not taking the action when $\tilde{\delta} = x_G$, and thus x_G satisfies

$$-\beta + (1 - \theta)\omega_1(E_{ns}(x_G) - E_s(x_G)) + \omega_2 x_G = 0. \quad (10)$$

As already observed, if L is not present, then the equilibrium cutoff point in both models is equal to β/ω_2 . That is, in this benchmark case the action is taken for $\tilde{\delta} \leq x_B = \beta/\omega_2$ in Model B and for $\tilde{\delta} \geq x_G = \beta/\omega_2$ in model G. Note also that in both models, shareholders are better off the lower is the equilibrium cutoff point, because when $\tilde{\delta}$ is below the cutoff point in both models, M is not acting in their best interests. The discipline L is able to exert on M 's actions is thus measured in both models by how low the equilibrium cutoff is. The following result characterizes the equilibrium for both models and compares Model B^a with Model G^a in terms of L 's effectiveness in disciplining M . The proofs of this and other results are found in the appendix.

Proposition 1: *In both Model B^a and Model G^a there exists a unique equilibrium, and the equilibrium is always disciplining.*

- (i) *In Model B^a equilibrium is characterized by a cutoff $x_B < \beta/\omega_2$ such that the manager takes the action if and only if $\tilde{\delta} \leq x_B$ and the large shareholder sells her shares if the manager takes the action. The cutoff x_B solves (5), where $E_s(\cdot)$ and $E_{ns}(\cdot)$ satisfy (4)*
- (ii) *In Model G^a equilibrium is characterized by a cutoff $x_G < \beta/\omega_2$ such that the manager takes the action if and only if $\tilde{\delta} \geq x_G$ and the large shareholder sells her shares if the manager does not take the action. The cutoff x_G solves (10), and $E_s(\cdot)$ and $E_{ns}(\cdot)$ satisfy (9).*

- (iii) Fixing β , ω_1 , ω_2 , and the distribution of $\tilde{\delta}$, both x_B and x_G are increasing in θ .
- (iv) Fixing β , θ , ω_1 , ω_2 , and the distribution of $\tilde{\delta}$, $x_G < x_B$. That is, if all else is equal, the large shareholder is more effective in disciplining the manager in Model G^a than in Model B^a .

This proposition states that when L observes \tilde{a} privately and no investor observes $\tilde{\delta}$ until period 2, the credible threat that L will exit if M does not act in shareholders' interests is an effective disciplining tool and reduces the agency cost. L 's impact is decreasing in θ , the probability that she is subject to a liquidity shock, because a higher value of θ makes her trades less informative and reduces her ability to "punish" M for not acting in shareholders' interests.

Perhaps surprisingly, since the two models appear as mirror images of one another, part (iv) of Proposition 1 states that, fixing all the model's parameters, L is more effective in Model G^a than she is in Model B^a . To understand this result, consider the limit case $\theta = 0$, where L is not subject to a liquidity shock. In Model B^a , the action is taken and L sells her shares if and only if $\tilde{\delta}$ has a relatively *low* realization, and $E_s(x) - E_{ns}(x) = E(\tilde{\delta} \mid \tilde{\delta} \leq x)$ for any possible cutoff point x . By contrast, in Model G^a the action is taken for relatively *large* realizations of $\tilde{\delta}$, and if $\theta = 0$ then $E_{ns}(x) - E_s(x) = E(\tilde{\delta} \mid \tilde{\delta} \geq x)$. Clearly, for any candidate cutoff point x and any distribution for $\tilde{\delta}$, $E(\tilde{\delta} \mid \tilde{\delta} \leq x) < E(\tilde{\delta} \mid \tilde{\delta} \geq x)$. The result that $|E_{ns}(x) - E_s(x)|$ is larger in Model G^a than in Model B^a for any x holds also when $0 < \theta < 1$, and the information communicated about \tilde{a} by L 's sales is imperfect. This means that L has a larger impact on M 's compensation for any cutoff level x in Model G^a than she has in Model B^a , which implies that for the equilibrium cutoffs we have $x_G < x_B$.

The difference between Models B^a and G^a is most apparent when we consider L 's disciplining tool, $|E_s(x) - E_{ns}(x)|$, as the cutoff x goes to zero, i.e., as we approach the best situation (in both models) from shareholders perspective. In Model B^a , this is equal to $E(\tilde{\delta} \mid \tilde{\delta} \leq x)$, which goes to zero as x vanishes. Thus, as M 's preferences get better aligned with those of shareholders, the tool that L can use to discipline M vanishes in Model B^a . This is not true in Model G^a , where L always has a non-trivial disciplining tool, because the difference between the price when L does not sell and the price when L sells always remains bounded away from zero. If $\theta = 0$, $E_{ns}(x) - E_s(x) = E(\tilde{\delta} \mid \tilde{\delta} \geq x) > E(\tilde{\delta}) > 0$. More generally, $E_{ns}(x) - E_s(x)$ has a positive lower bound that depends on θ . Thus, L is better able to exert discipline in Model G^a than in Model B^a .¹²

It is interesting to examine whether and to what extent L can have a positive disciplinary

¹² The reader might wonder whether our results might change if the specification of the possible liquidity trades is different. For example, would the results continue to hold if there is a positive probability that a liquidity shock would lead L to buy shares? We have analyzed Model B^a and Model G^a with this assumption. One complication that arises when there is a positive probability that L will buy shares due to a liquidity shock is that multiple equilibria may arise. Nevertheless, it can be shown that the equilibria in both models are disciplining and that the lowest equilibrium cutoff for Model G^a is smaller than the lowest equilibrium cutoff for Model B^a . Details of this analysis are available upon request.

impact when $\omega_2 = 0$, i.e., M 's compensation does not depend on the long-term price P_2 . Note first that if $\omega_2 = 0$ and L is not present, the agency problem is particularly severe since M 's compensation is independent of his action. Thus, M will always act against the shareholders' interests, taking the action for every $\tilde{\delta}$ in Model B and never taking the action in Model G. Thus, any equilibrium of Model B^a in which there is a positive probability that M does not take the action is a disciplining equilibrium, and similarly any equilibrium of Model G^a in which M takes the action with positive probability is also disciplining.

Now note that when $\omega_2 = 0$ and L is present, M 's compensation still does not depend on the actual realization of $\tilde{\delta}$ but only on the absolute difference $|E_s - E_{ns}|$ that depends on L 's trading strategy. This means that it is no longer true that in every equilibrium of Model B^a the manager takes the action for realizations of $\tilde{\delta}$ below a cutoff point and that in every equilibrium of Model G^a he must take the action for all realizations of $\tilde{\delta}$ above a cutoff point, and that M is only indifferent between taking and not taking the action at the equilibrium cutoff point. In fact, when $\omega_2 = 0$, similar logic to that of (5) and (10) implies that in any disciplining equilibrium M must be indifferent between taking and not taking the action for *every* realization of $\tilde{\delta}$.

The analysis of the case $\omega_2 = 0$ turns out to be quite complicated, but it leads to some interesting results, which are summarized in the next proposition.

Proposition 2: *Assume that $\omega_2 = 0$ and $\omega_1 > 0$.*

- (i) *In both Models B^a and G^a, the equilibria described in Proposition 1 have well defined limits as ω_2 vanishes. The limit of the equilibria in each case is an equilibrium for the models with $\omega_2 = 0$.*
- (ii) *In any disciplining equilibrium of either model, $(1 - \theta)\omega_1|E_s - E_{ns}| = \beta$, and the manager is indifferent between taking and not taking the action for all realizations of $\tilde{\delta}$.*
- (iii) *If there are multiple equilibria in Model B^a, then*
 - (a) *the equilibrium with the lowest agency cost (i.e., the best equilibrium for the shareholders) is one in which the manager takes the action for all $\tilde{\delta} \geq x$ for some x ;*
 - (b) *the equilibrium obtained in the limit as ω_2 vanishes, where the action is taken for $\tilde{\delta} \leq x_B$ for a cutoff x_B , is the one where the agency cost is the largest among all the equilibria, i.e., where the large shareholder is least effective in disciplining the manager.*
- (iv) *If there are multiple equilibria in Model G^a, then the equilibrium obtained as ω_2 vanishes, where the manager takes the action when $\tilde{\delta} \geq x_G$ for a cutoff x_G , is the one where the agency cost is the smallest among all the equilibria, i.e., where L is the most effective in disciplining the manager.*

This result shows that L can have a disciplinary impact on M 's action when $\omega_2 = 0$, but that there are often multiple equilibria in this case.¹³ Models B^a and G^a may now have equilibria

¹³ In fact, in addition to multiple, possibly a continuum, of pure-strategy equilibria, these models

that have similar forms, but the models continue to produce very different results. Unlike the case $\omega_2 > 0$, it is no longer true in general that L is more effective in Model G^a than she is in Model B^a in *all* of the equilibria.

The proposition states that when $\omega_2 = 0$, the best equilibria from shareholders' perspective in *both* Model B^a and Model G^a have the form we have encountered earlier for Model G^a , where M takes the action for realizations of $\tilde{\delta}$ *above* a cutoff. Note that in Model G^a , these relatively large realizations of $\tilde{\delta}$ represent the most beneficial ones from shareholders' perspective, while in Model B^a these realizations are the most harmful from shareholders' perspective. Thus, perhaps surprisingly, the equilibrium with the lowest agency cost in Model B^a involves M taking the action for the worst realizations of $\tilde{\delta}$. The key to this result is that, although the realizations for which the action is taken are the worst for shareholders, in this equilibrium the probability that M takes the action is the smallest among all the equilibria.

4. Is L More Effective with More Private Information?

We now assume that L is able to observe privately not only whether M has taken the action, i.e., the realization of \tilde{a} , but also the realization of $\tilde{\delta}$. We continue to assume that neither \tilde{a} nor $\tilde{\delta}$ is observed by other investors until period 2. As will become clear, the resulting models, denoted $B^{a,\delta}$ and $G^{a,\delta}$, produce dramatically different results from one another with this information structure, and we will therefore discuss them separately.

Consider Model $B^{a,\delta}$ first. Note that with $\omega_2 > 0$, any equilibrium must still involve M taking the action only if $\tilde{\delta}$ falls below a cutoff point. This follows immediately from considerations similar to those in the previous section. Again it is easy to see that in any equilibrium L prefers not to sell her shares if M does not take the action. However, when L observes $\tilde{\delta}$ it is no longer the case that she always prefers to sell when the action is taken by M . If $\tilde{\delta} < E_s$, where E_s is the market's conditional expectation of $\tilde{a}\tilde{\delta}$ given that L sells her shares, then even if M takes the action, the price response to the sale, measured by E_s , is more severe than the loss to the value of the firm that L will incur if she retains her shares, which is equal to $\tilde{\delta}$. Thus, L will not want to sell if $\tilde{\delta} < E_s$.

The above implies that if M takes the action for $\tilde{\delta} \leq x$, then we must have¹⁴

$$E_s(x) = \frac{\theta \Pr(\tilde{\delta} \leq x) E(\tilde{\delta} \mid \tilde{\delta} \leq x) + (1 - \theta) \Pr(\tilde{\delta} \in [E_s(x), x]) E(\tilde{\delta} \mid \tilde{\delta} \in [E_s(x), x])}{\theta + (1 - \theta) \Pr(\tilde{\delta} \in [E_s(x), x])}, \quad (11)$$

possess a continuum of mixed-strategy equilibria, where M mixes between taking the action and not taking it.

¹⁴ Note that $E_s(x)$ must be smaller than x , since it is equal to the weighted average of $\Pr(\tilde{\delta} \leq x) E(\tilde{\delta} \mid \tilde{\delta} \leq x)$ and $E(\tilde{\delta} \mid \tilde{\delta} \in [E_s(x), x])$, both of which are smaller than x .

and

$$E_{ns}(x) = \frac{\Pr(\tilde{\delta} < E_s(x))E(\tilde{\delta} | \tilde{\delta} < E_s(x))}{1 - \Pr(\tilde{\delta} \in [E_s(x), x])}. \quad (12)$$

Note that a sale by L communicates that either the action was taken and L chose to sell, i.e., $\tilde{a} = 1$ and $\tilde{\delta} \in [E_s(x), x]$, or L was subject to a liquidity shock. No sale by L communicates that either the action was not taken, or that it was taken and $\tilde{\delta} < E_s(x)$. An equilibrium cutoff point x_B will satisfy the indifference condition for M given again by

$$\beta - (1 - \theta)\omega_1(E_s(x_B) - E_{ns}(x_B)) - \omega_2x_B = 0, \quad (13)$$

The following result states that an equilibrium for Model $B^{a,\delta}$ with the characterization described above exists, and that L always has a disciplinary impact in equilibrium.

Proposition 3: *There exists at least one equilibrium in Model $B^{a,\delta}$, and every equilibrium is disciplining. For every equilibrium there is a cutoff $x_B < \beta/\omega_2$ such that the manager takes the action if and only if $\tilde{\delta} \leq x_B$, and the large shareholder sells her shares if the action is taken and $\tilde{\delta} \geq E_s(x_B)$, where $E_s(\cdot)$ and $E_{ns}(\cdot)$ are given by (11) and (12) and x_B solves (13). The probability that the large shareholder sells her shares in equilibrium is positive when $\theta > 0$ and vanishes as θ goes to zero.*

Note that L has a disciplinary impact in this model even though in equilibrium an actual exit may be observed quite rarely. This is unlike Model B^a , where L sells her shares in equilibrium whenever $\tilde{\delta} \leq x_B$. In fact, when the probability of a liquidity shock θ is very small, L is extremely unlikely to exit in the equilibrium of Model $B^{a,\delta}$. Nevertheless, L can have a significant disciplinary impact, despite this low probability of exit. When L exits, the market concludes that, except for the possibility of a liquidity shock, $\tilde{a}\tilde{\delta} \in [E_s(x_B), x_B]$ and therefore that not only was the action likely to have been taken, but that, if the action was taken, then $\tilde{\delta}$ is in the relatively more “harmful” range of values. In the limit when θ vanishes, L only exits when $\tilde{\delta} = x_B$, i.e., with probability zero. A sale in this case communicates that the action was taken and that $\tilde{\delta}$ is equal to the *worst* value for which the action is taken in equilibrium, namely x_B .¹⁵ Note, however, while the price impact of a sale is more pronounced, the information content of L not

¹⁵ To understand this intuitively note that when $\theta = 0$, since L knows both \tilde{a} and $\tilde{\delta}$ and has no liquidity motivation for trade, it is not possible that in equilibrium L sells for more than one realization of $\tilde{\delta}$ since $E_s(x_B)$ would have then had to be above one of the possible realizations of $\tilde{\delta}$ for which L sells. This is a contradiction because it entails L selling for some realizations below $E_s(x_B)$, which is suboptimal. When $\theta > 0$ then it is possible that $E_s(x_B) < x_B$ and so there is an interval of realizations of $\tilde{\delta}$, between $E_s(x_B)$ and x_B for which L sells in equilibrium.

selling, is diminished. In particular, in the limit case when $\theta = 0$, since L sells with probability zero, no information about \tilde{a} is communicated if L does not exit, and thus $E_{ns}(x_B)$ is equal the unconditional expectation of \tilde{a} given M 's strategy of taking the action for $\tilde{\delta} \leq x_B$, namely $\Pr(\tilde{\delta} \leq x_B)E(\tilde{\delta} | \tilde{\delta} \leq x_B)$.

