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, Revised October 2017

This paper consolidates and substantially extends the theory parts of two prior working papers: Back, Li and Ljungqvist (2015) and Collin-Dufresne and Fos (2015b). We are grateful to Jean-Noel Barrot, Alon Brav, Bradyn Breon-Drish, Carole Comerton-Forde, Quinn Curtis, David De Angelis, Alex Edmans, Vivian Fang, Dirk Hackbarth, Stacey Jacobsen, Wei Jiang, Barbara Katz, Doron Levit, Andrey Malenko, Nadya Malenko, Ernst Maug, and Konstantinos Zachariadis for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Activism, Strategic Trading, and Liquidity
Kerry Back, Pierre Collin-Dufresne, Vyacheslav Fos, Tao Li, and Alexander Ljungqvist
NBER Working Paper No. 22893
December 2016, Revised October 2017
JEL No. G3

ABSTRACT

We analyze dynamic trading 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. Variation in parameters can produce either positive or negative relations between market liquidity and economic efficiency, depending on the activism technology and model parameters. Two results that contrast with the previous literature are that (a) the relation between market liquidity and economic efficiency is independent of the activist's initial stake for a broad set of activism technologies and (b) an increase in noise trading can reduce market liquidity, because it increases uncertainty about the activist's trades (the activist trades in the opposite direction of noise traders) and thereby increases information asymmetry about the activist's intentions.

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

Tao Li
City University of Hong Kong
taoli3@cityu.edu.hk

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

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

Vyacheslav Fos
Carroll School of Management
Boston College
140 Commonwealth Avenue
Fulton Hall 330
Chestnut Hill, MA 02467
vyacheslav.fos@bc.edu

1. Introduction

Activist shareholders—who seek to alter corporate policies and thereby affect share values—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 was 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.[1] Existing shareholders can turn from being passive to being active when they recognize an opportunity for enhancing the value of their holdings.[2] And activist short-sellers can take actions so as to reduce firm value to benefit their short positions.[3] 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.

Differing views have been expressed regarding the effect of market liquidity 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 greater market liquidity enables a potential activist who does not already own a sizeable initial toehold to accumulate more shares and eventually become active. Of course, both market liquidity and activism are endogenous. Whether an exogenous shock moves the two in the same direction or in opposite directions depends on the nature of the shock. It also depends, as we show, on 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 in a non-binary effect on firm value. When an activist seeks

[1] Prominent examples include William Ackman, Carl Icahn, Daniel Loeb, and Nelson Peltz.

[2] CALPERS and the Norwegian sovereign wealth fund are well known examples of this form of activism.

[3] 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.

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 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. Other examples (replacing a CEO, replacing directors, changing governance rules) are similar. We find that there are significant differences between the binary model and continuous models. For example, a key result of Maug (1998)—an increase in noise trading increases activism if and only if the initial stake of the activist lies below a certain threshold—is true in the binary model but not in continuous models. For example, if the activism technology is such that the value the activist chooses to create depends convexly on her block size, then an increase in noise trading always increases activism.

The link between liquidity and activism is bidirectional: just as market liquidity affects activism, so also does the potential for activism affect market liquidity. The latter direction has received little attention in the literature. One insight we obtain by studying the effect of activism on liquidity is that increases in noise (liquidity) trading do not necessarily increase market liquidity. This stands in contrast to the classic remark of Treynor (1971)—which is encapsulated in Kyle's lambda (Kyle, 1985)—that "the liquidity of a market ... is inversely related to the average rate of flow of new information ... and directly related to the volume of liquidity motivated transactions." We show that an activist will generally trade in the opposite direction of noise traders, buying when they depress the price by selling and selling when they inflate the price by buying. Consequently, more noise trading produces more uncertainty about the size of the activist's eventual blockholding, increasing the ability of the activist to trade profitably before becoming active. This increase in information asymmetry due to liquidity trading can more than offset the direct effect of liquidity trading on market liquidity, causing market liquidity to fall.

We study liquidity and activism by generalizing the dynamic version of the Kyle (1985) model to a strategic 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 the Kyle model to non-Gaussian distributions, it affords tractability.[4]

[4]In Appendix D, we study the one-period model. This yields explicit results only for simple special cases that are

We are able to describe the equilibrium trading strategy in a way that is independent of the cost-of-effort function. The equilibrium trade depends on the number of shares owned by the trader as well as on the total number of shares bought by the trader and noise traders. The potential activist is more likely to become active the more shares she owns, so the value of an additional share is higher the more shares she already owns. Consequently, the number of shares she buys is higher the more shares she already owns. This contrasts with the standard Kyle model, in which the equilibrium trade is independent of the number of shares owned by the strategic trader, given the total number of shares bought by the trader and noise traders.

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 strategic traders who may intervene in corporate governance. These include Kyle and Vila (1991), Admati et al. (1994), Bolton and von Thadden (1998), Kahn and Winton (1998), Ravid and Spiegel (1999), Bris (2002), Noe (2002), and Faure-Grimaud and Gromb (2004).[5] 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. 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

uninteresting for the main question we investigate.

[5]See Edmans (2014) for a survey of the literature.

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), Edmans and Manso (2011), and Dasgupta and Piacentino (2015). 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.

Our paper is also related to two papers that study the relation between investment efficiency and information aggregation in large auctions (Atakan and Ekmekci (2014) and Axelson and Makarov (2017)). In these papers, the fact that the payoff to the winner of the auction depends on her future action can lead to equilibria where large auctions fail to aggregate the dispersed information and prices become inefficient. The setup of these papers is quite different from ours as in their model no investor possesses all relevant information regarding the action she is to take, but instead investment efficiency depends on the information about the economic state revealed by the auction. Further, these papers only consider a single round of trading.[6]

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

[6] Rostek and Weretka (2012) also investigate how the size of the market affects how auctions aggregate dispersed information and the consequences for market liquidity and price informativeness.

generally, about required disclosure rules for the trading positions of significant blockholders (e.g., Fishman and Hagerty, 1992, 1995).

3. Model

We analyze a Kyle model in which the strategic trader is a potential activist who can undertake costly effort to influence the management of a firm and hence affect the value of its stock. The 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.[7] We assume there is some fixed date T at which the trader must act in order to influence management (if she acts at all). Naturally, any action quickly becomes public information, so we assume that the market observes at T whether and to what extent the activist acts. Prior to T, the market is uncertain how many shares the trader owns and hence is uncertain about the trader's intentions. The uncertainty about the trader's intentions is resolved at T. Therefore, the trader has no private information after T and cannot profitably trade after T. Consequently, we model trading as stopping at T. Prior to T, the trader can profitably trade on her private information about the number of shares she owns, as we will show. We assume trading is continuous during the time interval $[0, T]$.[8] In some applications, T might be a choice variable of the activist. That would be an interesting extension, but it is beyond the scope of this paper.

