# NBER WORKING PAPER SERIES

# ACTIVISM, STRATEGIC TRADING, AND LIQUIDITY

Kerry Back Pierre Collin-Dufresne Vyacheslav Fos Tao Li Alexander Ljungqvist

Working Paper 22893 http://www.nber.org/papers/w22893

NATIONAL BUREAU OF ECONOMIC RESEARCH 1050 Massachusetts Avenue Cambridge, MA 02138 December 2016

The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2016 by Kerry Back, Pierre Collin-Dufresne, Vyacheslav Fos, Tao Li, and Alexander Ljungqvist. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Activism, Strategic Trading, and Liquidity Kerry Back, Pierre Collin-Dufresne, Vyacheslav Fos, Tao Li, and Alexander Ljungqvist NBER Working Paper No. 22893 December 2016 JEL No. G3

# ABSTRACT

We analyze dynamic trading in an anonymous market by an activist investor who can expend costly effort to affect firm value. We obtain the equilibrium in closed form for a general activism technology, including both binary and continuous outcomes. The optimal continuous trading strategy is independent of the activism technology. Activism, prices, and liquidity are jointly determined in equilibrium. Variation in noise trading volatility can produce either positive or negative effects on both efficiency and liquidity, depending on the activism technology and model parameters, because future effort depends on the realized amount of noise trading. The `lock in' effect emphasized in previous literature (e.g., Maug (1998)) holds only for special forms of the activism technology. Reducing the uncertainty about the activist's position improves market liquidity, but the effect on efficiency depends on the specification of the effort cost function. Variation in the activist's productivity produces a negative cross-sectional relation between efficiency and liquidity as the possibility of more activism exacerbates the risk of adverse selection.

Kerry Back Rice University Kerry.E.Back@rice.edu

Pierre Collin-Dufresne Ecole Polytechnique Federale de Lausanne CDM SFI SFI-PCD EXTRA 209 CH-1015 Lausanne Switzerland pierre.collin-dufresne@epfl.ch

Vyacheslav Fos Carroll School of Management Boston College 140 Commonwealth Avenue Fulton Hall 330 Chestnut Hill, MA 02467 vyacheslav.fos@bc.edu Tao Li City University of Hong Kong taoli3@cityu.edu.hk

Alexander Ljungqvist Stern School of Business New York University 44 West Fourth Street, #9-160 New York, NY 10012 and NBER aljungqv@stern.nyu.edu

### 1. Introduction

Activist shareholders play an important role in modern corporate governance. The Economist describes them as "capitalism's unlikely heroes" and reports that between 2010 and 2014, half the companies in the S&P 500 index had an activist shareholder and one in seven were the target of an activist campaign (*The Economist*, February 7th 2015). Activism comes in many forms. Perhaps the best known involves hedge funds accumulating stakes in firms with the intention to create value by influencing management.<sup>1</sup> Existing shareholders can turn from being passive to being active when they recognize an opportunity for enhancing the value of their holdings.<sup>2</sup> And activist short-sellers can take actions so as to reduce firm value to benefit their short positions.<sup>3</sup> The profitability of activism for investors hinges on their ability to trade before stock prices reflect their intention to become active. Thus, there is a fundamental link between market conditions (liquidity), activism, and firm value.

Liquidity can have both harmful and beneficial effects on activism. Coffee (1991) and Bhide (1993) argue that higher liquidity should be associated with lower economic efficiency, because liquid markets make it easy for large shareholders to 'take the Wall Street walk' (i.e., sell down their positions) rather than engage actively in a firm's corporate governance when intervention might increase firm value. A certain level of illiquidity might then be desirable, to 'lock in' large shareholders. Maug (1998) argues the opposite, pointing out that more noise trading enables a potential activist who does not already own a sizeable initial toehold to accumulate more shares and eventually become active. Thus, the size of the initial toehold seems to play a key role in the relation between liquidity and activism.

In this paper, we revisit the classic question of the relation between liquidity and economic efficiency by generalizing the activism technology. The existing literature considers only binary forms of activism, in which the activist's action leads to a fixed increase in firm value. In reality, there are many different types of activism, including those with non-binary effort and those resulting

<sup>&</sup>lt;sup>1</sup>Prominent examples include William Ackman, Carl Icahn, Daniel Loeb, and Nelson Peltz.

<sup>&</sup>lt;sup>2</sup>CALPERS and the Norwegian sovereign wealth fund are well known examples of this form of activism.

 $<sup>^{3}</sup>$ For an example, see Bloomberg Business on how "Hedge funds found a new way to attack drug companies and short their stock" (March 20th 2015), describing how some activist hedge funds challenge pharmaceutical patents in court to reduce the value of the firms owning these patents, presumably benefitting from previously established short positions.

in a non-binary effect on firm value. When an activist seeks to increase payouts (e.g., Carl Icahn and Apple), it arguably requires more effort to induce a larger change in payout policy, which leads to a larger effect on firm value. When an activist risk arbitrageur wants to influence whether an M&A deal is completed, the outcome is likely to be binary but the effort expended by the activist is continuous (Jiang, Li and Mei, 2016). The probability that the activist is successful is an increasing function of her continuous effort. When an activist seeks to fire a CEO, the outcome is binary though effort may not be. When an activist runs a proxy contest hoping to replace incumbent directors with nominees of her choosing, the number of directors the activist can get elected arguably depends on her effort. Given a wide range of activism technologies, it is an open question whether and how the relation between market liquidity and economic efficiency depends on the activism technology.

We extend the dynamic version of the Kyle (1985) model to a large trader (the potential activist) who can affect the firm's liquidation value by expending costly effort, but who can also decide to walk away. We work in continuous time, because similar to Back's (1992) extension of Kyle's model to non-Gaussian distributions, it affords tractability.<sup>4</sup> We obtain the equilibrium in closed form for a general form of activism technology, which leads to several new insights.

First, the underlying nature of the effort cost function, in particular whether it is binary or continuous, plays a crucial role in the relation between liquidity and activism. In our model, an increase in noise trading volatility increases the variance of the activist's terminal stake (because the activist on average buys if the price is depressed due to liquidity-motivated selling and on average sells if the price is inflated due to liquidity-motivated buying). As a result, it leads to higher (lower) expected effort if future effort is convex (concave) in the stake accumulated by the activist, *irrespective of the activist's initial stake size*.<sup>5</sup> This does not apply to the binary case, in which effort is a step function and hence neither convex nor concave. In the binary case, the relation between liquidity and activism is similar to that obtained by Maug (1998) in a one-period model. Thus, the effect of noise trading on activism can differ considerably depending on whether

<sup>&</sup>lt;sup>4</sup>In Appendix A, we study the one-period model. This yields explicit results only for simple special cases that are uninteresting for the main question we investigate.

 $<sup>^{5}</sup>$ When future effort is convex in the activist's stake, the unconditional expected effort is increasing in the variance of the stake in the same way that the value of a call option is increasing in the variance of its underlying (Merton, 1973).

the activism technology is binary or continuous.

A second new insight is that an increase in noise trading volatility does not necessarily improve market liquidity. This contrasts with the standard Kyle (1985) model in which market depth is an increasing function of noise trading volatility and a decreasing function of private information. The reason for the difference is that private information and noise trading volatility are linked in our model. An increase in noise trading volatility increases the variance of the activist's terminal stake due to her endogenous trading, as discussed above. Because the size of the activist's stake is private information, an increase in noise trading volatility increases adverse selection for market makers. This effect on adverse selection can more than offset the natural increase in market depth stemming from increased noise trading. Specifically, an increase in noise trading volatility reduces market depth in the binary model when the mean of the activist's initial stake is either very small or very large, and the same occurs in the continuous model when the derivative of the firm value as a function of the activist's stake size is convex in the stake size. Overall, that activism also affects market liquidity is a link that has not been emphasized in the literature, where liquidity is typically taken to be exogenous and one to one with the amount of noise trading.<sup>6</sup>

A third new insight is that, because both market liquidity and activism are jointly determined, the cross-sectional relation between them depends on the source of variation in either.<sup>7</sup> The model parameters that can be varied to change market liquidity and activism are the noise trading volatility, the activist's productivity, and the mean and variance of the activist's initial stake. Changes in each of these parameters are also naturally related to various policy measures (such as position disclosure rules, length of the trading window, or the legal environment). The comparative statics of our model are hence informative about the potential implications of changes in these policy measures.

To illustrate how the relation between market liquidity and activism depends on the source of variation in each, consider three examples. As noted earlier, variation in noise trading volatility

<sup>&</sup>lt;sup>6</sup>For example, Maug (1998) defines liquidity as the exogenous fraction of shares sold by uninformed investors. This is also consistent with the terminology used in Edmans' (2016) survey.

<sup>&</sup>lt;sup>7</sup>Our model can be used to interpret empirical results for samples of activism events by regarding each event as a separate example of the model. The events can all be consistent with our model but differ due to differences in parameter values. Throughout the paper, we discuss the implications of the model for a sample (cross-section) of events viewed as instances of the model with possibly different parameter values.

can produce either a positive or a negative cross-sectional relation between liquidity and activism, depending on the nature of the activism technology and, possibly, the size of the initial stake. Similarly, how greater uncertainty about the activist's position affects activism depends on the nature of the activism technology (though it always reduces market liquidity by increasing adverse selection risk). Finally, variation in the activist's productivity induces a negative cross-sectional relation between activism and liquidity, because a greater scope for value-changing activity by the activist increases the importance of private information regarding the activist's intentions (which depend endogenously on her holdings of the firm's stock) and hence reduces liquidity. Thus, which parameter is varied (or which policy change is considered) has different implications for the crosssectional relation between activism and market liquidity.

A surprising feature of our model is that the large trader's equilibrium trading strategy does not depend on the cost function. Unlike in the standard Kyle model, the equilibrium trade size depends on the number of shares currently owned as well as on aggregate orders. Purchases are larger when the number of shares already owned is larger. This reflects the increasing returns to scale inherent in activism—the aggregate benefit depends on the number of shares owned but the cost does not. Hence, the value to the large trader of an additional share is higher the more shares she already owns. Nevertheless, the cost-of-effort function plays no role in determining the equilibrium trading strategy.

# 2. Related Literature

DeMarzo and Urošević (2006) also analyze a dynamic market with a blockholder whose actions affect corporate value. A key distinction between their paper and ours is that they assume a fully revealing rational expectations equilibrium. In contrast, we follow Kyle (1985) by assuming there is some additional uncertainty in the market (namely, noise trading) that provides camouflage for the blockholder's trading. This allows the market's forecast of the blockholder's plans to sometimes deviate from what the blockholder herself regards as most likely, producing profitable trading opportunities.

There are several papers, in addition to Maug (1998), that analyze single-period market microstructure models involving one or more large traders who may intervene in corporate governance. These include Kyle and Vila (1991), Kahn and Winton (1998), Ravid and Spiegel (1999), Bris (2002), and Noe (2002). The papers most closely related to ours are Kyle and Vila's and Kahn and Winton's. Kahn and Winton's model structure is quite similar to Maug's (1998). In their comparison of their work with Maug's, they state that they complement Maug by focusing on issues other than the effect of liquidity on governance. Kyle and Vila's conclusion regarding the effect of noise trading on activism (a value-enhancing takeover in their case) is similar to the result we obtain with a binary value distribution (and similar to Maug's result). Our paper contributes to this literature by developing a model with a general activism technology. As our paper indicates, generalizing the activism technology leads to fundamental changes in the relation between liquidity and economic efficiency.