Is L 's impact larger in Model $B^{a,\delta}$, where she has more information, than in Model B^a ? It turns out that the answer to this is ambiguous in general. We will consider first the limit case where $\theta = 0$, and then discuss the general case $\theta > 0$. Let x be a candidate cutoff point for M 's strategy. If M takes the action when $\tilde{\delta} \leq x$, then in Model B^a , $E_s(x) = E(\tilde{\delta} | \tilde{\delta} \leq x)$ and $E_{ns}(x) = 0$. As noted above, in Model $B^{a,\delta}$, we have $E_s(x) = x$, and $E_{ns}(x) = \Pr(\tilde{\delta} \leq x)E(\tilde{\delta} | \tilde{\delta} \leq x)$. This is summarized in the table below:

	$E_s(x)$	$E_{ns}(x)$
Model B^a with $\theta = 0$	$E(\tilde{\delta} \tilde{\delta} \leq x)$	0
Model $B^{a,\delta}$ with $\theta = 0$	x	$\Pr(\tilde{\delta} \leq x)E(\tilde{\delta} \tilde{\delta} \leq x)$

M 's decision whether to take the action depends on $E_s(x) - E_{ns}(x)$. Note that for any x , both $E_s(x)$ and $E_{ns}(x)$ are larger in Model $B^{a,\delta}$ than they are in Model B^a . When L exits, the action is likely to be in a relatively more harmful range in Model $B^{a,\delta}$ than in Model B^a , but when L does not exit, the action may have still been taken in Model $B^{a,\delta}$ but not in Model B^a . Thus, it is not clear which of the models produces a larger difference $E_s(x) - E_{ns}(x)$ at the equilibrium cutoff point.

Figure 1 shows $E_s(x) - E_{ns}(x)$ when $\theta = 0$ in Models B^a and $B^{a,\delta}$ for three different distributions of $\tilde{\delta}$ from the Beta(p,q) family. Note that for the uniform distribution (Beta(1,1)), $E(\tilde{\delta} | \tilde{\delta} \leq x) = x/2$, while $x - \Pr(\tilde{\delta} \leq x)E(\tilde{\delta} | \tilde{\delta} \leq x) = x - x^2/2$, which is always larger. Thus, if $\tilde{\delta}$ has a uniform distribution and $\theta = 0$ (or very small), then L is always more effective in Model $B^{a,\delta}$ than in Model B^a .

For the other distributions in Figure 1, we see that for some values of x , $E_s(x) - E_{ns}(x)$ is larger in Model B^a than it is in Model $B^{a,\delta}$.¹⁶ For these distributions, whether L is more effective in Model $B^{a,\delta}$ or in Model B^a depends on the parameters β , ω_1 and ω_2 .¹⁷ As an example, consider the case where $\beta = 0.5$, $\omega_1 = 0.25$, and $\omega_2 = 0.5$. The table below gives the equilibrium cutoff, x_B , and the agency cost associated with the action, $E(\tilde{\delta}\tilde{a})$, in the two models. In the case of the uniform or Beta(1,1) distribution the agency cost is almost 10% less when

¹⁶ It can be shown that if $\theta = 0$, a sufficient condition for L to be more effective in Model $B^{a,\delta}$ than in Model B^a is that the density of $\tilde{\delta}$ is non-increasing over the support of $\tilde{\delta}$. This condition is satisfied by the uniform distribution but not by the other distributions in Figure 1.

¹⁷ Note that when $\tilde{\delta}$ is distributed Beta(3,2), $E_s(x) - E_{ns}(x)$ is not monotone for Model $B^{a,\delta}$. Thus, it is possible that there are multiple solutions to the equation $\beta - \omega_2 x = \omega_1(E_s(x) - E_{ns}(x))$, which would lead to multiple equilibria. The equilibrium is unique in the case of a uniform distribution, i.e., the Beta(1,1) distribution, as well as for the Beta(2,2) distribution. In the example discussed below the parameters are such that the equilibrium is also unique for the Beta(3,2) distribution, but this would not necessarily be true for other parameters.

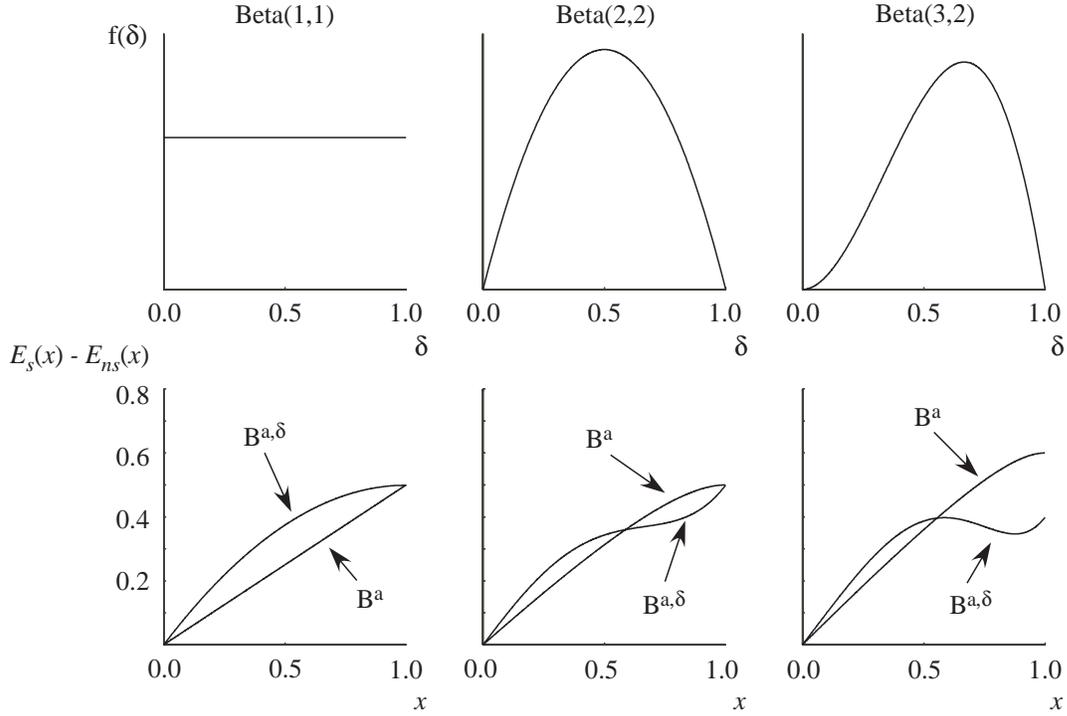


Figure 1: A comparison of $E_s(\cdot) - E_{ns}(\cdot)$ in Model B^a and Model $B^{a,\delta}$ when $\theta = 0$.

L observes both \tilde{a} and $\tilde{\delta}$ than it is when L observes only \tilde{a} . Thus, the additional information produces a gain of approximately 10%. However, in the case of the other distributions, L is less effective when she observes more information; the additional information results in an *increase* in the agency cost over the case where L has less information.

	x_B in B^a	x_B in $B^{a,\delta}$	$E(\tilde{a}\tilde{\delta})$ in B^a	$E(\tilde{a}\tilde{\delta})$ in $B^{a,\delta}$	% Gain in $B^{a,\delta}$
Beta(1,1)	0.800	0.764	0.320	0.292	9.65%
Beta(2,2)	0.776	0.804	0.390	0.412	-5.34%
Beta(3,2)	0.744	0.824	0.371	0.471	-21.09%

When $\theta > 0$, L 's trades in both models are less informative because of the possibility that a sale is due to a liquidity shock. To see how this affects L 's disciplinary impact, we consider an example where $\tilde{\delta}$ is distributed uniformly over $[0, 1]$, $\beta = 0.4$, $\omega_1 = 1$ and $\omega_2 = 0.5$. In this example, if L is not present then M takes the action if and only if $\tilde{\delta} \leq \beta/\omega_2 = 0.8$. Figure 2 shows the equilibrium cutoffs x_B as well as the price impact of exit, measured by $E_s(x_B)$ in the equilibrium of both models for all values of θ . We see that x_B is increasing in θ for both models, which is intuitive. When θ is small, L is more effective in Model $B^{a,\delta}$ than she is in Model B^a ,

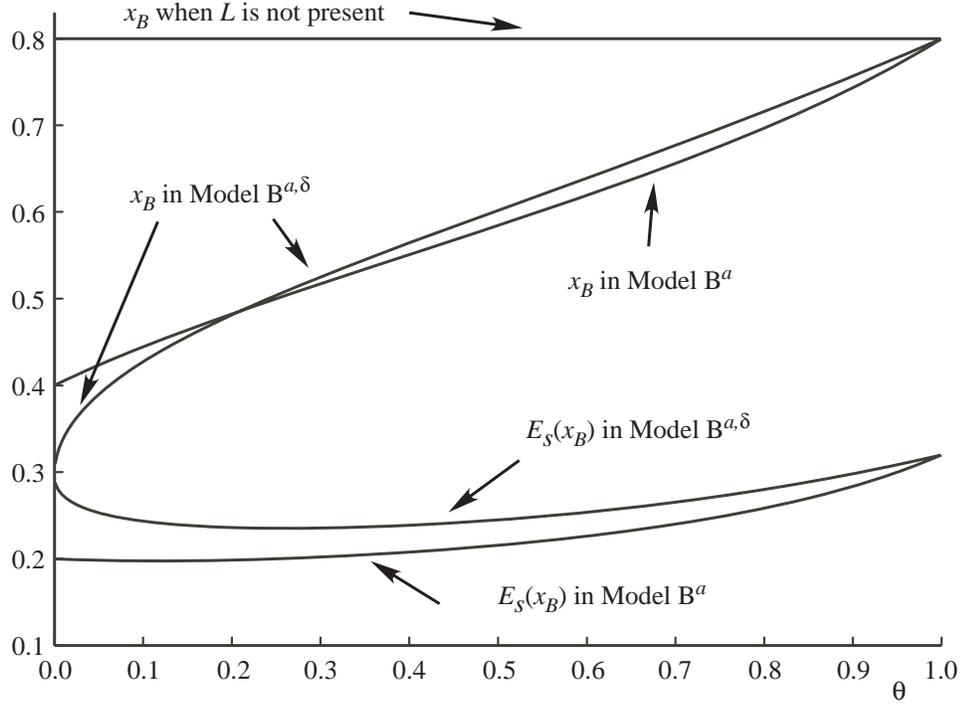


Figure 2: Example of equilibrium in Model B^a and Model $B^{a,\delta}$ for the uniform distribution.

which is consistent with our earlier discussion of the uniform distribution example. However, for values of θ above 0.25, the reverse is true, i.e, L is less effective when she has private information about $\tilde{\delta}$ in addition to \tilde{a} . It turns out this conclusion holds for *every* distribution of $\tilde{\delta}$ and the model's other parameters as θ grows towards 1. This is stated in the following proposition and explained intuitively below.

Proposition 4: *For any given distribution of $\tilde{\delta}$ and parameters β , ω_1 and ω_2 , there exists $\hat{\theta}$ such that if $\theta > \hat{\theta}$, then the large shareholder is more effective in disciplining the manager in the equilibrium of Model B^a than in any equilibrium of Model $B^{a,\delta}$.*

To understand this result note that in both models, as θ becomes large $E_s(x)$ converges to $E(\tilde{\delta}|\tilde{a}) = \Pr(\tilde{\delta} \leq x)E(\tilde{\delta} | \tilde{\delta} \leq x)$, because the probability that L sells grows to 1 and thus a sale becomes less informative about $\tilde{a}\tilde{\delta}$. However, in Model $B^{a,\delta}$, even as θ grows, $E_{ns}(x)$ remains strictly positive, while in Model B^a , $E_{ns}(x)$ is always equal to zero. Intuitively, as θ becomes large, exit carries the same (diminishing) information in both models, but the event in which L does not sell has very different information content in the two models. In Model B^a it guarantees that the action was not taken by M , while in Model $B^{a,\delta}$ it is also consistent with the action being taken but $\tilde{\delta} < E_s(x)$. This means that for θ sufficiently large, $E_s(x) - E_{ns}(x)$ is larger in Model B^a than it is in Model $B^{a,\delta}$ for all positive values of x , which implies that L is more

effective in Model B^a , where she does *not* have the private information about $\tilde{\delta}$.

We now turn to the Model $G^{a,\delta}$. In stark contrast to the previous discussion, we will show that L is *never* more effective in Model $G^{a,\delta}$ than she is in Model G^a , and, moreover, it is possible that the only equilibrium in Model $G^{a,\delta}$ is non-disciplining even though the equilibrium for Model G^a is disciplining for the same parameters. Thus, no additional disciplinary impact ever arises from the additional information of $\tilde{\delta}$ in Model $G^{a,\delta}$ relative to Model G^a , and the additional information may actually make L completely ineffective in disciplining M . We will also see that, unlike Model $B^{a,\delta}$, which possesses at least one and possibly multiple equilibria, Model $G^{a,\delta}$ has at most one equilibrium and it is possible that an equilibrium does not exist for this model. This is summarized in the following result and discussed and illustrated further below.

Proposition 5: *Model $G^{a,\delta}$ has at most one equilibrium. When an equilibrium exists for this model either (i) it is identical to the equilibrium of Model G^a , and $E_s(x_G) \leq x_G$, or (ii) it is non-disciplining and has $x_G = \beta/\omega_2$.*

To obtain some intuition, first observe that for reasons similar to those discussed earlier, in any equilibrium of Model $G^{a,\delta}$, M 's strategy must involve taking the action if and only if $\tilde{\delta} \geq x$ for some cutoff point x . Consider $E_s(x)$, the market's expectation of $\tilde{\delta}$ given a sale by L . If an equilibrium cutoff x_G exists, there are two possibilities: either $E_s(x_G) \leq x_G$ or $E_s(x_G) > x_G$. Suppose first that $E_s(x_G) \leq x_G$. Then we claim that L 's strategy must be the same as her strategy in the equilibrium of Model G^a , namely to sell if M does not take the action and retain her shares (unless subject to a liquidity shock) if M takes the action. This follows because, if $\tilde{\delta} \geq x_G$, then $\tilde{\delta} \geq E_s(x_G)$, and L does not want to sell her shares, while for $\tilde{\delta} < x_G$, L clearly prefers to sell because the action was not taken and $E_s(x_G) > 0$.¹⁸ Thus, in every equilibrium of Model $G^{a,\delta}$ for which $E_s(x_G) \leq x_G$, the equilibrium is the same as the one obtained in Model G^a , and L only uses information about $\tilde{\delta}$ but does not use the additional information about $\tilde{\delta}$. Moreover, this argument establishes that *if in the equilibrium of Model G^a we have $E_s(x_G) \leq x_G$, then this is also an equilibrium in Model $G^{a,\delta}$.*

Now consider the possibility of an equilibrium for Model $G^{a,\delta}$ in which $E_s(x_G) > x_G$. In this case when $\tilde{\delta} = x_G$, L sells her shares whether M takes the action or not, which means that when $\tilde{\delta} = x_G$, the first-period price P_1 is the same whether M takes the action or not. Thus, when $\tilde{\delta} = x_G$, L cannot have an impact on M 's decision. It follows that x_G must be equal to β/ω_2 , which is the cutoff when L is not present. We conclude that the only possible equilibrium of Model $G^{a,\delta}$ in which $E_s(x_G) > x_G$ is a non-disciplining equilibrium with $x_G = \beta/\omega_2$. Note that in this equilibrium L will typically use her private information about $\tilde{\delta}$ to sell her shares when $\beta/\omega_2 < \tilde{\delta} < E_s(\beta/\omega_2)$ even though M takes the action. This, of course, has to be factored into the determination of $E_s(\beta/\omega_2)$ in this model, but it does not have an impact on M 's decision.

