Activist Trading Dynamics*

Doruk Cetemen† Gonzalo Cisternas‡ Aaron Kolb§
S. Viswanathan¶

January 21, 2023

Abstract

Two activists with correlated private positions in a firm's stock trade sequentially before simultaneously exerting effort to determine the firm's value. A novel linear equilibrium exists in which trades have positive sensitivity to initial positions but are nonzero on average: the leader strategically moves the price to induce the follower to acquire more shares and thus add more value. We examine this equilibrium's implications for market outcomes and its connection with the prominent phenomenon of "wolf-pack" activism, whereby *multiple* hedge funds target the same firm. We also explore the possibility of equilibria in which activists trade against their initial positions.

Keywords: activism, insider trading, noisy signaling, hedge funds

*The views expressed in this paper are those of the authors and do not necessarily represent the position of the Federal Reserve Bank of New York or the Federal Reserve System. We thank Peter DeMarzo, Nathan Kaplan, Erik Madsen, and Debraj Ray for useful discussions, as well as Saketh Prazad for excellent research assistance.

†Department of Economics, City University of London, doruk.cetemen@city.ac.uk.

‡Research and Statistics Group, Federal Reserve Bank of New York, gonzalo.cisternas@ny.frb.org.

§Kelley School of Business, Indiana University, kolba@indiana.edu.

¶Fuqua School of Business, Duke University, viswanat@duke.edu.

1 Introduction

Activist shareholders play a central role in the way modern corporations are run. To improve performance, these types of "blockholders" actively shape firms' capital structure (e.g., dividend and equity issuance), business strategy (e.g., cost reductions, selling divisions), and corporate governance (e.g., executive compensation, board composition), in part ameliorating the losses that stem from the separation of ownership and control. Activist campaigns have become ubiquitous in recent years, increasingly targeting large-capitalization firms in addition to the low- and mid-"cap" counterparts that have been the more traditional targets in the past. Importantly, activism has become an established investment strategy within the business models of some investors—most notably, a select group of hedge funds.[1]

Two distinctive features of activism are that it: (i) involves minority shareholders, i.e., those whose blocks are not large enough to fully control management; and (ii) requires significant outlays, beyond the costs of block acquisition.[2] Consequently, it is critical for any activist to induce other blockholders to "come along" too. However, as has been pointed out (e.g., Edmans and Holderness, 2017), much of the theoretical literature has focused on settings with an activist in isolation or on multiple activists with fixed blocks. Thus, the fundamental question of how investors build stakes in anticipation of future activism, with other investors having skin in the game too, remains much less understood. This issue is also of great empirical relevance, as interventions by multiple activists have become extremely frequent (Becht et al., 2017), and the strength of any intervention is necessarily linked to block size, which is an endogenous variable.

In this paper, we examine a market-based mechanism through which activists attempt to steer other investors to add value to firms. Specifically, two activists decide how much stake to (de-)accumulate in a Kyle (1985) type of market structure, where: (i) private information is about initial blocks; and (ii) firm value is determined by effort choices, as in the single-player model of Back et al. (2018). We add two natural ingredients to this baseline setting. First, initial positions exhibit correlation. Second, trading is sequential: in the first period, a *leader* activist acts as the unique informed trader, anticipating that a *follower* will play that role in the second period. In the third period, both activists simultaneously exert effort to determine firm value. Thus, the leader behaves in a "Stackelberg" manner, anticipating how her actions will influence the firm's value via the follower's subsequent trading opportunities.

Sequential stake-building and endogenous fundamentals have important consequences. Indeed, Proposition 2 establishes the existence of an equilibrium in which the leader ac-

[1]See Brav et al. (2021b) for a comprehensive review of hedge fund activism. The authors document that almost 900 hedge funds have been involved in more than 4,600 "events" in the U.S. from 1994 to 2018.

[2]We review the key institutional details of activism in Section 6.

tivist's orders are nonzero on average. This is in stark contrast to the ubiquitous equilibrium in Kyle-type models, in which trades are based solely on the difference (or "gap") between the insider's and the market maker's belief about the fundamentals—which proxies for the potential gains from arbitrage—and are therefore zero in expectation.[3] Specifically, if positions are positively (negatively) correlated, the leader manipulates the price downward (upward) to induce the follower to acquire a larger position and ultimately exert higher effort.[4]

This finding is driven by the interplay between dynamic incentives and endogenous costs. Concretely, while activists' actions are substitutes in the firm's value—highlighting the free-riding problem at play in practice—trading and effort choices are strategic complements intertemporally, as any added value is applied to all shares. In particular, if leaders with higher initial blocks expect to have higher terminal positions, these *types* benefit more from inducing effort by the follower. As a result, in a linear equilibrium in which trading strategies attach a positive weight to initial positions, if correlation is positive (negative) the leader lowers (increases) the aforementioned weight relative to a traditional "Kyle" setting. Through this deviation, all types make the market maker more pessimistic about the follower's position, thereby making the exploitation of arbitrage opportunities more attractive for the follower. Further, leaders with larger blocks effectively deviate more in absolute terms.

This type of behavior has signaling implications that take us to the second part of the argument: the endogeneity of costs. As is usual, higher market-maker beliefs in our setting reduce the extent of mispricing and thus the amount of trading. These "limits to arbitrage" are, however, related to information transmission in a non-trivial way. Consider the case of positive correlation: with less information conveyed due to the reduced signaling, price impact falls for a fixed degree of correlation. The leader then finds it less costly to reduce her purchases in response to an increase in the prior belief, manifested in the weight attached to the prior belief in the leader's strategy becoming more negative than in a traditional "Kyle" setting. With all leader types adjusting downward along both dimensions of information—private and public—the leader sells on average. Similarly, when correlation is negative, the leader buys more aggressively than in settings where activism is not at play.

The nature of correlation between positions then matters for behavior, and hence for market outcomes. Thus, it is important to identify market characteristics that favor (dis)similar positions in a statistical sense, as these may shed light on the types of activism events for which our predictions are most plausible. Similarity among activists is one factor: if the

[3]Our use of the word "arbitrage" in this paper is in the sense of exploitation of superior information within a market, as opposed to exploiting price discrepancies across different markets.

[4]Our use of the term "manipulation" is in the sense of an agent distorting her behavior relative to a benchmark (here, gap strategies) to ultimately affect the actions of another agent via the channel of influencing the latter's beliefs. We review this and other forms of manipulation in the literature review.

activists involved follow similar business models, this reduces the odds of very different—or even opposite—positions. Market capitalization is another: as large-cap firms tend to have more shares outstanding, similar blocks (at least of moderate size in absolute terms) are more likely there than in mid- or low-cap counterparts. Finally, the presence/absence of investors with large short positions is indicative of negative/positive correlation being at play.

Section 5 then derives predictions for market outcomes—order flows, market liquidity, and most notably, firm values—that can be linked with these characteristics. In particular, if correlation is negative, we show that ex ante firm value is larger than in a benchmark where positions do not change on average, which in turn is larger than ex ante firm value if positions are positively correlated. Also, ex ante firm value in our model is higher than if a single activist acted in isolation, for all levels of correlation. Because the market price is simply the firm's expected value given publicly available information, these rankings translate to predictions about *average prices* during activism events: prices should be higher in settings that favor lower/negative correlation, and that feature more than one activist. Some empirical studies provide support for these findings. Specifically, around activism events, buy-and-hold abnormal returns are: inversely related to market capitalization (Brav et al., 2021b); higher for firms featuring traders with large short positions (Li et al., 2022); and higher in multiplayer engagements compared to single-activist attacks (Becht et al., 2017).

That said, our findings do rely on an activist being willing to act as a leader, so it is natural to explore conditions that facilitate Stackelberg-like behavior. We argue first that positions being not too negatively correlated is a necessary requirement: otherwise, the leader may benefit from having a contemporaneous "competitor" simply because the latter is likely to act as a supplier, with trade taking place at low price impact. Second, when correlation is positive, as correlation grows, a hypothetical leader becomes more inclined to act as such as the number of followers increases: competition effects can lead the followers to trade more aggressively, which amplifies the value of influencing the continuation game as correlation grows. From a real-world standpoint, therefore, our mechanism is more likely at play when: (i) activists seek arbitrage opportunities; (ii) the activists involved are similar, in that their positions are not too negatively correlated; and (iii) (if the setting favors sufficient positive correlation) more followers who act in a non-cooperative way are present.

Section 6 then takes an institutional perspective by relating these insights to the evidence on hedge-fund activism—in particular, the so-called *wolf-pack* activism, whereby multiple hedge funds engage with a target firm following a leader hedge fund that builds a stake in it. Indeed, not only are hedge funds the quintessential example of exploitation of arbitrage opportunities, but this intrinsic similarity stemming from their business models is enhanced by the fact that they tend to hold small to moderate stakes in target firms. In addition, these investors are likely to act in a non-cooperative fashion due to regulatory and legal costs faced otherwise, with leader hedge funds completing their blocks quickly after key regulatory ownership thresholds are crossed, likely to avoid competition. From this perspective, our model can shed light on forces that could be at play in this type of activism. In particular, the traditional free-riding problem is likely to be exacerbated in settings that combine the presence of multiple fellow activists with positively correlated positions, such as when large-cap firms are targets; by contrast, leader hedge funds are predicted to act more aggressively—thus mitigating the free-riding problem—in settings where correlation is lower and attracting other investors is less critical, such as when smaller firms are attacked.

We conclude the paper by turning to the task of characterizing the set of all linear equilibria. Specifically, when the order flow is highly volatile and hence the market prone to be liquid, creating mispricing for the follower may come at the expense of large trading losses. A "coordination" equilibrium can emerge, where at least one of the activists trades against its initial position: e.g., if the first activist is initially long and the second activist is initially short, the first activist would move towards a short position. Our main equilibrium can co-exist with this coordination one when correlation is positive, but it may cease to exist when the correlation is too negative: in the former case, large trades are costly both because of price impact and of the negative effect on the follower's effort; by contrast, with negative correlation, the value of manipulation incentivizes more aggressive trading, which goes against price impact and introduces convexity in the leader's problem. Reductions in order flow volatility then play a *dual* role: for any degree of correlation, they increase the leader's ability to manipulate the continuation game, making the coordination equilibrium less plausible; but since they also increase price impact, they also restore concavity of the leader's problem when correlation is negative. Indeed, we show that reducing the volatility of noise trading not only makes our main equilibrium re-emerge, but it also eliminates all other equilibria—in other words, market illiquidity refines the equilibrium under study.

Related literature. The free-riding problem that arises when improving firms' governance is a costly activity and ownership is dispersed has been recognized as early as Berle and Means (1932). Since then, the literature has focused on two forms of activism as mechanisms through which "improved monitoring" can arise: "voice," where a blockholder takes actions that directly affect firm value (e.g., Shleifer and Vishny, 1986, Kahn and Winton (1998) and Maug, 1998); and "exit," by which a blockholder can discipline a firm's management via the ex post threat of selling shares (e.g., Admati and Pfleiderer, 2009 and Edmans, 2009). Ours is a model of voice, as our activists exert effort to shape firm value; disposal of shares can instead happen in equilibrium to induce subsequent activists to govern through voice.

To study such steering dynamics in a tractable way, we follow Back et al. (2018), who introduce private information about positions and a one-time terminal effort choice in the single-player model of Kyle (1985); their focus is on the interplay between activism technologies and market liquidity. This modeling approach is also adopted by Doidge et al. (2021), where a group of activists trade non-cooperatively only once, to later on act as a coalition (in the sense of cooperative games) at the effort stage, ameliorating the free-riding problem. Away from this framework, some papers have studied how competition among multiple blockholders can have positive effects on activism: Edmans and Manso (2011) show that exit is a stronger disciplinary threat, while Brav et al. (2021a) show that an incentive to appear as skilled can lead hedge funds to exert effort when there is competition for investor funds.

On the empirical side, following Brav et al. (2008), the traditional approach for assessing the impact of activism campaigns consists of examining measures of "abnormality:" stock price appreciation and trading volume around activism events in excess of a benchmark. In our model, prices depart from the no-distortion benchmark in which traders do not change their position on average (e.g., Kyle, 1985), which is a natural proxy for "normal times".[5] As we already stated, we can then connect those price departures with studies that document abnormality measures for sub-samples of firms defined by characteristics that we can link to our key parameters of study (correlation and number of activists).

From an institutional viewpoint, our model is related to models of manipulation in which trading operates as a vehicle for influencing actions that can have real consequences: in Goldstein and Guembel (2008), short-selling can be a profitable strategy for a speculator when it induces a manager to forgo an investment decision; in Attari et al. (2006), a passive fund may dump shares to insure the value of the remaining block, as activism by a second investor has positive return only when a firm's fundamentals are low; in Khanna and Mathews (2012) a blockholder buys more shares to counter a speculator's attempt to lower firm value; and in Yang and Zhu (2021), Boleslavsky et al. (2017), and Ahnert et al. (2020), strategic trading can trigger interventions by governments. Chakraborty and Yılmaz (2004), Brunnermeier (2005) and Williams and Skrzypacz (2020) feature models of manipulation in financial markets that instead focus on how the institutional details of the economic environment can lead to novel patterns of trading.