Another strand of the literature on trading and activism that is tangentially related to our paper is the literature on "governance by exit," which includes the papers by Admati and Pfleiderer (2009), Edmans (2009), and Edmans and Manso (2011). In these models, a blockholder has access to private information about firm value and may sell her block on negative information. The blockholder's ability to trade on negative information and the manager's concern with the short-term stock price cause the manager to be more concerned than he otherwise would be about the impact of his actions on firm value and thereby improves governance. The focus of these papers is on trading by an insider who has private information about firm value that is exogenous to her trading. In contrast, in our model, the blockholder has no private information about exogenous elements of corporate value. Instead, we study strategic trading by an investor who can become active. Moreover, exit models all have a single round of trading, so they cannot analyze feedback from prices to blockholder actions.

The relation between market liquidity and economic efficiency is related to the on-going debate about the optimal duration of the pre-disclosure period for 13D filers (e.g., Bebchuk, Brav, Jackson and Jiang, 2013). Specifically, shortening the period in which an activist can trade anonymously has the effect of reducing cumulative noise trading during the period in which the activist can trade anonymously. The relation is also central to the debate about insider trading rules and, more generally, about required disclosure rules for trading positions of significant blockholders (e.g., Fishman and Hagerty, 1992, 1995).

## 3. Model

We analyze a Kyle model in which the large trader can undertake costly effort to influence the management of a firm and hence affect the value of its stock. The large trader has no private information about the exogenous value of the stock but has private information about her own position in the stock and thus is better informed about the value she will create. Trading is continuous during a time interval [0, T]. Denote the number of shares owned by the large trader at each date t by  $X_t$ . We assume that  $X_0$  is known only to the large trader.<sup>8</sup> Immediately after T, the large trader can expend effort to affect the value of the stock. Afterwards, positions are liquidated frictionlessly at some common value v.

Denote by C(v) the cost to the large trader of achieving a common value of v. Given  $X_T = x$ , the large trader chooses effort to maximize vx - C(v). The optimal value to the large trader is

$$G(x) \stackrel{\text{def}}{=} \sup_{v} \left\{ vx - C(v) \right\}.$$

Let V(x) denote the value of v at which the supremum is attained. This is the common value of shares to all traders after the large trader's expenditure of effort. By the envelope theorem,  $\partial G(x)/\partial x = V(x)$ . Thus,

$$G(x) = G(0) + \int_0^x V(a) \,\mathrm{d}a \,. \tag{1}$$

Assume G is a convex function, so V is increasing. A sufficient condition for G to be convex is that C be convex (G is the Fenchel conjugate of C).

In addition to the large trader, there are noise traders in the market. Let  $Z_t$  denote the cumulative number of shares purchased by noise traders through date t, with  $Z_0 = 0$ . Assume Z is a Brownian motion with zero drift and instantaneous standard deviation  $\sigma$ . Aggregate purchases by the large trader and noise traders are  $Y_t = X_t - X_0 + Z_t$ .

All orders are submitted to risk-neutral competitive market makers. The market makers therefore observe Y. They compete to fill orders, pushing the price to the expected value of  $V(X_T)$ 

<sup>&</sup>lt;sup>8</sup>This assumption is consistent with U.S. rules on the disclosure of ownership stakes. For instance, Schedule 13F requires quarterly disclosures of long positions in U.S. stocks held by institutional investment managers with more than \$100m in assets under management. In between these quarterly filings, only the investment managers themselves know their positions. Smaller investment managers do not have to disclose their positions at all. Moreover, there is at present no mandatory disclosure, for any type of investor, of short positions.

conditional on the history of orders. Let  $\mathcal{F}_t^Y$  denote the information conveyed by the history of orders through date t. We assume market makers are uncertain about the number of shares  $X_0$  that the large trader initially owns and view it as normally distributed with mean  $\mu_x$  and standard deviation  $\sigma_x$ .

In principle, the price at each date could depend on the entire history of orders up to that date, but, as in Kyle (1985), we search for an equilibrium in which the cumulative order process  $Y_t$  serves as a state variable. This means that the price at each date t is  $P(t, Y_t)$  for some function P. Also, we look for an equilibrium in which the large trader's trades are of order dt, meaning that  $dX_t = \theta_t dt$  for some stochastic process  $\theta$ .<sup>9</sup> Given  $P(\cdot)$ , the large trader chooses the trading strategy  $\theta$  to maximize

$$\mathsf{E}\left[G(X_T) - \int_0^T P(t, Y_t)\theta_t \,\mathrm{d}t \mid X_0\right] \,. \tag{2}$$

The large trader's value function is

$$J(t, x, y) \stackrel{\text{def}}{=} \sup_{\theta} \mathsf{E} \left[ G(X_T) - \int_t^T P(u, Y_u) \theta_u \, \mathrm{d}u \, \middle| \, X_t = x, Y_t = y \right].$$
(3)

We define an equilibrium to be a pair  $(P, \theta)$  such that the trading strategy  $\theta$  maximizes (2) given P and such that

$$P(t, Y_t) = \mathsf{E}\left[V(X_T) \mid \mathcal{F}_t^Y\right] \tag{4}$$

for each t, given  $\theta$ . This is the standard definition of equilibrium in a Kyle model, except for the fact that the value V depends on  $X_T$  in our model.

We can use the envelope condition to rewrite the objective function (2) as

$$G(X_0) + \mathsf{E}\left[\int_0^T \left(V(X_t) - P(t, Y_t)\right)\theta_t \,\mathrm{d}t \mid X_0\right].$$
(5)

This form of the objective function shows that both the market makers and the large trader agree on the marginal value  $V(X_t)$  of every share accumulated despite the fact that the large trader bears

<sup>&</sup>lt;sup>9</sup>This assumption is without loss of generality, because, as shown by Back (1992), if there are jumps or nonzero quadratic variation in the large trader's holdings X, the large trader pays bid-ask spread costs on these components of the order flow of a size similar to those paid by noise traders. It is suboptimal for the large trader to pay these costs, and they can be avoided by taking X to be continuous and of finite variation.

the entire cost of effort. This condition seems intuitively very important for an equilibrium to exist, since both the market makers and the large trader are risk-neutral.

To ensure that there are no doubling-type strategies available to the large trader (Back, 1992), we adopt the following regularity condition:

$$\mathsf{E}\left[V\left(\frac{\Lambda(X_0 - Z_T) - \mu_x}{\Lambda - 1}\right)^2\right] < \infty,\tag{6}$$

and we define a trading strategy  $\theta$  to be admissible if and only if

$$\mathsf{E} \int_{0}^{T} V(\mu_{x} + \Lambda(X_{t} - X_{0} + Z_{T}))^{2} \, \mathrm{d}t < \infty \,, \tag{7}$$

where it is convenient to define a specific measure of the 'signal to noise' ratio that appears frequently below:

$$\Lambda \equiv 1 + \sqrt{1 + \frac{\sigma_x^2}{\sigma^2 T}} \,. \tag{8}$$

# 4. Equilibrium

Theorem 1. The pricing rule

$$P(t,y) = \mathsf{E}\left[V(\mu_x + \Lambda Z_T) \mid Z_t = y\right] \tag{9}$$

and trading strategy

$$\theta_t = \frac{1}{T-t} \left( \frac{X_t - \mu_x - \Lambda Y_t}{\Lambda - 2} \right) \tag{10}$$

constitute an equilibrium. In this equilibrium, the distribution of Y given market makers' information is that of a Brownian motion with zero drift and standard deviation  $\sigma$ . Moreover,  $P(T, Y_T) = V(X_T)$  with probability 1. The value function is

$$J(t,x,y) = \frac{\Lambda - 1}{\Lambda} \mathsf{E}\left[G\left(\frac{\Lambda(x - Z_T) - \mu_x}{\Lambda - 1}\right) \middle| Z_t = y\right] + \frac{1}{\Lambda} \mathsf{E}\left[G(\mu_x + \Lambda Z_T) \middle| Z_t = y\right].$$
(11)

The equilibrium price evolves as  $dP(t, Y_t) = \lambda(t, Y_t) dY_t$ , where Kyle's lambda is

$$\lambda(t,y) = \frac{\partial P(t,y)}{\partial y} \,. \tag{12}$$

Furthermore,  $\lambda(t, Y_t)$  is a martingale on  $[0, T - \delta]$  for every  $\delta > 0$ , relative to the market makers' information set.

We note the surprising finding that the large trader's trading strategy can be fully specified without specifying the cost function C. Thus, the trading strategy is independent of the cost function, at least as expressed as a function of the cumulative noise trading and the large trader's accumulated shares.<sup>10</sup> Instead, the cost function determines the equilibrium price process and market liquidity. Thus, expressed as a function of the price process, which may seem more natural, the trading strategy will look different. The following corollary summarizes some of the properties of the large trader's position.

**Corollary.** The large trader's equilibrium position at time T is

$$X_T = \mu_x + \frac{\Lambda}{\Lambda - 1} (X_0 - \mu_x - Z_T).$$
(13)

It follows that  $X_T$  is normally distributed with unconditional mean  $\mathsf{E}[X_T] = \mu_x$  and unconditional variance  $\mathsf{V}[X_T] = (\sigma\sqrt{T} + \sqrt{\sigma^2 T + \sigma_x^2})^2$ .

As in a standard Kyle model, the large trader's trades are not forecastable. On average, market makers do not expect the large trader to trade in one direction or the other. Consequently,  $E[X_T] = \mu_x$ , as stated in the corollary. However, the variance of the large trader's terminal position is different than in the standard Kyle model, because liquidity shocks have an amplifying effect. The variance of the accumulated position increases more than linearly in noise trading. The foundation for this amplifying effect is the increasing returns to scale inherent in activism. The cost of activism is fixed in the sense that it does not depend directly on the number of shares the large trader owns. The more shares she owns, the more valuable is value-creating activism.

In the remainder of this section, we sketch the proof of Theorem 1. The complete proof is in Appendix B. There are two standard features of continuous-time Kyle models that we used to guess the form of the equilibrium in Theorem 1. The first feature is that the large trader trades in such a way as to ensure that the share price equals the marginal value at the terminal date. Otherwise,

<sup>&</sup>lt;sup>10</sup>This feature is due to continuous trading. Indeed, we show in Appendix A that in a one-period model, the trading strategy is not independent of the cost function.

she is clearly leaving money on the table. In our model, this means that  $P(T, Y_T) = V(X_T)$ with probability 1. The other feature is that informed orders are unpredictable to market makers, meaning that the drift of Y is zero on its own filtration; that is, Y is a martingale on its own filtration.<sup>11</sup> Because Y has the same quadratic variation as Z, this martingale property implies that Y must actually be a Brownian motion with the same standard deviation as Z.

For convenience, let h(y) denote P(T, y). This is a function we need to find. The property of inconspicuous strategic trading and the risk neutrality of market makers imply that the price at all dates t < T is the expectation of  $P(T, Y_T)$  treating Y as a Brownian motion with standard deviation  $\sigma$ . Therefore, we know the equilibrium pricing rule if we know h.