It is also possible that no equilibrium exists in Model $G^{a,\delta}$. This occurs when $E_s(x_G)$ in

¹⁸ In the limit case $\theta = 0$, $E_s(x_G) = 0$ and L is indifferent between selling and not selling.

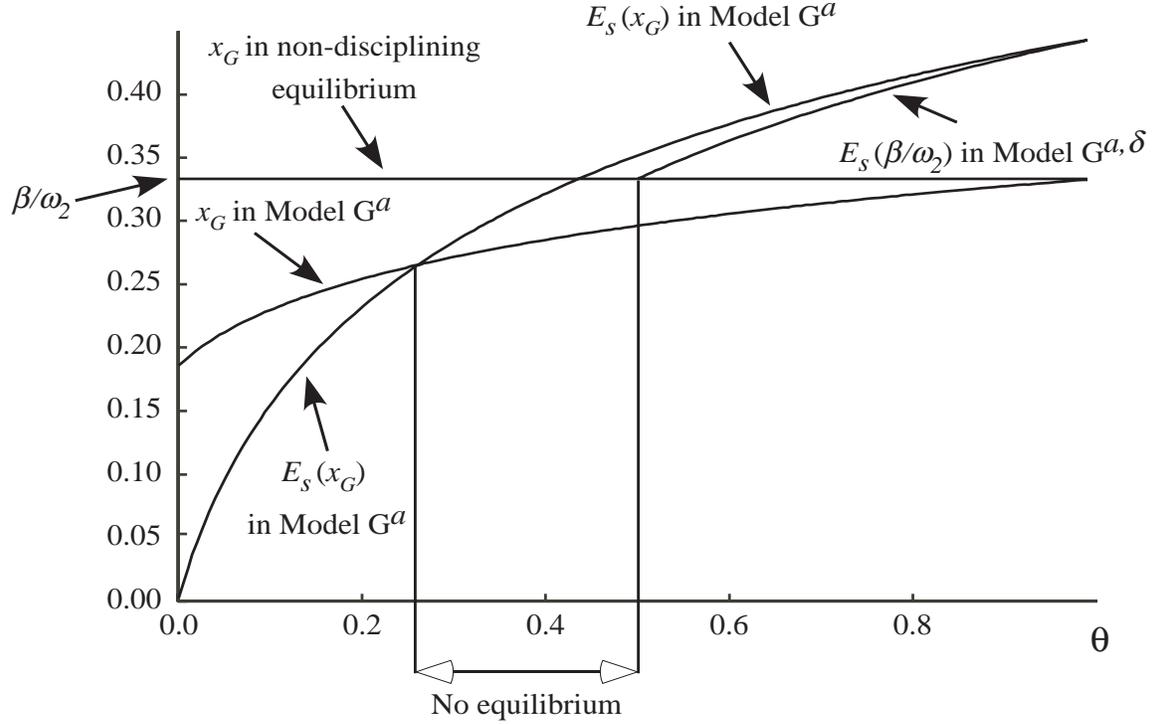


Figure 3: An example of existence and non-existence of equilibrium in Model $G^{a,\delta}$.

Model G^a exceeds x_G and at the same time $x_G = \beta/\omega_2$ is not an equilibrium for Model $G^{a,\delta}$ because $E_s(\beta/\omega_2) < \beta/\omega_2$. Such a case is illustrated in Figure 3. For this example we assume that $\tilde{\delta}$ is uniformly distributed on $[0, 1]$, $\beta = 0.33$, $\omega_1 = 0.25$ and $\omega_2 = 1$. For small θ , the equilibrium in Model $G^{a,\delta}$ is the same as that in Model G^a since for small enough θ , $E_s(x_G) < x_G$ in Model G^a . For large enough θ , the equilibrium in Model $G^{a,\delta}$ is non-disciplining and has $x_G = \beta/\omega_2$, since $E_s(\beta/\omega_2) > \beta/\omega_2 = x_G$, and thus L sells her shares for every $\tilde{\delta} \leq \beta/\omega_2$. Note that in the cases where θ is large and L has no effect on M 's behavior, L would have been effective in disciplining M if she only observed \tilde{a} ; in these cases x_G in Model G^a is strictly less than β/ω_2 . It is when L is more informed and knows both \tilde{a} and $\tilde{\delta}$ that she becomes ineffective.

The figure shows that there is an intermediate region ($0.265 < \theta < 0.5$) where an equilibrium does not exist in this example. In this region $E_s(x_G) > x_G$ in the equilibrium of Model G^a , and so the equilibrium of Model G^a is not an equilibrium for Model $G^{a,\delta}$. At the same time, the only other equilibrium candidate, $x_G = \beta/\omega_2$, produces a contradiction because if M takes the action for $\tilde{\delta} \geq \beta/\omega_2$ and L exits whenever the action is not taken (which is what she does in Model G^a), then $E_s(\beta/\omega_2) < \beta/\omega_2$ and so L does not want to sell. Since there is no equilibrium with $E_s(x_G) > x_G$ and there is no equilibrium with $E_s(x_G) \leq x_G$, an equilibrium does not exist.¹⁹

¹⁹ One might think that this nonexistence problem would disappear if both L and M were permitted to use mixed strategies. This is not the case since L can only be indifferent between selling and not selling when $\tilde{\delta} = E_s(x_G)$, and M can only be indifferent between taking the action and not taking the action when $\tilde{\delta} = x_G$. Since we assume that $\tilde{\delta}$ has a continuous distribution,

In summary, we have seen that if, in addition to observing whether the action is taken, L has private information about the consequences of the action, her disciplinary impact may be enhanced, but in many cases it will actually be weakened substantially relative to the case where she does not have this additional information. In Model $B^{a,\delta}$, L still has disciplinary impact. Exit is a more powerful threat, having a larger price impact because it occurs only for the relatively more harmful consequences, but, since L may retain her shares even if the action is taken, the information content of no exit by L is reduced relative to the model where only \tilde{a} is observed by L . Whether the overall disciplinary impact is enhanced by the additional information is ambiguous in general. However, when the probability of a liquidity shock is high, L 's effectiveness tends to be lower when she has the information about $\tilde{\delta}$. We also saw that additional information about $\tilde{\delta}$ never enhances L 's disciplinary impact in Model $G^{a,\delta}$ relative to Model G^a . At best, the equilibrium of the two models is the same. In other cases, Model $G^{a,\delta}$ only possesses a non-disciplining equilibrium or no equilibrium at all.

5. Can L 's Presence Exacerbate the Agency Problem?

In the models we have analyzed to this point we have shown that L 's presence generally has a disciplinary impact on M and the worst equilibrium in terms of the agency cost is one where L has no impact at all. We have not encountered a dysfunctional equilibrium, one in which L 's impact is negative. This will change below. We will consider an information structure where \tilde{a} is public in period 1, i.e., all investors observe whether M takes the action, and L 's private information consists of the realization of $\tilde{\delta}$. Again, and quite dramatically, our two models will produce very different results. We first show that the equilibrium of Model $B_a^{a,\delta}$ is disciplining; in fact, the agency cost in this model is lower than that in either Model B^a or Model $B^{a,\delta}$. (However, this does not imply that L has a higher impact in this model, because the benchmark case where L is not present is different when \tilde{a} is public than when it is not.) By contrast, we show that in Model $G_a^{a,\delta}$ the equilibrium is dysfunctional, and L 's presence *increases* the agency cost relative to the case where she is not present.

In the models analyzed so far, where \tilde{a} is not observed by investors until period 2, if L is not present, then discipline is only provided by the impact of the action on P_2 , and so the equilibrium of Model B is that M takes the action when $\tilde{\delta} \leq \beta/\omega_2$. Now consider Model B_a , where L is not present and \tilde{a} is public. Since $\omega_2 > 0$, equilibrium must again involve a cutoff x such that M takes the action if and only if $\tilde{\delta} \leq x$. If M is observed taking the action, investors conclude that $\tilde{\delta} \leq x$. Without any additional information, the expected value of $\tilde{a}\tilde{\delta}$ is $E(\tilde{\delta} \mid \tilde{\delta} \leq x)$, and thus $P_1 = \nu - E(\tilde{\delta} \mid \tilde{\delta} \leq x)$. Since $P_1 = \nu$ if M does not take the action, the equilibrium cutoff x_B is determined by

$$\beta - \omega_1 E(\tilde{\delta} \mid \tilde{\delta} \leq x_B) - \omega_2 x_B = 0. \quad (14)$$

using a mixed strategy for any particular realization of $\tilde{\delta}$ will not affect the equilibrium.

Note that this is the same as the equilibrium cutoff in Model B^a when $\theta = 0$, i.e., where L observes \tilde{a} privately and she is never subject to a liquidity shock (see Proposition 1). This is intuitive, since in this case L exits if and only if the action is taken, and so in equilibrium investors know exactly when the action is taken in both Model B^a with $\theta = 0$ and in Model B_a .

Now consider Model $B_a^{\alpha, \delta}$, where \tilde{a} is public and L observes $\tilde{\delta}$ privately. If the action is not taken, then the price in period 1 is ν independent of L 's trade, and L sells only when she is subject to the liquidity shock. If the action is taken, then, as in Model $B^{\alpha, \delta}$, L sells whenever $\tilde{\delta} \geq E_s(x)$, where x is the cutoff value of $\tilde{\delta}$ below which the action is taken. This means that we must have

$$E_s(x) = \frac{\theta \Pr(\tilde{\delta} \leq x) E(\tilde{\delta} \mid \tilde{\delta} \leq x) + (1 - \theta) \Pr(\tilde{\delta} \in [E_s(x), x]) E(\tilde{\delta} \mid \tilde{\delta} \in [E_s(x), x])}{\theta \Pr(\tilde{\delta} \leq x) + (1 - \theta) \Pr(\tilde{\delta} \in [E_s(x), x])}, \quad (15)$$

and

$$E_{ns}(x) = E(\tilde{\delta} \mid \tilde{\delta} < E_s(x)). \quad (16)$$

In equilibrium, M must again be indifferent between taking and not taking the action at the cutoff $\tilde{\delta} = x_B$. This means that any equilibrium cutoff x_B in Model $B_a^{\alpha, \delta}$ must satisfy

$$\beta - \omega_1 E_s(x_B) - \omega_2 x_B = 0. \quad (17)$$

Note that $E_{ns}(x_B)$, which measures the price impact of no sale, does not affect the determination of x_B in this model, because if M takes the action when $\tilde{\delta} = x_B$, then L sells her shares for sure, while if M does not take the action, $P_1 = \nu$ independent of L 's trading. In other words, since M can be sure that $P_1 = \nu$ if he does not take the action, and since L always exits when $\tilde{\delta} = x_B$, the inference investors would make if L retains her shares is irrelevant to M 's decision when $\tilde{\delta} = x_B$.

The next result confirms that there exists a unique equilibrium to Model $B_a^{\alpha, \delta}$ and compares the agency cost associated with the action in this model to that in the equilibria of Models B^a and $B^{\alpha, \delta}$ analyzed in previous sections.

Proposition 6: *There exists a unique equilibrium in Model $B_a^{\alpha, \delta}$ and the equilibrium is disciplining. In equilibrium*

- (i) *the manager takes the action if and only if $\tilde{\delta} \leq x_B$, where x_B is determined by (17);*

- (ii) *the large shareholder sells her shares if the action is taken and $\tilde{\delta} \geq E_s(x_B)$, where $E_s(\cdot)$ is determined by (15);*
- (iii) *the agency cost is lower than the agency cost in the unique equilibrium of Model B^a and is also lower than the agency cost in any equilibrium of Model $B^{a,\delta}$.*

The key to understanding how L 's presence affects M 's behavior is to examine M 's incentives at the cutoff realization $\tilde{\delta} = x$ below which the action is taken. In general, this depends on the difference between the (expected) first-period price, P_1 , when M takes the action and when he does not. While the short-term compensation difference between taking and not taking the action is only a function of L 's trading decisions in Models B^a and $B^{a,\delta}$, this is no longer true when M 's action is publicly observable in period 1. The following table shows the first-period price in the three models when M does and does not take the action. All prices are given as a function of the cutoff point x under the assumption that $\theta = 0$.

Model	P_1 if M Does	P_1 if M	Difference
	Not take Action	Takes Action	
B^a with $\theta = 0$	ν	$\nu - E(\tilde{\delta} \mid \tilde{\delta} \leq x)$	$E(\tilde{\delta} \mid \tilde{\delta} \leq x)$
$B^{a,\delta}$ with $\theta = 0$	$\nu - \Pr(\tilde{\delta} \leq x)E(\tilde{\delta} \mid \tilde{\delta} \leq x)$	$\nu - x$	$x - \Pr(\tilde{\delta} \leq x)E(\tilde{\delta} \mid \tilde{\delta} \leq x)$
$B_a^{a,\delta}$ with $\theta = 0$	ν	$\nu - x$	x

First consider the comparison between Model $B^{a,\delta}$ and Model $B_a^{a,\delta}$. When $\theta = 0$ and $\tilde{\delta}$ is at the cutoff x , if M takes the action, L exits and this reveals perfectly that $\tilde{\delta} = x$ in both models, since this is the only realization of $\tilde{\delta}$ for which L exits. If M does not take the action, however, then in Model $B_a^{a,\delta}$, $P_1 = \nu$ since this will be observable to investors, but in Model $B^{a,\delta}$ investors only observe that L is not selling, which provides no information about $\tilde{\delta}$ and the resulting first-period price is $\nu - E(\tilde{\delta}) = \nu - \Pr(\tilde{\delta} \leq x)E(\tilde{\delta} \mid \tilde{\delta} \leq x)$. In other words, in both of these models M will suffer the “maximal hit” to his compensation when he takes the action at the cutoff point, but the consequences of not taking the action are quite different. In Model $B^{a,\delta}$ investors do not learn anything from L not exiting, since they do not observe $\tilde{\delta}$ and since L only exits when $\tilde{\delta} = x$. However, since $\tilde{\delta}$ is public in Model $B_a^{a,\delta}$, the first-period price will reflect the fact that the action was not taken. This means that the consequences of taking the action are greater in Model $B_a^{a,\delta}$ and thus the agency cost in the equilibrium of Model $B_a^{a,\delta}$ is lower than that in any equilibrium of Model $B^{a,\delta}$.