After the activist does or does not exert effort at T, the stock trades at a share price v that incorporates the market makers' expected value of the activist's effort. Of course, the activist's efforts may only affect the company's operations with some lag, but we assume the market correctly discounts the future cash flows into the share price v at date T. Denote by $C(v)$ the cost to the activist of achieving a share price of v.[9] We assume C is lower semicontinuous and takes values in

[7] We assume the trader has no private information about aspects of the firm other than her own blockholding in order to focus on the effect of activism on market liquidity. However, in one example (Example 1 in Section 6), other private information can easily be incorporated (Collin-Dufresne and Fos, 2015b).

[8] Kyle (1985) shows that his discrete-time equilibrium converges to the equilibrium of his continuous-time model as the time periods become shorter and the number of time periods grows. We have not checked this convergence result for our model, but we conjecture that it would be true. Hence, we believe that our results are relevant in discrete time as well. However, the continuous-time model is substantially more tractable than the discrete-time model, so we work in continuous time.

[9] A generalization of our model would be to allow the cost to depend on the blockholding x as $C(v, x)$. Such a model was studied in one of our previous papers (Back, Li and Ljungqvist, 2015). In that model, the asset value is binary with possible values $v_L < v_H$, the cost of generating the high value is c, and activism cannot be successful

$[0, \infty]$. We define $C(v) = \infty$ when a value v is infeasible. Furthermore, we assume

$$\lim_{v \to -\infty, +\infty} \left| \frac{C(v)}{v} \right| = \infty \,. \tag{1}$$

Thus, C grows more than linearly at extreme values of v (or such values are infeasible). Given $X_T = x$, the activist chooses effort to maximize $vx - C(v)$. The optimal value to the activist is

$$G(x) \stackrel{\text{def}}{=} \sup_v \{vx - C(v)\} \,. \tag{2}$$

Because G is the supremum of a collection of affine functions, G is convex. By assumption (1), we can restrict attention to compact sets in the maximization problem in (2) and G is everywhere finite. This fact and the lower semincontinuity of C imply that the supremum in (2) is attained at some value of v. Let

$$V(x) \in \operatorname{argmax}_v \{vx - C(v)\} \,. \tag{3}$$

By a standard argument, $V(x)$ is a subgradient of G at x.[10] Because G is convex, it is differentiable almost everywhere. Thus, $V(x) = G'(x)$ almost everywhere. Given a blockholding of size x at the end of trading, it is optimal for the activist to exert effort to move the stock value to $V(x)$. Thus, $V(x)$ is the common value of shares to all traders after the activist's optimal expenditure of effort. Because $V(x)$ is a subgradient of G at x, the common value of shares $V(x)$ is also the marginal value of shares for the activist.

Denote the number of shares owned by the strategic trader at each date t by X_t. We assume that X_0 is known only to the trader.[11] Other market participants regard X_0 as being normally

unless the block size is at least B, where B is exogenously given. So, $C(v_H, x) = \infty$ if $x < B$ and $C(v_H, x) = c$ if $x \geq B$. We found that the minimum block size B played only a minor role in the equilibrium. If the blockholder desires to be active based on the block holding X_{T-} just before T and the block is not large enough for activism, then she will place a discrete order for $B - X_{T-}$ at T. In equilibrium, the market knows in this circumstance that the blockholder will become active, so the block is priced at v_H and there is no gain or loss on the trade at T. More generally, when C depends on x, we conjecture that the activist will purchase enough shares via a block trade at T to drive down the cost of activism as far as possible, meaning $\partial C(v, X_T)/\partial x = 0$.

[10] Because $V(x)$ attains the supremum in (2), we have $G(x) = xV(x) - C(V(x))$, and by the definition of G, $G(a) \geq aV(x) - C(V(x))$ for all a. Combining these two facts yields $G(x) \leq xV(x) + G(a) - aV(x)$ for all a, which is the definition of $V(x)$ being a subgradient of G at x. This is an instance of the envelope theorem: ignoring the possibility of nondifferentiability, we can differentiate the identity $G(x) = V(x)x - C(V(x))$ and use the first-order condition $x = C'(v)$ for the optimization problem in (2) to deduce that $G'(x) = V(x)$.

[11] This assumption is consistent with U.S. rules on the disclosure of ownership stakes. Investors with activist intentions are required to submit a Schedule 13D filing to the SEC only once their ownership reaches 5% of a targeted

distributed with mean μ_x and standard deviation σ_x. In addition to the strategic 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 σ. Aggregate purchases by the strategic trader and noise traders are $Y_t = X_t - X_0 + Z_t$. An important 'signal to noise ratio' parameter in the equilibrium is:

$$\Lambda \stackrel{\text{def}}{=} 1 + \sqrt{1 + \frac{\sigma_x^2}{\sigma^2 T}}\,. \tag{4}$$

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 $V(X_T)$ conditional on the history of orders. Let $\mathcal{F}_t^Y$ denote the information conveyed by the history of orders through date t. 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 strategic trader's trades are of order $\mathrm{d}t$, meaning that $\mathrm{d}X_t = \theta_t\,\mathrm{d}t$ for some stochastic process θ.[12] Given $P(\cdot)$, the strategic trader chooses the trading strategy θ to maximize

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

We assume the trader's information set at each date t consists of X_0 and the history of noise trades until date t. Of course, the trader knows her own trades, so she also knows the history of X until date t. The assumption that the trader knows the history of noise trades is motivated by the fact that the trader should be able to observe the price and, as we will show, price changes reveal

firm's outstanding shares, though they may have to disclose their stakes earlier, to the extent that they are subject to Section 13(f) of the Securities Exchange Act of 1934. Section 13(f) requires quarterly disclosures of long positions in U.S. stocks and options 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, and those holding short positions, do not have to disclose their positions at all.