Finally, from a modeling viewpoint, our setting relates to models of belief manipulation employing Gaussian fundamentals and/or shocks in settings other than financial markets, e.g., Holmström (1999), Cisternas (2018), Bonatti and Cisternas (2020), Cetemen (2020), and Ekmekci et al. (2020). A key distinction of our setting is that noisier signals (here, order

[5] This can happen if both activists trade simultaneously once, or if the leader is myopic, or if there is a single trader; but also if traders are not interested in changing the firm's value (in which case no trading takes place, as private information is about positions).

flows) can lead to more manipulation, despite beliefs (here, prices) becoming less responsive. Indeed, a larger trade by the leader resulting from higher volatility leads to a bigger terminal block all else equal; but this necessarily incentivizes more dampening due to the follower's added value being applied to even more shares, a strategic effect that can be strong.

2 Model

Setup. A *leader* activist (she) and a *follower* counterpart (he) hold initial positions of X_0^L and X_0^F shares in a firm, respectively. Each activist's *block* is her/his private information, and such *types* are normally distributed with mean μ, variance ϕ, and covariance $\rho \in [-\phi, \phi]$.

Actions unfold in three periods. In period 1, the leader acts as a single informed trader in a Kyle (1985) market structure. Specifically, she submits an order for $\theta^L \in \mathbb{R}$ units of the firm's stock to a competitive market maker who executes it at a public price P_1 after observing the total order flow of the form

$$\Psi_1 = \theta^L + \sigma Z_1.$$

In this specification, Z_1 is standard normal random variable independent of the initial positions that captures noise traders, and the volatility $\sigma > 0$ is a commonly known scalar.

Having observed the first-period price, in period 2 the follower replaces the leader as the single informed trader in an otherwise identical round of trading: he orders $\theta^F \in \mathbb{R}$ units from the same market maker who in turn executes at a (public) price P_2 after observing the total order flow

$$\Psi_2 = \theta^F + \sigma Z_2,$$

where Z_2 is standard normal and independent of (X_0^L, X_0^F, Z_1). Let $(\mathcal{F}_t^\Psi)_{t=0,1,2}$ denote the public filtration, i.e., the information generated by the prior and the order flows $(\Psi_t)_{t=1,2}$.

Finally, in period 3, the activists simultaneously take actions that determine the firm's fundamentals. Specifically, activist i exerts effort $W^i \in \mathbb{R}$ at a cost $\frac{1}{2}(W^i)^2$, $i \in \{L, F\}$, resulting in each share of the firm having a *true value* of

$$W = W^L + W^F.$$

In other words, the firm's fundamental value in the absence of any activism has been normalized to zero.[6]

[6]Note that our model allows for negative effort, which can be seen as value destruction. Bliss et al. (2019) provide some specific examples of negative activism, where blockholders take costly actions to reduce firm value; these include exerting effort to provide negative information about firm fraud, challenging firm

Payoffs. Let us introduce two preliminary pieces of notation. First, to link players with their corresponding trading periods, define $t(L) := 1$, $t(F) := 2$, $i(1) := L$, and $i(2) := F$. Second, we denote the activists' *terminal* positions by

$$X_T^i = X_0^i + \theta^i, \; i \in \{L, F\}. \tag{1}$$

Activist i maximizes the value of its holdings net of trading and effort costs:

$$\sup_{\theta^i, W^i} \mathbb{E}\left[\left(W^i + W^{-i}\right) X_T^i - P_{t(i)}\theta^i - \frac{1}{2}(W^i)^2 | X_0^i, \mathcal{F}_{t(i)-1}^{\Psi}, \theta^i\right]. \tag{2}$$

Clearly, the first-order condition with respect to effort W^i implies that

$$W^i = X_T^i, \; i \in \{L, F\}. \tag{3}$$

Hence, activist i's objective (2) is effectively

$$\sup_{\theta^i} \mathbb{E}\left[(X_T^i + X_T^{-i})X_T^i - P_{t(i)}\theta^i - \frac{1}{2}(X_T^i)^2 | X_0^i, \mathcal{F}_{t(i)-1}^{\Psi}, \theta^i\right]. \tag{4}$$

Let us highlight a few noteworthy features of the model:

1. *Flexible correlation.* Correlation in our model can take any value. Since a mix of a "long" (e.g., positive position) and a "short" (e.g., negative position) activist is more likely when correlation is negative, we would expect the latter case to be more representative of real-world activism cases where the target firm features traders with large "short" positions. Conversely, a prevalence of similar activists holding "long" positions is indicative of positive correlation. Importantly, however, since in practice there is a fixed number of shares, an element of negative correlation is always at play (if an activist's position is too large, others are necessarily small, and vice-versa). This tension is then likely to be less acute in large-cap firms, followed by mid-cap, and lastly small-cap, if shares outstanding on average decrease as market capitalization falls—Section C in the Online Appendix documents this pattern. The upshot is that positive/negative correlation likely become more plausible as market capitalization increases/decreases.

2. *Free-riding and alternative activism technologies.* The perfect substitutability of effort choices in the determination of the firm's value offers a stark representation of the traditional free-rider problem at play: all shareholders benefit from activism undertaken

patents or blocking favorable acquisitions by the firm.

by any individual blockholder. In this line, while we have chosen a specification of fundamentals that varies continuously with total effort, the model can also represent a linear approximation of engagements with binary outcomes (e.g., proxy fight), and where the probability of success is likewise increasing in aggregate effort.

3. *Endogenous fundamentals and incentives to trade.* The firm's endogenous value $X_T^i + X_T^{-i}$ implies that, relative to single-player setup with exogenous fundamentals, the *static* incentives to trade are modified through two channels: a higher fundamental value due to the extra effort exerted (X_T^i term), reflecting that larger blocks do translate to stronger interventions; and a higher or lower fundamental value depending on what the other activist will do (X_T^{-i}), which is linked to how positions are correlated initially. As we will see, these *direct* effects result in steeper/weaker incentives that are priced in the form of stronger/weaker price impact. On the other hand, the value of the the holdings for each activist is given by $(X_T^i + X_T^{-i})X_T^i$, $i \in \{L, F\}$, so a *dynamic* complementarity between orders and terminal positions across players is at play. For the leader, the higher her position, the more she benefits from inducing a higher position for the follower—this *strategic* effect is a key driver of our findings.

Linear Strategies and Equilibrium Concept. A trading strategy for a player is *linear* if it conditions on the history of signals observed by that player in a linear way. That is,

$$\theta^L = \alpha_L X_0^L + \delta_L \mu, \tag{5}$$

for the leader, while the follower can also condition on the first-period price:

$$\theta^F = \alpha_F X_0^F + \beta_F P_1 + \delta_F \mu. \tag{6}$$

Similarly, a pricing rule is linear if $P_{t(i)}$ is affine in $\Psi_{t(i)}$, $i = L, F$. As is traditional, we will be looking for linear equilibria: (i) the activists' linear strategies are mutual best-responses when taking as given a linear pricing rule set by the market maker, and (ii) the market maker's linear pricing rule satisfies $P_{t(i)} = \mathbb{E}[W^L + W^F | \mathcal{F}_{t(i)}^{\Psi}]$.

Our main goal will be to characterize linear equilibria exhibiting $\alpha_L > 0$ and $\alpha_F > 0$, i.e., trading strategies exhibiting *positive block sensitivity* (PBS). Thus, larger leader/follower blockholders acquire relatively more stock than their smaller counterparts, which means that trading only increases the relative strength of engagement across types. The question of other linear equilibria beyond the PBS class is relegated to Section 7.

PBS equilibria also conform with the linear equilibria usually examined in the literature, where informed traders place orders that have positive sensitivity to their private information.

From this perspective, it is of special interest whether $\mathbb{E}[\theta^L|\mathcal{F}_1^\Psi] = \mathbb{E}[\theta^F|\mathcal{F}_1^\Psi] = 0$, that is, the activists behave in an unpredictable manner as in Kyle (1985) and its generalizations.

3 Learning and Pricing

We begin by characterizing learning and pricing, fixing conjectured strategies (5)-(6). We frequently use the projection theorem for Gaussian random variables: if x and y are jointly normally distributed, then $\mathbb{E}[y|x] = \mathbb{E}[y] + \frac{\text{Cov}(x,y)}{\text{Var}(x)}(x - \mathbb{E}[x])$ and $\text{Var}(y|x) = \text{Var}(y) - \frac{\text{Cov}^2(x,y)}{\text{Var}(x)}$. Supporting details and expressions are in the Online Appendix.

3.1 Initial beliefs

First-period quoted price. We begin by characterizing the market maker's ex ante expectation of firm value, $P_0 = \mathbb{E}[X_T^L + X_T^F]$, which corresponds to the price quoted to the leader before placing an order, and is needed for calculating execution prices. Using (5)-(6),

$$P_0 = \mathbb{E}[(1 + \alpha_L)X_0^L + \delta_L \mu + (1 + \alpha_F)X_0^F + \beta_F P_1 + \delta_F \mu].$$

Since $\mathbb{E}[P_1] = P_0$, we can solve for P_0 as a function of μ as long as $\beta_F \neq 1$. We show in Remark 1 that this must hold in any linear equilibrium, so we assume it in what follows and verify ex post that our candidate equilibrium satisfies it.

Players' private beliefs. Correlation in privately known initial positions implies that the players have private beliefs about each others' positions. Throughout, we use Y_t^i to denote player i's private (mean) belief about the position of player $-i$ following period t. Therefore,

$$Y_0^i = \mu + \frac{\rho}{\phi}(X_0^i - \mu), \qquad \nu_0^i := \text{Var}(X_0^{-i}|X_0^i) = \phi - \frac{\rho^2}{\phi}.$$

3.2 First-period updating

The market maker's belief updating. After observing the first-period total order flow, Ψ_1, the market maker updates beliefs about both activists' positions. We begin with the corresponding (public) belief about the leader's *initial* position, which reads

$$\mathbb{E}[X_0^L|\mathcal{F}_1^\Psi] = \mu + \frac{\alpha_L \phi}{\alpha_L^2 \phi + \sigma^2} \{\Psi_1 - \mu(\alpha_L + \delta_L)\}. \tag{7}$$

Now, letting (M_1^L, M_1^F) denote the posterior belief about the *contemporaneous* positions (X_T^L, X_0^F), we get $M_1^L = (1+\alpha_L)\mathbb{E}[X_0^L|\mathcal{F}_1^\Psi] + \delta_L\mu$ after using (5). Similarly,

$$M_1^F := \mathbb{E}[X_0^F|\mathcal{F}_1^\Psi] = \mu + \frac{\alpha_L\rho}{\alpha_L^2\phi + \sigma^2}\left\{\Psi_1 - \mu(\alpha_L + \delta_L)\right\} \tag{8}$$

where the only difference is the presence of the covariance term ρ. In particular, using (6), $\mathbb{E}[X_T^F|\mathcal{F}_1^\Psi] = (1+\alpha_F)M_1^F + \beta_F P_1 + \delta_F\mu$.

Let $\begin{pmatrix} \gamma_1^L & \rho_1 \\ \rho_1 & \gamma_1^F \end{pmatrix}$ denote the posterior covariance matrix of the market maker's beliefs about (X_T^L, X_0^F) after period one (see (B.2)-(B.4) in the Online Appendix). Intuitively, while at this stage price impact will naturally depend on the extent of initial uncertainty about positions, in the next stage the updated uncertainty about the follower's initial position will determine his informational advantage relative to the market maker.

First-period pricing. The market maker sets a first-period execution price according to $P_1 = \mathbb{E}[X_T^L|\mathcal{F}_1^\Psi] + \mathbb{E}[X_T^F|\mathcal{F}_1^\Psi]$. By the projection theorem,

$$P_1 = P_0 + \Lambda_1\left\{\Psi_1 - \mu(\alpha_L + \delta_L)\right\}, \text{ with} \tag{9}$$

$$\Lambda_1 := \frac{\alpha_L\phi}{\alpha_L^2\phi + \sigma^2} \times \frac{1 + \alpha_L + \rho(1+\alpha_F)/\phi}{1-\beta_F}. \tag{10}$$

That is, the price responds to unexpected realizations of the order flow, with the intensity of the response given by Λ_1, usually referred to as *price impact.*

In the expression for Λ_1, the first fraction is well-known: it is the price impact that arises when the firm's value is normally distributed with variance ϕ. The second fraction in turn reflects the endogeneity of such fundamentals. Specifically, the numerator encodes how different types take different actions that influence the firm: the term α_L captures that large unanticipated total orders are now even more indicative of higher fundamentals because, as higher leader types purchase more units, they will also exert more effort in correspondence with their trade; $\rho(1+\alpha_F)/\phi$ in turn captures that more or less firm value can also originate from the follower's effort depending on how types correlate.