We now compare Model B^a , where L observes $\tilde{\delta}$ privately and no investor observes $\tilde{\delta}$, and Model $B_a^{a,\delta}$, where $\tilde{\delta}$ is public and L observes $\tilde{\delta}$ privately. In both of these models, although for different reasons, $P_1 = \nu$ if M does not take the action. In Model B^a this is because L retains her shares only if the action is not taken, while in Model $B_a^{a,\delta}$ this is because it is publicly observed that the action was not taken. The difference in discipline comes about because if M takes the action at the cutoff $\tilde{\delta} = x$, then in Model B^a L sells her shares and investors only know that $\tilde{\delta} \leq x$, and thus $P_1 = \nu - E(\tilde{\delta} \mid \tilde{\delta} \leq x)$, while in Model $B_a^{a,\delta}$ investors know from the fact that

L exits that $\tilde{\delta} = x$. It follows that at any possible cutoff point x the difference between P_1 when M does not take the action and when he does is again greater in Model $B_a^{a,\delta}$ than it is in B^a , which implies the equilibrium cutoff point in Model $B_a^{a,\delta}$ is always lower than that in Model B^a , and thus again that the agency cost is smaller.

The above discussion, with appropriate modifications, applies to the model with $\theta > 0$. Model $B_a^{a,\delta}$ produces the best outcome from shareholders' perspectives among the models considered so far, because, as in Model B^a , M is able to obtain the highest compensation when he does not take the action and at the same time, as in Model $B^{a,\delta}$, he suffers the most severe consequences when he does take the action at the cutoff point. Note, however, that this does *not* imply that L has a larger disciplinary *impact* in Model $B_a^{a,\delta}$ than she does in the other models, because the benchmark situation where L is not present is different when \tilde{a} is public information than when \tilde{a} is not public information in period 1. In fact, it can be shown that the L 's impact can be higher or lower in Model $B_a^{a,\delta}$ relative to Models B^a and $B^{a,\delta}$.

Let us now turn to Model $G_a^{a,\delta}$, and again let us start with Model G_a where L is not present and \tilde{a} is public. If M is observed taking the action and L is not present, the expected value of $\tilde{\delta}$ will be $E(\tilde{\delta} \mid \tilde{\delta} \geq x)$, and thus the price in period 1 will be $\nu + E(\tilde{\delta} \mid \tilde{\delta} \geq x)$, where x is the cutoff such that M takes the action if and only if $\tilde{\delta} \geq x$. Since the price if M does not take the action is ν , the equilibrium cutoff x_G is determined by

$$-\beta + \omega_1 E(\tilde{\delta} \mid \tilde{\delta} \geq x_G) + \omega_2 x_G = 0. \quad (18)$$

Note that again this is the same as the equilibrium cutoff point in Model G^a when $\theta = 0$, i.e., where L observes \tilde{a} privately and is never subject to a liquidity shock. This is because a sale by L in Model G^a with $\theta = 0$ communicates perfectly that the action was not taken. Now consider Model $G_a^{a,\delta}$. Since \tilde{a} is publicly observed, if the action is not taken, then again $P_1 = \nu$ independent of whether L sells. If the action is taken for $\tilde{\delta} \geq x$, then L exits when $\tilde{\delta} \leq E_s(x)$. Note that we must have $x \leq E_s(x)$ because if $\tilde{a} = 1$ then it is publicly known that $\tilde{a}\tilde{\delta} \geq x$. This means that

$$E_s(x) = \frac{\theta \Pr(\tilde{\delta} \geq x) E(\tilde{\delta} \mid \tilde{\delta} \geq x) + (1 - \theta) \Pr(\tilde{\delta} \in [x, E_s(x)]) E(\tilde{\delta} \mid \tilde{\delta} \in [x, E_s(x)])}{\theta \Pr(\tilde{\delta} \geq x) + (1 - \theta) \Pr(\tilde{\delta} \in [x, E_s(x)])}. \quad (19)$$

The equilibrium cutoff x_G in model $G_a^{a,\delta}$ must solve

$$-\beta + \omega_1 E_s(x_G) + \omega_2 x_G = 0. \quad (20)$$

Note that, analogous to the case of Model $B_a^{a,\delta}$, if the action is taken for $\tilde{\delta} \geq x$, then $E_{ns}(x) = E(\tilde{\delta} \mid \tilde{\delta} > E_s(x))$, but $E_{ns}(\cdot)$ does not affect the determination of the equilibrium cutoff x_G ,

because when $\tilde{\delta}$ is equal to the cutoff point, which is the lowest possible value of $\tilde{\delta}$ for which the beneficial action is taken, L strictly prefers to sell her shares whether the action is taken or not, and thus $P_1 = \nu + E_s(x_G)$ with probability 1.

The next result states the existence of a unique equilibrium for Model $G_a^{a,\delta}$. Most interestingly, it shows that when \tilde{a} is public information in period 1, having L privately observe $\tilde{\delta}$ and potentially trade on this information actually *reduces* the discipline placed on M . In other words, the agency cost in Model $G_a^{a,\delta}$ is higher than that in Model G_a where L is not present, which means that the equilibrium of Model $G_a^{a,\delta}$ is *dysfunctional*.

Proposition 7: *There exists a unique equilibrium in Model $G_a^{a,\delta}$ and the equilibrium is dysfunctional. In equilibrium*

- (i) *the manager takes the action if and only if $\tilde{\delta} \geq x_G$, where x_G is determined by (20),*
- (ii) *the large shareholder sells her shares if the action is taken and $\tilde{\delta} \leq E_s(x_G)$, where E_s is determined by (19), and*
- (iii) *the agency cost is higher than that in Model G_a , i.e., large shareholder's impact on the agency costs is negative.*

To understand why L 's presence is harmful in this model, consider again M 's incentives at a cutoff point x . When L is not present, if M takes the action, he is rewarded by increase in his first-period compensation equal to $\omega_1 E(\tilde{\delta} \mid \tilde{\delta} \geq x)$, because investors have no information about $\tilde{\delta}$ other than what is revealed by the fact that M has chosen to take the action, which means that $\tilde{\delta} \geq x$. In Model $G_a^{a,\delta}$, L exits when $\tilde{\delta} = x$, since the selling price, $\nu + E_s(x)$, is larger than the value of the firm given that M takes the action, given by $\nu + x$. Exit by L communicates that the realization of $\tilde{\delta}$ is relatively low among the values of $\tilde{\delta}$ for which the action is taken, since L only chooses to sell when $\tilde{\delta} \in [x, E_s(x)]$. (In the extreme case in which $\theta = 0$, L only sells when $\tilde{\delta} = x$, and thus $E_s(x) = x$.) Thus, when $\tilde{\delta} = x$, L 's trading causes the market to revise *downward* its expectation of $\tilde{\delta}$ relative to the expectation based only on M 's willingness to take the action. For any potential cutoff x , this lowers the differential compensation for M between taking the action and not taking the action in Model $G_a^{a,\delta}$ relative to Model G_a . It follows that L 's presence reduces M 's incentives to take the value-enhancing action and shareholders would be better off if L was not present.

Note that this implies that for $\theta = 0$, the agency cost in Model $G_a^{a,\delta}$ is strictly higher than that in Model G^a , because in this case Model G^a is identical to Model G_a ; in both of these models investors know perfectly when the action is taken but nothing more about $\tilde{\delta}$. Since we already observed that Model $G_a^{a,\delta}$ produces the same or a worse outcome than Model G^a , it follows that, for $\theta = 0$, Model G^a (equivalently Model G_a) produces a lower agency cost than either Model $G_a^{a,\delta}$ or Model $G_a^{a,\delta}$. When $\theta > 0$ the comparison between the agency cost in Models $G_a^{a,\delta}$ and G^a is ambiguous since \tilde{a} is publicly observed in Model $G_a^{a,\delta}$, while this information is only communicated with some noise in Model G^a . The lowest agency cost for Model G is obtained

in Model G_a where \tilde{a} is publicly observed and L is not present.

6. The Model with Action-Only Uncertainty

For completeness, we now discuss the one remaining information structure in which L 's trading potentially has an impact on M 's actions. Here all investors are assumed to observe $\tilde{\delta}$ in period 1, while L observes, in addition, whether M has taken the action. In other words, there is no uncertainty about $\tilde{\delta}$, but only L observes \tilde{a} in period 1. As we will show, this is actually an information structure where Model B and Model G behave like mirror images of one another and, for the same parameters, produce the same type of results in terms of L 's impact. This highlights the critical role played by the inferences investors need to make about $\tilde{\delta}$ for some of our results so far.

When investors know the realization of $\tilde{\delta}$, equilibrium will depend on the realized value δ . The prior distribution of $\tilde{\delta}$ will not play any role in determining the equilibrium. However, to the extent that $\tilde{\delta}$ is drawn from a particular distribution, one can still discuss the agency costs averaging over the possible values of δ that investors observe in period 1. The next result characterizes the equilibrium and the agency cost in Models $B_\delta^{a,\delta}$ and $G_\delta^{a,\delta}$. Note that even with $\omega_2 > 0$, equilibrium necessarily involves a mixed strategy for M for a range of values of δ .

Proposition 8: *In each of Model $B_\delta^{a,\delta}$ and Model $G_\delta^{a,\delta}$ there exists a unique equilibrium, which is always disciplining.*

- (i) *In Model $B_\delta^{a,\delta}$ equilibrium is characterized by a function $m_B(\delta)$ such that for a given δ the manager takes the action with probability $m_B(\delta)$ and the large shareholder sells her shares if the manager takes the action. The mixing probability is given by:*

$$m_B(\delta) = \begin{cases} 1, & \text{if } \delta < \frac{\beta}{(1-\theta)\omega_1 + \omega_2} \\ \left(\frac{\theta}{1-\theta}\right) \left(\frac{\beta - \omega_2\delta}{(\omega_1 + \omega_2)\delta - \beta}\right), & \text{if } \frac{\beta}{(1-\theta)\omega_1 + \omega_2} \leq \delta \leq \frac{\beta}{\omega_2} \\ 0, & \text{otherwise.} \end{cases} \quad (21)$$

- (ii) *In Model $G_\delta^{a,\delta}$ equilibrium is characterized by a function $m_G(\delta)$ such that for a given δ the manager takes the action with probability $m_G(\delta)$ and the large shareholder sells her shares if the manager does not take the action. Furthermore, $m_G(\delta) = 1 - m_B(\delta)$ where $m_B(\delta)$ is given in (21).*
- (iii) *Holding fixed $\theta, \beta, \omega_1, \omega_2$ and the distribution of $\tilde{\delta}$, the impact of the large shareholder's presence on the agency cost is equal in Model $B_\delta^{a,\delta}$ and Model $G_\delta^{a,\delta}$.*

Figure 4 shows examples of equilibria in Models $B_\delta^{a,\delta}$ and $G_\delta^{a,\delta}$. We assume that $\beta = 0.4$, $\omega_1 = 0.75$, and $\omega_2 = 0.5$ and that the support of $\tilde{\delta}$ is the unit interval. The top panel shows

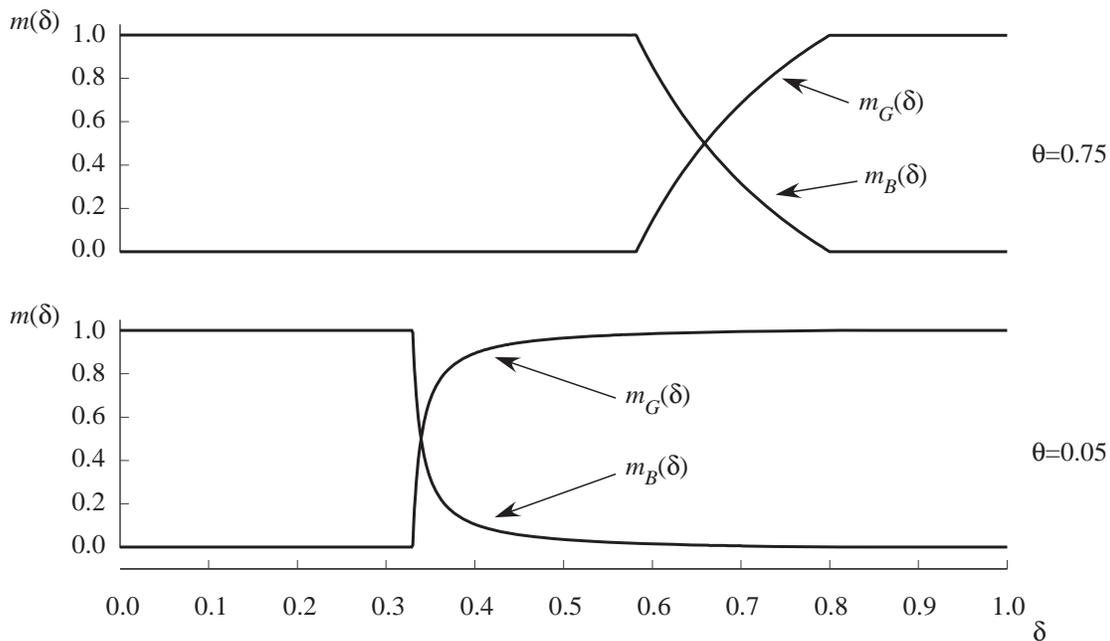


Figure 4: Examples of equilibrium in Models $B_\delta^{a,\delta}$ and $G_\delta^{a,\delta}$. Parameters are $\beta = 0.4$, $\omega_1 = 0.75$, and $\omega_2 = 0.5$. In the top panel $\theta = 0.75$ and in the bottom $\theta = 0.05$.

$m_B(\delta)$ and $m_G(\delta)$ for $\theta = 0.75$, while the bottom shows the same for $\theta = 0.05$. If L were not present, M would only act in the shareholders interest if $\delta > 0.8$. This means that in the absence of L the manager takes the action for all $\delta \leq 0.8$ in Model B_δ and for $\delta \geq 0.8$ in Model G_δ . In Model $B_\delta^{a,\delta}$, L 's presence decreases the probability that M takes the action by $1 - m_B(\delta)$ for a range of δ below 0.8. Analogously, in Model $G_\delta^{a,\delta}$, L 's presence increases the probability that M takes the action by $m_G(\delta)$ for the same range of δ below 0.8. Since $m_G(\delta) = 1 - m_B(\delta)$, the disciplinary impact is the same. We also see that, as in previous models, L is less effective when θ is larger.

Note that this is the only information structure we have studied in which there is complete symmetry between Model B and Model G, as one might expect in a standard principal-agent setting. This suggests that the results in previous sections were driven by the inference investors must make in period 1 about the consequences of the action, $\tilde{\delta}$, based on L 's trading behavior.

7. Discipline when Exit is Costly

So far we have assumed that L incurs no costs in exiting other than the price impact of her sale, which is due to the information revealed by L 's willingness to sell. In this section we assume that exit entails additional costs to L and possibly to other shareholders. For example,

there may be a transactions cost that L must pay when she trades. Alternatively, L 's exit may lower the value of the firm, because exit results in a loss of future benefits to the firm that would be realized through L 's continued presence as a shareholder. Additional exit costs, no matter what the source, generally affect L 's willingness to exit, and this can affect discipline by limiting L 's threat of exit. The two types of costs are different, however, in that a simple transactions cost is borne only by L , whereas a loss in the value of the firm created by L 's exit is borne by all shareholders and enters M 's compensation directly. As we will see, these two types of exit costs will lead to different results.

Below we analyze Model B^a in the presence of each of these two types of exit costs. (To focus on the different issues that arise in the two cases, we analyze each type of exit cost separately.) Relative to the case analyzed in Section 3, we show the following: (i) L 's disciplinary impact is either the same or lower when she must incur a transactions cost to exit; (ii) with either type of exit cost, equilibrium may involve L using a mixed strategy where she exits with a probability smaller than 1 if M takes the action; (iii) L 's impact may be increasing in the probability of a liquidity shock θ ; and (iv) L may be more effective in disciplining M if her exit lowers the value of the firm than in the base case where it does not.