In the standard Kyle model,  $h(Y_T) = v$  in equilibrium, where v is the exogenous value of the asset. This equality occurs because the large trader in the Kyle model trades in such a way that  $Y_T = h^{-1}(v)$ , or equivalently,  $X_T = X_0 + h^{-1}(v) - Z_T$ . Thus, the large trader offsets noise trades one-for-one and also purchases (or sells if negative)  $h^{-1}(v)$  shares. This contrasts with our expression (13), in which the coefficient on  $-Z_T$  in the formula for  $X_T$  is  $\Lambda/(\Lambda - 1) > 1$ . As mentioned before, in our model, the large trader offsets noise trades by more than one-for-one, due to the increasing returns inherent in activism.

The large trader achieves the equality  $Y_T = h^{-1}(v)$  in the standard Kyle model by causing Y to be a Brownian bridge terminating at  $h^{-1}(v)$ . In our model, the asset value depends on  $X_T$ , because the large trader's incentives to be active depend on  $X_T$ , and the equality  $h(Y_T) = V(X_T)$  should hold in equilibrium (price equals marginal value). This equality implies a link between  $Y_T$  and  $Z_T$  that does not occur in a Brownian bridge. The following lemma generalizes the concept of a Brownian bridge and is key to our equilibrium construction. The first term on the right-hand side of equation (14) below is the large trader's equilibrium order  $dX_t = \theta_t dt$ .

**Lemma 1.** Let  $\varepsilon$  be a standard normal random variable that is independent of Z. Let b be a nonnegative constant, and set  $a = \sigma \sqrt{(2b+1)T}$ . Then, the solution Y of the stochastic differential equation

$$dY_t = \frac{a\varepsilon - bZ_t - (b+1)Y_t}{T-t} dt + dZ_t$$
(14)

<sup>&</sup>lt;sup>11</sup>Cho (2003) calls this 'inconspicuous insider trading.' It is a consequence of the Hamilton-Jacobi-Bellman equation and is therefore a necessary condition for equilibrium. See Back (1992). In Appendix A, we show that it is also a necessary condition in one-period model, when the activist can condition her demand on the noise trading process.

on the time interval [0,T) has the following properties:  $Y_T \stackrel{\text{def}}{=} \lim_{t\to T} Y_t$  exists a.s., Y is a Brownian motion with zero drift and standard deviation  $\sigma$  on its own filtration on [0,T], and, with probability 1,

$$Y_T = \frac{a\varepsilon - bZ_T}{b+1} \,. \tag{15}$$

The proof of the lemma is provided in Appendix C. The stochastic differential equation of a Brownian bridge is equation (14) with b = 0, so the process Y defined by equation (14) is a generalization of a Brownian bridge. The distribution of a Brownian bridge, conditional on  $\varepsilon$ , is the distribution of Z conditioned to end at  $a\varepsilon$ , and the unconditional distribution of the Brownian bridge is the same as that of Z. As stated in the lemma, the unconditional distribution of the generalized Brownian bridge is also the same as that of Z. Thus, the property of inconspicuous strategic trading holds. Note that for the unconditional distribution to be the same as that of Z, the right-hand side of equation (15) must have variance equal to  $\sigma^2 T$ . This is equivalent to the condition  $a = \sigma \sqrt{(2b+1)T}$  specified in the lemma.

In the standard Kyle model, the standard normal random variable  $\varepsilon$  in equation (14) is a transformation of the exogenous asset value v. Assuming v is continuously distributed, we have that  $\varepsilon = \mathsf{N}^{-1}(F(v))$ , where  $\mathsf{N}$  is the standard normal distribution function and F is the distribution function of v.<sup>12</sup> In our model, the large trader's private information concerns her initial position  $X_0$ . Set  $\varepsilon = (X_0 - \mu_x)/\sigma_x$ , which is a standard normal random variable. Substitute  $Z_T = Y_T - (X_T - X_0)$  and the definition of  $\varepsilon$  into equation (15) and rearrange to obtain

$$Y_T = \frac{a(X_0 - \mu_x)/\sigma_x + b(X_T - X_0)}{2b + 1} \,. \tag{16}$$

The random variable  $Y_T$  cannot depend directly on  $X_0$ , because  $X_0$  is not observed by market makers. In order to cancel the  $X_0$  terms in equation (16), we need  $a = b\sigma_x$ . Combining the two conditions on a, we find that b must satisfy the equation:

$$\sigma\sqrt{(2b+1)T} = b\sigma_x \,. \tag{17}$$

<sup>&</sup>lt;sup>12</sup>The function  $N^{-1} \circ F$  is the function  $h^{-1}$  in the standard Kyle model.

This has a unique positive solution  $b = 1/(\Lambda - 2)$ . With this formula for b, we have<sup>13</sup>

$$Y_T = \frac{b(X_T - \mu_x)}{2b + 1} \quad \Longleftrightarrow \quad X_T = \mu_x + \Lambda Y_T \,, \tag{18}$$

Therefore,  $V(X_T) = V(\mu_x + \Lambda Y_T)$ . This equals  $h(Y_T)$ —and hence price equals marginal value at the end of trading—if we define  $h(y) = V(\mu_x + \Lambda y)$  as in equation (9). To prove the equilibrium result, all that remains is to verify that expression (10) is an optimal trading strategy for the large trader. We do that in Appendix C. It is straightforward to verify optimality because every trading strategy is optimal if it implies that price equals marginal value at the end of trading.

As in Back (1992), the large trader's value function can be interpreted as the expected profit achieved by not trading until maturity T, at which time she trades along the residual supply curve of the asset, buying or selling shares until price equals marginal value. We use that characterization in the proof of the theorem to derive the formula (11) for the value function.

### 5. Liquidity and Activism

This section presents some general results regarding the effects of model parameters on economic efficiency and market liquidity. The examples in the next section illustrate these results. We measure economic efficiency by the initial price P(0,0), which incorporates the value expected to be created by activism. Let  $\overline{P}$  denote P(0,0) as a function of the model parameters. There are at least two different ways in which we could measure liquidity. First, we could use the expected average lambda:

$$\frac{1}{T} \mathsf{E} \int_0^T \lambda(t, Y_t) \, \mathrm{d}t \, .$$

Theorem 1 shows that  $\lambda$  is a martingale (up to times arbitrarily close to time T), so the expected average lambda is equal to the initial lambda  $\lambda(0,0)$ , which we denote  $\overline{\lambda}$  as a function of the parameters. Lambda measures the absolute price impact of trades. In examples in which absolute price changes are stationary over time,  $\lambda(0,0)$  is the natural measure of liquidity. However, there are other examples in which percentage price changes are stationary, and in those examples, some

<sup>&</sup>lt;sup>13</sup>Equation (18) implies the corollary stated before. From (18) and the definition  $Y_T = X_T - X_0 + Z_T$ , we obtain  $X_T = \mu_x + \Lambda(X_T - X_0 + Z_T)$ , which can be rearranged to yield (13).

measure of the percentage price impact of trades is more natural. Example 3 in the next section, for instance, uses the percentage price impact at date 0 as the measure of liquidity.

The following theorem shows that the effects of the model parameters on efficiency and liquidity depend in several cases on whether V is convex or concave. The domain of V is the entire real line (the large trader's terminal block size  $X_T$  can take any real value), so a convex V must be unbounded above, and a concave V must be unbounded below. We give one example in Section 6 in which V is affine and hence unbounded in both directions. In general, convexity is more reasonable than concavity, because concavity (unbounded below) implies that the possible value destruction must be unlimited, not even respecting limited liability. We give several examples in Section 6 of a convex V. We also give some examples (including the binary case) in which V is bounded both above and below and hence neither convex nor concave. The proof of Theorem 2 is in Appendix D.

# Theorem 2.

- 1. An increase in the amount of noise trading increases economic efficiency  $(\partial \overline{P}/\partial \sigma \ge 0)$  if V is convex and reduces economic efficiency  $(\partial \overline{P}/\partial \sigma \le 0)$  if V is concave.
- 2. An increase in the expected initial block size
  - (a) increases economic efficiency  $(\partial \overline{P}/\partial \mu_x \ge 0)$ , and
  - (b) reduces market liquidity  $(\partial \overline{\lambda} / \partial \mu_x \ge 0)$  if V is convex and increases market liquidity  $(\partial \overline{\lambda} / \partial \mu_x \le 0)$  if V is concave.
- 3. An increase in uncertainty about the initial block size
  - (a) increases economic efficiency  $(\partial \overline{P}/\partial \sigma_x \ge 0)$  if V is convex and reduces economic efficiency  $(\partial \overline{P}/\partial \sigma_x \le 0)$  if V is concave, and
  - (b) reduces market liquidity  $(\partial \overline{\lambda} / \partial \sigma_x \ge 0)$  if the following regularity condition is satisfied:

$$\lim_{\epsilon \to \infty} V' \left( \mu_x + \Lambda \sigma \sqrt{T} \epsilon \right) \, \epsilon \, \mathrm{e}^{-\epsilon^2/2} = 0 \,. \tag{*}$$

To understand the role of convexity or concavity in Theorem 2, note that Theorem 1 implies

$$\overline{P} = \mathsf{E}[V(\mu_x + \Lambda Y_T)], \qquad (19)$$

and

$$\overline{\lambda} = \Lambda \mathsf{E}[V'(\mu_x + \Lambda Y_T)], \qquad (20)$$

where we regard  $Y_T$  as normally distributed with mean zero and variance equal to  $\sigma^2 T$ . The standard deviation of  $\Lambda Y_T$  is  $\Lambda \sigma \sqrt{T}$ , which is an increasing function of both  $\sigma$  and  $\sigma_x$ . Thus, increases in those parameters create mean-preserving spreads in the distribution of  $\mu_x + \Lambda Y_T$ , which cause  $\overline{P}$  to rise when V is convex and to fall when V is concave. Likewise, increases in those parameters cause  $\overline{\Lambda}$  to rise when V' is convex.<sup>14</sup>

Theorem 2 does not provide a general result regarding the effect of a change in the volatility of noise trading  $\sigma$  on market liquidity  $\overline{\lambda}$ , because a change in  $\sigma$  has two effects on  $\overline{\lambda}$ , and they can be in opposite directions. First, an increase in  $\sigma$  causes the factor  $\Lambda$  in equation (20) to fall. Second, an increase in  $\sigma$  is a mean-preserving spread in the distribution of  $\mu_x + \Lambda Y_T$ , so it causes the expectation in equation (20) to rise when V' is convex. Depending on which of these two effects is stronger, an increase in  $\sigma$  can cause  $\overline{\lambda}$  either to fall (as in Kyle, 1985) or to rise. The latter occurs for certain parameter values in Example 3 in Section 6. (However, as remarked before, it is more natural in that example to measure liquidity by the percentage price impact.) An increase in noise trading can also reduce market liquidity when neither V nor V' is convex. This occurs for some parameter values in Examples 4 and 5 in Section 6. The reason that greater noise trading produces lower market liquidity in those examples is that greater uncertainty about  $Z_T$  implies greater uncertainty about  $X_T$  (see equation (13)) and hence increases information asymmetry about the ultimate asset value. This phenomenon does not occur in the standard Kyle model in which the asset value is independent of  $X_T$ .