The denominator $1-\beta_F$ encodes how the first-period price affects the firm's value via the channel of the follower's trade: an increase in the order flow that leads the market maker to believe the firm has higher value affects firm's fundamentals by β_F, which further affects the market maker's pricing of the firm, thereby influencing the follower's trade again by β_F, and so forth. As long as the slope β_F is different from 1 (as it must be in equilibrium — see Remark 1), the price is always well defined once accounting for this amplification mechanism.

The follower's posterior belief. To set up the follower's best response problem, we need the follower's updated belief about the leader's terminal position given the first-period price:

$$Y_1^F := (1+\alpha_L)\left[Y_0^F + \frac{\alpha_L \nu_0^F}{\alpha_L^2 \nu_0^F + \sigma^2}\left\{\frac{P_1 - P_0}{\Lambda_1} + \alpha_L(\mu - Y_0^F)\right\}\right] + \delta_L \mu. \tag{11}$$

Via Y_0^F, (11) is a function of the follower's state variables (X_0^F, P_1, μ), as desired.[7]

3.3 Second-period updating

Second-period pricing. Observing Ψ_2, the market maker sets a second-period execution price of $P_2 = \mathbb{E}[X_L^T + X_F^T | \mathcal{F}_2^\Psi]$. Using that $M_T^L := \mathbb{E}[X_T^L|\mathcal{F}_2^\Psi]$ and $M_T^F := \mathbb{E}[X_T^F|\mathcal{F}_2^\Psi]$ can be written as linear functions of μ, P_1, and Ψ_2, we obtain

$$P_2 = P_1 + \Lambda_2[\Psi_2 - \alpha_F M_1^F - \beta_F P_1 - \delta_F \mu], \text{ with} \tag{12}$$

$$\Lambda_2 = \frac{\alpha_F \gamma_1^F}{\alpha_F^2 \gamma_1^F + \sigma^2} \times [1 + \alpha_F + \rho_1/\gamma_1^F]. \tag{13}$$

Equations (12)–(13) admit the same interpretation as (9)–(10). Notice that there is no $(1+\alpha_L)$ term accompanying ρ_1/γ_1^F in the price impact wedge because ρ_1 carries it implicitly, as ρ_1 denotes the correlation between the leader's terminal position and the follower's initial one. There is also no denominator because P_2 does not affect the firm's value.[8]

Finally, while the leader *could* update about the follower using P_2 (or Ψ_2), this is payoff-irrelevant. This is because (i) she does not trade again, and (ii) each activist's optimal effort is independent of the other's.

4 Equilibrium Trading

Using (4), the best-response problem of player $i \in \{L, F\}$ reads

$$\sup_\theta \; -\theta\, \mathbb{E}_i[P_{t(i)-1} + \Lambda_{t(i)}\{\Psi_{t(i)} - \mathbb{E}[\Psi_{t(i)}|\mathcal{F}_{t(i)-1}^\Psi]\}|\theta] + \frac{(X_0^i + \theta)^2}{2} + (X_0^i + \theta)\mathbb{E}_i[X_T^{-i}|\theta], \tag{14}$$

where $\mathbb{E}_i[\cdot] := \mathbb{E}[\cdot|\mathcal{F}_{t(i)-1}^\Psi, X_0^i]$ is player i's conditional expectation operator at the beginning of period $t(i)$.

[7] The follower needs to use the order flow Ψ_1 to form his posterior belief in (11). Since $\Lambda_1 \neq 0$ in any linear equilibrium (see Remark 1), he can infer Ψ_1 from P_1 via (9).

[8] Note, again, that as Ψ_1 can be inverted from P_1, M_1^F in (12) is ultimately an affine function of (μ, P_1). Thus, (X_0^F, P_1, μ) are sufficient for the follower's best response problem.

The only structural difference between the players' problems lies in each activist's ability to influence the other's terminal position, which is captured by the last term, $\mathbb{E}_i[X_T^{-i}|\theta]$. From this perspective, since the leader has already moved when the follower gets to trade, this latter term is exogenous in the follower's problem, so his first-order condition reads

$$0 = -\mathbb{E}_F[P_1 + \Lambda_2\{\Psi_2 - \mathbb{E}[\Psi_2|\mathcal{F}_1^\Psi]\}|\theta] - \theta\Lambda_2 + (X_0^F + \theta) + Y_1^F. \tag{15}$$

On the other hand, the leader's counterpart is

$$0 = -\mathbb{E}_L[P_0 + \Lambda_1\{\Psi_1 - \mathbb{E}[\Psi_1]\}|\theta] - \theta\Lambda_1 + (X_0^L + \theta) + \mathbb{E}_L[X_T^F|\theta] + (X_0^L + \theta)\frac{\partial \mathbb{E}_L[X_T^F|\theta]}{\partial\theta}, \tag{16}$$

where the last term captures the leader's ability to affect the follower's terminal position by influencing follower's trade in the second period.

The second-order conditions (SOCs) for the players also have similar forms:

$$1 - 2\Lambda_1(1-\beta_F) < 0, \text{ for } i = L, \tag{17}$$

$$1 - 2\Lambda_2 < 0, \text{ for } i = F. \tag{18}$$

The scalar 1 in (17)-(18) reflects a convexity in the players' payoffs that arises from the interaction between endogenous terminal effort and earlier trades, which is in contrast to a standard *static* setup with exogenous fundamentals. On the other hand, $(1-\beta_F)\Lambda_1$ in (17) reflects the leader's effective cost for the last unit traded: the direct impact on the price net of the change in the asset value due to the sensitivity of the follower's effort to the market price. As the expression for Λ_1 in (10) shows, however, these steeper or weaker incentives arising from the latter channel are fully anticipated by the market maker and hence perfectly priced, which results in the effective cost being independent of β_F.

Remark 1. *The second-order conditions (17)-(18) must hold given any linear pricing rules where the sensitivities Λ_1 and Λ_2 are general scalars. Thus, $\beta_F \neq 1$ must hold in a candidate equilibrium for part (i) of the equilibrium concept to be satisfied.*

4.1 The follower's trading

Finding an equilibrium is challenging because first-period variables depend on second-period ones by backward induction, and the latter depend on the former via learning; further, all players' beliefs must be correct. To simplify the exposition, we describe the follower's and leader's behavior separately, beginning with the follower.

Proposition 1. *In a PBS equilibrium:* $\alpha_F = \sqrt{\sigma^2/\gamma_1^F}$*;* $\beta_F < 1$*, with* $sign(\beta_F) = -sign(\rho)$*; and* $\delta_F < 0$*. Further, in belief space, the follower's trade admits the representation*

$$\theta^F = \alpha_F(X_0^F - M_1^F). \tag{19}$$

Hence, the follower's trade is zero in expectation: $\mathbb{E}[\theta^F] = \mathbb{E}[\theta^F|\mathcal{F}_1^\Psi] = 0$*. In the particular case of* $\rho = 0$*, both players trade according to (19) with* $\gamma_1^F = \phi$ *and* $M_1^F = \mu$ *(and this constitutes the unique linear equilibrium).*

It is expected that the weight on the prior, δ_F, is negative — we defer a formal explanation to the next section, where we discuss the leader's counterpart. To understand why β_F and ρ have different signs, it is useful to consider the representation (19). Consider the case of positive correlation: a high price is indicative of a leader with a high type, which leads the market maker to update positively on the follower's position. As the informational wedge in (19) falls, the follower buys less; in other words, higher prices lead to lower purchases by the follower, so $\beta_F < 0$. Conversely, with negative correlation, high first-period prices map to low market maker's beliefs about the follower, and hence to more aggressive buying by the latter agent: β_F must be positive.

There are two noteworthy aspects of (19). First, trades are a function of an information wedge but not explicitly a function of mispricing, i.e., the difference between the firm's true value and the market maker's perception of it. The reason is that, with linear trading and effort strategies and Gaussian learning, fundamental mispricing, $\mathbb{E}[W^L + W^F|\mathcal{F}_1^F] - \mathbb{E}[W^L + W^F|\mathcal{F}_1^*]$, is proportional to $X_0^F - M_1^F$; thus, the latter proxies for the extent of mispricing.[9]

Second, the intensity of trading, $\alpha_F = \sqrt{\sigma^2/\gamma_1^F}$, is exactly as in a one-shot (Kyle) counterpart with exogenous fundamentals, despite price impact Λ_2 in (13) exhibiting the novel term $1 + \alpha_F + \rho_1/\gamma_1^F \neq 1$. The reason is that this wedge effectively reflects the pricing done by the market maker in response to the change in the follower's marginal benefit to trade relative to a single-player setting with exogenous fundamentals: follower's effort complementing his trading, and the leader affecting the firm's value in a correlated manner (see Section 2). With trading costs that adjust perfectly to the change in benefits, the traditional trading intensity encountered in the literature is recovered.

Finally, regarding the last part of the proposition, if the initial positions are i.i.d. the market maker learns nothing about the follower from the first-period trade, so $M_1^F = \mu$ and $\gamma_1^F = \phi$. But this means that the continuation game is unresponsive to the leader's behavior,

[9] It is easy to see that $\mathbb{E}[W^F|\mathcal{F}_1^F] - \mathbb{E}[W^F|\mathcal{F}_1^*] \propto X_0^F - M_1^F$ and $\mathbb{E}[W^L|\mathcal{F}_1^F] - \mathbb{E}[W^L|\mathcal{F}_1^*] \propto \mathbb{E}[X_0^L|\mathcal{F}_1^F] - \mathbb{E}[X_0^L|\mathcal{F}_1^*]$. With Gaussian learning, however, the follower's private belief about the leader's initial position combines his type X_0^F and the first-period order flow, Ψ_1, *linearly*. Thus, the market maker's belief is a linear combination of M_1^F and Ψ_1 with the same weights, so $\mathbb{E}[X_0^L|\mathcal{F}_1^F] - \mathbb{E}[X_0^L|\mathcal{F}_1^*] \propto X_0^F - M_1^F$.

and hence static behavior is optimal for her too. In what follows, we assume $\rho \neq 0$.

4.2 The leader's trading and PBS equilibrium

We now present a central result of this paper. To this end, let $\alpha^K := \sqrt{\sigma^2/\phi}$ denote the traditional (Kyle) trading intensity when the prior variance is ϕ.

Proposition 2. *Fix $\sigma > 0$. There exists $\underline{\rho} \in (-\phi, 0)$ such that for all $\rho \in [\underline{\rho}, \phi]$, there exists a PBS equilibrium. In any such equilibrium, the leader trades according to $\theta^L = \alpha_L X_0^L + \delta_L \mu$, where $\alpha_L > 0$ and $\delta_L < 0$. Moreover, if $\rho > 0$, then*

$$\alpha_L < \alpha^K < -\delta_L,$$

and the reverse inequalities hold if $\rho \in (-\underline{\rho}, 0)$. In turn, the follower trades as in (19). There also exists $\rho_0 \in [\underline{\rho}, 0)$ such that there is a unique PBS equilibrium for all $\rho \in [\rho_0, \phi]$, and α_L is decreasing in ρ on this interval.

In a PBS equilibrium, the leader's strategy departs from the traditional ones in the microstructure literature in that the weights attached to the type and prior diverge from α^K in opposite directions, with a ranking that depends on the correlation of positions. Let us now explain the economics behind this result, deferring a detailed discussion about the lower bound $\underline{\rho}$ to Section 7.[10]

The result stems from a combination of dynamic incentives and endogenous costs. Regarding the former, recall from the leader's first-order condition (16) that her incentives are distorted by $X_T^L \frac{\partial \mathbb{E}_L[X_T^F|\theta]}{\partial \theta^L}$ relative to the follower's. This term captures the leader's *value of manipulation*, i.e., the component of her continuation value that relates to the follower's behavior. Using that $\theta^F = \alpha_F X_0^F + \beta_F P_1 + \delta_F \mu$, this term reads

$$X_T^L \frac{\partial \mathbb{E}_L[X_T^F|\theta]}{\partial \theta^L} = X_T^L \beta_F \frac{\partial P_1}{\partial \Psi_1} = X_T^L \beta_F \Lambda_1. \tag{20}$$

To illustrate, consider the positive correlation case. There, $\beta_F < 0$, suggesting that the leader would like to engage in a *downward deviation* from a traditional gap strategy. Intuitively, high/low first-period order flows Ψ_1 (and hence first-period prices) are indicative of a high/low type of the follower, so the market maker's belief about the follower M_1^F satisfies $\partial M_1^F/\partial \Psi_1 > 0$. Thus, by (19), manipulating M_1^F downwards implies that a larger arbitrage opportunity is created for the follower, so the latter would build up his position

[10] Numerically, we have not found multiple PBS equilibria in the region $\rho \in [\underline{\rho}, \rho_0)$.

more. But with a bigger block, the follower would exert more effort, resulting in more firm value that the leader can enjoy.