Suppose first that in Model B^a, L must incur a transactions cost of τ whenever she sells her shares. Assume that L still sells whenever she is subject to a liquidity shock and that a liquidity-motivated sale occurs with probability θ . With $\omega_2 > 0$, M again takes the action when $\tilde{\delta} \leq x$ for some x . If L is not subject to a liquidity shock, she will sell only if the difference between $E(\tilde{\delta} \mid \tilde{\delta} \leq x)$, the expectation of $\tilde{\delta}$ given her information, and $E_s(x)$, the price impact of her sale, exceeds the transactions cost τ . We will refer to the difference $E(\tilde{\delta} \mid \tilde{\delta} \leq x) - E_s(x)$ as L 's *information advantage*.

The information advantage L has depends on how much of her information is revealed in equilibrium. If L sells whenever the action is taken, then it is straightforward to show that her information advantage is given by

$$E(\tilde{\delta} \mid \tilde{\delta} \leq x) - E_s(x) = \frac{\theta (1 - \Pr(\tilde{\delta} \leq x)) E(\tilde{\delta} \mid \tilde{\delta} \leq x)}{\theta + (1 - \theta)\Pr(\tilde{\delta} \leq x)}. \quad (22)$$

Note that L only has an information advantage when $\theta > 0$. If $\theta = 0$, then a sale by L reveals that the action was taken with probability 1, and thus $E_s(x) = E(\tilde{\delta} \mid \tilde{\delta} \leq x)$. It follows that when $\theta = 0$, L will not be willing to pay the transactions cost, and the equilibrium must therefore be non-disciplining. When θ is positive, L has an information advantage since she knows whether she is trading for information or liquidity reasons. This information advantage may, however, be insufficient to cover the transactions cost. In such a case there may be no equilibrium in pure strategies. To see this, let $\bar{x} = \beta/\omega_2$ be the cutoff in a non-disciplining equilibrium and let x_B be

the equilibrium cutoff if L exits whenever the action is taken. Assume that

$$\left(1 - \Pr(\tilde{\delta} \leq \bar{x})\right) E(\tilde{\delta} \mid \tilde{\delta} \leq \bar{x}) > \tau > \frac{\theta \left(1 - \Pr(\tilde{\delta} \leq x_B) E(\tilde{\delta} \mid \tilde{\delta} \leq x_B)\right)}{\theta + (1 - \theta) \Pr(\tilde{\delta} \leq x_B)}. \quad (23)$$

The term on the left of (23) is the information advantage L has in a non-disciplining equilibrium. The term on the right of (23) is L 's information advantage in equilibrium assuming she sells whenever M takes the action. If (23) holds, then L strictly prefers to trade on her information when it is assumed she does not trade on her information, and L strictly prefers not to trade on her information when it is assumed that she trades on it whenever she is not subject to a liquidity shock. There is therefore no equilibrium in which L uses a pure strategy.

To analyze mixed-strategy equilibria, suppose that when M takes the action and L is not subject to a liquidity shock, L sells with probability ψ . Then for a given cutoff point x and a mixing probability ψ , $E_s(x)$ and $E_{ns}(x)$ are given by:

$$E_s(x) = \frac{(\theta + (1 - \theta)\psi) \Pr(\tilde{\delta} \leq x) E(\tilde{\delta} \mid \tilde{\delta} \leq x)}{\theta + (1 - \theta)\psi \Pr(\tilde{\delta} \leq x)}; \quad (24)$$

$$E_{ns}(x) = \frac{(1 - \psi) \Pr(\tilde{\delta} \leq x) E(\tilde{\delta} \mid \tilde{\delta} \leq x)}{1 - \psi \Pr(\tilde{\delta} \leq x)}.$$

Note that when $\psi < 1$, investors can no longer conclude from L retaining her shares that the action was not taken. Thus, unlike Model B^a without transactions costs, $E_{ns}(x)$ is not equal to zero. L 's information advantage when she uses a mixing probability ψ is given by

$$\frac{\theta(1 - \Pr(\tilde{\delta} \leq x)) E(\tilde{\delta} \mid \tilde{\delta} \leq x)}{\theta + (1 - \theta)\psi \Pr(\tilde{\delta} \leq x)}. \quad (25)$$

Note that this is decreasing in the mixing probability ψ .

In a mixed-strategy equilibrium, the equilibrium cutoff for M , x_B , and the equilibrium mixing strategy for L , ψ , are such that (i) M is indifferent between taking and not taking the action when $\tilde{\delta} = x_B$ and (ii) L is indifferent between selling and not selling when M takes the action. For the latter to occur, the information advantage given in (25) must equal τ .

We do not offer a general existence result for the model with a transactions cost. However, Figure 5 illustrates some examples, where we assume that $\tilde{\delta}$ has a uniform distribution on $[0, 1]$, $\omega_1 = \omega_2 = 0.4$, and $\beta = 0.2$. Three different values for the transactions cost are considered: $\tau = 0$ (the base case with no transactions cost), $\tau = 0.05$, and $\tau = 0.10$. When $\tau = 0$, the equilibrium is the one we discussed in Section 3. In this case L has an information advantage for

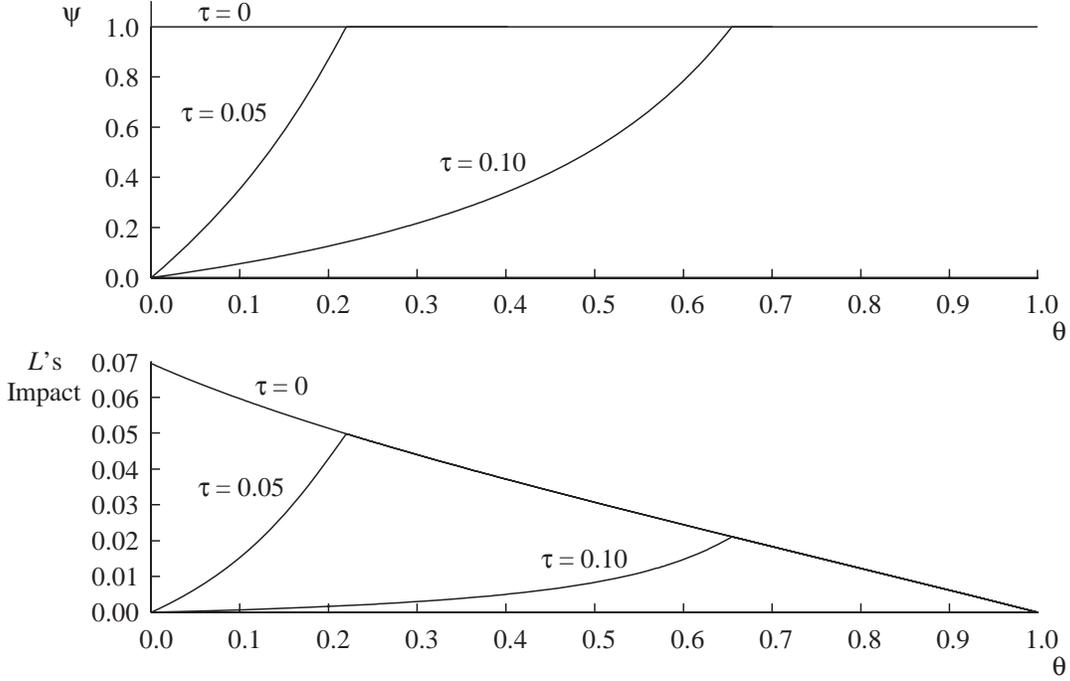


Figure 5: Example of equilibrium in Model B^a with transactions cost τ . Parameters are $\beta = 0.2$ and $\omega_1 = \omega_2 = .4$. Transactions cost levels are set at 0.00, 0.05 and 0.10.

all positive θ , and she never uses a mixed strategy, so $\psi = 1$ for all θ . If $\tau > 0$, then L uses a mixed strategy when θ is sufficiently small. The top panel of Figure 5 shows that the equilibrium mixing probability ψ is an increasing function of θ . That is, the higher the probability of liquidity shock, the higher is the probability that (if she is not subject to a liquidity shock) L will sell when M takes the action. Intuitively, an increase in θ means that L does not need to refrain from trading on her information as much to create an information advantage sufficiently large to cover the transactions cost.

The bottom panel of Figure 5 shows L 's disciplinary impact, i.e., the reduction in the agency cost brought about by L 's presence. Note first that, as can be expected, L 's impact is lower when $\psi < 1$. If θ is such that $\psi = 1$ (L sells whenever the action is taken), then her disciplinary impact is the same no matter what the exact level of the transactions cost. Second, while L 's disciplinary impact is decreasing in the probability of a liquidity shock θ when she uses the pure strategy, this is not true over the range of θ for which L is using a mixed strategy. When the mixed strategy is used, L 's disciplinary impact is actually increasing in θ . This is because when θ increases, L 's information advantage increases and L trades more aggressively on her information, i.e., ψ is larger. This makes the difference between $E_s(x)$ and $E_{ns}(x)$ larger and ultimately increases the disciplinary impact that L 's trading has. Thus, over the range where L is mixing, L 's disciplinary impact is actually enhanced when θ is larger. Once θ is large enough that L sells with probability 1 when the action is taken, an increase in θ only has the effect of lowering the informativeness of a sale, which reduces L 's disciplinary impact.

We now turn to a different type of exit cost, one that arises when L 's continued presence is beneficial to the firm, due to any type of monitoring that L employs in later periods. We assume specifically that if L exits in period 1 for any reason, the value of the firm is reduced by π , independent of both \tilde{a} and $\tilde{\delta}$. (We assume that exit entails no other costs.) From L 's immediate perspective, π is no different than the transactions cost τ , because it is a cost incurred in selling that must be factored into her trading decision. However, unlike the case of a transactions cost, the loss of π associated with L 's exit affects the value of the firm to *all* investors and thus enters M 's compensation directly.

It is not intuitively clear whether L is more or less effective in disciplining M when her exit reduces the value of the firm. On the one hand, as in the case of the simple transactions cost, to the extent that the exit “punishment” will be used less frequently, this would generally tend to reduce L 's disciplinary impact. However, when L does choose to exit, the negative effect this has on M 's compensation is larger when π is positive than when it is zero, which would generally tend to increase L 's disciplinary impact. Below we examine the nature of the tradeoff between these two effects.

Consider Model B^a with an exit cost of π , and suppose that, absent a liquidity shock, L sells with probability ψ if M takes the action. In this case, if M takes the action for $\tilde{\delta} \leq x$, his expected utility when $\tilde{\delta} = \delta$ is

$$\begin{aligned} \beta + \omega_1(\nu - (\theta + (1 - \theta)\psi)(E_s(x) + \pi) - (1 - \theta)(1 - \psi)E_{ns}(x)) \\ + \omega_2(\nu - \delta - (\theta + (1 - \theta)\psi)\pi), \end{aligned} \quad (26)$$

where $E_s(x)$ and $E_{ns}(x)$ are given in (24). If M does not take the action, his expected utility is

$$\omega_1(\nu - \theta(E_s(x) + \pi) - (1 - \theta)E_{ns}(x)) + \omega_2(\nu - \theta\pi). \quad (27)$$

It follows that the equilibrium cutoff point x_B must solve

$$\beta - (1 - \theta)\psi(\omega_1 + \omega_2)\pi - (1 - \theta)\psi\omega_1(E_s(x_B) - E_{ns}(x_B)) - \omega_2x_B = 0. \quad (28)$$

This condition is quite similar to that obtained for Model B^a in Section 3. The main difference is that the private benefit here is effectively lowered by $(1 - \theta)\psi(\omega_1 + \omega_2)\pi$. The reason is that the loss to the value of the firm that accompanies L 's exit is felt directly by M . This works to alleviate the agency problem and potentially to enhance L 's disciplinary impact.

An illustration of equilibria in the model where $\pi > 0$ is provided in Figure 6. We assume again that $\tilde{\delta}$ is uniformly distributed on $[0, 1]$, $\beta = 0.2$, $\omega_1 = \omega_2 = 0.4$. Note that for relatively small values of π , the middle panel shows that $\psi = 1$, i.e., L sells with probability 1 if the action is taken. The range of π values where this is true is larger when θ is relatively large, because

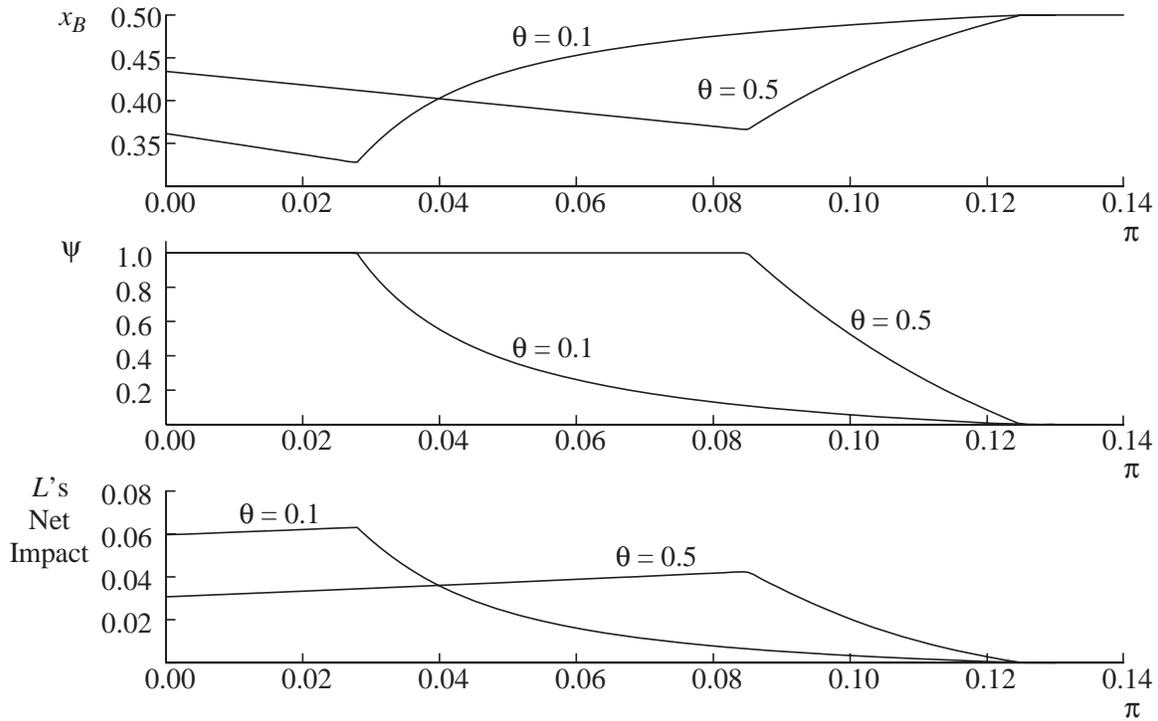


Figure 6: Example of equilibrium in Model B^a where L 's exit leads the value of the firm to be reduced by π . The example assumes that $\tilde{\delta}$ is distributed uniform on $[0, 2]$, $\beta = 0.2$, and $\omega_1 = \omega_2 = .4$.