[12] This assumption is without loss of generality, because, as shown by Back (1992), if there are jumps or nonzero quadratic variation in the strategic trader's holdings X, the 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 strategic trader to pay these costs, and they can be avoided by taking X to be continuous and of finite variation, that is, by submitting very small trades $\mathrm{d}X$ ('very small' meaning of order $\mathrm{d}t$).

aggregate orders $\mathrm{d}Y$, from which the trader can infer $\mathrm{d}Z$ simply by subtracting $\mathrm{d}X$. The strategic trader's value function is

$$J(t,x,y) \stackrel{\text{def}}{=} \sup_{\theta} \mathsf{E}\left[G(X_T) - \int_t^T P(u,Y_u)\theta_u \,\mathrm{d}u \,\middle|\, X_t = x, Y_t = y\right]. \tag{6}$$

Here, the supremum is taken over strategies θ that are adapted to the trader's information and that satisfy the mild 'no-doubling' strategies condition (9) stated below. We define an equilibrium to be a pair (P,θ) such that the trading strategy θ maximizes (5) given P subject to (9) and such that

$$P(t,Y_t) = \mathsf{E}\left[V(X_T) \mid \mathcal{F}_t^Y\right] \tag{7}$$

for each t, given θ. 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.

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

$$\mathsf{E}\left[V\left(\frac{\Lambda(X_0 - Z_T) - \mu_x}{\Lambda - 1}\right)^2\right] < \infty\,, \tag{8}$$

where Λ is defined in (4), and we define a trading strategy θ to be admissible if and only if[13]

$$\mathsf{E}\int_0^T V(\mu_x + \Lambda(X_t - X_0 + Z_T))^2 \,\mathrm{d}t < \infty\,. \tag{9}$$

4. Equilibrium

Theorem 1. *The pricing rule*

$$P(t,y) = \mathsf{E}\left[V(\mu_x + \Lambda\, Z_T) \mid Z_t = y\right] \tag{10}$$

[13]This condition is used in verifying the optimality of the potential activist's trading strategy. It implies that the local martingale $\int_0^t P\,\mathrm{d}Z$ is actually a martingale. For the connection to doubling strategies, see Back (1992), who points out that the martingale property means that noise traders would not lose money on average if they could trade "at the midpoint of the bid-ask spread."

and trading strategy

$$\theta_t = \frac{1}{T-t}\left(\frac{X_t - \mu_x - \Lambda Y_t}{\Lambda - 2}\right) \tag{11}$$

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 σ. Moreover, $P(T, Y_T) = V(X_T)$ with probability 1. The value function is

$$J(t,x,y) = \frac{\Lambda - 1}{\Lambda}\mathsf{E}\left[G\left(\frac{\Lambda(x - Z_T) - \mu_x}{\Lambda - 1}\right)\,\middle|\, Z_t = y\right] + \frac{1}{\Lambda}\mathsf{E}\left[G(\mu_x + \Lambda Z_T)\mid Z_t = y\right]. \tag{12}$$

The equilibrium price evolves as $\mathrm{d}P(t, Y_t) = \lambda(t, Y_t)\,\mathrm{d}Y_t$, where Kyle's lambda is

$$\lambda(t,y) = \frac{\partial P(t,y)}{\partial y}. \tag{13}$$

Furthermore, $\lambda(t, Y_t)$ is a martingale on $[0, T-\delta]$ for every $\delta > 0$, relative to the market makers' information set. The strategic trader's equilibrium position at time T is

$$X_T = \mu_x + \frac{\Lambda}{\Lambda - 1}(X_0 - \mu_x - Z_T). \tag{14}$$

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

We note the surprising finding that the strategic 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 trader's accumulated shares. The strategy (11) is in fact linear in X_t and Y_t. We show in Appendix D that the trading strategy in a single-period model is linear when V is linear but not when V is nonlinear. Therefore, the trading strategy is not independent of the cost function in a single-period model. Of course, local linearity in continuous time distinguishes continuous time from discrete time in other contexts as well (e.g., the CCAPM). The cost function does affect the equilibrium pricing rule (10). Thus, if the trading strategy in the continuous-time model is expressed as a function of the price process, then it may depend on the cost function.

In the remainder of this section, we sketch the proof of Theorem 1 with the aim of providing

some intuition for the results. 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 strategic trader trades in such a way that the share price equals the marginal value at the terminal date. Otherwise, she is clearly leaving money on the table. In our model, the 'price equals marginal value' condition is that $P(T, Y_T) = V(X_T)$. The other feature is that the strategic trader's trades are not forecastable. On average, market makers do not expect the trader to trade in one direction or the other. Cho (2003) calls this 'inconspicuous insider trading.'[14] Consequently, $\mathsf{E}[X_T] = \mu_x$. Furthermore, the unpredictability of informed orders means that the drift of Y is zero on its own filtration; that is, Y is a martingale on its own filtration. 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 σ. Therefore, we know the equilibrium pricing rule if we know h.

In the basic continuous-time Kyle model, $h(Y_T) = v$ in equilibrium, where v is the exogenous value of the asset. This equality occurs because the strategic 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 trader offsets noise trades one-for-one and also purchases (or sells if negative) $h^{-1}(v)$ shares. This contrasts with our expression (14), in which the coefficient on $-Z_T$ in the formula for X_T is $\Lambda/(\Lambda-1) > 1$. In both the standard Kyle model and in our model, the strategic trader responds to noise trades by, for example, buying after they have sold and depressed the price. However, the trader buys more in our model than in the standard Kyle model, because the more shares she owns the more likely she is to become active and hence the more valuable are additional shares. This can be seen from the 'price equals marginal value' condition. In the standard model, marginal value is the exogenous value v, and 'price equals marginal value' takes the form $h(X_T - X_0 + Z_T) = v$. Thus, any change in Z_T must be offset by a one-for-one change in the opposite direction for X_T, keeping price constant. In

[14] Inconspicuous insider trading is a consequence of the Hamilton-Jacobi-Bellman equation and is therefore a necessary condition for equilibrium. See Back (1992). In Appendix D, we show that it is also a necessary condition in a single-period model, when the strategic trader can condition her demand on the noise trades.

our model, 'price equals marginal value' takes the form $h(X_T - X_0 + Z_T) = V(X_T)$. If, for example, we reduce Z_T by 1 and increase X_T by 1, then we will keep price constant, but the marginal value $V(X_T)$ will have increased, because owning an additional share will cause the activist to exert more effort, increasing the value of shares. Thus, the activist has the incentive to buy more shares, raising the price further.

The strategic trader achieves the equality $h(Y_T) = v$ in the standard Kyle model by causing Y to be a Brownian bridge terminating at $h^{-1}(v)$. In our model, the 'price equals marginal value' condition $h(Y_T) = V(X_T)$ 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 (15) below is the strategic trader's equilibrium order $\mathrm{d}X_t = \theta_t \, \mathrm{d}t$.