The proposition's ranking of the leader's strategy coefficients precisely encodes this type of deviation. To see why, notice first that in the value of manipulation, leaders with higher terminal positions benefit more from reducing their purchases, as the additional value stemming from the follower's extra effort is applied to more units. With the coefficient α_L on the leader's type being positive, higher types indeed end up holding larger blocks; but since $\alpha_L < \alpha^K$, these types effectively end up scaling back more at the same time.

To rationalize $\delta_L < -\alpha^K$, i.e., an increased sensitivity to the prior in the leader's strategy, we need to incorporate the endogenous cost aspect of the analysis: price impact. Specifically, it is easy to show that δ_L satisfies

$$\delta_L = \frac{1}{(1-\beta_F)\Lambda_1} \times \frac{\partial}{\partial \mu}(\mathbb{E}_L[W_L + W_F] - P_1), \tag{21}$$

i.e., it corresponds to the sensitivity of mispricing to the prior, scaled by the effective marginal cost of trading. The derivative is always negative: in forecasting the firm's value, the market maker relies more on the prior than the leader does, in a reflection of the market maker's (leader's) informational disadvantage (advantage). As μ grows, therefore, *all types* scale back because their arbitrage opportunities shrink. But with a lower signaling coefficient α_L, there is less price impact for each fixed $\rho > 0$ than with α^K, holding everything else fixed (see Λ_1 in (10)). Further scaling back in response to an increase in μ is then less costly, as the trading losses become smaller. Thus, we conclude that all types deviate downwards on both dimensions of information, private and public.[11]

We conclude with two observations. First, the logic is identical with negative correlation: an unexpectedly *high* first-period order flow is now a signal of the follower having a lower initial position, and the market maker's belief falls—all leader types then find it optimal to buy more aggressively, i.e., $\alpha_L > \alpha^K$, and hence $-\alpha^K < \delta_L$ via the price impact channel. More generally, the signaling coefficient α_L is decreasing in ρ, reflecting that the value of manipulation across leader types is larger when initial positions exhibit a stronger statistical linkage: as $|\rho|$ grows, the market maker relies more on the first-period order flow to learn about the follower, so it is profitable for the leader to manipulate more.

Second, in equilibrium the follower may also buy more because the market becomes more

[11] The analogous expression for the signaling coefficient is $\alpha_L = \frac{1}{(1-\beta_F)\Lambda_1}\frac{\partial}{\partial X_0^L}(\mathbb{E}_L[W_L+W_F]-P_1)+\frac{\beta_F}{1-\beta_F}$, which is an *equation* for α_L (in contrast to (21), where δ is absent in the right-hand side due to canceling out in the difference). The derivative is now positive by the same logic, while the last term stems from the value of manipulation, e.g., $\beta_F < 0$ when $\rho > 0$, and there is downward pressure on α_L; the denominator in turn captures the amplification effect discussed in Section 3 applied to all inframarginal units.

liquid: any informational advantage $X_0^F - M_1^F$ gets amplified by $\sqrt{\sigma^2/\gamma_1^F} > \sqrt{\sigma^2/\phi}$. This, however, is an effect that arises in equilibrium only, as deviations by the leader are hidden.[12]

5 Predictions

In this section, we first explore the implications of the PBS equilibrium for market outcomes: order flows, market liquidity, and firm value. We then assess the plausibility of this equilibrium from the lens of first-mover advantages: what factors—namely correlation, liquidity, and number of followers—incentivize an activist to act as a leader? The answers to these questions pave the way for our main application in Section 6.

5.1 Market Outcomes

Let $\mathbb{E}[\cdot]$ denote the expectation operator with respect to the prior distribution. Note that absent any trading, the firm would take value $\mathbb{E}[X_0^L + X_0^F] = 2\mu$—hence, we assume $\mu > 0$ in what follows. The next result characterizes average order flows and firm values.

Proposition 3. *In any PBS equilibrium,*

(i) Order flow: $\mathbb{E}[\Psi_1] < 0$ *if and only if* $\rho > 0$, *while* $\mathbb{E}[\Psi_2] = \mathbb{E}[\Psi_2|\mathcal{F}_1^\Psi] = 0$ *for all* ρ.

(ii) Firm value and average prices: $\mathbb{E}[W^L + W^F] = \mathbb{E}[P_1] = \mathbb{E}[P_2] = (2 + \alpha_L + \delta_L)\mu$, *which is (ii.1) less than* 2μ *if and only if* $\rho > 0$, *and (ii.2) always greater than* μ. *Moreover, for* $\rho \geq \rho_0$, *ex ante firm value (and hence, average prices) is decreasing in* ρ.

(iii) Price impact: $\partial\Lambda_1/\partial\rho > 0$ *in a neighborhood of* $\rho = 0$.

The existence of a statistical link between positions has sharp implications for outcomes. With positive correlation, for instance, the resulting downward deviation manifests in first-period *selling pressure*: leader types sell on average, and the expected order flow, $\mathbb{E}[\Psi_1]$, is negative; obviously, the opposite occurs when correlation is negative, as buying pressure emerges in this case. By contrast, the second-period order flow satisfies $\mathbb{E}[\Psi_2|\mathcal{F}_1^\Psi] = 0$, and hence P_2 updates in the direction of the order flow as is traditional in the literature.

Consequently, when $\rho < 0$, the manipulation motive increases the firm's ex ante value relative to a world in which blockholders do not change their positions on average (or simply do not trade), and the opposite occurs when $\rho > 0$ (part (ii.1)). Importantly, by the law of

[12] The form of manipulation uncovered is reminiscent of *encouragement effects* in teams, e.g., Bolton and Harris (1999) and Cetemen et al. (2019). With positive correlation, a key distinction is that our mechanism operates via inducing pessimism about the underlying fundamentals: lowering the firm's price, corresponding to the market maker's belief about the firm, and also the follower's belief about the leader's contribution.

iterated expectations, these conclusions map to predictions regarding prices: cross-sectional averages of prices during activism events should differ from those that arise in "normal times" defined as situations where the manipulation motive is absent, such as when activism itself is absent. Moreover, average stock prices should be "abnormally" higher during activism events where negative correlation is likely at play, and vice-versa.[13]

Part (ii.2) then establishes that, despite the increased inefficiencies when $\rho > 0$, the presence of a leader is still desirable: for any level of initial correlation (positive or negative), the firm's ex ante value is higher than its counterpart value when the follower acts as a lone activist (which corresponds to $\mathbb{E}[X_0^F + \alpha^K(X_0^F - \mu)] = \mathbb{E}[X_0^F] = \mu$). To understand why, notice that since the follower's average contribution to the firm is μ, the leader lowers the firm's value if and only if $\mathbb{E}[X_T^L] = \mathbb{E}[X_0^L + \alpha_L X_0^L + \delta_L \mu] = (1 + \alpha_L + \delta_L)\mu < 0$, i.e., when the leader ends up *reversing her initial position*. But the leader's efforts to transfer costs to the follower work so long as that they preserve or induce correlation in terminal positions ($X_T^L X_T^F$ term)—it is intuitive that an equilibrium exhibiting manipulation motives, where the goal is to steer a counterparty's behavior in a desired direction, should not exhibit such a reversal by the agent performing the manipulation.[14]

Part (iii) states that price impact is increasing in ρ, at least around $\rho = 0$, which is opposite the relationship between the signaling coefficient α_L and ρ (see Proposition 2 and Figure 1). This is a familiar finding: in equilibrium, the extent of insider trading, and hence of information transmission, is naturally disciplined by the strength of price impact. We note that (ii) seems to hold for all values of ρ, as seen in Figure 1; away from $\rho = 0$, the difficulty is purely technical in that α_L satisfies a non-linear equation (see (23) in Section 7).

Proposition 3 provides theoretical support for some empirical findings in the literature. First, Becht et al. (2017) show that activism by multiple hedge funds performs "strikingly better" than single-activist engagements; this is consistent with (ii) in Proposition 3, which holds for all non-trivial correlations. Second, Li et al. (2022) show that activists' buy-and-hold abnormal returns—the canonical proxy measure for the performance of blockholder activism—is larger for firms featuring traders with large *short* positions; in our setting, a

[13]Lower ex ante firm value or prices when $\rho > 0$ is a rather strong prediction because we are averaging across all block types, and activism could be subject to selection effects in practice. Our result, however, uncovers how the sign of correlation defines the direction of a force that alters the free-riding at play—the way in which the leader transfers activism costs to the follower—and that ultimately has real consequences.

[14]This is transparent when correlation is positive. With negative correlation, notice that a sufficiently negative leader type would sell in a PBS equilibrium, thereby lowering the price and inducing the follower to trade less aggressively due to $\beta_F > 0$—in other words, such a leader is effectively trying to bring the follower to her (short) side. Of course, these are properties specific to a PBS equilibrium, whose distinctive feature is the manipulation motive. In this regard, in Section 7 we explore other linear equilibria in which one or two activists place a negative weight on their initial positions—there, large types, either positive or negative, do reverse their positions, but because of a coordination motive.

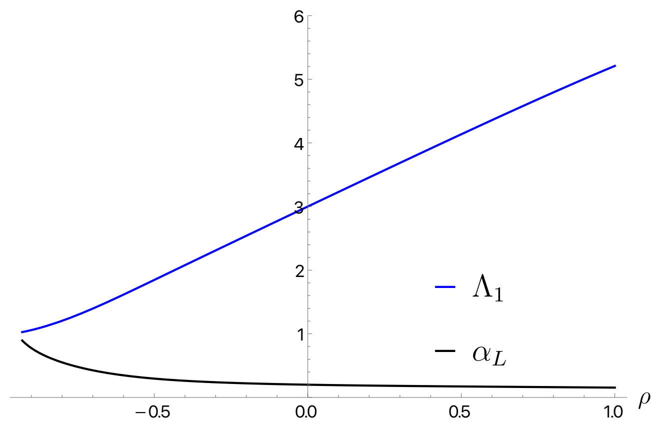

Figure 1: Price impact and the leader's signaling coefficient as functions of covariance in initial positions. Parameter values: $\mu = \phi = 1$, $\sigma = .2$.

mix of activists holding long and short positions is more likely when $\rho < 0$, and it is precisely there that both the extent of insider trading (as measured by α) and average prices are higher; further, both measures increase as $\rho < 0$ decays. Relatedly, Cookson et al. (2022) show that greater disagreement among investors, measured using posts on a social media platform for investors, leads to more informed trading by activists and more short selling, suggesting that negative correlation between blockholders is plausible in certain circumstances. Finally, Brav et al. (2021b) show that abnormal return measures are highest for small-cap firms, followed by mid- and then large-; but if negative correlation is more likely to arise in firms with smaller market capitalization as argued, our results also conform with this finding.

We conclude this analysis of market outcomes with a discussion of noise trading volatility, σ. To this end, consider the following figure:

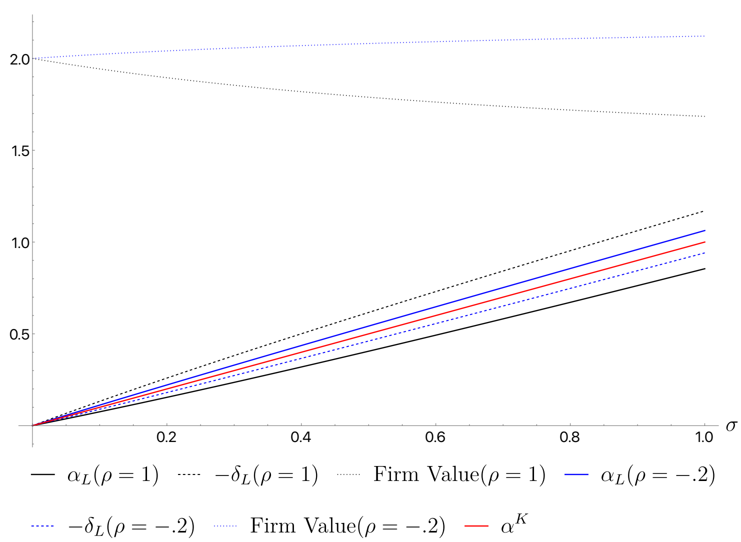

Figure 2: Leader's strategy and ex ante firm value as a function of σ. Parameter values: $\mu = \phi = 1$.