When V is convex (and satisfies the regularity condition), cross-sectional variation in either  $\mu_x$ or  $\sigma_x$  produces a negative cross-sectional relation between market liquidity and economic efficiency: efficiency is higher in less liquid markets. The reason is that a greater likelihood for activism (due to changes in  $\mu_x$  or  $\sigma_x$ ) increases the importance of asymmetric information regarding the large trader's intentions and makes the market less liquid. This direction of causality (activism  $\rightarrow$ liquidity) is the opposite of that with which the literature has been concerned.

<sup>&</sup>lt;sup>14</sup>Note that V' cannot be concave, because it is nonnegative and hence bounded below.

Cross-sectional variation in efficiency and liquidity can also be due to cross-sectional variation in the cost function C. In the examples in the next section, each cost function depends on a productivity parameter. An increase in the activist's productivity generally increases economic efficiency and generally reduces market liquidity (because asymmetric information about the activist's intentions is more important when the activist is more productive). Thus, cross-sectional variation in productivity also generally leads to a negative cross-sectional relation between economic efficiency and liquidity. Again, this is not the direction of causality emphasized in the literature.

#### 6. Examples

We consider five examples. The equilibria are presented in Table 1 in terms of the functions C, G, V, h, P, and  $\lambda$  and the parameters  $\mu_x, \sigma_x, \sigma$ , and  $\Lambda$  defined in Sections 3 and 4. The examples are distinguished by their cost functions C(v). The cost functions include an additional productivity parameter  $\psi$  (and a second productivity parameter  $\Delta$  in Examples 4 and 5). Comparative statics with respect to all parameters are presented in Tables 2 and 3.

In the first example, V is affine. In the second and third, V is bounded below and convex. In the fourth and fifth, V is bounded both above and below and hence is neither convex nor concave.

**Example 1** (Quadratic Cost). This example is from Collin-Dufresne and Fos (2015). Effort is continuous and cost is quadratic. The cost function is  $C(v) = (v - v_0)^2/(2\psi)$  for constants  $v_0$  and  $\psi > 0$ . Thus, value can be either destroyed or created by the activist. The parameter  $\psi$  measures the activist's productivity (for either value creation or value destruction). The value V(x) is affine in x, so it is both convex and concave. By Theorem 2, this implies that economic efficiency is independent of the parameters  $\sigma$  and  $\sigma_x$ . Intuitively, it is independent because, whatever effects those parameters have on possible value creation, they have the same effects on possible value destruction. Also, market liquidity is independent of  $\mu_x$ . Therefore, the only parameter that can produce cross-sectional variation in both efficiency and liquidity in this example is the productivity parameter  $\psi$ , and variation in it produces a negative cross-sectional relation between efficiency and liquidity.

This symmetric quadratic example closely resembles the classic Kyle model in which the terminal value is normally distributed. As in that model, the equilibrium price process is a Brownian

motion (on its own filtration) and Kyle's lambda is constant and increasing in the signal-to-noise ratio  $\sigma_x/\sigma$ . Kyle's lambda is also increasing in the activist's productivity  $\psi$ . In fact, and unlike in the Kyle model, the limit of lambda when the signal-to-noise ratio goes to zero is strictly positive:  $\lim_{\lambda\to 0} \lambda = \psi$ . This illustrates the difference between the two models. Even if there is very little private information at the start of the model, there is private information later in the model because only the activist knows her own trades, which determine her incentives for activism and so ultimately determine the asset value. The importance of this private information depends on the activist's productivity  $\psi$ , which is the lower bound on lambda.

**Example 2** (Asymmetric Quadratic Cost). In this example, value can be created  $(v > v_0)$  but cannot be destroyed. The cost function is

$$C(v) = \begin{cases} \infty & \text{if } v < v_0 \,, \\ (v - v_0)^2 / (2\psi) & \text{if } v \ge v_0 \,, \end{cases}$$

for constants  $v_0$  and  $\psi > 0$ . Again,  $\psi$  measures the activist's productivity. The value V(x) is convex in x, so Theorem 2 implies that economic efficiency is improved by increases in either  $\sigma$ or  $\sigma_x$ . In this example, a change in the amount of noise trading causes economic efficiency and liquidity  $(1/\overline{\lambda})$  to move in the same direction. However, changes in  $\mu_x$ ,  $\sigma_x$ , or  $\psi$  cause economic efficiency and liquidity to move in opposite directions.

Even though the large trader can only create and cannot destroy value in this example, the trading strategy (expressed as a function of cumulative order flow and the large trader's position) is identical to that in Example 1 (and in fact is the same in all examples). The price and Kyle's lambda do, however, depend on the cost function. To illustrate the differences between Examples 1 and 2, we plot two (randomly generated) paths of uninformed order flow and the corresponding informed orders, the resulting equilibrium price, and Kyle's lambda in Figures 1 and 2 below. Figure 1 shows a case where the uninformed noise traders are net cumulative buyers of the stock, whereas Figure 2 shows a path where cumulative trades by noise traders are sales.

Independent of the (symmetric or asymmetric) cost function, the large trader trades in the opposite direction of the noise traders with an amplification as discussed before. The figures illustrate the amplification. When the large trader accumulates a positive number of shares (Figure 2), prices ultimately reflect the positive value creation; thus, the prices with symmetric and asymmetric cost functions converge to the same value. Also, in that case, Kyle's lambda in the asymmetric model converges to the constant price impact that prevails throughout in the symmetric cost function model.

However, when the large trader accumulates a large short position, the price and price impact processes look very different in the two models. In the asymmetric model, the market infers the short position from the net short order flow and price converges to  $v_0$  as the market correctly expects the large trader not to expend any effort. Correspondingly, Kyle's lambda converges to zero, because, given the large negative position the large trader is anticipated to hold, a marginal increase in her position would not be expected to lead to significant positive value creation. However, in the symmetric model, the market infers from the net cumulative short position that the large trader will destroy value at maturity. The market impounds this negative value in the price. Kyle's lambda remains constant and strictly positive in the symmetric model.

**Example 3** (Exponential). This is another example of a convex V in which value can be created but not destroyed. For parameters  $v_0 > 0$  and  $\psi > 0$ , the cost of effort is

$$C(v) = \begin{cases} \frac{1}{\psi} v \log\left(\frac{v}{v_0}\right) - \frac{1}{\psi} (v - v_0) & \text{if } v > v_0, \\ \infty & \text{if } v \le v_0. \end{cases}$$

This implies  $V(x) = v_0 e^{\psi x}$ . Again,  $\psi$  measures the activist's productivity. In general, when  $\sigma_x$  is small, the partial derivative  $\partial \Lambda / \partial \sigma$  is small. In this example, V' is convex, and when  $\sigma_x$  is small, the effect of a change in  $\sigma$  on  $\Lambda$  is less than the effect of a change in  $\sigma$  on the expectation in (20). Consequently, an increase in noise trading  $\sigma$  causes market liquidity (as measured by  $1/\overline{\lambda}$ ) to fall.<sup>15</sup> However, as remarked in Section 5, it is more natural to measure market liquidity in this example by the percentage price impact,  $\lambda/P$ . In fact, the percentage price impact is constant in this example, and it is a decreasing function of  $\sigma$ . Measuring liquidity in this way, cross-sectional variation in

$$\sigma_x < \Lambda \sigma^2 T \sqrt{\frac{\psi(\Lambda - 1)\Lambda}{2(1 + \psi \Lambda^2 \sigma^2 T)}} \,.$$

<sup>&</sup>lt;sup>15</sup>The precise condition for this to occur is that



Figure 1: Informed, uninformed and total order flow, prices and Kyle's lambda in the symmetric and asymmetric quadratic cost function examples: noise traders are net buyers.

 $\sigma$  produces a positive cross-sectional relation between efficiency and liquidity, and cross-sectional variation in  $\sigma_x$  or  $\psi$  produces a negative cross-sectional relation between efficiency and liquidity. Increases in  $\mu_x$  increase efficiency but have no effect on liquidity.

Even though the prior uncertainty is normal and thus the large trader's cumulative holdings are normally distributed, the endogenous terminal value of the stock in this example is lognormally distributed. The stock price follows a geometric Brownian motion process as in the Black and



Figure 2: Informed, uninformed and total order flow, prices and Kyle's lambda in the symmetric and asymmetric quadratic cost function examples: noise traders are net sellers.

Scholes (1973) model. The parameters of the process are endogenously determined by the primitives of the model (the signal-to-noise ratio  $\sigma_x/\sigma$ , the noise trader volatility, and the productivity parameter). This model is similar to the Kyle model with an exogenous lognormally distributed terminal value presented in Back (1992). As with an exogenous lognormal value, the percentage price impact is constant. However, the constant percentage price impact ('return impact') is not the same as when the value is exogenous. Indeed, in our model, price impact depends not only on the signal-to-noise ratio but also on the large trader's productivity. As discussed for price impact in Example 1, when the signal-to-noise ratio goes to zero, the percentage price impact in this example remains strictly greater than zero.

**Example 4** (Binary). This example is from Back, Li and Ljungqvist (2015). The model of activism is the same as that studied in the context of a single-period Kyle model by Maug (1998). The outcome is binary (success or failure). Success comes at an effort cost of c. The value of the stock is  $v_0$  in the absence of effort and  $v_0 + \Delta$  for a constant  $\Delta > 0$  if effort is exerted. It is optimal to exert effort if  $X_T \Delta \ge c$ . The value V(x) is a step function, equal to  $v_0$  for  $x < c/\Delta$  and equal to  $v_0 + \Delta$  for  $x \ge c/\Delta$ . Therefore, it is neither convex nor concave. In the equilibrium price and in Kyle's lambda, the parameter c appears only in the ratio  $c/\Delta$ . It is convenient to define  $\psi = \Delta/c$ , which is the value creation per unit cost. Then,  $\psi$  and  $\Delta$  measure the activist's productivity.

In this example, cross-sectional variation in either of the productivity parameters  $\psi$  or  $\Delta$  produces a negative cross-sectional relation between efficiency and liquidity, because higher productivity increases both efficiency and adverse selection. However, cross-sectional variation in either  $\mu_x$  or  $\sigma_x$  produces a negative cross-sectional relation between efficiency and liquidity if and only if  $\Delta \mu_x < c$ . The condition  $\Delta \mu_x < c$  means that the expected initial stake  $\mu_x$  is too small on its own to justify the cost of activism. In this case, a marginal increase in  $\mu_x$  increases adverse selection, because it moves the probability of activism from below 50% towards 50%. Hence, it reduces liquidity (while increasing economic efficiency). Also, when  $\Delta \mu_x < c$ , a marginal increase in  $\sigma_x$  increases economic efficiency, because it makes the expected initial stake  $\mu_x$  a less reliable predictor of the actual initial stake  $X_0$ . An increase in  $\sigma_x$  always reduces market liquidity in this example, so it causes liquidity and efficiency to move in opposite directions when  $\Delta \mu_x < c$ .