L 's information advantage is increasing in θ . For this range we see in the top panel that the equilibrium cutoff point x_B is decreasing as π increases, and this means that over this range L 's disciplinary impact is greater the larger is the continuation value π . When π is large enough relative to θ so that L uses a mixed strategy in equilibrium, then L becomes less effective in disciplining M when π increases, and the cutoff point x_B increases.

Note that the highest possible information advantage obtains when the equilibrium is non-disciplining, i.e., when $x_B = \beta/\omega_2 = 0.5$ (or, equivalently, when $\theta = 1$), and this information advantage is equal to 0.125. The figure shows that the probability that L trades on her information vanishes when the continuation value π exceeds 0.125. For such large continuation values, exit is so costly that it is never done in equilibrium and thus the equilibrium becomes non-disciplining. Note also that if π is relatively low (below about 0.04), then L is more effective when $\theta = 0.1$ than when $\theta = 0.5$ (i.e., when the probability of a liquidity sale is higher), but the reverse is true when π is relatively large, as long as discipline is still possible, i.e., for $\pi < 0.125$. To see this, note that when π increases, L must at some point start mixing, which reduces her effectiveness. This occurs for lower values of π when θ is low because, other things equal, the information advantage increases in θ . Thus, for intermediate values of π , L 's disciplinary impact is increasing

in θ .

The bottom panel of Figure 6 plots L 's “net impact,” which is defined as the decrease in agency costs associated with the action (in period 1) minus the expected exit cost associated with L 's informed trading in equilibrium, which is given by $(1 - \theta)\psi\Pr(\tilde{\delta} \leq x_B)\pi$. If the net impact is positive, then shareholders are better off when L disciplines M through exit in period 1 even though exit leads to a reduction in the value of the firm. In this example L 's net impact is always non-negative, and it is increasing in π over range where L is not mixing (and where x_B is decreasing). For other distributions than the uniform, however, it can be shown that L 's net impact can actually be negative. In those cases the loss of π that occurs when L exits is larger than the savings in agency cost brought about by the discipline she exerts in period 1.

In summary, we have seen that exit costs may not eliminate, and in some cases can even enhance, L 's ability to discipline through exit. We have also seen that with exit costs, an increase in the probability of a liquidity shock θ can enhance L 's disciplinary impact.

8. On Extensions with Additional Uncertainty

One of the assumptions made throughout our analysis is that β , the private benefit or cost, is constant and its value is common knowledge across all agents. If β is random and known only to M , then the inference of $\tilde{\alpha}\tilde{\delta}$ by L and the other investors is more complicated, because M 's decision clearly depends on the value of β and not just on $\tilde{\delta}$. One special case that is covered by our analysis so far is one where $\tilde{\beta} = \gamma_0 + \gamma_1\tilde{\delta}$ with $\gamma_0 > 0$ and $\gamma_1 < \omega_2$. For example, consider Model B^a and recall that when $\tilde{\delta} = \delta$, M takes the action in Model B^a when $\beta - (1 - \theta)\omega_1(E_s - E_{ns}) - \omega_2\delta \geq 0$. If $\beta = \gamma_0 + \gamma_1\delta$, this becomes

$$\gamma_0 - (1 - \theta)\omega_1(E_s - E_{ns}) - (\omega_2 - \gamma_1)\delta \geq 0. \quad (29)$$

Similarly, in Model G^a the condition is

$$-\gamma_0 + (1 - \theta)\omega_1(E_{ns} - E_s) + (\omega_2 - \gamma_1)\delta \geq 0. \quad (30)$$

Conditions (29) and (30) are the same as those in Section 3 for the case of fixed β , with γ_0 playing the role of the “fixed” value of β of Section 3, and $\omega_2 - \gamma_1$ playing the role of the coefficient ω_2 in Section 3. Since our analysis in Section 3 as well as in the rest of the paper relied on the assumption that $\omega_2 > 0$, the analysis applies fully to the model with $\tilde{\beta} = \gamma_0 + \gamma_1\tilde{\delta}$ if have $\gamma_1 < \omega_2$.²⁰

²⁰ If $\gamma_1 > \omega_2$, then some of our results would be “switched,” because in Model B the manager will now take the action for all realizations of $\tilde{\delta}$ *above* a cutoff, and in Model G he will take the action for realizations *below* a cutoff value. The general observations, however, e.g., that additional information may lead to lower disciplinary impact and that L 's presence may exacerbate the agency problem, will not change.

We have also assumed implicitly throughout our analysis that the nature of the agency problem is common knowledge, i.e., that investors know whether Model B or Model G captures the situation at hand. We now consider the possibility that investors are uncertain about whether Model B or Model G is appropriate, i.e., whether they would like to encourage or discourage M to take the action. Specifically, assume that with probability α the agency problem is that of Model B, with a (known) private benefit denoted by β_B a loss of value to the firm denoted by $\tilde{\delta}_B$, and with probability $1 - \alpha$ the agency problem is described by Model G, with a (known) private benefit denoted by β_G a loss of value to the firm denoted by $\tilde{\delta}_G$. Suppose that L can observe whether M has taken the action and that she also knows whether the action it is good or bad. However, as in the models of Section 3, no investor has information about the realization of $\tilde{\delta}_B$ or $\tilde{\delta}_G$. Then, under the maintained assumption that L can only exit or retain her shares, it can be shown that L can only provide discipline for one of the two types of agency problems, but not both.

To obtain some intuition, denote by \tilde{a}_B the random variable indicating that M has taken a “bad” action (i.e., that the appropriate model is Model B and M has taken the action) and \tilde{a}_G to be the random variable indicating that M has taken a “good” action. Define E_s and E_{ns} be the expectation of $\tilde{a}_G\tilde{\delta}_G - \tilde{a}_B\tilde{\delta}_B$ conditional on L selling or not selling respectively. It is easy to see that if $E_s > 0$ in equilibrium, then L will exit in all cases except when M takes a good action. This implies that L provides no discipline when the action is bad, since she exits whether the bad action is taken or not. Similarly, if the equilibrium value of E_s is negative, then L will exit only when a bad action is taken. In this case L provides no discipline when the action is good, since she retains her shares no matter what M does.²¹

Interestingly, it can be shown that “model uncertainty” of the type discussed above can enhance the disciplinary impact of L in the following sense. Define $x_B(\alpha)$ to be the cutoff in the case where the agency problem is described by Model B as a function of the probability that this is the model, α . In the equilibria where $E_s < 0$ and L provides discipline only for Model B, $x_B(\alpha)$ can be *increasing* in α , especially for α near 1. The same can occur for $x_G(\alpha)$. When $E_s > 0$ and M provides discipline for Model G, $x_G(\alpha)$ can be decreasing in α , especially for α near zero. Both of these effects occur because the possibility that the agency problem is of the other type can increase $E_s - E_{ns}$ relative to the case where it is common knowledge what the agency problem is.²²

9. Concluding Remarks

We have shown that there is a scope for a large shareholder to affect managerial decisions

²¹ It is possible that there are two equilibria, one in which $E_s > 0$ and one in which $E_s < 0$.

²² It should also be noted that the conclusion that L can only discipline in of the two agency problems is due to L having only two possible trades, namely exit or retain her shares. If L had a third possible trade, e.g., to increase her stake by buying more shares, then it would in principle be possible for her to provide discipline in both the Model B and Model G. A full examination of this is beyond the scope of this paper.

simply by her ability to employ the “Wall Street walk” and sell her shares whenever it is optimal to do so based on her information. No commitment on the part of the large shareholder, whom we call “ L ,” is required for this mechanism. In many of the settings that we have considered, L ’s threat of exit can be quite effective in alleviating the agency problem between managers and shareholders. We have found, however, that while the L ’s effectiveness in disciplining the manager, whom we call “ M ,” is based on her having private information about M ’s actions, L is not always more effective when she has more private information. In fact, we have shown that L ’s effectiveness could be reduced if she has more private information. Even worse, we have shown that in some cases L ’s presence can actually exacerbate the agency problem. This occurs when the trading strategy of L interferes with the discipline that investors are able to exert based on public information.

In our analysis we have assumed that L acquires her private information costlessly. If information is costly, however, then in our model L would only acquire information that would lead to a reduction in agency costs (i.e., to what we called a disciplining equilibrium). This follows because L ’s *ex ante* expected trading profits in our model are zero. That is, before L observes her private information and before it is known whether she is subject to a liquidity shock, the expected trading profits L obtains when she trades on private information is just offset by the expected losses she suffers when she is subject to a liquidity shock. Thus, the value of any private information acquired by L can only be due to the increase in the value of her initial shares brought about by her possible disciplinary impact. This implies that situations where L becomes less effective by acquiring more information or where L ’s trading on information actually exacerbates the agency problem (leading to what we called a dysfunctional equilibrium) would not arise if private information is costly.

Our model also assumes that L ’s trade has an immediate and direct effect on the price, i.e., that she does not trade anonymously, and that M ’s compensation depends on the price at which L trades. Suppose instead that L trades anonymously in a “noisy” market, i.e., one that includes others potentially subject to liquidity shocks. Then two seemingly conflicting effects emerge. First, if L is able to trade anonymously in a more liquid market, then the direct impact of her trade on the price would generally be lower than when the price is based on L ’s trade alone. Other things equal, this would seem to reduce L ’s ability to discipline M , since her trade would have less impact on M ’s compensation. Second, if L is able to trade anonymously in a more liquid market, then her information advantage is larger and, in fact, her *ex ante* expected trading profits are positive and not zero as it is in the case where L ’s trade is directly observed. In general, this would increase L ’s willingness to gather information and trade on it, and this can potentially increase her disciplinary impact.

In fact, if M ’s compensation is sensitive to prices beyond the price at the time L makes her trading decision (but before all her information becomes public), then L ’s disciplinary impact would not necessarily be reduced even if her trade has a lower price impact due to the presence of other liquidity traders. This would occur, for example, if L ’s trade (and possibly the motive

for her trade) becomes public, e.g., as a result of trade disclosure. In this case there will be further price impact at the point where information about L 's trade becomes public, which would potentially affect M 's compensation. If, as observed above, trading in a more liquid market would lead L to collect private information that enhances her disciplinary impact, then we would expect her disciplinary impact overall to be larger.

To see how disciplining through the exit strategy would work in this case suppose, for example, that M 's short-term compensation is based on the price prevailing at the end of the calendar year. Suppose L exits anonymously in October, and then in November it becomes known or is disclosed that she has exited. Then the price in December would fully reflect the fact that L has exited and this would affect M 's compensation. At the same time, L is able to benefit from her information advantage by trading in October, allowing her to recover the cost of information or any other costs exit entails. Of course, when L trades anonymously she generally imposes a trading cost on other investors who may be subject to a liquidity shock. The additional cost on other investors should also be taken into account in a full analysis of this case.

Our model is consistent with a number of activities that large shareholders are known to engage in and which may be more common than many overt forms of shareholder activism. These include the 'targeting' of firms and various forms of behind-the-scene communication and negotiation between large shareholders and managers, often referred to as 'jawboning.' Such activities have been reported to be successful in affecting managerial decisions, for example, in Carleton, Nelson and Weisbach (1998). Note that in our model there is no need for the manager and the large shareholder to communicate. However, the model implicitly assumes that the manager is aware of the presence of the large shareholder and understands the price impact of the large shareholder's trade, which means that he knows what type of information the large shareholder has. If the manager is not fully aware of the large shareholder's presence or the type of information she possesses, then achieving the disciplinary impact may require that the large shareholder communicate with the manager prior to the "onset" of our model. The threat to exit that may be implicit or explicit when the large shareholder communicates with the manager is indeed likely to be more credible than threats to engage in proxy fights or to submit shareholder proposals, both of which are costly and often unlikely to succeed.

In summary, our analysis suggests that the ability of large shareholders to vote with their feet and exit via the "Wall Street Walk" does not necessarily weaken corporate governance. If managers care about the market value of their firm, then the impact of an informed large shareholder on the price has the potential to align managerial decisions with shareholders' preferences and reduce agency costs. However, while there is this potential for the large shareholder to have a disciplinary impact, we have shown that the large shareholder's ability to have such an impact, as well as the degree to which she is effective in disciplining the manager through potential exit, depend critically on the nature of the agency problem as well as on the information structure.

Appendix

Proof of Proposition 1: Let $A(x) = \int_x^{\bar{\delta}} \delta dF(\delta)$, $B(x) = \int_0^x \delta dF(\delta)$, and $C(x) = \int_x^{\bar{\delta}} dF(\delta)$. Consider first Model B. If $x_B < \bar{\delta}$ is an equilibrium cutoff, then $\beta - (1 - \theta)\omega_1 (E_s^B(x_B) - E_{ns}^B(x_B)) - \omega_2 x_B = 0$, where

$$E_s^B(x) = \frac{B(x)}{\theta + (1 - \theta)(1 - C(x))}, \quad E_{ns}^B(x) = 0. \quad (A1)$$

The equilibrium will be unique if $E_s^B(x) - E_{ns}^B(x)$ is nondecreasing in x . We have

$$E_s^B(x) - E_{ns}^B(x) = \frac{B(x)}{1 - C(x)} \left(\frac{1 - C(x)}{1 - (1 - \theta)C(x)} \right). \quad (A2)$$

The first part of the product on the right hand side of (A2), namely $B(x)/(1 - C(x))$, is equal to $E(\tilde{\delta} \mid \delta \leq x)$ which is nondecreasing in x . The second part is nondecreasing in x since $C'(x) \leq 0$ and $0 < \theta < 1$. It is easy to see that $x_B < \beta/\omega_2$, which means that the equilibrium must be disciplining.

Now consider Model G^a. If $x_G < \bar{\delta}$ is an equilibrium cutoff, then

$$-\beta + (1 - \theta)\omega_1 (E_{ns}^G(x_G) - E_s^G(x_G)) + \omega_2 x_G = 0, \quad (A3)$$

where

$$E_s^G(x) = \frac{\theta A(x)}{\theta + (1 - \theta)(1 - C(x))}, \quad E_{ns}^G(x) = \frac{A(x)}{C(x)}. \quad (A4)$$

In this case the equilibrium will be unique if $E_{ns}^G(x) - E_s^G(x)$ is nondecreasing in x . We have

$$E_{ns}^G(x) - E_s^G(x) = \frac{A(x)}{C(x)} \left(\frac{1 - C(x)}{1 - (1 - \theta)C(x)} \right). \quad (A5)$$

The first part of the product on the right hand side of (A5), namely $A(x)/C(x)$, is equal to $E(\tilde{\delta} \mid \tilde{\delta} \geq x)$, which is nondecreasing in x . The second part is the same as the second part in (A2) and is nondecreasing in x . To show that $x_G < x_B$ whenever $x_B < \bar{\delta}$, it is sufficient to show that for all x in the support of $\tilde{\delta}$, $E_{ns}^G(x) - E_s^G(x) > E_s^B(x_B) - E_{ns}^B(x_B)$, or

$$\frac{A(x)}{C(x)} \left(\frac{1 - C(x)}{1 - (1 - \theta)C(x)} \right) > \frac{B(x)}{1 - C(x)} \left(\frac{1 - C(x)}{1 - (1 - \theta)C(x)} \right). \quad (A6)$$

This follows immediately since $A(x)/C(x) = E(\tilde{\delta} \mid \tilde{\delta} \geq x)$, which is clearly greater than $B(x)/(1 - C(x)) = E(\tilde{\delta} \mid \tilde{\delta} \leq x)$.