Lemma 1. *Let ε 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*

$$\mathrm{d}Y_t = \frac{a\varepsilon - bZ_t - (b+1)Y_t}{T-t}\,\mathrm{d}t + \mathrm{d}Z_t \tag{15}$$

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

$$Y_T = \frac{a\varepsilon - bZ_T}{b+1}\,. \tag{16}$$

The proof of the lemma is provided in Appendix A. The stochastic differential equation of a Brownian bridge is equation (15) with $b = 0$, so the process Y defined by equation (15) is a generalization of a Brownian bridge. A Brownian bridge is a Brownian motion conditioned to end at a particular point. With $b = 0$, the Brownian bridge Y in (15) ends at $a\varepsilon$. Because (when $b = 0$) the ending point is normally distributed with zero mean and variance equal to $\sigma^2 T$, the unconditional distribution of Y is that of a Brownian motion with the same law as Z. In other words, Y is a Brownian motion on its own filtration. The lemma states that this is also true when $b \neq 0$. 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 (16) 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 ε in equation (15) is a transformation of the exogenous asset value v. Assuming v is continuously distributed, we have that $\varepsilon = \mathsf{N}^{-1}(F(v))$, where N is the standard normal distribution function and F is the distribution function of v. In our model, the strategic 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 ε into equation (16) and rearrange to obtain

$$Y_T = \frac{a(X_0 - \mu_x)/\sigma_x + b(X_T - X_0)}{2b+1}\,. \tag{17}$$

X_0 vanishes from this equation if $a = b\sigma_x$. The two conditions on a are satisfied if and only if $b = 1/(\Lambda - 2)$. With this formula for b, we have

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

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 (10). 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 (14).

As in Back (1992), the strategic 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 (12) for the value function. We show in Appendix B that the function J defined in (12) satisfies the Hamilton-Jacobi-Bellman equation:

$$0 = \sup_{\theta}\left\{-P\theta + J_t + J_x\theta + J_y\theta + \frac{1}{2}\sigma^2 J_{yy}\right\}\,. \tag{19}$$

In fact, we show that the following hold:

$$-P + J_x + J_y = 0\,, \tag{20a}$$

$$J_t + \frac{1}{2}\sigma^2 J_{yy} = 0\,. \tag{20b}$$

The first of these two equations states that the coefficient on θ in the optimization problem in (19) is zero. Therefore, any θ achieves the optimum. There are in fact many optima to the trader's problem (taking the price process as given), as is also true in the basic continuous-time Kyle model (Back, 1992). Any strategy is optimal provided only that the share price equals marginal value at the terminal date.

To prove the theorem, all that remains is to verify that (11) is an optimal trading strategy. That is straightforward to verify since any trading strategy is locally optimal and the strategy (11) implies that price equals marginal value at T, as the lemma shows.

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 $P(0,0)$. As remarked before, we assume the market correctly discounts the eventual effects on firm cash flows of the activist's effort at T into the date–T price $V(X_T)$, and $P(0,0)$ is the expected value of $V(X_T)$, so $P(0,0)$ incorporates the value per share expected to be created by activism. Let $\overline{P}$ denote $P(0,0)$ as a function of the model parameters. We calculate the value per share instead of the aggregate market value of shares outstanding, because (as in Kyle (1985)) the number of shares outstanding is not a parameter of the model. That is, given the per-share value function $V(\cdot)$, the number of shares outstanding does not affect the equilibrium trading strategy or price process. On the other hand, the cost function $C(\cdot)$ specifies the total cost to the activist, rather than the cost per share outstanding (which would be irrelevant to the activist). Therefore, we have measures of value and cost that are in different units, the former being per share and the latter being total. Hence, we cannot compute the benefit of activism net of the activist's costs. As a result, we will measure economic efficiency by the value per share $P(0,0)$ rather than as value net of costs. Fortunately, this should do little damage, because

costs per share should in practice be quite small compared to value. We base this argument on the fact that activists recoup their costs from shareholdings that are typically small percentages of the total number of shares outstanding.[15]

We generally measure market illiquidity by the expected average lambda:

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

Theorem 1 shows that λ 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 illiquidity. However, there are other examples in which percentage price changes are stationary, and in those examples, some 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 illiquidity.

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. Recall that, when G is differentiable, $V(x) = G'(x)$, so the convexity or concavity of V is determined by the third derivative of G. The function G is convex, but we cannot in general sign its third derivative. However, because the domain of V is the entire real line (the activist's terminal block size X_T can take any real value), a convex V must be unbounded above, and a concave V must be unbounded below. In general, convexity seems more reasonable than concavity, because concavity (unbounded below) implies that the possible value destruction must be unlimited, not respecting limited liability.[16] 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 C.

[15] Collin-Dufresne and Fos (2015a) report that the average activist holds 7.51% of the target's outstanding shares when making her first public disclosure through a Schedule 13D filing. Collin-Dufresne et al. (2107) show that when Schedule 13D filers use derivatives to increase their overall economic exposure to the stock, their average exposure increases to 8.70%. Given that after a 13D disclosure, stock prices incorporate the expected effects of activism on firm value, the average Schedule 13D filer thus expects to recoup her costs of activism from a 7%-8% toehold.

[16] Of course, the standard Kyle (1985) model also implies a Gaussian distributed firm value which does not satisfy limited liability either.

Theorem 2.

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

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

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

$$\overline{P} = \mathsf{E}[V(\mu_x + \Lambda Y_T)]\,, \qquad (21)$$

and

$$\overline{\lambda} = \Lambda\mathsf{E}[V'(\mu_x + \Lambda Y_T)]\,, \qquad (22)$$

where we regard Y_T as normally distributed with mean zero and variance equal to $\sigma^2 T$. The standard deviation of ΛY_T is $\Lambda\sigma\sqrt{T}$, which is an increasing function of both σ and σ_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.[17] We can in fact say a bit more about how changes in parameters affect the distribution of $V(\mu_x + \Lambda Y_T)$. Because V is monotone, the distribution of $V(a + \Lambda Y_T)$ first-order stochastically dominates the distribution of $V(b + \Lambda Y_T)$ when $a > b$. This is a stronger statement than (2a) of Theorem 2. Also, when V is concave, a reduction in either σ

[17] Note that V' cannot be concave, because it is nonnegative and hence bounded below.

or σ_x leads to second-order stochastic dominance.[18]

Theorem 2 does not provide a general result regarding the effect of a change in the volatility σ of noise trading on market illiquidity $\overline{\lambda}$, because a change in σ has two effects on $\overline{\lambda}$, and they can be in opposite directions. First, an increase in σ causes the factor Λ in equation (22) to fall. Second, an increase in σ is a mean-preserving spread in the distribution of $\mu_x+\Lambda Y_T$, so it causes the expectation in equation (22) to rise when V' is convex. Depending on which of these two effects is stronger, an increase in σ 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 (14)) 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 μ_x or σ_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 μ_x or σ_x) increases the importance of asymmetric information regarding the potential activist'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.