Cross-sectional variation in  $\sigma$  can produce either a negative or a positive cross-sectional relation between economic efficiency and market liquidity. There are four possible outcomes of a change in noise trading volatility, depending on the inequalities shown in Tables 2 and 3. The four possibilities are illustrated in Figure 3. An increase in the standard deviation  $\sigma$  of noise trading increases economic efficiency if and only if  $\Delta \mu_x < c$ . In that case, the large trader must on average acquire shares in the market to make activism worthwhile. An increase in noise trading volatility makes it easier to acquire the necessary shares. On the other hand, if the expected initial stake is high, higher noise trading volatility makes it easier for the large trader to unwind her stake and exit rather than incurring the cost to become active, so an increase in noise trading reduces economic efficiency. These are the effects described by Maug (1998).

However, unlike in Maug's one-period model, the effect of a change in noise trading volatility on market liquidity depends on the absolute size of the expected initial stake relative to a threshold (shown in Table 3) that depends on  $\sigma$ ,  $\sigma_x$ , and T. When the absolute expected initial stake is large, it is unlikely that the large trader will trade enough to change the profitability of activism: if  $\mu_x - c/\Delta$  is positive and large, it is unlikely that she will sell enough shares so that  $X_T < c/\Delta$ ; and if  $\mu_x - c/\Delta$  is negative and large in absolute value, it is unlikely that she will buy enough shares so that  $X_T > c/\Delta$ . Thus, Kyle's lambda is low—the market is highly liquid. In this circumstance, if noise trading increases, the probability that the large trader will trade out of an existing position or into a new position increases, and it increases so much that market liquidity actually falls.



Figure 3: Effect of an Increase in Liquidity Trading in the Binary Example. Increasing noise trading increases economic efficiency  $(\partial \overline{P}/\partial \sigma > 0)$  when  $\mu_x > c/\Delta$  and reduces economic efficiency when  $\mu_x < c/\Delta$ . Increasing noise trading increases market liquidity  $(\partial \overline{\lambda}/\partial \sigma < 0)$  when  $|\mu_x - c/\Delta|$  is below a threshold depending on  $\sigma$  that is specified in Table 3 and reduces market liquidity when  $|\mu_x - c/\Delta|$  is above the threshold. In this example,  $\sigma_x = 0.1$  and T = 1.

The equilibrium price in this example is the base value  $v_0$  plus the value  $\Delta$  of activism multi-

plied by the conditional probability that activism will occur. Activism occurs if and only if

$$Y_T \ge \frac{c/\Delta - \mu_x}{\Lambda} \,.$$

Market makers compute the probability of activism at each date t based on  $Y_T$  being normally distributed with mean  $Y_t$  and standard deviation  $\sigma\sqrt{T-t}$ . Equations (13) and (16) imply

$$Y_T \ge \frac{c/\Delta - \mu_x}{\Lambda} \quad \iff \quad X_T \ge \frac{c}{\Delta}$$
$$\iff \quad Z_T \le X_0 - \mu_x + \frac{\Lambda - 1}{\Lambda} \left( \mu_x - \frac{c}{\Delta} \right) \,.$$

Of course, the condition  $X_T \ge c/\Delta$  is necessary and sufficient for exerting effort to be optimal for the large trader. The last condition shows that the large trader exerts effort if and only if noise traders sell enough shares (or do not buy too many shares). Selling by noise traders makes the asset cheaper for the large trader and hence induces her to buy shares and become active.

**Example 5** (Probabilistic Binary). Many activist campaigns have a specific objective, and the outcome can be expressed as success or failure. For example, activists may attempt to block a merger, to force a company to be put up for sale, to oust a CEO, to remove anti-takeover provisions, to initiate a dividend, etc. However, it may be unrealistic to assume, as in Example 4, that the amount of effort required to achieve success is known. To capture uncertainty about the outcome, Example 5 models success as a random event, the probability of which depends on the large trader's effort. Because the large trader is risk neutral, she cares about the expected asset value, which is  $v_0 + \Delta p$ , where p denotes the probability of success. Thus, instead of modeling the stock value v as being either  $v_0$  or  $v_0 + \Delta$ , we model it as being  $v_0 + \Delta p$ , where p ranges continuously between 0 and 1. Assume that the cost of achieving a probability of success equal to p is

$$c[p + (1 - p)\log(1 - p)]$$

for a constant c > 0. Therefore, the cost of achieving an expected asset value equal to v is

$$C(v) = \begin{cases} \infty & \text{if } v < v_0, \\ c \left[ \frac{v - v_0}{\Delta} + \left( 1 - \frac{v - v_0}{\Delta} \right) \log \left( 1 - \frac{v - v_0}{\Delta} \right) \right] & \text{if } v_0 \le v < v_0 + \Delta, \\ \infty & \text{if } v \ge v_0 + \Delta. \end{cases}$$

The large trader's optimal effort implies a probability of success of  $1 - e^{-\Delta X_T/c}$ . The value V(x) is bounded below (by  $v_0$ ) and bounded above (by  $v_0 + \Delta$ ) and hence is neither uniformly convex nor uniformly concave. As in Example 4, the cost parameter c appears in the equilibrium price and in Kyle's lambda only through the ratio  $c/\Delta$ . As in Example 4, define  $\psi = \Delta/c$ , so the activist's productivity is measured by  $\psi$  and  $\Delta$ .

The equilibrium is described in Table 1 in terms of

$$d_1 \stackrel{\text{def}}{=} \frac{\mu_x + \Lambda y}{\Lambda \sigma \sqrt{T - t}}$$

and

$$d_2 \stackrel{\text{def}}{=} d_1 - \psi \Lambda \sigma \sqrt{T - t} \,.$$

The comparative statics are described in Tables 2 and 3 in terms of  $\overline{d}_1$ , which is  $d_1$  with t = y = 0and  $\overline{d}_2 = \overline{d}_1 - \psi \Lambda \sigma \sqrt{T}$ , and in terms of  $\mu_x^*$  defined as follows. Set g(x) = N(x)/n(x). It is well known that g is a strictly increasing function that maps the real line onto the positive reals. Define

$$\mu_x^* = \psi \Lambda^2 \sigma^2 T + \Lambda \sigma \sqrt{T} g^{-1} \left( \frac{1}{\psi \Lambda \sigma \sqrt{T}} \right) \,.$$

The comparative statics in this example are very similar to those in Example 4. There are only three differences. First, the condition  $\mu_x < c/\Delta$  that determines some of the signs in Example 4 is replaced by  $\mu_x < \mu_x^*$ . Second, the condition that determines when an increase in  $\sigma$  increases market liquidity takes different forms in the two examples. Third, an increase in the productivity parameter  $\psi$  does not always reduce market liquidity in this example. The condition under which it reduces market liquidity is shown in Table 3.

# 7. Conclusion

This paper revisits the classic question of the relation between liquidity and economic efficiency. We develop a dynamic version of the Kyle model in which an activist trader can affect the liquidation value of the firm by expending costly effort. We allow for a general activism technology and show that its specification is paramount to the relation between efficiency and liquidity. For example, the well-known 'lock-in' intuition developed in prior literature (e.g., Coffee (1991), Bhide (1993), Maug (1998)) does not extend to the case where the activist's optimal effort is convex in her accumulated stake. In that case, an increase in noise trading volatility leads to higher expected effort regardless of the activist's initial stake size. We present several examples to show how the relation between liquidity and economic efficiency depends on the type of activism technology.

The model also shows that, when the value of the firm depends on the activist's effort, an increase in noise trading volatility does not necessarily improve market liquidity and reduce price impact. This is because more uncertainty about noise trading induces more uncertainty about the activist's position and hence about her future effort, thus potentially increasing adverse selection risk. As a result, changes in noise trading have potentially ambiguous effects not only on economic efficiency but also on market liquidity.

In the model, both market liquidity (measured by price impact) and effort level are jointly determined. Their cross-sectional relation thus depends on which underlying model parameter is varied. In addition to noise trading volatility, we also analyze how changing the activist's productivity and the prior uncertainty about her position affects the cross-sectional relation between market liquidity and economic efficiency.

These comparative statics are important because they inform policy makers about the potential effects of policy changes on economic efficiency and market liquidity. Several policy changes that are currently under consideration can be linked directly to the parameters of the model. For example, a Tobin tax on stock transactions might reduce noise trading. A change in the number of days between crossing the 5% threshold and filing a Schedule 13D report can be viewed as a change in noise trading.<sup>16</sup> A change in disclosure rules more generally could change the uncertainty about the

<sup>&</sup>lt;sup>16</sup>This is because in the model, what matters is  $\sigma^2 T$ —the cumulative amount of noise trading over the entire trading period. So from the perspective of the large trader, reducing the trading horizon T is isomorphic to reducing noise trading volatility and keeping T fixed.

large trader's initial position. Finally, regulatory changes that would affect the ability to engage in activism (e.g., any new regulation that makes it more or less costly to take over firms or to challenge patents) can be interpreted as changes in the productivity parameters.

## Appendix A. The one-period model

Here, we consider the one-period model where the large trader starts with some position  $X_0$ , known only to her, and trades once to choose  $X_1 = X_0 + \theta$  so as to maximize her objective function

$$\mathsf{E}[G(X_1) - \theta P(Y) \,|\, X_0, Z] \,. \tag{A.1}$$

Recall that

$$G(x) \stackrel{\text{def}}{=} \sup_{v} \left\{ vx - C(v) \right\},$$

and that the supremum on the right-hand side is attained by  $V(x) = C'^{-1}(x) = \partial G(x)/\partial x$ . Further, the competitive market makers have a prior that  $X_0 \sim N(\mu_x, \sigma_x)$  and observe total order flow  $Y = \theta + Z$  where noise trading  $Z \sim N(0, \sigma^2)$ . For simplicity, we assume that  $X_0$  and Z are uncorrelated. The zero-profit condition for market makers implies that the price satisfies

$$P(Y) = \mathsf{E}[V(X_1) \,|\, Y].$$

Note that we assume that not only  $X_0$  but also Z is observed by the large trader when she chooses her optimal trading decision. As pointed out by Rochet and Vila (1994), this simplifies the analysis and is consistent with the continuous time model, where in equilibrium the large trader effectively observes noise trades. We point out in an example below how making the alternative assumption, that the large trader chooses her trades before observing Z, affects the equilibrium.

Assuming that the large trader conditions on both  $X_0$  and Z, her first-order condition is simply:

$$V(X_0 + \theta) - P(\theta + Z) - \theta P'(\theta + Z) = 0.$$
(A.2)

The second-order condition is:

$$V'(X_0 + \theta) - 2P'(\theta + Z) - \theta P''(\theta + Z) \le 0.$$
 (A.3)

This FOC defines an optimal trading strategy for the large trader  $\theta(X_0, Z)$  given an equilibrium pricing function  $P(\cdot)$ . In turn, given a conjectured optimal trading strategy of the form  $\theta(X_0, Z)$ ,

the equilibrium pricing function is given by:

$$P(y) = \mathsf{E}[V(X_0 - Z + y) | \theta(X_0, Z) + Z = y].$$
(A.4)

An equilibrium is then a pair of functions  $(\theta(x, z), P(y))$  that satisfy the three equations (A.2)-(A.4).

By using (A.2) in (A.4), we see that a necessary condition for an equilibrium is that the trading strategy be inconspicuous, that is,

$$0 = \mathsf{E}[\theta(X_0, Z) \,|\, Y] \,.$$

We now illustrate how to derive the equilibrium explicitly in the simplest case where V(x) is linear.