Consider first the increasing curves at the bottom, which depict the equilibrium coefficients in the leader's strategy as σ varies for two fixed levels of correlation, one positive and one negative. As a determinant of market liquidity, higher noise trading volatility suggests more aggressive trading: this is confirmed in the figure, where α_L is always increasing in σ. But higher liquidity also implies that it is more costly for the leader to steer the follower, as moving the first-period price requires larger trades. While this logic suggests less manipulation—i.e., behavior closer to Kyle's—the figure demonstrates that the wedge between α_L and δ_L actually *grows* with σ. The reason is that the benefits of manipulation grow too: as higher leader types acquire larger terminal positions through the liquidity channel than low types do, any change in the follower's behavior now affects more units, and the incentives to manipulate are steeper. Finally, if a more liquid market induces more manipulation, it must negatively affect ex ante firm value when correlation is positive, and the opposite must occur when correlation is negative; the top curves in the figure illustrate.

We confirm these findings by comparing the extreme cases of $\sigma = 0$ and $+\infty$ when $\rho > 0$, since a PBS equilibrium exists for all $\sigma > 0$ when correlation is non-negative.

Proposition 4. *Fix $\rho > 0$. In the unique PBS equilibrium,*

(i) $\lim_{\sigma\to 0} \alpha_L = \lim_{\sigma\to 0} \delta_L = 0$, *while* $\lim_{\sigma\to+\infty} \alpha_L = +\infty$ *and* $\lim_{\sigma\to+\infty} \delta_L = -\infty$.

(ii) $\lim_{\sigma\to 0} |\alpha_L - \alpha^K| = 0$ *and* $\lim_{\sigma\to+\infty} |\alpha_L - \alpha^K| > 0$.

In particular, by (ii), the benefit of manipulation survives as the market becomes infinitely liquid ($\sigma \to +\infty$), due to the leader's terminal position also growing without bound. This is in contrast with traditional models where signal noise reduces manipulation incentives due to beliefs becoming less responsive to unanticipated shocks (e.g., Holmström, 1999).

Some studies have documented a positive relationship between market liquidity and activism (e.g., Collin-Dufresne and Fos, 2015; Brav et al., 2021b). While our model does not speak to such market timing considerations, interestingly it uncovers that higher market liquidity can offer stronger incentives for price manipulation when firms' values are endogenous. Further, if positively correlated positions are more likely to arise in large-cap firms, our model suggests that these firms may suffer from exacerbated free-riding motives that lower firm value.

5.2 First-Mover Advantages

In this section, we examine conditions that favor an activist's willingness to act as a leader. To do so, we contrast our model with a simultaneous-move version in which both activists place orders at the same time in only one round of trading before ultimately exerting effort. The next result characterizes the type of equilibrium that emerges there, and leverages the tractability of the model around $\rho = 0$ for comparison.

Proposition 5. *With simultaneous moves, there exists $\rho_0^{sim} \in (-\phi, 0)$ such that for all $\rho \in [\rho_0^{sim}, \phi]$, there exists a unique symmetric PBS equilibrium.*[15] *In this equilibrium, the activists trade according to $\theta^i = \sqrt{\frac{\sigma^2}{2\phi}}(X_0^i - \mu)$, $i = L, F$. In a neighborhood of $\rho = 0$, the leader gets a higher ex ante payoff if she goes first.*

The presence of multiple contemporaneous activist traders necessarily raises the issue of *competition*, which can be clearly seen if types coincide ($X_0^L = X_0^F$): there, the activists' aggregate order is proportional to $2\sqrt{\sigma^2/2\phi}$, which is larger than $\sqrt{\sigma^2/\phi}$, the analogous coefficient if a single informed monopolist traded once. The coefficient is independent of ρ due to two forces that cancel each other out. Higher correlation effectively generates more price impact—a force that would dampen trading here—but it also incentivizes the leader to trade more aggressively by increasing her private inference about the follower's effort.

Regarding the payoff comparison, by setting $\rho = 0$ in the sequential-move game we shut down the leader's manipulation motive, which enables us to compare pure competition effects across settings. Proposition 5 then confirms that an activist wants to become a leader, and that this incentive remains such by continuity in the presence of mild manipulation motives.

Farther away from $\rho = 0$, whether a first-mover advantage arises will depend on the sign and magnitude of the correlation. Indeed, for large *negative covariance*, the presence of another trader can be beneficial because the activists are likely to be on opposite sides of the market, which means that trade can take place with little impact on the price. For $\rho > 0$, therefore, going first implies escaping from competition *and* enjoying an ability to manipulate the game; moreover, as ρ increases, the benefit is larger due to the market maker becoming more responsive to the outcome of the first-period and the downward deviation resulting in lower expenditures. By contrast, if $\rho < 0$ is sufficiently low, going first implies giving up the benefit of having a counterparty with which to trade in exchange for an ability to strategically influence firm value; further, the manipulation requires additional purchases. Figure 3 illustrates these points: ex ante payoffs for the leader in the sequential game are larger than those for an individual activist in the simultaneous-move version except when ρ is sufficiently negative.

Our mechanism is thus more likely in engagements involving activists whose stakes are not too extremely negatively correlated—this is a notion of *similarity* among activists. In reality, moreover, with a fixed float of shares available to be traded, the blocks of such

[15] The appearance of ρ_0^{sim} is analogous to that of ρ_0 in Proposition 2, which we discuss in Section 7.

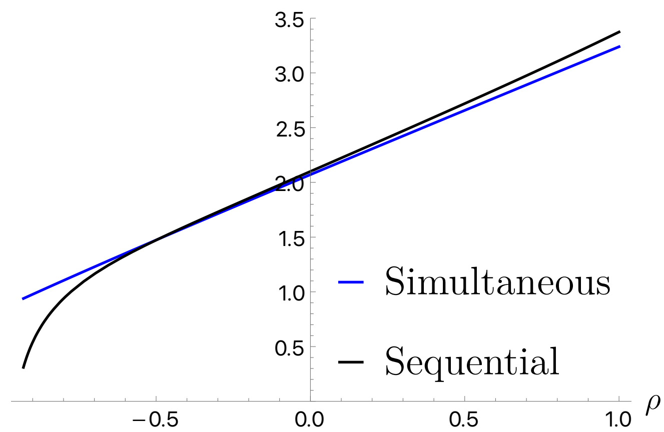

Figure 3: Leader's payoff comparison. Parameter values: $\mu = \phi = 1$, $\sigma = .2$.

similar activists cannot be too large either. Otherwise, if a prospective leader's trade is too large, it may rule out the possibility of a second activist accumulating a large positive stake; alternatively, it may reduce the possibility of sequentiality from the perspective of market makers, as larger blocks increase the odds of trade among activists. From the viewpoint of applications, therefore, our mechanism is more likely to arise in engagements where activists are all *similar* in the sense described above, and they have *small to moderate* blocks.

5.3 Number of Followers

In the last part of this section, we examine how the incentives for market leadership change as the number of followers varies. Specifically, we consider the case in which the initial stake of our original follower becomes diluted among N individuals: that is, there are N followers all with an identical initial position X_0^F, where the latter random variable is Gaussian with mean μ/N and variance ϕ/N^2, and such that $\text{Cov}(X_0^F, X_0^L) = \rho/N$. As before, the firm's value is $W^L + \sum_{i=1}^{N} W^{F,i}$, where $W^j = X_T^j$, is the effort exerted by activist j.

The reason for this normalization is twofold. First, notice that the aggregate position of the followers has mean μ, variance ϕ, and covariance ρ with the leader, just as in the baseline model; thus, the normalization rules out incentives to go first that stem from a mechanical increase in aggregate second-period uncertainty from the perspective of the leader and market maker, which would favor manipulation. Second, notice that each follower's average baseline effort—i.e., absent any trading—is decreasing in N, as his initial position has a shrinking mean. Put together, any stronger incentives to go first by the leader must necessarily come from strategic considerations in the trading game played among the followers.

In this context, we look for equilibria in which the followers play symmetric (linear) strategies in period 2: coupled with the symmetry in the followers' initial positions, we only need to keep track of the market maker's belief about a *single* follower's initial position given the observed first-period order flow; let M_1^F and γ_1^F denote the corresponding mean and variance, respectively. We concentrate on the case of positive correlation and defer a discussion of negative correlation to the end.

Proposition 6. *Fix any $\rho \in (0, \phi]$. In the unique PBS equilibrium, each follower trades according to $\theta^F = \alpha_F(X_0^F - M_1^F)$, where $\alpha_F = \sqrt{\frac{\sigma^2}{N\gamma_1^F}}$. In addition, α_F is increasing in N; both α_L and the firm's ex ante value are decreasing in N; and the leader's ex ante payoff grows in proportion to $\sqrt{N}$ asymptotically.*

The trading coefficient α_F generalizes that of Proposition 5 for the one-shot two-player case to account for N followers and an endogenous posterior variance γ_1^F. Importantly, the latter decays at rate $1/N^2$, fixing the leader's strategy. Consequently, the competition effect from Section 5.2—i.e., smaller individual trades that in total add up to more than the monopoly counterpart—is now exacerbated: since each follower's contribution to the firm is a smaller fraction of the total, the price responds less to each individual trade, prompting more aggressive behavior as N grows. With followers that are more sensitive to mispricing, the leader's manipulation incentive grows too, and so α_L decreases in N when $\rho > 0$.[16] In turn, since the followers' orders are zero on average, ex ante firm value is decreasing in N.[17]

The proposition also states that the leader's ex ante payoff is of the order $\sqrt{N}$ for N large, implying that the benefits of acting as a leader grow with the number of followers. The source of this is the interaction term $\mathbb{E}[X_T^L N X_T^F]$, which captures the value of the leader's block that is attributed to the followers' effort choices.[18] Indeed, it can be shown (Online Appendix B.3) that, for some scalar $C(N)$ that is uniformly bounded in N,

$$\mathbb{E}[(X_0^L + \theta_L)N(X_0^F + \alpha_F(X_0^F - M_1^F))] = C(N) + \alpha_F \rho (1 + \alpha_L)\frac{\sigma^2}{\alpha_L^2 \phi + \sigma^2}, \tag{22}$$

and hence payoffs grow in proportion to α_F as long as $\sigma > 0$ (recall that $\alpha_L > 0$). It is only when $\sigma = 0$ that the market maker learns the leader's type, implying that the leader and the market maker share the same belief about the follower, which effectively eliminates the possibility of creating arbitrage opportunities from the perspective of any leader type.

[16] While this decay in α_L raises γ_1^F all else equal, this effect cannot overturn the direct downward effect that larger N has on γ_1^F, as $\gamma_1^F \leq \phi/N^2$ for any linear strategy of the leader.

[17] Ex ante firm value is increasing in α_L as in the $N = 1$ case (see Proposition 3), so it is decreasing in N.

[18] No other terms depend explicitly on N or α_F. In the particular case of trading costs, for instance, it can be shown that price impact in (10) simplifies to $\frac{\text{Cov}(\Psi_1, X_T^L + X_T^F)}{\text{Var}(\Psi_1)} = \frac{\alpha_L[(1+\alpha_L)\phi + \rho]}{\alpha_L^2\phi + \sigma^2}$ in equilibrium, which is independent of N and α_F; this follow from the first-period order flow not carrying the followers' trades, and from their additional value to the firm being unpredictable from the market maker's perspective.

Two additional observations are instructive. First, the term $\alpha_F \rho$ in (22) uncovers a *complementarity* between the number of followers and the correlation among initial positions: when types are more correlated, the leader benefits from more followers because their increased trading intensity α_F leads to additional firm value that is more in line with the leader's. Figure 4 illustrates: for each fixed N, ex ante payoffs grow with ρ.

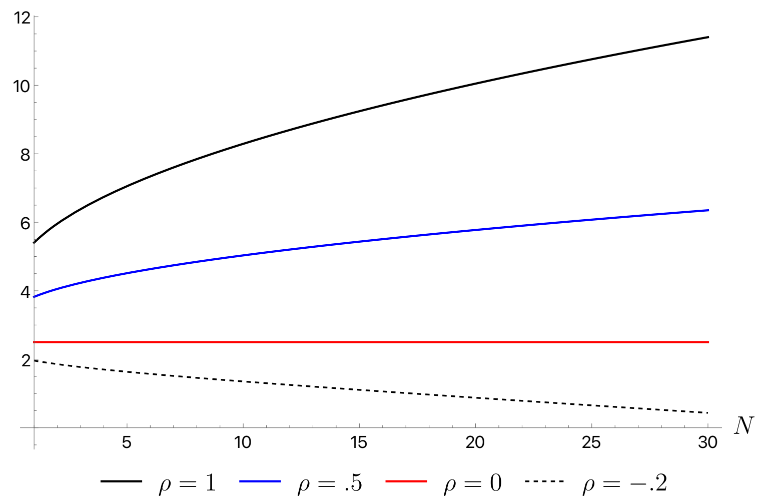

Figure 4: Leader's expected payoff as a function of the number of followers, for various levels of covariance. Other parameter values: $\phi = \mu = \sigma = 1$.