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

[18] This is a consequence of the following facts: (i) second-order stochastic dominance of ξ_A over ξ_B is equivalent to $\mathsf{E}[u(\xi_A)] \geq \mathsf{E}[u(\xi_B)]$ for all monotone concave u, (ii) the composition $u \circ V$ is concave, when u is monotone and concave and V is concave, and (iii) an increase in either σ or σ_x produces a mean-preserving spread in the distribution of $\mu_x + \Lambda Y_T$.

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 λ and the parameters μ_x, σ_x, σ, and Λ defined in Sections 3 and 4. The examples are distinguished by their cost functions $C(v)$. The cost functions include an additional productivity parameter ψ (and a second productivity parameter Δ 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 (2015b). 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 ψ 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 σ and σ_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 μ_x. Therefore, the only parameter that can produce cross-sectional variation in both efficiency and liquidity in this example is the productivity parameter ψ, 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 σ_x/σ. Kyle's lambda is also increasing in the activist's productivity ψ. In fact, and unlike in the Kyle model, the limit of lambda when the signal-to-noise ratio goes to zero is strictly positive: $\lim_{\sigma_x/\sigma \to 0} \lambda = \psi$. 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 ψ, 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

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

for constants v_0 and $\psi > 0$. Again, ψ 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 σ or σ_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 μ_x, σ_x, or ψ cause economic efficiency and liquidity to move in opposite directions.

Even though the activist can only create and cannot destroy value in this example, the trading strategy (expressed as a function of cumulative order flow and the strategic 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 noise trades and the corresponding activist trades, equilibrium price, and Kyle's lambda in Figures 1 and 2 below. Figure 1 shows a case where the 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 strategic trader trades in the opposite direction of the noise traders with an amplification as discussed before. The figures illustrate the amplification. When the strategic 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 strategic trader accumulates a large short position, the price and price

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.**

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 trader not to expend any effort. Correspondingly, Kyle's lambda converges to zero, because, given the large negative position the 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 activist will

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.**

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

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

This implies $V(x) = v_0 \mathrm{e}^{\psi x}$. Again, ψ measures the activist's productivity. In general, when σ_x is small, the partial derivative $\partial\Lambda/\partial\sigma$ is small. In this example, V' is convex, and when σ_x is small, the effect of a change in σ on Λ is less than the effect of a change in σ on the expectation in (22). Consequently, an increase in noise trading σ causes market liquidity (as measured by $1/\overline{\lambda}$) to fall.[19] However, as remarked in Section 5, it is more natural to measure market illiquidity in this example by the percentage price impact, λ/P. In fact, the percentage price impact is constant in this example, and it is a decreasing function of σ. Measuring liquidity in this way, cross-sectional variation in σ produces a positive cross-sectional relation between efficiency and liquidity, and cross-sectional variation in σ_x or ψ produces a negative cross-sectional relation between efficiency and liquidity. Increases in μ_x increase efficiency but have no effect on liquidity.

Even though the prior uncertainty is normal and thus the strategic 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 Scholes (1973) model. The parameters of the process are endogenously determined by the primitives of the model (the signal-to-noise ratio σ_x/σ, 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 σ_x/σ but also on the activist's productivity. As discussed for price impact in Example 1, when the signal-to-noise ratio goes to zero, the percentage price impact in

[19] The precise condition for this to occur is that

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

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 $X_T\Delta \geq c$. The value $V(x)$ is a step function, equal to v_0 for $x < c/\Delta$ and equal to $v_0 + \Delta$ for $x \geq c/\Delta$. 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/Δ. It is convenient to define $\psi = \Delta/c$, which is the value creation per unit cost. Then, ψ and Δ measure the activist's productivity.

In this example, cross-sectional variation in either of the productivity parameters ψ or Δ 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 μ_x or σ_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 μ_x is too small on its own to justify the cost of activism. In this case, a marginal increase in μ_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 σ_x increases economic efficiency, because it makes the expected initial stake μ_x a less reliable predictor of the actual initial stake X_0. An increase in σ_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 σ 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 σ of noise trading increases economic efficiency if and only if $\Delta\mu_x < c$. In that case, the potential activist 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 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 σ, σ_x, and T. When the absolute expected initial stake is large, it is unlikely that the potential activist 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 potential activist will trade out of an existing position or into a new position increases, and it increases so much that market liquidity actually falls.

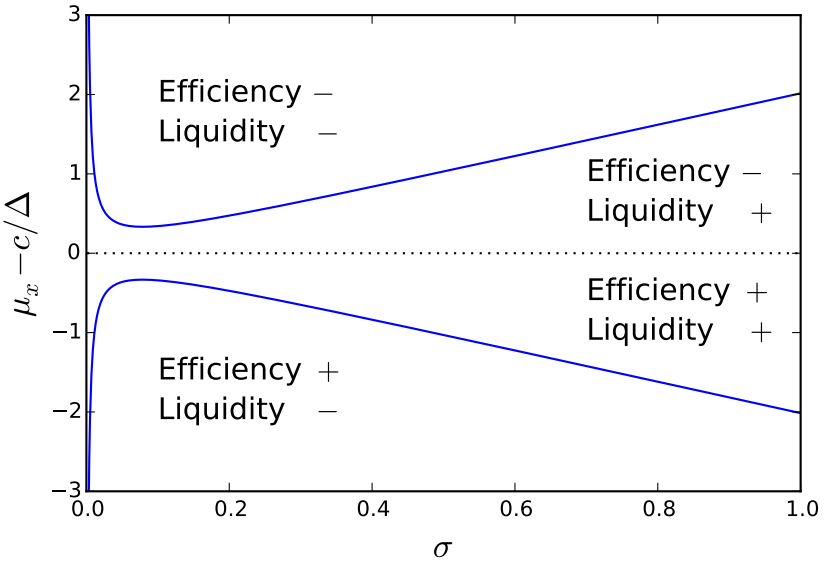

Figure 3: **Effect of an Increase in Noise Trading in the Binary Example.** The signs indicate the effect of an increase in noise trading σ on $\bar{P}$ (efficiency) and $1/\bar{\lambda}$ (liquidity). Increasing noise trading increases economic efficiency when $\mu_x > c/\Delta$ and reduces economic efficiency when $\mu_x < c/\Delta$. Increasing noise trading increases market liquidity when $|\mu_x - c/\Delta|$ is below a threshold depending on σ 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 Δ of activism multi-