Appendix A.1. The Linear V(x) Case

Assume  $C(v) = \frac{v^2}{2\psi}$ . Then,  $V(x) = \psi x$ . To solve for an equilibrium in this case, we guess that  $P(y) = \psi(p_0 + \Lambda y)$ . Then the FOC gives:

$$\theta = (X_0 - p_0 - \Lambda Z)/(2\Lambda - 1)$$

. Note that since  $\theta$  is inconspicuous, we can restrict ourselves to  $p_0 = \mu_x$ . The SOC is satisfied if

$$2\Lambda - 1 > 0.$$

Conversely, if we conjecture that the activist chooses a linear trading rule of the form  $\theta = \beta_x(X_0 - \mu_x) + \beta_z Z$ , we have

$$P(y) = \psi y + \psi \mathsf{E}[X_0 - Z \,|\, \beta_x (X_0 - \mu_x) + (\beta_z + 1)Z = y] = \psi(\mu_x + \Lambda y),$$

where  $\Lambda$  is given by

$$\Lambda = 1 + \frac{\beta_x \sigma_x^2 - (\beta_z + 1)\sigma_z^2}{\beta_x^2 \sigma_x^2 + (\beta_z + 1)^2 \sigma_z^2}$$

This follows from the linear projection theorem for Gaussian random variables:

$$\mathsf{E}[X_0 - Z \,|\, \beta_x (X_0 - \mu_x) + (\beta_z + 1)Z = y] = \mu_x + \frac{\beta_x \sigma_x^2 - (\beta_z + 1)\sigma_z^2}{\beta_x^2 \sigma_x^2 + (\beta_z + 1)^2 \sigma_z^2} y \,.$$

It follows that an equilibrium exists if there is a solution  $\Lambda$  that satisfies the SOC and the equation:

$$\Lambda = 1 + \frac{\beta_x \sigma_x^2 - (\beta_z + 1)\sigma_z^2}{\beta_x^2 \sigma_x^2 + (\beta_z + 1)^2 \sigma_z^2},$$
(A.5)

where  $\beta_x, \beta_z$  are given by:

$$\beta_x = \frac{1}{2\Lambda - 1},$$
$$\beta_z = \frac{-\Lambda}{2\Lambda - 1}.$$

There is one unique solution that satisfies the SOC, given by:

$$\Lambda = \frac{1}{2} \left( 1 + \sqrt{1 + 4\frac{\sigma_x^2}{\sigma_v^2}} \right) \,.$$

Appendix A.2. The Linear V(x) Case When Z Is Not Known to the Activist

When the large trader cannot condition her trading decision on Z because it is unknown to her at the time of trading, she chooses  $X_1 = X_0 + \theta$  so as to maximize her objective function:

$$\mathsf{E}[G(X_1) - \theta P(Y) | X_0].$$
 (A.6)

Her first-order condition (A.2) is replaced by:

$$V(X_0 + \theta) - \mathsf{E}[P(\theta + Z) - \theta P'(\theta + Z) | X_0] = 0, \tag{A.7}$$

and the second-order condition becomes:

$$V'(X_0 + \theta) - \mathsf{E}[2P'(\theta + Z) - \theta P''(\theta + Z) | X_0] \le 0.$$
(A.8)

Since the market makers' zero-profit condition is unchanged, an equilibrium has to satisfy (A.4) above as well. So, an equilibrium in this case will be a trading strategy  $\theta(X_0)$  that satisfies

equation (A.7) given a pricing function P(y) that satisfies

$$P(y) = \mathsf{E}[V(X_0 - Z + y) | \theta(X_0) + Z = y].$$

Consider the linear case, where  $V(x) = \psi x$ . As before, it is natural to conjecture that  $P(y) = \psi(p_0 + \Lambda y)$ . The FOC then gives:

$$\theta = \frac{X_0 - p_0}{2\Lambda - 1}.\tag{A.9}$$

Furthermore, because of the linearity of V(x), the FOC and equilibrium condition immediately imply that the trading strategy should be inconspicuous, that is,  $p_0 = \mu_x$ . Note that for cost functions where V(x) is not linear, the inconspicuousness of the trading strategy is no longer implied by the FOC when Z is not observed by the large trader.

If  $\theta = \beta_x (X_0 - \mu_x)$ , then the price function is:

$$P(y) = \psi y + \psi \mathsf{E}[X_0 - Z \mid \beta(X_0 - \mu_x) + Z = y] = \psi \Lambda y$$

with

$$\Lambda = \frac{\beta_x \sigma_x^2 - \sigma_z^2}{\beta_x^2 \sigma_x^2 + \sigma_z^2}$$

Thus, an equilibrium is a solution for  $\Lambda$  that satisfies this equation (and the SOC) with

$$\beta_x = \frac{1}{2\Lambda - 1} \,.$$

There is a unique equilibrium given by:

$$\Lambda = \frac{1}{2} (1 + \frac{\sigma_x}{\sigma_v}). \tag{A.10}$$

### Appendix A.3. Discussion

In general, unlike in the continuous time model, we do not know how to solve for the equilibrium explicitly for general cost functions outside the simple linear case. We can, however, prove that the linear trading strategy is not optimal in general. That is, unlike in the continuous time model, it is not optimal to adopt the same linear strategy of the form  $\theta = \beta_x(X_0 - \mu_x) + \beta_z Z$  for all convex cost functions C(v). Indeed, suppose the activist adopts such a trading strategy. Then, the market makers' zero-profit condition implies that

$$P(y) = \mathsf{E}[V(X_0 - Z + y) | \beta_x(X_0 - \mu_x) + (1 + \beta_z)Z = y] = \int V(u + y) \, \mathrm{n}\left(\frac{u - M(y)}{\sqrt{\Omega}}\right) \, \mathrm{d}u \,,$$

where n(x) is the standard Gaussian density and

$$M(y) = \mathsf{E}[X_0 - Z \mid \beta_x (X_0 - \mu_x) + (1 + \beta_z)Z = y] = \mu_x + (\Lambda - 1)y$$
(A.11)

$$\Omega = \mathsf{V}[X_0 - Z \,|\, \beta_x (X_0 - \mu_x) + (1 + \beta_z) Z = y]. \tag{A.12}$$

For this to be an equilibrium, the FOC should be satisfied:

$$V(X_0 + \theta) - P(\theta + Z) - \theta P'(\theta + Z) = 0.$$
(A.13)

Consider, for example, the exponential case  $V(x) = v_0 e^{\psi x}$ . Then  $P(y) = v_0 e^{\psi(\mu_x + \Lambda y) + \frac{\psi^2}{2}\Omega}$ . In order for the FOC to hold, we need:

$$e^{\psi(X_0+\theta)} - e^{\psi(\mu_x + \Lambda(\theta+Z)) + \frac{\psi^2}{2}\Omega} - \theta\psi\Lambda e^{\psi(\mu_x + \Lambda(\theta+Z)) + \frac{\psi^2}{2}\Omega} = 0,$$
(A.14)

or equivalently:

$$e^{\psi(X_0 - \mu_x - \frac{\psi}{2}\Omega) + \psi\theta(1 - \Lambda) - \psi\Lambda Z} - 1 - \theta\psi\Lambda = 0.$$
(A.15)

Clearly,  $\theta$  cannot be linear in  $X_0$  and Z.

# Appendix B. Proof of Theorem 1

We need to verify the optimality of the trading strategy (10). As in Section 4, define  $h(z) = V(\mu_x + \Lambda z)$ . Define

$$g(x,y) = \sup_{\overline{y}} \int_{y}^{\overline{y}} \left( V(x-y+z) - h(z) \right) \, \mathrm{d}z \,.$$

Because  $\overline{y} = y$  is feasible in this optimization problem, we have  $g(x,y) \ge 0$  for all (x,y). The solution to the optimization problem is given by the first-order condition  $V(u + \overline{y}^*(u)) = h(\overline{y}^*(u))$  as

$$\overline{y}^*(u) = \frac{u - \mu_x}{\Lambda - 1} \,. \tag{B.1}$$

Thus,

$$g(x,y) = \int_{y}^{\overline{y}^{*}(x-y)} \left( V(x-y+z) - h(z) \right) \, \mathrm{d}z \,. \tag{B.2}$$

Substituting the definition of h and V = G' in (B.2), it is straightforward to calculate that

$$g(x,y) = \frac{\Lambda - 1}{\Lambda} G\left(\frac{\Lambda(x-y) - \mu_x}{\Lambda - 1}\right) + \frac{1}{\Lambda} G(\mu_x + \Lambda y) - G(x).$$
(B.3)

This implies that

$$g_x(x,y) = V\left(\frac{\Lambda(x-y) - \mu_x}{\Lambda - 1}\right) - V(x),$$
  
$$g_y(x,y) = V(\mu_x + \Lambda y) - V\left(\frac{\Lambda(x-y) - \mu_x}{\Lambda - 1}\right).$$

Thus,

$$g_x(x,y) + g_y(x,y) = V(\mu_x + \Lambda y) - V(x) = h(y) - V(x).$$
 (B.4)

Furthermore, the monotonicity of V implies that  $g_x$  and  $g_y$  are bounded on bounded rectangles.

Define

$$J(T, x, y) = G(x) + g(x, y)$$

and, for t < T, set

$$J(t, x, y) = G(x) + \mathsf{E}[g(x, y + Z_T - Z_t) \mid \mathcal{F}_t^Z].$$
(B.5)

From this definition and (B.3), we see that J is as stated in (11). Because  $g_x$  and  $g_y$  are bounded on bounded rectangles, we can use the bounded convergence theorem to justify interchanging differentiation and expectation and thereby obtain

$$\begin{split} J_x(t,x,y) &= \mathsf{E}\left[ \left. V\left( \frac{\Lambda(x-y-Z_T+Z_t)-\mu_x}{\Lambda-1} \right) \right| \mathcal{F}_t^Z \right] \,, \\ J_y(t,x,y) &= \mathsf{E}\left[ \left. V\left( \mu_x + \Lambda(y+Z_T-Z_t)) \right| \mathcal{F}_t^Z \right] - \mathsf{E}\left[ \left. V\left( \frac{\Lambda(x-y-Z_T+Z_t)-\mu_x}{\Lambda-1} \right) \right| \mathcal{F}_t^Z \right] \right. \\ &= P(t,y) - \mathsf{E}\left[ \left. V\left( \frac{\Lambda(x-y-Z_T+Z_t)-\mu_x}{\Lambda-1} \right) \right| \mathcal{F}_t^Z \right] \,. \end{split}$$

Thus,

$$J_x(t, x, y) + J_y(t, x, y) = P(t, y).$$
(B.6)

Furthermore,

$$J(t, x, Z_t) = G(x) + \mathsf{E}[g(x, Z_T) \mid \mathcal{F}_t^Z]$$

which is an  $\mathbb{F}^Z$  martingale. Applying Itô's formula and equating the drift to zero gives

$$J_t(t, x, y) + \frac{1}{2}\sigma^2 J_{yy}(t, x, y) = 0.$$
(B.7)

Consider an arbitrary trading strategy. Using Itô's formula and substituting (B.6) and (B.7), we obtain

$$J(T, X_T, Y_T) = J(0, X_0, Y_0) + \int_0^T dJ$$
  
=  $J(0, X_0, Y_0) + \int_0^T P(t, Y_t) \theta_t \, dt + \int_0^T J_y(t, X_t, Y_t) \, dZ_t$ 

The no-doubling conditions (6) and (7) ensure that

$$\mathsf{E} \int_0^T J_y(t, X_t, Y_t) \, \mathrm{d}Z_t = 0$$

Therefore, rearranging and taking expectations yields

$$J(0, X_0, 0) = \mathsf{E}\left[J(T, X_T, Y_T) - \int_0^T P(t, Y_t)\theta_t \, \mathrm{d}t\right] \,.$$

Because  $g \ge 0$ , we have  $J(T, X_T, Y_T) \ge G(X_T)$ . Hence,

$$J(0, X_0, 0) \ge \mathsf{E}\left[G(X_T) - \int_0^T P(t, Y_t)\theta_t \,\mathrm{d}t\right] \,. \tag{B.8}$$

This shows that  $J(0, X_0, 0)$  is an upper bound on the large trader's expected value. The bound is achieved by a strategy if and only if  $g(X_T, Y_T) = 0$ .