Second, the figure confirms that the leader's payoff is in fact increasing in N for fixed positive ρ, but it also shows that the leader's payoff can be decreasing in N if positions are negatively correlated (lowest curve; see also the last term in (22)). This reflects that, as the market maker anticipates the leader's incentives, the followers' more aggressive behavior may result in terminal positions that are more negatively correlated as N grows. That said, the finding does not assert that the leader ceases to find it optimal to go first, as it is the *outside option* of trading simultaneously that matters; in light of the discussion of Section 5.2 for two activists, we would expect those incentives to be weaker nonetheless.[19]

Let us bring together our results so far and offer an overview and interpretation that can be useful for our main application next. Concretely, our results suggest that our model and proposed mechanism are more likely in engagements involving activists who: (i) are sensitive to arbitrage opportunities/mispricing; (ii) hold similar stakes in a statistical sense, in that blocks are not too negatively correlated; and (iii) have small to moderate stakes. Further, we would expect our mechanism to be reinforced by the presence of follower activists that

[19]In Lemma B.1 in the Online Appendix, we confirm that the outside option is indeed less attractive when $\rho = 1$ (which simplifies the updating in the simultaneous-move game): the leader's ex ante payoff in the sequential version net of the simultaneous-move counterpart also grows in proportion to $\sqrt{N}$ for N large.

compete with one another if there is a positive statistical relationship among positions, which is more likely when the target firm is of mid- or large-capitalization. By contrast, the presence of other follower activists is less critical for a leader to arise when the firm has small capitalization and thus there is an element of negative correlation in positions at play; but if such a leader emerges, her behavior would be very aggressive.

6 Application: Wolf-Pack Activism

From an institutional viewpoint, our model rests on three core assumptions: (i) activism is a costly endeavor after building a stake; (ii) activists act non-cooperatively; and (iii) follower activists respond to arbitrage opportunities. Studies in the empirical finance literature, as well as in the legalliterature studying corporate governance, are instructive in this regard.

Costly activism. Activists seek to accomplish a variety of outcomes in target firms: in governance, restructurings such as board changes, executive compensation, or even ousting a CEO; in business strategy, takeovers, spin-offs or even selling the firm; and in capital structure, modifications in payout policies, equity issuance, buybacks and so forth. While activists' approaches vary in their aggressiveness (ranging from simple communication to litigation), the planning and execution of these outcomes requires research, consultants and legal fees that are all very costly. For instance, Gantchev (2013) estimates that, on average, making direct demands costs \$2.94M; board representation costs \$1.83M; and a final proxy battle costs \$5.94M, for a total cost of \$10.71M. Moreover, even analyzing how to vote on a proposed change by an "insurgent" entails costs, reflected in the outsourcing of these duties to "proxy" advisors that lowers overhead costs.[20] With the additional share value created benefiting all shareholders, a well-recognized free-riding problem arises.

Non-cooperative behavior. There are substantial costs associated with being perceived as a "group" from the standpoint of Section 13(d)(3) of the Securities Exchange Act.[21] At the core of these is that any activist must disclose her position within 10 days of exceeding a total 5% ownership level — an organized group of activists is thus treated as a single entity that owns a block equal to the sum of its components, with all the identities revealed in the event of disclosure. From this perspective, there are potential legal fees if the target firm alleges a violation of disclosure requirements; in contrast, if these activists are below the 5% threshold and act non-cooperatively, then due to their anonymity the firm would not be aware of them. Also, there are costs associated with disclosure: since a group must disclose

[20] Coffee Jr and Palia (2016), p. 16.
[21] Ibid 24—26.

earlier *ceteris paribus*, it necessarily invites undesired competition that makes it costly to achieve any desired block size. Additionally, the target firm may bar the acquisition of more shares by the group members, which may preclude the success of any engagement.

That said, changes in SEC regulations since 1992 imply that activists can communicate in a limited manner without this being characterized as insider trading or trading as a group—unless an explicit agreement is in place, which is argued to be a rare phenomenon (e.g., Becht et al., 2017). Consequently, activists can be aware of each other's existence. The rise of hedge fund activism—which we discuss next—is partly attributed to the resulting improved knowledge regarding the economic environment.[22]

Sensitivity to arbitrage opportunities. The activist ecosystem is multifaceted, featuring blockholders that are active in expressing their voice by jawboning firms or breaking up firms; index funds that are largely passive in that they limit themselves to voting; and in between, blockholders that mainly trade but may make their voice heard (Edmans and Holderness, 2017). In the last decades, hedge fund activism has had a meteoric rise, demonstrating greater participation from the latter category of blockholders. For instance, Brav et al. (2021b) document that, in the U.S. alone, more than 900 hedge funds have targeted more than 3,000 firms, for a total of more than 4,600 events over the period 1994-2018.[23]

Two points are noteworthy here. First, hedge funds are the quintessential example of exploitation of arbitrage, or mispricing, opportunities—thus, they are natural candidates for our theory.[24] Second, there is important suggestive evidence of sequentiality in hedge fund activism: multiple hedge funds of small or moderate size that attack a firm in a parallel and seemingly non-cooperative manner after a lead hedge fund has built a stake in it—a phenomenon termed *wolf-pack activism* in the law literature examining corporate governance.

The evidence on multi-activist engagements traditionally comes from two sources. First, from public disclosures where more than one hedge fund reveals its attack on a firm—between 2000 and 2010, Becht et al. (2017) documents that more than a quarter of 1,740 engagements involved multiple hedge funds. Second, indirectly, from the abnormal stock behavior around activism events that has been documented extensively in the literature and that is particularly acute the day in which the 5% threshold is crossed (e.g., Brav et al., 2021b). In this line, Wong (2020) finds that trades of disclosing activists on the crossing

[22] For more on this topic, see Briggs (2007).

[23] As the authors argue, there are three key features of hedge funds that favor their rise relative to index funds: the steep incentives for performance that their managers face; the more concentrated portfolios they hold; and the ability to lock-in capital for longer periods due to their restrictions on redemptions.

[24] "What we do know is that the targets of hedge funds are not randomly distributed, but rather tend to have some common characteristics, including in most (but not all) studies a low Tobin's Q, below average leverage, a low dividend payout, and a "value," as opposed to "growth," orientation." Ibid, p. 5.

date explain only 25% of the abnormal turnover observed in the data, suggesting that other non-disclosing activists are involved.

The argument for sequentiality stems from incentives: hedge funds obviously benefit from less competitive environments, which means they want to act fast once the 5% threshold is crossed to avoid block acquisition becoming more costly.[25] Importantly, such fast completion is plausible because hedge funds face important costs above 10%, which means that less than half of their terminal position is acquired over the 10-day window.[26] Consistent with this logic, Bebchuk et al. (2013) find that the median stake of hedge fund activists is 6.3%, and that hedge fund leader of the pack trades primarily in the crossing day and the one after (see pp. 23-24); and Collin-Dufresne and Fos (2015) find that a filer's trades are mostly concentrated on the crossing day, where the average purchase is 1% of shares outstanding.

Applied relevance. These facts support our model and findings. First, the distinctive feature that wolf packs solely consist of a specific type of investor is a clear sign of similarity. Translated to our setting, while it is always possible that two hedge funds have extreme opposing positions in a firm, we think their similar business strategies, funding sources, and regulatory constraints pave the way for a not-too-negative statistical relationship in positions to hold, particularly in larger firms. Second, their blocks are typically small ($\sim 6\%$), so that identities are often not disclosed due to stakes remaining below 5%.[27] Third, since limited communication is permitted, a hedge fund evaluating an engagement may have a good idea of the potential size of the pack, which may prompt her to act as a leader.

From this perspective, one understudied aspect of multi-agent activism is how an activist induces other blockholders to buy shares in the target firm. Our paper offers a non-cooperative price mechanism through which followers can be influenced via the channel of exploiting arbitrage opportunities—to a first-order approximation, precisely the element unifying hedge funds' business models. Our model predicts that for large-cap firms, the success of such engagements is likely to be hindered by free-riding forces that are exacerbated by the incentives to transfer costs to other activists. By contrast, activists are predicted to act more aggressively when targeting small-cap firms via the inference made on subsequent activists' positions, which mitigates the free-riding problem. These considerations add to the well-known difficulties/advantages of targeting large-/small-cap firms.

[25] Di Maggio et al. (2019) provide compelling evidence on this risk: the best clients of brokers handling the order of an activist are much more likely to buy the associated stock during the 10-day window period.

[26] As an example, the short swing rule or Section 16(b) of the Securities Act gives the issuer the right to ask a hedge fund holding over 10% to return any profits from reversal trades over a 6 month period. Also, insider trader rules that putlimitations on trading arise above 10% ownership.

[27] For instance, see Coffee Jr and Palia (2016) for notable examples of attacks in which wolf packs are undisclosed.

7 Other Linear Equilibria and Refinement

So far we have explored linear equilibria exhibiting a positive weight on initial positions: that is, high types build their positions more (or de-accumulate less) than low types due to their more attractive arbitrage opportunities. In this section, we relax this restriction by exploring if and when one or both activists in fact trade *against* their private information. This can happen because the the activists exhibit coordination motives that are self-fulfilling.

To build intuition, suppose that the activists start "long" on the firm (i.e., $X_0^L, X_0^F > 0$), a likely outcome when types are positively correlated. Further, suppose that the leader expects the follower to acquire a substantial short position on the firm's value, i.e., $\alpha_F < 0$, potentially indicative of negative effort by the follower. The leader may then find it profitable to acquire a short position so as to build a negative terminal position as well, as this would yield a positive surplus due to both players exerting negative effort. By the same logic, the follower would choose $\alpha_F < 0$. Importantly, while this type of coordination can rely on the firm potentially taking a negative value, it should not be disregarded as implausible in practice.[28] Indeed, it simply reflects the idea that acquiring a negative position can be profitable if it triggers a mechanism that ends up reducing a firm's value, such as when it precludes undertaking a value-enhancing action.[29] For instance, if our leader were able to short-sell—i.e., sell *borrowed* shares—she could profit from a reduced, yet positive, value of the firm, an incentive that would be stronger if she expected others to do the same.

Formally, Proposition A.1 in the Appendix characterizes the set of linear equilibria as solutions to (i) a set of equations for the coefficients in the activists strategies and (ii) a set of inequalities that include conditions for concavity in both activists' problems. In particular, it is shown there that leader's and follower's signaling coefficients satisfy

$$\alpha_L = \frac{\sigma^2}{\phi\alpha_L} - \underbrace{\frac{\rho\alpha_F}{\phi(1+\alpha_L)+\rho(1+\alpha_F)}}_{=\frac{\beta_F}{1-\beta_F}} \quad \text{and} \quad \alpha_F^2 = \frac{\sigma^2}{\gamma_1^F}, \tag{23}$$

respectively. That is, the equation for the leader's coefficient α_L carries a "Kyle component" $\sigma^2/\phi\alpha_L$ plus a correction term stemming from the value of manipulation in any linear equilibrium.[30] As for the follower, the corresponding coefficient can only take the standard form,

[28] Specifically, if a leader with a positive initial position reverses it, her effort will be negative; and if this position is sufficiently large, her payoff can be positive.

[29] See Goldstein and Guembel (2008) in the context of a speculator and a manager who can make an investment decision.

[30] The equation for α_L is the analog of (21) for δ_L from in Section 4.2. The "Kyle component" terminology originates from the equation for α_L admitting the solution $\sqrt{\sigma^2/\phi}$ absent the manipulation term. Further, from footnote 11, we deduce that this component corresponds to $\frac{1}{(1-\beta_F)\Lambda_1}\frac{\partial}{\partial X_0^L}(\mathbb{E}_L[W_L+W_F]-P_1)$ while

$\sqrt{\sigma^2/\gamma_1^F}$, or its negative, $-\sqrt{\sigma^2/\gamma_1^F}$.

We are interested in conditions under which such equilibria exhibiting $\alpha_L < 0$ or $\alpha_F < 0$ can emerge. The next result offers a glimpse into this question.

Proposition 7. *(i) Positive correlation: If $\rho > 0$, then for sufficiently large $\sigma > 0$, there exists a linear equilibrium in which α_L and α_F are strictly negative.*

(ii) Perfect negative correlation: If $\rho = -\phi$, there is no linear equilibrium in which α_L and α_F have the same sign. A linear equilibrium in which $\alpha_L < 0 < \alpha_F$ exists for all $\sigma > 0$.

According to part (i), if correlation is positive, both activists can trade against their positions provided the volatility of the noise traders is large. That is, the possibility of coordination emerges precisely when the leader's manipulation ability is limited by the reduced responsiveness of the market maker's belief. Part (ii) then exploits the analytical convenience of the case of perfect negative correlation to demonstrate that the weights on initial positions naturally must have opposite signs across activists in that case: in particular, the leader trades against her initial position to go on the same side of the follower. Consequently, fixing the volatility of noise traders $\sigma > 0$, as ρ falls from ϕ to $-\phi$: equilibria with negative weights on positions for both players can co-exist with the PBS one when correlation is positive; as ρ falls into the negative domain, equilibria with different signs on initial positions can emerge; eventually, as ρ approximates $-\phi$, only the latter type of equilibria are possible.