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

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

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 (14) and (17) imply

$$\begin{aligned} Y_T \geq \frac{c/\Delta - \mu_x}{\Lambda} \quad &\iff \quad X_T \geq \frac{c}{\Delta} \\ &\iff \quad Z_T \leq X_0 - \mu_x + \frac{\Lambda - 1}{\Lambda}\left(\mu_x - \frac{c}{\Delta}\right)\,. \end{aligned}$$

Of course, the condition $X_T \geq c/\Delta$ is necessary and sufficient for exerting effort to be optimal for the strategic trader. The last condition shows that the 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 potential activist 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, success is instead viewed as a random event in this example, the probability of which depends on the activist's effort. Because the activist is risk neutral, she cares about the expected asset value, which is $v_0 + \Delta p$, where p denotes the probability of success and Δ is the value created by 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 activist's optimal effort implies a probability of success of $1 - \mathrm{e}^{-\Delta X_T/c}$. Thus,

$$V(x) = v_0 + \left(1 - \mathrm{e}^{-\Delta x/c}\right)\Delta\,.$$

The function V 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/Δ. As in Example 4, define $\psi = \Delta/c$, so the activist's productivity is measured by ψ and Δ.

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 $t = y = 0$ and $\overline{d}_2 = \overline{d}_1 - \psi\Lambda\sigma\sqrt{T}$, and in terms of μ_x^* defined as follows. Set $g(x) = \mathsf{N}(x)/\mathsf{n}(x)$. 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 σ increases market liquidity takes different forms in the two examples. Third, an increase in the productivity

parameter ψ does not always reduce market liquidity in this example. The condition under which it reduces market liquidity is shown in Table 3.

7. Discussion

There are many different types of activism technology, including binary forms of activism, those with non-binary effort, and those resulting in a non-binary effect on firm value. For example, when an activist seeks to increase payouts, 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 seeks 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. Agitating for the replacement of the CEO or the board of directors has similar features. In these cases, the probability that the activist is successful is an increasing function of her continuous effort. Large-sample evidence about heterogeneity in activism technology is reported by Brav et al. (2008), who classify hedge-fund activism campaigns by the activists' stated goals. These goals could be used to identify the properties of the activism technologies discussed in Theorem 2.

One implication of our model is that most of the comparative statics depend on these properties. This implication is important to regulators and empirical researchers alike. Consider, for example, a change in uncertainty about the activist's initial block size. The model shows that the effect of this uncertainty on economic efficiency can switch from positive to negative, depending on the activism technology. The model therefore suggests that regulators need to consider what types of activism technologies would be affected by a proposed change (e.g., activists' engagements in M&A deals or activists' campaigns to change payout policies). Moreover, the model shows that empirical research that pools observations for different activism technologies when evaluating the effects of changes in uncertainty about the initial block size could fail to find significant effects even if uncertainty matters.

Another implication of the model is that the role of noise trading is more nuanced than previously thought. Early models of corporate governance typically do not differentiate between noise trading and market liquidity. Our model, however, shows that an increase in noise trading may not lead to an increase in market liquidity (see Example 4). Similarly, the effect of noise trading on economic efficiency can be either positive or negative, depending on the activism technology. This

observation is important in assessing the impact of a regulatory change. The recent legal debate about changing the length of the trading period during which activists can trade anonymously is a case in point (e.g., Bebchuk et al. (2013)). Such a change can be viewed as a change in noise trading because in our model, what matters is $\sigma^2 T$—the cumulative amount of noise trading over the entire trading period. So from the perspective of a potential activist, reducing the trading horizon T is isomorphic to reducing noise trading volatility and keeping T fixed. Similarly, a Tobin tax on stock transactions might reduce the number of traders in the market and thus lead to a reduction in noise trading.[20]

Finally, the model shows how changes in disclosure rules that lead to changes in the precision of disclosed ownership information can affect economic efficiency and market liquidity. Consider, for example, Section 13(f) of the Securities Exchange Act of 1934 which requires quarterly disclosures of long (but not of short) positions in U.S. stocks and options 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 precise positions. Form 13(f) filings thus constitute at best noisy signals about an investment manager's block size. Changes to Section 13(f)—say, to the frequency of 13(f) filings or the inclusion of short positions—could therefore affect investors' uncertainty about an activist investor's initial toehold.

8. 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. Market liquidity affects activism, because it affects the ease with which the potential activist can either accumulate a stake or 'take the Wall Street walk.' One result that contrasts with the previous literature is that the relation between market liquidity and activism is independent of the activist's initial stake for a broad set of activism technologies.

In our setup, activism also affects market liquidity, because the activist's private information about her own intentions (which arises in our model because of the activist's private information about the size of her blockholding) creates adverse selection for market makers. This second

[20] A Tobin tax on stock transactions would likely have other effects on financial markets, the investigation of which is beyond the scope of our paper.

direction of causality has received little if any attention in the prior literature. One effect of this causality is that an increase in noise trading can reduce market liquidity, because it increases activist trading and hence increases information asymmetry regarding the activist's blockholding.

Appendix A. 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

$$\mathrm{d}W_t = \frac{1}{\sigma}\left(\mathrm{d}Y_t - \frac{1}{T-t}\hat{U}_t\,\mathrm{d}t + \frac{b+1}{T-t}Y_t\,\mathrm{d}t\right) = \frac{1}{\sigma}\left(\frac{1}{T-t}(U_t - \hat{U}_t)\,\mathrm{d}t + \mathrm{d}Z_t\right)\,. \tag{A.1}$$

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

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

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

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

The ODE (A.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 (A.3) is $-a^2/T$, and the right-hand side is

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

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 (A.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 (A.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/σ 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 σ 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 (16).