Now consider the strategy (10). For this strategy,  $V(X_T) = h(Y_T)$  which implies (from the definition of  $\overline{y}^*$  in equation (B.1)) that  $\overline{y}^*(X_T - Y_T) = Y_T$  and  $g(X_T, Y_T) = 0$ . Thus, the strategy is optimal.

## Appendix C. Proof of the Lemma

Define  $U_t = a\varepsilon - bZ_t$ . We use filtering to establish the proposition. As is customary, we use the symbol  $\hat{}$  to denote conditional expectations given  $\mathcal{F}_t^Y$ . We want to compute  $\hat{U}_t$ . Let  $\Sigma(t)$ denote the conditional variance of  $U_t$  given  $\mathcal{F}_t^Y$ . We have  $U_0 = a\varepsilon$ ,  $\hat{U}_0 = 0$ , and  $\Sigma(0) = a^2$ . The stochastic process U evolves as

$$\mathrm{d}U_t = -b\,\mathrm{d}Z_t\,.$$

The observation process is Y, and

$$\mathrm{d}Y_t = \frac{1}{T-t} U_t \,\mathrm{d}t - \frac{b+1}{T-t} Y_t \,\mathrm{d}t + \mathrm{d}Z_t \,.$$

The innovation process is W defined by  $W_0 = 0$  and

$$dW_t = \frac{1}{\sigma} \left( dY_t - \frac{1}{T - t} \hat{U}_t \, dt + \frac{b + 1}{T - t} Y_t \, dt \right) = \frac{1}{\sigma} \left( \frac{1}{T - t} (U_t - \hat{U}_t) \, dt + dZ_t \right) \,. \tag{C.1}$$

From Kallianpur (1980, Equation 10.5.9), the filtering equation is

$$d\hat{U}_t = \frac{1}{\sigma} \left( \frac{\Sigma(t)}{T-t} - b\sigma^2 \right) dW_t \,. \tag{C.2}$$

From Kallianpur (1980, Equation 10.5.10), the conditional variance evolves as

$$\frac{\mathrm{d}\Sigma(t)}{\mathrm{d}t} = -\frac{\Sigma(t)^2}{(T-t)^2\sigma^2} + \frac{2b\Sigma(t)}{T-t}\,.\tag{C.3}$$

The ODE (C.3) with initial condition  $\Sigma(0) = a^2$  is satisfied by  $\Sigma(t) = (T - t)a^2/T$ . For this function  $\Sigma(\cdot)$ , the left-hand side of (C.3) is  $-a^2/T$ , and the right-hand side is

$$-\frac{a^4}{\sigma^2 T^2} + \frac{2ba^2}{T} = -\frac{a^2}{T} \left(\frac{a^2}{\sigma^2 T} - 2b\right) = -\frac{a^2}{T},$$

using the definition  $a = \sigma \sqrt{(2b+1)T}$  for the last equality. Thus, the conditional variance of  $U_t$  is  $(T-t)a^2/T$ . Consequently, the filtering equation (C.2) simplifies to

$$\mathrm{d}\hat{U}_t = \frac{1}{\sigma} \left( \frac{a^2}{T} - b\sigma^2 \right) \,\mathrm{d}W_t = (b+1)\sigma \mathrm{d}W_t \,,$$

using the definition  $a = \sigma \sqrt{(2b+1)T}$  again for the last equality. Because  $\hat{U}_0 = W_0 = 0$ , this equation implies that  $\hat{U} = (b+1)\sigma W$ . Equation (C.1) for the innovation process now becomes

$$\mathrm{d}W_t = \frac{1}{\sigma} \left( \mathrm{d}Y_t + \frac{b+1}{T-t} (Y_t - \sigma W_t) \,\mathrm{d}t \right) \,.$$

This equation is satisfied by  $W = Y/\sigma$ . Thus,  $Y/\sigma$  is the innovation process. The innovation process is a standard Brownian motion on  $\mathbb{F}^Y$ , so Y is a Brownian motion with standard deviation  $\sigma$  on  $\mathbb{F}^Y$ . Moreover, we have

$$\mathrm{d}\hat{U}_t = (b+1)\sigma\mathrm{d}W_t = (b+1)\mathrm{d}Y_t\,,$$

so  $\hat{U}_t = (b+1)Y_t$ . Because Y is a Brownian motion on  $\mathbb{F}^Y$ , the limit  $Y_T = \lim_{t \to T} Y_t$  exists almost surely, and we have  $U_T = (b+1)Y_T$ , which is the same as (15).

# Appendix D. Proof of Theorem 2

First, we establish the comparative statics of economic efficiency. From (19), we have

$$\overline{P} = \mathsf{E}\left[V\left(\mu_x + \Lambda\sigma\sqrt{T}\varepsilon\right)\right],$$

where  $\varepsilon$  is a standard normal variable. It follows (since  $V'(x) \ge 0 \ \forall x$ ) that

$$\frac{\partial \overline{P}}{\mu_x} = \mathsf{E}\left[V'\left(\mu_x + \Lambda \sigma \sqrt{T}\epsilon\right)\right] \ge 0\,.$$

If V is convex, then, for all  $\epsilon \in (-\infty, \infty)$ , we have

$$\Lambda \sigma \sqrt{T} \epsilon V' \left( \mu_x + \Lambda \sigma \sqrt{T} \epsilon \right) \ge V \left( \mu_x + \Lambda \sigma \sqrt{T} \epsilon \right) - V(\mu_x) \,.$$

If V is concave, then we have the opposite inequality. Also, from the definition of  $\Lambda$ ,

$$\Lambda \sigma \sqrt{T} = \sigma \sqrt{T} + \sqrt{\sigma^2 T + \sigma_x^2} \,,$$

which is an increasing function of  $\sigma$  and also an increasing function of  $\sigma_x$ . Thus, when V is convex,

$$\begin{split} \frac{\partial \overline{P}}{\partial \sigma} &= \mathsf{E}\left[\epsilon V'\left(\mu_x + \Lambda \sigma \sqrt{T}\epsilon\right)\right] \left(\frac{\partial (\Lambda \sigma \sqrt{T})}{\partial \sigma}\right) \\ &\geq \frac{1}{\Lambda \sigma \sqrt{T}} \mathsf{E}\left\{\left[V\left(\mu_x + \Lambda \sigma \sqrt{T}\epsilon\right) - V(\mu_x)\right]\right\} \left(\frac{\partial (\Lambda \sigma \sqrt{T})}{\partial \sigma}\right) \geq 0\,, \end{split}$$

where the last inequality follows by Jensen's inequality. When V is concave, we obtain the opposite inequality. The same reasoning produces the results for  $\partial \overline{P} / \partial \sigma_x$ .

Now we establish the comparative statics of market liquidity. From (20),

$$\overline{\lambda} = \Lambda \int_{-\infty}^{+\infty} V' \left( \mu_x + \Lambda \sigma \sqrt{T} \epsilon \right) \mathbf{n}(\epsilon) \, \mathrm{d}\epsilon \, .$$

It follows that

$$\frac{\partial \overline{\lambda}}{\partial \mu_x} = \Lambda \int_{-\infty}^{+\infty} V'' \left( \mu_x + \Lambda \sigma \sqrt{T} \epsilon \right) \mathbf{n}(\epsilon) \, \mathrm{d}\epsilon$$

So,  $\partial \overline{\lambda} / \partial \mu_x \ge 0$  if V is convex, and the opposite inequality holds if V is concave. Furthermore,

$$\frac{\partial \overline{\lambda}}{\partial \sigma_x} = \frac{\partial \Lambda}{\partial \sigma_x} \int_{-\infty}^{+\infty} \left\{ V'\left(\mu_x + \Lambda \sigma \sqrt{T}\epsilon\right) + V''\left(\mu_x + \Lambda \sigma \sqrt{T}\epsilon\right) \Lambda \sigma \sqrt{T}\epsilon \right\} \mathbf{n}(\epsilon) \,\mathrm{d}\epsilon \,.$$

Note that

$$\int_{-\infty}^{+\infty} V'' \left( \mu_x + \Lambda \sigma \sqrt{T} \epsilon \right) \Lambda \sigma \sqrt{T} \epsilon \mathsf{n}(\epsilon) \, \mathrm{d}\epsilon = \int_{-\infty}^{+\infty} \epsilon \mathsf{n}(\epsilon) \frac{\mathrm{d}V' \left( \mu_x + \Lambda \sigma \sqrt{T} \epsilon \right)}{\mathrm{d}\epsilon} \, \mathrm{d}\epsilon \, .$$

Using this fact, integration by parts, and assumption (\*), we obtain

$$\int_{-\infty}^{+\infty} V'' \left( \mu_x + \Lambda \sigma \sqrt{T} \epsilon \right) \Lambda \sigma \sqrt{T} \epsilon \mathsf{n}(\epsilon) \, \mathrm{d}\epsilon = -\int_{-\infty}^{+\infty} V' \left( \mu_x + \Lambda \sigma \sqrt{T} \epsilon \right) \frac{\mathrm{d}[\epsilon \mathsf{n}(\epsilon)]}{\mathrm{d}\epsilon} \, \mathrm{d}\epsilon$$
$$= \int_{-\infty}^{+\infty} V' \left( \mu_x + \Lambda \sigma \sqrt{T} \epsilon \right) [\epsilon^2 - 1] \mathsf{n}(\epsilon) \, \mathrm{d}\epsilon.$$

Thus:

$$\frac{\partial \overline{\lambda}}{\partial \sigma_x} = \frac{\partial \Lambda}{\partial \sigma_x} \int_{-\infty}^{+\infty} V' \left( \mu_x + \Lambda \sigma \sqrt{T} \epsilon \right) \epsilon^2 \mathbf{n}(\epsilon) \,\mathrm{d}\epsilon \,.$$

Since  $V'(x) \ge 0 \ \forall x \text{ and } \partial \Lambda / \partial \sigma_x > 0$  it follows that

$$\frac{\partial \overline{\lambda}}{\partial \sigma_x} \ge 0 \,.$$

# References

- Admati, A.R., Pfleiderer, P., 2009. The 'Wall Street walk' and shareholder activism: Exit as a form of voice. Review of Financial Studies 22, 2645–2685.
- Back, K., 1992. Insider trading in continuous time. Review of Financial Studies 5, 387–409.
- Back, K., Li, T., Ljungqvist, A., 2015. Liquidity and governance. Working Paper.
- Bebchuk, L., Brav, A., Jackson, R., Jiang, W., 2013. Pre-disclosure accumulations by activist investors: Evidence and policy. Journal of Corporation Law 39.
- Bhide, A., 1993. The hidden costs of stock market liquidity. Journal of Financial Economics 34, 31–51.
- Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81, 637–659.
- Bris, A., 2002. Toeholds, takeover premium, and the probability of being acquired. Journal of Corporate Finance 8, 227–253.
- Cho, K.H., 2003. Continuous auctions and insider trading: Uniqueness and risk aversion. Finance & Stochastics 7, 47–71.
- Coffee, Jr, J.C., 1991. Liquidity versus control: The institutional investor as corporate monitor. Columbia Law Review 91, 1277–1368.
- Collin-Dufresne, P., Fos, V., 2015. Shareholder activism, informed trading, and stock prices. Working Paper.
- DeMarzo, P.M., Urošević, B., 2006. Ownership dynamics and asset pricing with a large shareholder. Journal of Political Economy 114, 774–815.
- Edmans, A., 2009. Blockholder trading, market efficiency and managerial myopia. Journal of Finance 64, 2481–2513.
- Edmans, A., Manso, G., 2011. Governance through trading and intervention: A theory of multiple blockholders. Review of Financial Studies 24, 2395–2428.