Proof of Proposition 2: Let p be the probability that M takes the action and let e be the expectation of $\tilde{\delta}$ conditional on M taking the action. Consider first Model B^a. Since $E_{ns}^B = 0$, the equilibrium condition for a disciplining equilibrium is $\beta - \omega_1 E_s^B = -\theta \omega_1 E_s^B$ or $(1 - \theta)E_s^B = \beta/\omega_1$, where $E_s^B = pe/(\theta + (1 - \theta)p)$. If we let $c = \beta/\omega_1$, the equilibrium condition becomes $(1 - \theta)pe/(\theta + (1 - \theta)p) = c$ or $p(e - c) = \theta c/(1 - \theta)$. Thus any strategy for M that results in a p and an e satisfying this relation is an equilibrium. In general there are many equilibria since there are many strategies for M that lead to p 's and e 's that satisfy the equilibrium condition. It will be useful to write the equilibrium condition in two ways:

$$pe = \frac{\theta c}{1 - \theta} + cp, \quad (\text{A7})$$

and

$$p = \frac{\theta c}{(1 - \theta)(e - c)}. \quad (\text{A8})$$

From the equilibrium condition given in (A7) we can conclude that pe , i.e., the expected loss in Model B^a, is increasing in p , the probability that M takes the action, since c is fixed and positive. From the equilibrium condition given in (A8), we can conclude p is decreasing in e , the expectation of $\tilde{\delta}$ given that M takes the action. Thus the expected loss is the lowest in that equilibrium that has the *highest* e .

Now consider Model G^a. The equilibrium condition for a disciplining equilibrium is $-\beta + \theta \omega_1 E_s^G = \theta \omega_1 E_s^G + (1 - \theta) \omega_1 E_{ns}^G$, which is equivalent to $(1 - \theta)(E_{ns}^G - E_s^G) = \beta/\omega_1$, where $E_s^G = \theta pe/(\theta + (1 - \theta)(1 - p))$ and $E_{ns}^G = e$. Again, letting $c = \beta/\omega_1$, we find that the equilibrium condition is $(1 - p)(e - c) = \theta c/(1 - \theta)$, or

$$p = \frac{1}{1 - \theta} \left(1 - \frac{\theta e}{e - c} \right). \quad (\text{A9})$$

Once again any strategy for M that results in a p and an e satisfying this relation is an equilibrium. Now in model G^a we want to *maximize* pe , which is the expected gain. It is clear from (A9) that p is increasing in e , which means that to maximize pe we want the equilibrium that has the highest e . Thus, in both Model B^a and Model G^a the best equilibrium is the one with the highest e .

We now show that for both Model B^a and Model G^a the best disciplining equilibrium (when such an equilibrium exists) is achieved when M follows an “upper” strategy, i.e., one in which

M takes the action if and only if $\tilde{\delta} > x$ for some x . First consider Model B^a where equilibrium requires $p(e - c) = \theta c / (1 - \theta)$. If M follows an upper strategy, then

$$p(e - c) = \int_x^{\tilde{\delta}} (\delta - c) dF(\delta). \quad (A10)$$

Note that $\int_x^{\tilde{\delta}} (\delta - c) dF(\delta)$ is maximized at $x = c$ and is decreasing for all $x > c$. Moreover the maximum attained at $x = c$ is the maximum attained across all strategies. Thus if a disciplining equilibrium exists, there exists an $x^\circ > c$ such that

$$p(e - c) = \int_{x^\circ}^{\tilde{\delta}} (\delta - c) dF(\delta) = \frac{\theta c}{1 - \theta}. \quad (A11)$$

Moreover, this upper strategy equilibrium will have the highest expectation of $\tilde{\delta}$ conditional on M taking the action, i.e. the highest e , and is therefore the best. (Note that there cannot be another upper strategy equilibrium with an x greater than x° since $\int_x^{\tilde{\delta}} (\delta - c) dF(\delta)$ is decreasing in x for $x > c$.) Now consider Model G^a . The equilibrium condition is $(1 - p)(e - c) = \theta c / (1 - \theta)$. Assume that M follows an upper strategy equilibrium in which he takes the action if and only if $\tilde{\delta} > x$ and let $p(x)$ and $e(x)$ be the values of p and e for this equilibrium as functions of x . Let x_c be such that $e(x_c) = c$. Such an x will exist and is unique as long as $c \leq \bar{\delta}$. (If $c > \bar{\delta}$ there is no disciplining equilibrium.) Now $(1 - p(x))(e(x) - c)$ is increasing in x for all $x > x_c$ and it attains its maximum $(\bar{\delta} - c)$ as x approaches $\bar{\delta}$. Moreover, this maximum is the maximum across all strategies M might follow. This means that if there is a disciplining equilibrium, then there is a unique $x^\circ > x_c$ such that $(1 - p(x^\circ))(e(x^\circ) - c) = \theta c / (1 - \theta)$. Since $e(x^\circ)$ is the highest e possible, this is the optimal strategy in Model G^a .

Proof of Proposition 3: Let $g(x) = \beta - (1 - \theta)\omega_1(E_s(x) - E_{n_s}(x)) - \omega_2 x$, where $E_s(x)$ and $E_{n_s}(x)$ are defined by (11) and (12). Since we assume that $\tilde{\delta}$ is continuously distributed, it follows that $E_s(x)$ and $E_{n_s}(x)$ are continuous functions of x . Moreover for all $x > 0$, $E_s(x) - E_{n_s}(x) > 0$. Since $g(0) > 0$ and $g(\beta/\omega_2) < 0$, there must be at least one $x_B \in (0, \beta/\omega_2)$ such that $g(x_B) = 0$ and this is an equilibrium cutoff for Model $B^{a,\delta}$. For any given value of x , $E_s(x)$ solves

$$\theta \int_0^x \delta dF(\delta) - \theta E_s(x) + (1 - \theta) \int_{E_s(x)}^x (\delta - E_s(x)) dF(\delta) = 0. \quad (A12)$$

As $\theta \rightarrow 0$, it is clear that $E_s(x) \rightarrow x$. Since L sells on the interval $[E_s(x), x]$, the probability of L selling vanishes as $\theta \rightarrow 0$.

Proof of Proposition 4: Let $x^* = \beta/\omega_2 < \bar{\delta}$, $x_1(\theta)$ be the highest type manager that takes the action in Model B^a for a given θ , and $x_2(\theta)$ be the same for Model $B^{a,\delta}$. We want to show that

for θ sufficiently close to 1, the gain produced by L in Model B^a is greater than that produced in Model $B^{a,\delta}$. The gain is related to the size of the interval of $\tilde{\delta}$ realizations that refrain from taking the action due to L 's presence. For Model B^a this is

$$x^* - x_1(\theta) = \frac{\beta}{w_2} - \frac{\beta - (1 - \theta)w_1 Q(x_1(\theta), \theta)}{w_2} = \frac{(1 - \theta)w_1 Q(x_1(\theta), \theta)}{w_2}, \quad (A13)$$

where

$$Q(x_1(\theta), \theta) = \frac{\int_0^{x_1(\theta)} \delta \, dF(\delta)}{\theta + (1 - \theta) \int_0^{x_1(\theta)} dF(\delta)}. \quad (A14)$$

Note that

$$\lim_{\theta \rightarrow 1} Q(x_1(\theta), \theta) = \int_0^{x^*} \delta \, dF(\delta). \quad (A15)$$

For Model $B^{a,\delta}$ this is

$$x^* - x_2(\theta) = \frac{\beta}{w_2} - \frac{\beta - (1 - \theta)w_1 R(x_2(\theta), \theta)}{w_2} = \frac{(1 - \theta)w_1 R(x_2(\theta), \theta)}{w_2}, \quad (A16)$$

where

$$R(x_2(\theta), \theta) = \frac{\theta \int_0^{x_2(\theta)} \delta \, dF(\delta) + (1 - \theta) \int_{y(\theta)}^{x_2(\theta)} \delta \, dF(\delta)}{\theta + (1 - \theta) \int_{y(\theta)}^{x_2(\theta)} dF(\delta)} - \frac{\int_0^{y(\theta)} \delta \, dF(\delta)}{1 - \int_{y(\theta)}^{x_2(\theta)} dF(\delta)}, \quad (A17)$$

and where $y(\theta)$ solves

$$y(\theta) = \frac{\theta \int_0^{x_2(\theta)} \delta \, dF(\delta) + (1 - \theta) \int_{y(\theta)}^{x_2(\theta)} \delta \, dF(\delta)}{\theta + (1 - \theta) \int_{y(\theta)}^{x_2(\theta)} dF(\delta)}. \quad (A18)$$

Note that

$$\lim_{\theta \rightarrow 1} R(x_2(\theta), \theta) = \int_0^{x^*} \delta \, dF(\delta) - \frac{\int_0^{y^*} \delta \, dF(\delta)}{1 - \int_{y^*}^{x^*} dF(\delta)}, \quad (A19)$$

where $y^* = \int_0^{x^*} \delta \, dF(\delta)$. We now take the limit of $(x^* - x_1(\theta))/(x^* - x_2(\theta))$ as $\theta \rightarrow 1$. This is

$$\lim_{\theta \rightarrow 1} \frac{x^* - x_1(\theta)}{x^* - x_2(\theta)} = \lim_{\theta \rightarrow 1} \frac{Q(x_1(\theta), \theta)}{R(x_2(\theta), \theta)} = \frac{\int_0^{x^*} \delta \, dF(\delta)}{\int_0^{x^*} \delta \, dF(\delta) - \frac{\int_0^{y^*} \delta \, dF(\delta)}{1 - \int_{y^*}^{x^*} dF(\delta)}}, \quad (\text{A20})$$

which is strictly greater than 1 since $y^* > 0$. This means that for θ sufficiently close to 1, the cutoff below which the action is taken in Model B^a is strictly smaller than that in Model B^{a,δ}.

Proof of Proposition 5: If there are two equilibria to Model G^{a,δ}, one must be the equilibrium of Model G^a and the other must be non-disciplining. Let x be the cutoff in the equilibrium of Model G^a and let z be the cutoff in the non-disciplining equilibrium, i.e., $\beta/\omega 2$. Then, for the equilibrium of Model G^a to be an equilibrium of Model G^{a,δ} we must have

$$E_s = \frac{\theta \int_x^{\bar{\delta}} \delta \, dF(\delta)}{\theta + (1 - \theta) \int_x^{\bar{\delta}} dF(\delta)} \leq x < z. \quad (\text{A21})$$

The second inequality must be strict if there are two distinct equilibria. Now consider the condition for the non-disciplining equilibrium. Let $y = E_s$ in the non-disciplining equilibrium. We have

$$y = \frac{\theta \int_z^{\bar{\delta}} \delta \, dF(\delta) + (1 - \theta) \int_z^y \delta \, dF(\delta)}{\theta + (1 - \theta) \int_0^y dF(\delta)}. \quad (\text{A22})$$

Now let θ_m be the minimum value of θ that is consistent with a non-disciplining equilibrium. At this value $y = z$ and we have

$$z = \frac{\theta_m \int_z^{\bar{\delta}} \delta \, dF(\delta)}{\theta_m + (1 - \theta_m) \int_0^z dF(\delta)}. \quad (\text{A23})$$

This means that for a non-disciplining equilibrium to exist, θ must be such that

$$\theta \geq \theta_m = \frac{z \int_0^z dF(\delta)}{\int_z^{\bar{\delta}} \delta \, dF(\delta) + z \int_0^z dF(\delta) - z}. \quad (\text{A24})$$

Now, since $z > x$, it follows that

$$\frac{\theta \int_z^{\bar{\delta}} \delta \, dF(\delta)}{\theta + (1 - \theta) \int_0^z dF(\delta)} < \frac{\theta \int_x^{\bar{\delta}} \delta \, dF(\delta)}{\theta + (1 - \theta) \int_0^x dF(\delta)}. \quad (\text{A25})$$

Using the condition for equilibrium in Model G^a to be an equilibrium for Model $G^{a,\delta}$, i.e., (A21), we have

$$\frac{\theta \int_z^{\bar{\delta}} \delta \, dF(\delta)}{\theta + (1 - \theta) \int_0^z dF(\delta)} < \frac{\theta \int_x^{\bar{\delta}} \delta \, dF(\delta)}{\theta + (1 - \theta) \int_0^x dF(\delta)} \leq x < z. \quad (\text{A26})$$

This implies that

$$\theta < \frac{z \int_0^z dF(\delta)}{\int_z^{\bar{\delta}} \delta \, dF(\delta) + z \int_0^z dF(\delta) - z} = \theta_m, \quad (\text{A27})$$

which contradicts (A24), which is necessary for the existence of a non-disciplining equilibrium. Thus we cannot have both types of equilibria.

Proof of Proposition 6: Let

$$Q^a(x) = \frac{\theta \int_0^x \delta \, dF(\delta) + (1 - \theta) \int_{Q^a(x)}^x \delta \, dF(\delta)}{\theta \int_0^x dF(\delta) + (1 - \theta) \int_{Q^a(x)}^x dF(\delta)};$$

$$Q^b(x) = \frac{\int_0^x \delta \, dF(\delta)}{\theta + (1 - \theta) \int_0^x dF(\delta)}; \quad (\text{A28})$$

$$Q^c(x) = \frac{\theta \int_0^x \delta \, dF(\delta) + (1 - \theta) \int_{Q^c(x)}^x \delta \, dF(\delta)}{\theta + (1 - \theta) \int_{Q^c(x)}^x dF(\delta)}.$$

To show that Model $B_a^{a,\delta}$ produces the highest ex ante value for the firm, it is sufficient to show that for all x in the support of the distribution of δ , $Q^a(x)$ is at least as large as $Q^b(x)$ and $Q^c(x)$. To see why this is sufficient first note that any equilibrium cutoff x for Model $B_a^{a,\delta}$ must solve

$$\beta - w_2 x = w_1 Q^a(x), \quad (\text{A29})$$

while equilibrium cutoffs for Model B^a and Model B^{a,δ} solve respectively:

$$\beta - w_2x = (1 - \theta)w_1Q^b(x), \quad (\text{A30})$$

and

$$\beta - w_2x = (1 - \theta)(w_1Q^c(x) - E_{ns}(x)). \quad (\text{A31})$$

If $Q^a(x)$ is at least as large as $Q^b(x)$ and $Q^c(x)$, then there is an equilibrium cutoff solving (A29) that is no greater than any solving (A30) and (A31). Moreover, if $Q^a(x)$ is strictly larger than $Q^b(x)$ and $Q^c(x)$ (which will generally be the case), then there is an equilibrium cutoff solving (A29) that is strictly less than any solving (A30) and (A31).