Appendix B. Proof of Theorem 1

We need to verify the optimality of the trading strategy (11). As in Section 4, define $h(z) = V(\mu_x + \Lambda z)$. Define

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

Because $\overline{y} = y$ is feasible in this optimization problem, we have $g(x,y) \geq 0$ 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,

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

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

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

This implies that

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

Thus,

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

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

The function J defined in (12) is given by

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

and, for $t < T$, set

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

From this definition and (B.3), we see that J is as stated in (12). 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

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

Thus,

$$J_x(t,x,y) + J_y(t,x,y) = P(t,y)\,. \tag{B.6}$$

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

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

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

$$\begin{aligned}
J(T, X_T, Y_T) &= J(0, X_0, Y_0) + \int_0^T dJ \\
&= J(0, X_0, Y_0) + \int_0^T P(t, Y_t)\theta_t \,\mathrm{d}t + \int_0^T J_y(t, X_t, Y_t)\,\mathrm{d}Z_t\,.
\end{aligned}$$

The no-doubling conditions (8) and (9) 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 $g \geq 0$, we have $J(T, X_T, Y_T) \geq G(X_T)$. Hence,

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

This shows that $J(0, X_0, 0)$ is an upper bound on the strategic 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 (11). For this strategy, the lemma implies—see (18)— that $X_T = \mu_x + \Lambda Y_T$, which implies from the definition of h that $V(X_T) = h(Y_T)$. In turn, this 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 (11) is optimal.

Appendix C. Proof of Theorem 2

First, we establish the comparative statics of economic efficiency. From (21), we have

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

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

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

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

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

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

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

which is an increasing function of σ and also an increasing function of σ_x. Thus, when V is convex,

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

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 (22),

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

It follows that

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

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

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

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

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

Thus:

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

Since $V'(x) \geq 0\ \ \forall x$ and $\partial\Lambda/\partial\sigma_x > 0$ it follows that

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

Appendix D. 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

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

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:

$$V(X_0 + \theta) - P(\theta + Z) - \theta P'(\theta + Z) = 0. \tag{D.2}$$

The second-order condition is:

$$V'(X_0 + \theta) - 2P'(\theta + Z) - \theta P''(\theta + Z) \leq 0. \tag{D.3}$$

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:

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

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

By using (D.2) in (D.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 D.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 θ 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 Λ 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 Λ that satisfies the SOC and the equation:

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

where β_x, β_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 D.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:

$$\mathsf{E}\left[G(X_1) - \theta\, P(Y) \,|\, X_0\right]. \tag{D.6}$$

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

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

and the second-order condition becomes:

$$V'(X_0 + \theta) - \mathsf{E}[2P'(\theta + Z) - \theta P''(\theta + Z) \,|\, X_0] \leq 0. \tag{D.8}$$

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

equation (D.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:

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

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:

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

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 Λ that satisfies this equation (and the SOC) with

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

There is a unique equilibrium given by:

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

Appendix D.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 $\mathrm{n}(x)$ is the standard Gaussian density and

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

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

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

$$V(X_0 + \theta) - P(\theta + Z) - \theta P'(\theta + Z) = 0. \tag{D.13}$$

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:

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

or equivalently:

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

Clearly, θ cannot be linear in X_0 and Z.

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.

Admati, A.R., Pfleiderer, P., Zechner, J., 1994. Large shareholder activism, risk sharing, and financial market equilibrium. Journal of Political Economy 102, 1097–1130.

Atakan, A., Ekmekci, M., 2014. Auctions, actions, and the failure of information aggegation. The American Economic Review 2014-2048.

Axelson, U., Makarov, I., 2017. Informational black holes in financial markets. Working paper London School of Economics .

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.

Bolton, P., von Thadden, E.L., 1998. Blocks, liquidity, and corporate control. The Journal of Finance 53, 1–25.

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

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., 2015a. Do prices reveal the presense of informed trading? The Journal of Finance 70, 1555–1582.

Collin-Dufresne, P., Fos, V., 2015b. Shareholder activism, informed trading, and stock prices. Working Paper.

Collin-Dufresne, P., Fos, V., Muravyev, D., 2107. Informed trading and option prices: Evidence from activist trading. Working Paper.

Dasgupta, A., Piacentino, G., 2015. The wall street walk when blockholders compete for flows. The Journal of Finance 70, 2853–2896.

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., 2014. Blockholders and corporate governance. Annu. Rev. Financ. Econ. 6, 23–50.

Edmans, A., Manso, G., 2011. Governance through trading and intervention: A theory of multiple blockholders. Review of Financial Studies 24, 2395–2428.

Faure-Grimaud, A., Gromb, D., 2004. Public trading and private incentives. Review of financial Studies 17, 985–1014.

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.

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.

Rostek, M., Weretka, M., 2012. Price inferece in small markets. Econometrica 80, 687–711.

Treynor, J., 1971. The only game in town. Financial Analysts Journal 27, 12–14+22.

Table 1: **Equilibrium in Five Examples.** The cost functions $C(v)$ and productivity parameters ψ and Δ are defined in Section 6. The parameters μ_x, σ_x, σ, and Λ are defined in Section 3. The functions G, V, h, P, and λ are calculated from the cost function as explained in Sections 3 and 4.