- Fishman, M.J., Hagerty, K.M., 1992. Insider trading and the efficiency of stock prices. The RAND Journal of Economics , 106–122.
- Fishman, M.J., Hagerty, K.M., 1995. The mandatory disclosure of trades and market liquidity. Review of Financial Studies 8, 637–676.
- Jiang, W., Li, T., Mei, D., 2016. Influencing control: Jawboning in risk arbitrage. Working paper.
- Kahn, C., Winton, A., 1998. Ownership structure, speculation, and shareholder intervention. Journal of Finance 53, 99–129.
- Kallianpur, G., 1980. Stochastic Filtering Theory. Springer-Verlag, New York.
- Kyle, A., 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1335.
- Kyle, A.S., Vila, J.L., 1991. Noise trading and takeovers. RAND Journal of Economics 22, 54-71'.
- Maug, E., 1998. Large shareholders as monitors: Is there a trade-off between liquidity and control? The Journal of Finance 53, 65–98.
- Merton, R.C., 1973. Theory of rational option pricing 4, 141–183.
- Noe, T.H., 2002. Investor activism and financial market structure. Review of Financial Studies 15, 289–318.
- Ravid, S.A., Spiegel, M., 1999. Toehold strategies, takeover laws and rival bidders. Journal of Banking and Finance 23, 1219–1242.

Table 1: Equilibrium in Five Examples. The cost functions C(v) and productivity parameters  $\psi$  and  $\Delta$  are defined in Section 6. The parameters  $\mu_x$ ,  $\sigma_x$ ,  $\sigma$ , and  $\Lambda$  are defined in Section 3. The functions G, V, h, P, and  $\lambda$  are calculated from the cost function as explained in Sections 3 and 4.

2. Asymmetric Quadratic Cost	$\begin{cases} \infty & \text{if } v < v_0 \end{cases}$	$(v - v_0)^2/(2\psi)$ if $v \ge v_0$	$v_0 x + \psi(x^+)^2/2$	$v_0 + \psi x^+$	$v_0 + \psi \left( \mu_x + \Lambda y \right)^+$	$v_0 + \psi(\mu_x + \Lambda y) N\left(rac{\mu_x + \Lambda y}{\Lambda\sigma\sqrt{T-t}} ight) + \psi\Lambda\sigma\sqrt{T-t}n\left(rac{\mu_x + \Lambda y}{\Lambda\sigma\sqrt{T-t}} ight)$	$\psi \Lambda N \left( rac{\mu x + \Lambda y}{\Lambda \sigma \sqrt{T-t}}  ight)$								5. Probabilistic Binary	$\int \infty \qquad \qquad \text{if } v < v_0 ,$	$\left  \left\{ z \left[ \frac{v - v_0}{\Delta} + \left( 1 - \frac{v - v_0}{\Delta} \right) \log \left( 1 - \frac{v - v_0}{\Delta} \right) \right] \right. \text{ if } v_0 \le v < v_0 + \Delta$	$\lim_{n \to \infty} v \ge v_0 + \Delta.$	$\int v_0 x \qquad \qquad \text{if } x < 0$	$\Big\{(v_0+\Delta)x-z\left(1-\mathrm{e}^{-\Delta x/z} ight) \hspace{0.2cm} ext{if}\hspace{0.2cm} x\geq 0$	$\int v_0 \qquad \text{if } x < 0$	$\Big( v_0 + \Delta \left( 1 - \mathrm{e}^{-\Delta x/z}  ight) ~~\mathrm{if}~x \geq 0$	$\int v_0 \qquad \qquad \text{if } y < -\mu_x/\Lambda$	$\Big[ v_0 + \Delta \left( 1 - \mathrm{e}^{-\Delta x/z}  ight)  \mathrm{if} \; y \geq -\mu_x/\Lambda$	$v_0 + \Delta N(d_1) - \Delta N(d_2) \mathrm{e}^{-\psi(\mu_x + \Lambda y) + \psi^2 \Lambda^2 \sigma^2(T-t)/2}$	i + i + i + i + i + i + i + i + i + i +
1. Quadratic Cost	$(v-v_0)^2/(2\psi)$		$v_0x + \psi x^2/2$	$v_0 + \psi x$	$v_0 + \psi \mu_x + \psi \Lambda y$	$v_0 + \psi \mu_x + \psi \Lambda y$	$V \phi$	3. Exponential	$rac{1}{\psi} v \log\left(rac{v}{v_0} ight) - rac{1}{\psi} v  .$	$v_0(\mathrm{e}^{\psi x}-1)/\psi$	$v_0 e^{\psi x}$	$v_0 e^{\psi(\mu_x + \Lambda y)}$	$v_0 \mathrm{e}^{\psi\left(\mu_x + \Lambda y + rac{1}{2}\Lambda^2 \sigma^2(T-t) ight)}$	$\psi \Lambda P(t,y)$	4. Binary	$\int \infty  \text{if } v \notin \{v_0, v_0 + \Delta\}$	$\begin{cases} 0 & \text{if } v = v_0 \end{cases}$	$\begin{bmatrix} c & \text{if } v = v_0 + \Delta \end{bmatrix}$	$\int v_0 x \qquad \text{if } x < c/\Delta$	$\left( (v_0 + \Delta)x - c  \text{if } x \ge c/\Delta \right)$	$\int v_0  \text{if } x < c/\Delta$	$\int v_0 + \Delta   ext{if } x \geq c/\Delta$	$\int v_0 \qquad \text{if } y < \frac{c/\Delta - \mu_x}{\Lambda}$	$v_0 + \Delta$ otherwise	$v_0 + \Delta N \left( rac{\mu_x + \Lambda y - c/\Delta}{\Lambda \sigma \sqrt{T - t}}  ight)$	$\Lambda = (\mu_x + \Lambda_y - c/\Delta) / (\mu_x + \Lambda_y - c/\Delta)$
	C(v)		G(x)	V(x)	h(y)	P(t,y)	$\lambda(t,y)$		C(v)	G(x)	V(x)	h(y)	P(t,y)	$\lambda(t,y)$			C(v)			(n)D	(~)/1	$(r) \Lambda$	h(a)	(6)11	P(t,y)	1/1 2.1

Table 2: Economic Efficiency Comparative Statics. The signs are the signs of the partial derivatives of  $\overline{P}$  with respect to the parameters. A 0 indicates that the partial derivative is 0. The partial derivative  $\partial \overline{P}/\partial \mu_x$  is positive in all cases, by Theorem 1, so that partial derivative is omitted from the table. The value function V is affine in Example 1 and convex in Examples 2 and 3, so the signs of the partial derivatives of  $\overline{P}$  with respect to  $\sigma$  and  $\sigma_x$  are given by Theorem 2 for those examples.

	1. Quadratic Cost	2. Asymmetric Quadratic Cost
$\sigma$	0	+
$\sigma_x$	0	+
$\psi$	+	+
	3. Exponential	
$\sigma$	+	
$\sigma_x$	+	
$\psi$	+	
	4. Binary	4. Probabilistic Binary
σ	$\int +  \text{if } \psi \mu_x < 1$	$\int +  \text{if } \mu_x < \mu_x^*$
	$\int - \text{ if } \psi \mu_x > 1$	$\int - \text{ if } \mu_x > \mu_x^*$
σ	$\int + \text{ if } \psi \mu_x < 1$	$\int + \text{ if } \mu_x < \mu_x^*$
$v_x$	$- \text{ if } \psi \mu_x > 1$	$\int - \text{ if } \mu_x > \mu_x^*$
$\psi$	+	+
$\Delta$	+	+

Table 3: Market Liquidity Comparative Statics. The signs are the signs of the partial derivatives of  $\overline{\lambda}$  with respect to the parameters, except for Example 3 (exponential). For Example 3, the signs are the signs of the partial derivatives of  $\overline{\lambda}/\overline{P}$  with respect to the parameters.

	1. Quadratic Cost	2. Asymmetric Quadratic Cost
$\mu_x$	0	+
$\sigma$	-	_
$\sigma_x$	+	+
$\psi$	+	+
	3. Exponential	
$\mu_x$	0	
$\sigma$	_	
$\sigma_x$	+	
$\psi$	+	
	4. Binary	5. Probabilistic Binary
	$\int + \text{ if } \psi \mu_x < 1$	$\int + \text{ if } \mu_x < \mu_x^*$
$\mu_x$	$\int - \text{ if } \psi \mu_x > 1$	$\left\{-\text{if }\mu_x>\mu_x^*\right.$
σ	$\int +  \text{if } (\mu_x - 1/\psi)^2 > T\sigma^2 \Lambda^2 (\Lambda - 1)$	$\int +  \text{if } \left(2 - \Lambda + \frac{\Lambda^2 \sigma^2 T}{c^2}\right) N(\overline{d}_2) > \left(\overline{d}_1 + \frac{\Lambda \sigma \sqrt{T}}{c}\right) n(\overline{d}_2)$
	$\int - \text{ if } (\mu_x - 1/\psi)^2 < T\sigma^2 \Lambda^2 (\Lambda - 1)$	$\int - \operatorname{if} \left( 2 - \Lambda + \frac{\Lambda^2 \sigma^2 T}{c^2} \right) N(\overline{d}_2) < \left( \overline{d}_1 + \frac{\Lambda \sigma \sqrt{T}}{c} \right) n(\overline{d}_2)$
$\sigma_x$	+	+
ψ	+	$\begin{cases} + & \text{if } \psi \Lambda \sigma \sqrt{T} n(\overline{d}_2) < \left(1 - \psi \Lambda \sigma \sqrt{T} \ \overline{d}_2\right) N(\overline{d}_2) \\ - & \text{if } \psi \Lambda \sigma \sqrt{T} n(\overline{d}_2) > \left(1 - \psi \Lambda \sigma \sqrt{T} \ \overline{d}_2\right) N(\overline{d}_2) \end{cases}$
$\Delta$	+	+