We will first show that $Q^a(x) \geq Q^b(x)$ for all x . Define $\Pi_a^b = \int_a^b dF(\delta)$ and $\Delta_a^b = \int_a^b \delta dF(\delta)$, and let Q^a be shorthand for $Q^a(x)$. It is straightforward to show that the sign of $Q^a(x) - Q^b(x)$ is the same as the sign of

$$\theta^2 (1 - \Pi_0^x) \Delta_0^x + \theta(1 - \theta) (\Delta_{Q^a}^x - \Pi_{Q^a}^x \Delta_0^x) - (1 - \theta)^2 (\Pi_0^x \Delta_{Q^a}^x - \Pi_{Q^a}^x \Delta_0^x). \quad (\text{A32})$$

The first term in (A32) is clearly nonnegative. Consider now the last term. Observe that

$$\begin{aligned} \Pi_0^x \Delta_{Q^a}^x - \Pi_{Q^a}^x \Delta_0^x &= \left(\Pi_{Q^a}^x \Delta_{Q^a}^x + \Pi_0^{Q^a} \Delta_{Q^a}^x \right) - \left(\Pi_{Q^a}^x \Delta_{Q^a}^x + \Pi_{Q^a}^x \Delta_0^{Q^a} \right) \\ &= \Pi_0^{Q^a} \Delta_{Q^a}^x - \Pi_{Q^a}^x \Delta_0^{Q^a}. \end{aligned} \quad (\text{A33})$$

Now since $x \geq Q^a \geq 0$ we have

$$\begin{aligned} \Pi_0^{Q^a} \Delta_{Q^a}^x &\geq Q^a \Pi_0^{Q^a} \Pi_{Q^a}^x \\ \Pi_{Q^a}^x \Delta_0^{Q^a} &\leq Q^a \Pi_0^{Q^a} \Pi_{Q^a}^x. \end{aligned} \quad (\text{A34})$$

This means that the last expression in (A32) is nonnegative (and strictly positive if $x > Q^a > 0$). It is easy to see that the non-negativity of the last expression in (A32) implies that the second expression in (A32) is also nonnegative. This is because $\Delta_{Q^a}^x \geq \Pi_0^x \Delta_{Q^a}^x$.

Now consider $Q^c(x)$. First note that if $\theta = \theta \int_0^x dF(\delta)$, $Q^a(x) = Q^c(x)$. We need only consider

the cases where $\theta > \theta \int_0^x dF(\delta)$, i.e., $\theta > 0$ and $\int_0^x dF(\delta) < 1$. From the definition of $Q^c(x)$ we know that $Q^c(x)$ is equal to y such that y solves

$$y - \frac{\theta \int_0^x \delta dF(\delta) + (1 - \theta) \int_y^x \delta dF(\delta)}{\theta + (1 - \theta) \int_y^x dF(\delta)} = 0. \quad (\text{A35})$$

Using the definition of $Q^a(x)$, one can easily see that

$$Q^a(x) - \frac{\theta \int_0^x \delta dF(\delta) + (1 - \theta) \int_{Q^a(x)}^x \delta dF(\delta)}{\theta + (1 - \theta) \int_{Q^a(x)}^x dF(\delta)} > 0. \quad (\text{A36})$$

when $\theta > \theta \int_0^x dF(\delta)$. To show that $Q^a(x) \geq Q^c(x)$ it is sufficient to show that there is a unique solution to (A35) and that the left hand side of (A35) is increasing in y at the solution. Let $S(y)$ be the left hand side of (A35). It is straightforward to show that

$$S'(y) = 1 + \frac{(1 - \theta)f(y)}{\theta + (1 - \theta) \int_y^x dF(\delta)} S(y). \quad (\text{A37})$$

One can see from (A37) that $S(y)$ is increasing for all values of y that solve (A35). Since $S(y)$ is continuous, this means that there is a unique y that solves (A35). From this and (A36) it follows that $Q^a(x) \geq Q^c(x)$.

Proof of Proposition 7: First we prove that an equilibrium exists and is unique. To do this we first show that for all x in the support of the distribution of $\tilde{\delta}$, there is a unique $E_s(x)$ that solves the defining equation given by (19) in Section 5, and that $E_s(x)$ is increasing in x . From (19) we see that $E_s(x)$ is any value of y that solves

$$y = \frac{\theta \int_x^{\tilde{\delta}} \delta dF(\delta) + (1 - \theta) \int_x^y \delta dF(\delta)}{\int_x^{\tilde{\delta}} dF(\delta) + \int_x^y dF(\delta)}, \quad (\text{A38})$$

or, equivalently, which solves

$$S(y) = \theta \int_x^{\tilde{\delta}} (y - \delta) dF(\delta) + (1 - \theta) \int_x^y (y - \delta) dF(\delta) = 0. \quad (\text{A39})$$

It is easy to see that $S(x) < 0$, $S(\tilde{\delta}) > 0$, and $S'(y) > 0$ for all $y \in (x, \tilde{\delta})$. This means that there is a unique $y > x$ that solves (A39).¹ Using the fact that (A39) implicitly defines y as a function

¹ Note that we are assuming that $\theta > 0$. In the limit case where $\theta = 0$, we have $y = x$ as a solution to (A39).

of x , it is easy to show that

$$y'(x) = \frac{(y(x) - x)f(x)}{\theta \int_x^{\bar{\delta}} dF(\delta) + (1 - \theta) \int_x^y dF(\delta)} > 0. \quad (\text{A40})$$

Thus $E_s(x)$ is an increasing function of x . Now for an interior equilibrium in Model $G_\delta^{a,\delta}$ (i.e., one for which $0 < x_G < \bar{\delta}$), we must have $-\beta + \omega_1 E_s(x_G) + \omega_2 x_G = 0$. Since we have shown that $E_s(x)$ is monotone increasing in x , it follows that if there is an interior equilibrium, it is unique. (Otherwise, the equilibrium is not interior and either $x_G = 0$ if $-\beta + \omega_1 E_s(0) > 0$ or $x_G = \bar{\delta}$ if $-\beta + \omega_1 E_s(\bar{\delta}) + \omega_2 \bar{\delta} < 0$.)

Define $E_a(x) = E(\tilde{\delta} \mid \tilde{\delta} \geq x)$. As discussed in Section 5, when L is not present, the equilibrium cutoff x is determined by $-\beta + \omega_1 E_a(x) + \omega_2 x = 0$, while if L is present, the equilibrium cutoff point is determined by $-\beta + \omega_1 E_s(x) + \omega_2 x = 0$. To show that there is less discipline when L is present, it is sufficient to show that $E_a(x) > E_s(x)$ for all $x < \bar{\delta}$. We have

$$\begin{aligned} E_a(x) - E_s(x) &= \frac{\int_x^{\bar{\delta}} \delta dF(\delta)}{\int_x^{\bar{\delta}} dF(\delta)} - \frac{\theta \int_x^{\bar{\delta}} \delta dF(\delta) + (1 - \theta) \int_x^{E_s(x)} \delta dF(\delta)}{\int_x^{\bar{\delta}} dF(\delta) + \int_x^{E_s(x)} dF(\delta)} \\ &= \frac{(1 - \theta) \int_x^{E_s(x)} dF(\delta)}{\int_x^{\bar{\delta}} dF(\delta) + \int_x^{E_s(x)} dF(\delta)} \left(\frac{\int_x^{\bar{\delta}} \delta dF(\delta)}{\int_x^{\bar{\delta}} dF(\delta)} - \frac{\int_x^{E_s(x)} \delta dF(\delta)}{\int_x^{E_s(x)} dF(\delta)} \right) > 0 \end{aligned} \quad (\text{A41})$$

The inequality in (A41) follows since $E_s(x) < \bar{\delta}$ if $x < \bar{\delta}$ and

$$\left(\frac{\int_x^{\bar{\delta}} \delta dF(\delta)}{\int_x^{\bar{\delta}} dF(\delta)} - \frac{\int_x^{E_s(x)} \delta dF(\delta)}{\int_x^{E_s(x)} dF(\delta)} \right) = E(\tilde{\delta} \mid x \leq \tilde{\delta}) - E(\tilde{\delta} \mid x \leq \tilde{\delta} \leq E_s(x)) > 0. \quad (\text{A42})$$

Proof of Proposition 8: Consider first Model $B_\delta^{a,\delta}$. It is clear that in any equilibrium, L exits whenever M takes the action and retain them whenever M does not take the action (unless subjected to a liquidity shock). Let $m_B(\delta)$ be the probability that M takes the action in Model $B_\delta^{a,\delta}$ when the realization of $\tilde{\delta}$ is δ . This means that

$$E_s(\delta) = \frac{m_B(\delta)\delta}{\theta + (1 - \theta)m_B(\delta)}; \quad E_{ns}(\delta) = 0. \quad (\text{A43})$$

M is indifferent between taking the action and not taking it if and only if

$$\beta - \omega_1 E_s(\delta) - \omega_2 \delta = -\omega_1 \left(\theta E_s(\delta) + (1 - \theta) E_{ns}(\delta) \right). \quad (\text{A44})$$

or,

$$\beta - (1 - \theta)\omega_1 \left(\frac{m_B(\delta)\delta}{\theta + (1 - \theta)m_B(\delta)} \right) - \omega_2\delta = 0. \quad (\text{A45})$$

Now consider Model $G^{a,\delta}$, and denote by $m_G(\delta)$ be the probability that M takes the action for a given δ . In this case in any equilibrium L will retain her shares when M takes the action (unless she is subject to a liquidity shock) and sell if M does not take the action. This means that when $\tilde{\delta} = \delta$ we have

$$E_s(\delta) = \frac{\theta m_G(\delta)\delta}{\theta + (1 - \theta)(1 - m_G(\delta))}, \quad E_{ns}(\delta) = \delta. \quad (\text{A46})$$

It follows that in Model $G_\delta^{a,\delta}$ M is indifferent between taking the action and not taking it if and only if

$$-\beta + \omega_1(\theta E_s(\delta) + (1 - \theta)E_{ns}(\delta)) + \omega_2\delta = \omega_1 E_s(\delta), \quad (\text{A47})$$

or,

$$-\beta + (1 - \theta)\omega_1 \left(\frac{(1 - m_G(\delta))\delta}{\theta + (1 - \theta)(1 - m_G(\delta))} \right) + \omega_2\delta = 0. \quad (\text{A48})$$

Note that (A45) defines the function $m_B(\delta)$ over the range of δ for which $m_B(\delta) \in (0, 1)$ and (A48) does the same for $m_G(\delta)$.

REFERENCES

- Admati, Anat, Paul Pfleiderer and Josef Zechner (1994): "Large Shareholder Activism and Financial Market Equilibrium," *Journal of Political Economy* 102, 1097-1130.
- Aghion, Philippe, Patrick Bolton and Jean Tirole (2004): "Exit Options in Corporate Finance: Liquidity versus Incentives," *Review of Finance* 8, 327-353.
- Attai, Mukarram, Suman Benerjee, and Thomas Noe (2006): "Crushed by Rational Stampede: Strategic Share Dumping and Shareholder Insurrections," *Journal of Financial Economics* 79, 181-222.
- Bainbridge, Stephen (2005): "Shareholder Activism and Institutional Investors," UCLA School of Law, Law and Economics Research Paper No. 05-20.
- Bebchuk, Lucian A. (2005): "The Case for Increasing Shareholder Power," *Harvard Law Review* 118, 833-914.
- Black, Bernard S. (1990): "Shareholder Passivity Reexamined," *Michigan Law Review* 89, 520-608.
- Black, Bernard S., and John C. Coffee (1994): "Haile Britannia?: Institutional Investor Behavior under Limited Regulation," *Michigan Law Review* 92, 1997-2087.
- Bolton, Patrick and Ernst-Ludwig von Thadden (1998): "Blocks, Liquidity, and Corporate Control," *Journal of Finance* 53, 1-25.
- Burkart, Mike, Denis Gromb and Fausto Panunzi (1997): "Large Shareholder, Monitoring and the Value of the Firm," *Quarterly Journal of Economics* 112, 693-728.
- Carleton, Willard T., James M. Nelson, and Michael S. Weisbach (1998): "The Influence of Institutions on Corporate Governance through Private Negotiations: Evidence from TIAA-CREF," *Journal of Finance*, 1135-1014.
- Coffee, John C (1991): "Liquidity versus Control: The Institutional Investor as Corporate Monitor," *Columbia Law Review*, 91, 1277-1368.
- Faure-Grimaud, Antoine, and Denis Gromb (2004): "Public Trading and Private Incentives," *Review of financial Studies*, 985-1014.
- Gillan, Stuart L. and Laura T. Starks (1998): "A Survey of Shareholder Activism: Motivation and Empirical Evidence," *Contemporary Finance Digest* 2, 310-334.
- Gopalan, Rahdakrishnan (2005): "Large Shareholder Trading and Takeovers: The Disciplinary Role of Voting With Your Feet," Working paper, Ross School of Business, University of Michigan, November.
- Grundfest, Joseph (1993): "Just Vote No: A Minimalist Strategy for Dealing with Barbarians

- Inside the Gates,” *Stanford Law Review* 45, 857-903.
- Hartzell, Jay C. and Laura T. Starks (2003): “Institutional Investors and Executive Compensation,” *Journal of Finance*, 2351-2374.
- Holmstrom, Bengt and Jean Tirole (1993): “Market Liquidity and Performance Monitoring,” *Journal of Political Economy* 101, 678-709.
- Huddart, Steven (1993): “The Effect of a Large Shareholder on Corporate Value,” *Management Science*, 1407-1421.
- Kahn, Charles and Andrew Winton (1998): “Ownership Structure, Speculation, and Shareholder Intervention,” *Journal of Finance*, 99-129.
- Macey, Jonathan R. (1997): “Institutional Investors and Corporate Monitoring: A Demand-Side Perspective,” *Managerial and Decision Economics*, 18, 601-610.
- Maug, Ernst (1998): “Large Shareholders as Monitors: Is There a Trade-Off between Liquidity and Control?,” *Journal of Finance*, 65-98.
- Mello, Antonio S., and Rafael Repullo (2004): “Shareholders Activism is Non-Monotonic in Market Liquidity,” *Financial Research Letters* 1, 2-10.
- Noe, Thomas (2002): “Investor Activism and Financial Market Structure,” *Review of Financial Studies* 15, 289-315.
- Palmiter, Alan R. (2002): “Mutual Fund Voting of Portfolio Shares: Why Not Disclose?,” *Cardozo Law Review*, Vol. 23, No. 4, 1419-1491.
- Parrino, Robert, Richard W. Sias, and Laura T. Starks (2003): “Voting With Their Feet: Institutional Ownership Changes Around Forced CEO Turnover,” *Journal of Financial Economics* Vol. 68, 3-46.
- Roe, Mark J. (1994): *Strong Managers, Weak Owners: The Political Roots of American Corporate Finance*, Princeton University Press, Princeton, NJ .
- Romano, Roberta (1993): “Public Pension Fund Activism in Corporate Governance Reconsidered,” *Columbia Law Review*, Vol. 93, 797-853.
- Romano, Roberta (2001): “Less is More: Making Institutional Investor Activism a Valuable Mechanism of Corporate Governance,” *Yale Journal on Regulation*, Vol. 18, 174-252.
- Sias, Richard, Laura Starks and Sheridan Titman (2001): “The Price Impact of Institutional Trading,” Working paper, Department of Finance, University of Texas, August.
- Tirole, Jean (2001): “Corporate Governance,” *Econometrica*, Vol. 69, 1-35.
- Yung, Chris (2005): “Insider Trading with Private Information and Moral Hazard,” *Rinance Research Letters*, Vol. 2, 51-57.