	1. Quadratic Cost	**2. Asymmetric Quadratic Cost**
$C(v)$	$(v-v_0)^2/(2\psi)$	$\begin{cases}\infty & \text{if } v < v_0\\ (v-v_0)^2/(2\psi) & \text{if } v \geq v_0\end{cases}$
$G(x)$	$v_0 x + \psi x^2/2$	$v_0 x + \psi (x^+)^2/2$
$V(x)$	$v_0 + \psi x$	$v_0 + \psi x^+$
$h(y)$	$v_0 + \psi\mu_x + \psi\Lambda y$	$v_0 + \psi\left(\mu_x + \Lambda y\right)^+$
$P(t,y)$	$v_0 + \psi\mu_x + \psi\Lambda y$	$v_0 + \psi(\mu_x+\Lambda y)\mathsf{N}\left(\frac{\mu_x+\Lambda y}{\Lambda\sigma\sqrt{T-t}}\right) + \psi\Lambda\sigma\sqrt{T-t}\,\mathsf{n}\left(\frac{\mu_x+\Lambda y}{\Lambda\sigma\sqrt{T-t}}\right)$
$\lambda(t,y)$	$\psi\Lambda$	$\psi\Lambda\mathsf{N}\left(\frac{\mu_x+\Lambda y}{\Lambda\sigma\sqrt{T-t}}\right)$
	3. Exponential	
$C(v)$	$\frac{1}{\psi} v \log\left(\frac{v}{v_0}\right) - \frac{1}{\psi}(v-v_0)$	
$G(x)$	$v_0(\mathrm{e}^{\psi x}-1)/\psi$	
$V(x)$	$v_0 e^{\psi x}$	
$h(y)$	$v_0 e^{\psi(\mu_x+\Lambda y)}$	
$P(t,y)$	$v_0 \mathrm{e}^{\psi\left(\mu_x+\Lambda y+\frac{1}{2}\Lambda^2\sigma^2(T-t)\right)}$	
$\lambda(t,y)$	$\psi\Lambda\, P(t,y)$	
	4. Binary	**5. Probabilistic Binary**
$C(v)$	$\begin{cases}\infty & \text{if } v \notin \{v_0, v_0+\Delta\}\\ 0 & \text{if } v = v_0\\ c & \text{if } v = v_0+\Delta\end{cases}$	$\begin{cases}\infty & \text{if } v < v_0\,,\\ z\left[\frac{v-v_0}{\Delta} + \left(1-\frac{v-v_0}{\Delta}\right)\log\left(1-\frac{v-v_0}{\Delta}\right)\right] & \text{if } v_0 \leq v < v_0+\Delta\,,\\ \infty & \text{if } v \geq v_0+\Delta\,.\end{cases}$
$G(x)$	$\begin{cases}v_0 x & \text{if } x < c/\Delta\\ (v_0+\Delta)x - c & \text{if } x \geq c/\Delta\end{cases}$	$\begin{cases}v_0 x & \text{if } x < 0\\ (v_0+\Delta)x - z\left(1-\mathrm{e}^{-\Delta x/z}\right) & \text{if } x \geq 0\end{cases}$
$V(x)$	$\begin{cases}v_0 & \text{if } x < c/\Delta\\ v_0+\Delta & \text{if } x \geq c/\Delta\end{cases}$	$\begin{cases}v_0 & \text{if } x < 0\\ v_0 + \Delta\left(1-\mathrm{e}^{-\Delta x/z}\right) & \text{if } x \geq 0\end{cases}$
$h(y)$	$\begin{cases}v_0 & \text{if } y < \frac{c/\Delta-\mu_x}{\Lambda}\\ v_0+\Delta & \text{otherwise}\end{cases}$	$\begin{cases}v_0 & \text{if } y < -\mu_x/\Lambda\\ v_0 + \Delta\left(1-\mathrm{e}^{-\Delta x/z}\right) & \text{if } y \geq -\mu_x/\Lambda\end{cases}$
$P(t,y)$	$v_0 + \Delta\mathsf{N}\left(\frac{\mu_x+\Lambda y-c/\Delta}{\Lambda\sigma\sqrt{T-t}}\right)$	$v_0 + \Delta\mathsf{N}(d_1) - \Delta\mathsf{N}(d_2)\mathrm{e}^{-\psi(\mu_x+\Lambda y)+\psi^2\Lambda^2\sigma^2(T-t)/2}$
$\lambda(t,y)$	$\Delta\mathsf{n}\left(\frac{\mu_x+\Lambda y-c/\Delta}{\Lambda\sigma\sqrt{T-t}}\right)\Big/\,\sigma\sqrt{T-t}$	$\psi\Delta\Lambda\mathsf{N}(d_2)\mathrm{e}^{-\psi(\mu_x+\Lambda y)+\psi^2\Lambda^2\sigma^2(T-t)/2}$

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 σ and σ_x are given by Theorem 2 for those examples.

	1. Quadratic Cost	**2. Asymmetric Quadratic Cost**
σ	0	$+$
σ_x	0	$+$
ψ	$\begin{cases} + & \text{if } \mu_x > 0 \\ - & \text{if } \mu_x < 0 \end{cases}$	$+$
	3. Exponential	
σ	$+$	
σ_x	$+$	
ψ	$+$	
	4. Binary	**4. Probabilistic Binary**
σ	$\begin{cases} + & \text{if } \psi\mu_x < 1 \\ - & \text{if } \psi\mu_x > 1 \end{cases}$	$\begin{cases} + & \text{if } \mu_x < \mu_x^* \\ - & \text{if } \mu_x > \mu_x^* \end{cases}$
σ_x	$\begin{cases} + & \text{if } \psi\mu_x < 1 \\ - & \text{if } \psi\mu_x > 1 \end{cases}$	$\begin{cases} + & \text{if } \mu_x < \mu_x^* \\ - & \text{if } \mu_x > \mu_x^* \end{cases}$
ψ	$+$	$+$
Δ	$+$	$+$

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**
μ_x	0	$+$
σ	$-$	$-$
σ_x	$+$	$+$
ψ	$+$	$+$
	3. Exponential	
μ_x	0	
σ	$-$	
σ_x	$+$	
ψ	$+$	
	4. Binary	**5. Probabilistic Binary**
μ_x	$\begin{cases} + & \text{if } \psi\mu_x < 1 \\ - & \text{if } \psi\mu_x > 1 \end{cases}$	$\begin{cases} + & \text{if } \mu_x < \mu_x^* \\ - & \text{if } \mu_x > \mu_x^* \end{cases}$
σ	$\begin{cases} + & \text{if } (\mu_x - 1/\psi)^2 > T\sigma^2\Lambda^2(\Lambda-1) \\ - & \text{if } (\mu_x - 1/\psi)^2 < T\sigma^2\Lambda^2(\Lambda-1) \end{cases}$	$\begin{cases} + & \text{if } \left(2-\Lambda+\frac{\Lambda^2\sigma^2T}{c^2}\right)\mathsf{N}(\overline{d}_2) > \left(\overline{d}_1+\frac{\Lambda\sigma\sqrt{T}}{c}\right)\mathsf{n}(\overline{d}_2) \\ - & \text{if } \left(2-\Lambda+\frac{\Lambda^2\sigma^2T}{c^2}\right)\mathsf{N}(\overline{d}_2) < \left(\overline{d}_1+\frac{\Lambda\sigma\sqrt{T}}{c}\right)\mathsf{n}(\overline{d}_2) \end{cases}$
σ_x	$+$	$+$
ψ	$+$	$\begin{cases} + & \text{if } \psi\Lambda\sigma\sqrt{T}\mathsf{n}(\overline{d}_2) < \left(1-\psi\Lambda\sigma\sqrt{T}\,\overline{d}_2\right)\mathsf{N}(\overline{d}_2) \\ - & \text{if } \psi\Lambda\sigma\sqrt{T}\mathsf{n}(\overline{d}_2) > \left(1-\psi\Lambda\sigma\sqrt{T}\,\overline{d}_2\right)\mathsf{N}(\overline{d}_2) \end{cases}$
Δ	$+$	$+$