This brings us to the topic of the lower bound $\underline{\rho} < 0$ in Proposition 2, which guarantees the existence of our main equilibrium under study. Recall that in a standard one-shot Kyle model, the only force limiting a trader's orders—i.e., putting limits to arbitrage—is price impact. In the current model, however, there is also the possibility of manipulation. With positive correlation, more aggressive trading carries the extra cost of lowering the follower's contribution to the firm. By contrast, with negative correlation, trading more aggressively is beneficial in that it encourages the follower to exert effort, a force going against price impact.

Consequently, aside from the extra convexity stemming from the complementarity between the leader's terminal position and effort, the leader's problem is more "concave" than traditional ones when $\rho > 0$, and so a PBS equilibrium always exists. By contrast, the problem gains *convexity* when $\rho < 0$. This latter issue is severe: fixing $\sigma > 0$, when ρ becomes sufficiently close to $-\phi$, the leader's second-order condition cannot be satisfied by positive (α_L, α_F) pairs. The threshold $\underline{\rho} < 0$ in Proposition 2 ensures second-order conditions hold.

The dual role that order flow volatility plays is now apparent. First, for any level of covariance, lowering σ increases the leader's ability to manipulate the continuation game,

the manipulation correction is $\frac{\beta_F}{1-\beta_F}$.

making the coordination equilibrium less likely to arise. Second, for negative covariance, lowering σ increases price impact due to the order flow becoming more informative, which introduces concavity in the leader's problem and thus makes our PBS equilibrium more likely to arise. The next result offers a strong "refining" result in this respect.

Proposition 8. *Suppose that $\rho \in (-\phi, \phi)$. Then for sufficiently small but positive σ, a PBS equilibrium exists and is the unique equilibrium within the linear class.*

Thus, market illiquidity not only refines our PBS equilibrium in regions where it exists, but it also expands its range of existence without other equilibria emerging.[31]

8 Conclusion

We have developed a model of activism where first-mover advantages shape firm values and financial markets through the channel of strategic trading. This is an important question because multiple blockholders influence management in practice, and their blocks—which determine willingness to intervene—are endogenous. Hence, games of influence emerge. From this perspective, a key advantage of the game we propose is its simplicity, which parallels Stackelberg treatments of oligopolistic markets that have become benchmarks for understanding core incentives in industrial organization. But at the same time, our model proves rich enough to connect with the vast literature examining activism empirically. We now discuss our modeling choices, while shedding light on potential future work.

The endogeneity of the firm's fundamentals is key for our manipulation strategy to arise, a fact that is supported by the extensive literature on Kyle models that predict equilibrium strategies depending on gaps only. However, with sequential trading over two rounds, endogeneity is not enough, as the market maker would not necessarily learn about the follower's position from a first-period order flow that only carries leader's trades. Non-trivial correlation among initial positions opens this latter channel: as argued, this assumption is economically meaningful (e.g., hedge funds with similar trading strategies resulting in similar blocks in a statistical sense) and clearly the generic case. Thus, our Stackelberg model is "minimal" for uncovering the type of equilibrium we have studied.[32]

[31] It is important to stress that this result does not undermine our equilibrium in light of the positive relationship between activism and market liquidity that some studies have documented; in fact, our equilibrium exists for all $\sigma > 0$ when $\rho > 0$. The result only asserts that our equilibrium can be guaranteed to be the *unique prediction* when the market is sufficiently illiquid.

[32] If the leader could not commit to not trading again in the second period, the manipulation motive would still be present; but it may tamed if the follower scales back his purchases in response to the leader's presence (Proposition 5).

We would also expect our mechanism to be at play in a fully dynamic setting with repeated rounds of simultaneous trading among multiple activists, even if positions are initially independent. In fact, from an incentives standpoint, the presence of all activists in every round implies that the market maker will rely on the public history to forecast each activist's terminal position at all times; but this means that each activist will have an incentive to manipulate the market maker's belief about the other activists' positions so as to induce them to acquire larger positions.[33] In the process, positions will naturally acquire correlation as the activists condition their trades on the observed prices.[34]

Finally, in line with most of the literature, we have not forced the leader to reveal her position; but as argued, activists must disclose their blocks—and further, their intended actions—when ownership exceeds 5%. Our model is still relevant for three reasons. First, in many large firms activists generally do not accumulate 5%, yet still attack: in 2021, for instance, such "under the threshold" campaigns were a majority in the U.S., featuring targets whose average market capitalization was substantially higher than targets of campaigns that had to be disclosed.[35] Second, as stated in Section 6, since filing can occur with a delay of as much as 10 days, other activists can (and do) trade in the interim—that our game ends in the "third" period can then be understood as a subsequent disclosure act that reveals activists' actions, and hence firm value. Third, methods such as total return swaps and over-the-counter derivatives can be used to circumvent filing. That said, we would expect our leader to randomize by "noising up" her trade as in Huddart et al. (2001) if disclosure were mandatory for all ranges as in their model, thus preserving her ability to manipulate the market maker's belief.[36] We leave this and other questions for future research.

A Appendix: Proofs

A.1 Preliminaries

We begin by stating a proposition that characterizes equilibria via a system of equations and inequality conditions derived from the players' first and second order conditions and the pricing equations. The first half of the proposition below provides necessary conditions

[33] A similar analog between fully dynamic and Stackelberg analyses arises in the oligopoly model of Bonatti et al. (2017), where manipulation via overproduction is reminiscent of leader-follower Cournot incentives.

[34] Activists' beliefs would also be private in such a dynamic version due to individual trades depending on privately observed positions and the order flow carrying the sum of activists' trades. The ensuing "beliefs about beliefs" problem can be handled using the techniques in Foster and Viswanathan (1996) nonetheless.

[35] https://www.cnbc.com/2022/01/15/activist-hedge-funds-launched-89-campaigns-in-2021-heres-how-they-fared.html.

[36] See Ordonez-Calafi and Bernhardt (2022) for a model of disclosure thresholds that studies the tradeoff between transparency and an activist's ability to discipline management.

for equilibrium. The second half of the proposition is a strong converse: it shows that we can focus on the system of equations for the signaling coefficients (α_F, α_L); these coefficients determine price impact and therefore pin down the remaining coefficients.

Proposition A.1. *The tuple* $(\alpha_F, \beta_F, \delta_F, \alpha_L, \delta_L)$ *together with a pricing rule defined by* (10)-(13) *characterize an equilibrium only if* $\Lambda_1 \neq 0$*,* $\Lambda_2 \neq 0$*,* $\beta_F \neq 1$*,* $\phi(1+\alpha_L)+\rho \neq 0$*, and*

$$\alpha_F^2 = \sigma^2/\gamma_1^F, \tag{A.1}$$

$$\beta_F = -\frac{\rho}{\phi(1+\alpha_L)+\rho}\alpha_F, \tag{A.2}$$

$$\delta_F = \frac{(\alpha_L+\delta_L)\rho - \alpha_L\phi - (\phi-\rho)}{\phi(1+\alpha_L)+\rho}\alpha_F, \tag{A.3}$$

$$\alpha_L = \frac{\sigma^2}{\phi\alpha_L} - \frac{\rho\alpha_F}{\phi(1+\alpha_L)+\rho(1+\alpha_F)}, \tag{A.4}$$

$$\delta_L = -\frac{\sigma^2}{\phi\alpha_L}, \tag{A.5}$$

$$0 \geq \sigma^2 - \alpha_L^2\phi - 2\alpha_L[\rho(1+\alpha_F)+\phi], \tag{A.6}$$

$$0 \geq -\alpha_F[\sigma^2(\phi+\rho(1+\alpha_L)) + \alpha_L^2(\phi^2-\rho^2)]. \tag{A.7}$$

Further, if $\rho \neq 0$*, one of the following conditions must hold:*

$$\alpha_F = \alpha_{F,1}(\alpha_L) := \sqrt{\frac{\sigma^4+\alpha_L^2\sigma^2\phi}{\sigma^2\phi+\alpha_L^2(\phi^2-\rho^2)}} = \frac{(\rho+\phi+\phi\alpha_L)(\alpha_L^2\phi-\sigma^2)}{\rho[\sigma^2-\alpha_L(1+\alpha_L)\phi]} \quad \text{or} \tag{A.8}$$

$$\alpha_F = \alpha_{F,2}(\alpha_L) := -\sqrt{\frac{\sigma^4+\alpha_L^2\sigma^2\phi}{\sigma^2\phi+\alpha_L^2(\phi^2-\rho^2)}} = \frac{(\rho+\phi+\phi\alpha_L)(\alpha_L^2\phi-\sigma^2)}{\rho[\sigma^2-\alpha_L(1+\alpha_L)\phi]}. \tag{A.9}$$

Conversely, suppose (α_F, α_L) *satisfy* (A.6) *and* (A.7)*, either* (A.8) *or* (A.9)*, and* $\phi(1+\alpha_L)+\rho \neq 0$*. Then (i)* $(\beta_F, \delta_F, \delta_L)$ *are well defined via* (A.2)*,* (A.3)*, and* (A.5)*, with* $\beta_F \neq 1$*; (ii)* $\Lambda_1 \neq 0$ *and* $\Lambda_2 \neq 0$ *are well defined via* (10) *and* (13)*; and (iii) the associated strategies and pricing rule constitute an equilibrium.*

Proof. We first establish necessity, starting with the follower's conditions. Expanding the follower's FOC (15) at the candidate strategy (6) yields an expression that is linear in (X_0^F, P_1, μ), which must be identically zero over $(X_0^F, P_1, \mu) \in \mathbb{R}^3$. Hence, the coefficients on each variable (X_0^F, P_1, μ) must be zero, delivering the following three equations:

$$0 = \frac{\tilde{\Lambda}_2}{\gamma_1^F}(\sigma^2 - \alpha_F^2\gamma_1^F), \tag{A.10}$$

$$0 = -\frac{\tilde{\Lambda}_2}{\gamma_1^F}\left[\frac{\rho\sigma^2(1-\beta_F)}{\phi(1+\alpha_L)+\rho(1+\alpha_F)} + \beta_F\alpha_F\gamma_1^F\right], \tag{A.11}$$

$$0 = \frac{\tilde{\Lambda}_2}{\gamma_1^F}\left[-\sigma^2 + \frac{(2+\alpha_F+\alpha_L+\delta_F+\delta_L)\rho\sigma^2}{\phi(1+\alpha_L)+\rho(1+\alpha_F)} - \alpha_F\delta_F\gamma_1^F\right], \tag{A.12}$$

where $\tilde{\Lambda}_2 := \frac{\gamma_1^F}{\alpha_F^2\gamma_1^F+\sigma^2} \times [1+\alpha_F+\rho_1/\gamma_1^F]$. We argue that in any linear equilibrium, the right hand sides are well defined and $\tilde{\Lambda}_2 \neq 0$. First, $\gamma_1^F > 0$ for any (finite) α_F. Second, (18) implies $\Lambda_2 \neq 0$, so $\tilde{\Lambda}_2$ is well defined and nonzero. Third, $\Lambda_1 \neq 0$ implies $\phi(1+\alpha_L)+\rho(1+\alpha_F) \neq 0$ in the denominators in (A.11) and (A.12).

We can now derive (A.1)-(A.3) and (A.7). Since $\tilde{\Lambda}_2 \neq 0$ is necessary for equilibrium, (A.10) reduces to (A.1). (Note that this implies $\alpha_F \neq 0$.) Using this fact to write $\alpha_F\gamma_1^F = \sigma^2/\alpha_F$, (A.11) reduces to

$$\begin{aligned}
0 &= -\frac{\tilde{\Lambda}_2}{\gamma_1^F}\left[\frac{\rho\sigma^2(1-\beta_F)}{\phi(1+\alpha_L)+\rho(1+\alpha_F)} + \beta_F\frac{\sigma^2}{\alpha_F}\right] \\
&= -\frac{\tilde{\Lambda}_2\sigma^2}{\gamma_1^F\alpha_F[\phi(1+\alpha_L)+\rho(1+\alpha_F)]}\left[\rho\alpha_F + \beta_F[\phi(1+\alpha_L)+\rho]\right].
\end{aligned} \tag{A.13}$$

We claim that $\phi(1+\alpha_L)+\rho \neq 0$ in equilibrium. By way of contradiction, if $\phi(1+\alpha_L)+\rho = 0$, then (A.13) implies $\alpha_F = 0$ or $\rho = 0$. Equation (A.1) rules out $\alpha_F = 0$. And if $\rho = 0$, we have $\alpha_L = -1$, and thus $\Lambda_1 = 0$, violating the leader's SOC. Hence, $\phi(1+\alpha_L)+\rho \neq 0$, and (A.13) reduces to (A.2). Analogous arguments yield (A.3) from (A.12). Lastly, using (A.1) to eliminate α_F^2 terms, the follower's SOC (18) reduces to (A.7).

Next, we derive the leader's conditions (A.4)-(A.5) and (A.6). For the leader, the FOC (16) must hold for all $(X_0^L, \mu) \in \mathbb{R}^2$. Setting the coefficients on these variables to 0 and using (A.1) and (A.2), it is straightforward to show that the leader's FOC reduces to (A.4)-(A.5) where $\alpha_L \neq 0$ in equilibrium since the leader's SOC implies $\Lambda_1 \neq 0$. The leader's SOC is equivalent to (A.6).

To obtain (A.8) or (A.9), first note that the positive and negative values of α_F solving (A.1) are $\pm\sqrt{\frac{\sigma^4+\alpha_L^2\sigma^2\phi}{\sigma^2\phi+\alpha_L^2(\phi^2-\rho^2)}}$. Next, solve for α_F in (A.4) by multiplying through by the denominators on the right hand side and rearrange terms to obtain

$$\alpha_F\rho[\sigma^2-\alpha_L(1+\alpha_L)\phi] = [\phi(1+\alpha_L)+\rho](\alpha_L^2\phi-\sigma^2). \tag{A.14}$$

We claim that $\sigma^2-\alpha_L(1+\alpha_L)\phi \neq 0$ in any solution to (A.14). Indeed, since $\phi(1+\alpha_L)+\rho \neq 0$, $\sigma^2-\alpha_L(1+\alpha_L)\phi = 0$ would imply $\alpha_L^2\phi-\sigma^2 = 0$, but these two equations cannot hold simultaneously. Thus, if $\rho \neq 0$, (A.14) implies

$$\alpha_F = \frac{(\rho + \phi + \phi\alpha_L)(\alpha_L^2\phi - \sigma^2)}{\rho[\sigma^2 - \alpha_L(1 + \alpha_L)\phi]}. \tag{A.15}$$

Since the solutions to (A.1) are $\alpha_F = \alpha_{F,1}$ and $\alpha_F = \alpha_{F,2}$, we obtain (A.8) and (A.9).

For the sufficiency half of the proposition, take (α_F, α_L) as in the statement. Clearly, either $\alpha_F = \alpha_{F,1}$ or $\alpha_F = \alpha_{F,2}$ implies (A.1). Now given $\phi(1 + \alpha_L) + \rho \neq 0$, we can multiply through (A.8) or (A.9) by $\rho[\sigma^2 - \alpha_L(1+\alpha_L)\phi]$ to recover (A.14). To recover (A.4) from (A.14), simply note that (A.6) can be rewritten as $\sigma^2 + \alpha_L^2\phi - 2\alpha_L[\rho(1+\alpha_F) + \phi(1+\alpha_L)] \leq 0$, which implies $\alpha_L \neq 0$ and $\phi(1+\alpha_L) + \rho(1+\alpha_F) \neq 0$. Given that $\phi(1+\alpha_L) + \rho \neq 0$ by supposition, (β_F, δ_F) are well defined by (A.2)-(A.3). Further, $\phi(1 + \alpha_L) + \rho(1 + \alpha_F) \neq 0$ implies that $1 \neq -\frac{\rho\alpha_F}{\phi(1+\alpha_L)+\rho} = \beta_F$. This establishes (i). It follows that Λ_1 and Λ_2 are well defined by (10) and (13), respectively. Moreover, by construction, (A.6)-(A.7) imply (17)-(18), so $\Lambda_1 \neq 0$ and $\Lambda_2 \neq 0$, establishing (ii).

For part (iii) of the sufficiency claim, observe that since the players' best responses problems are quadratic, it suffices to check first and second order conditions. Given that the inequalities $\Lambda_1 \neq 0$, $\Lambda_2 \neq 0$, $\beta_F \neq 1$, $\phi(1 + \alpha_L) + \rho \neq 0$ are satisfied, the equations (A.1)-(A.5) imply the FOCs (15) and (16) by construction, and as noted for part (ii), the SOCs (17) and (18) are satisfied. □

Define $\hat{\alpha} := \frac{-\phi+\sqrt{\phi^2+4\sigma^2\phi}}{2\phi} > 0$ to be the positive root of the denominator on the right side of (A.8). Note that $\alpha^K > \hat{\alpha}$.

A.2 Proof of Proposition 1

By Proposition A.1, α_F must satisfy (A.1), so either $\alpha_F = \alpha_{F,1} := \sqrt{\frac{\sigma^2}{\gamma_1^F}}$ or $\alpha_F = \alpha_{F,2} := -\sqrt{\frac{\sigma^2}{\gamma_1^F}}$. Since $\alpha_F > 0$ in any PBS equilibrium (by definition), $\alpha_F = \alpha_{F,1}$, and then (β_F, δ_F) are characterized by (A.2)-(A.3).

When $\rho = 0$, note that (A.7) becomes $-\alpha_F[\sigma^2\phi + \alpha_L^2] \leq 0$, which is satisfied by $\alpha_F = \alpha_{F,1}$ and not $\alpha_F = \alpha_{F,2}$. Equation (A.4) yields $\alpha_L = \pm\alpha^K$. Of these, only $\alpha_L = \alpha^K$ satisfies (A.6). Given $(\alpha_F, \alpha_L) = (\alpha^K, \alpha^K)$, $(\beta_F, \delta_F, \delta_L) = (0, -\alpha^K, -\alpha^K)$ is the unique solution to (A.2), (A.3), and (A.5). This characterizes the unique linear equilibrium for $\rho = 0$.

For the rest of the proof, consider $\rho \neq 0$. To sign β_F, recall that $\alpha_F, \alpha_L > 0$ and $|\rho| \leq \phi$, so $sign(\beta_F) = -sign(\rho)$ via (A.2). Similarly, from (A.3), $sign(\delta_F) = sign((\alpha_L + \delta_L)\rho - \alpha_L\phi - (\phi - \rho))$. This is unambiguously negative, since $(\alpha_L + \delta_L)\rho \leq 0$ by Proposition 2 (which does not rely on the current result), and since $\alpha_L\phi > 0$ and $\phi - \rho \geq 0$ by assumption.

We now establish that $\beta_F < 1$. For $\rho > 0$, this is immediate since $\beta_F < 0$. For $\rho < 0$, recall the equation (23) for α_L stemming from the leader's FOC. As we show in the proof of Proposition 2 (again, without circularity), in a PBS equilibrium, $\alpha_L > \alpha^K$ when $\rho < 0$. Hence, when $\rho < 0$, we have $\alpha_L > \frac{(\alpha^K)^2}{\alpha_L} = \frac{\sigma^2}{\phi\alpha_L}$, and thus (23) implies $\frac{\beta_F}{1-\beta_F} > 0$. This, in turn, implies $\beta_F \in (0, 1)$. For the case $\rho = 0$, we already showed above that $\beta_F = 0$ in the unique equilibrium, also satisfying the inequality $\beta_F < 1$.

Next, we verify that in any linear equilibrium (PBS or otherwise), the follower's strategy has the form (19) for $\alpha_F = \alpha_{F,1}$ or $\alpha_F = \alpha_{F,2}$. First, express M_1^F in terms of P_1 and μ by using (9) to replace the surprise term $\Psi_1 - \mu(\alpha_L + \delta_L)$ in (8):

$$M_1^F = \mu + \frac{\alpha_L \rho}{\alpha_L^2 \phi + \sigma^2} \frac{P_1 - P_0}{\Lambda_1}, \tag{A.16}$$

where P_0 is linear in μ (see (B.1) in the Online Appendix). Substituting (A.16) into (19) then yields an expression for the follower's strategy in which the coefficient on X_0^F is $\alpha_{F,i}$, and the coefficients on (P_1, μ) equal $(\beta_{F,i}, \delta_{F,i})$ when (A.2)-(A.3) hold.

Lastly, given any first-period order flow, $\mathbb{E}[\theta^F | \mathcal{F}_1^\Psi] = \mathbb{E}[\alpha_F (X_0^F - M_1^F) | \mathcal{F}_1^\Psi] = \alpha_F (M_1^F - M_1^F) = 0$. And the law of iterated expectations, $\mathbb{E}[\theta^F] = \mathbb{E}[\mathbb{E}[\theta^F | \mathcal{F}_1^\Psi]] = 0$.

A.3 Proof of Proposition 2

The proof is as follows. First, we consider $\rho \in (0, \phi]$, for which we establish existence of a PBS equilibrium and uniqueness within the PBS class (Proposition A.2). Second, we show that for $|\rho| > 0$ sufficiently small (allowing for positive or negative ρ), there exists a unique equilibrium within the whole linear class, and it is a PBS equilibrium (Proposition A.3). (Note that Proposition 1 already covers $\rho = 0$.) For both cases we prove the inequalities stated in the proposition. Third, we show that a PBS equilibrium fails to exist if ρ is sufficiently low (Proposition A.4), and we construct $\underline{\rho}, \rho_0 \in (-\phi, 0)$ presented in the proposition.

Proposition A.2. *If $\rho \in (0, \phi]$, there is a unique PBS equilibrium, and $\alpha_L < \alpha^K < -\delta_L$.*

Proof. By Proposition A.1, (A.8) is a necessary condition for (α_F, α_L) to be part of PBS equilibrium. Let $L(\alpha_L)$ and $R(\alpha_L)$ denote the left and right sides of (A.8). L is positive and strictly increasing in α_L for $\alpha_L \geq 0$. Meanwhile, R is continuous on $[0, \hat{\alpha}) \cup (\hat{\alpha}, +\infty)$ and satisfies $R(\hat{\alpha}-) = -\infty$, $R(\hat{\alpha}+) = +\infty$, and $R(\alpha^K) = 0$. Further, for $\alpha_L \in [0, \hat{\alpha}) \cup (\hat{\alpha}, +\infty)$,

$$R'(\alpha_L) = -\phi \frac{(\alpha_L^2 \phi - \sigma^2)^2 + (\rho + \phi)(\alpha_L^2 + \sigma^2) + 2\alpha_L^3 \phi^2}{\rho[\sigma^2 - \alpha_L(1 + \alpha_L)\phi]^2}, \tag{A.17}$$

which is unambiguously strictly negative when $\rho > 0$. Thus, R is strictly decreasing on $(\hat{\alpha}, +\infty)$, so there exists a solution to (A.8) on $(\hat{\alpha}, \alpha^K)$ and this is the only solution on $(\hat{\alpha}, +\infty)$. Since $L(0) > 0$, while $R(0) = -(\rho + \phi)/\rho < 0 < L(0)$ (given $\rho > 0$), there is no solution on $[0, \hat{\alpha})$, so this solution is the unique among $\alpha_L \geq 0$. And by (A.5), $\alpha_L < \alpha^K$ implies $\alpha^K < -\delta_L$ (and $\delta_L < 0$).

Given a unique candidate for PBS equilibrium, we now verify SOCs. For the leader, note that since $\alpha_L, \alpha_F > 0$, (A.6) is bounded above by $\sigma^2 - \alpha_L^2\phi - \alpha_L\phi$, which is negative since $\alpha_L > \hat{\alpha}$. For the follower, (A.7) holds by inspection for $\rho > 0$ since $\alpha_L > 0$ and $\alpha_F > 0$. □

Next, we turn to $|\rho| > 0$ close to 0.

Proposition A.3. *If $|\rho| > 0$ is sufficiently small, there exists a unique linear equilibrium, and it is a PBS equilibrium. If $\rho > 0$, $\alpha_L < \alpha^K < -\delta_L$, and if $\rho < 0$, $\alpha_L > \alpha^K > -\delta_L$.*

Proof. Assume throughout that $\rho \neq 0$. Let us call any pair (α_L, α_F) satisfying (A.8) or (A.9) a *candidate signaling pair*. We construct two candidate signaling pairs (α_L^*, α_F^*) and $(\alpha_L^\flat, \alpha_F^\flat)$. We then show that for small $|\rho|$, there are no other candidate signaling pairs satisfying the leader's second order condition, and of these two pairs, only (α_L^*, α_F^*) satisfies the follower's SOC. We then invoke the converse part of Proposition A.1 to establish existence of a unique equilibrium based on (α_L^*, α_F^*).

We claim that if $\rho < 0$, there exists $\alpha_L^* \in (\alpha^K, \infty)$ solving (A.8) and $\alpha_L^\flat \in (\hat{\alpha}, \alpha^K)$ solving (A.9). Analogous arguments for the case $\rho > 0$ establish the existence of $\alpha_L^* \in (\hat{\alpha}, \alpha^K)$ and $\alpha_L^\flat \in (\alpha^K, \infty)$; we omit this case for brevity. In either case, we will ultimately show that α_L^* is the unique equilibrium value of α_L for small $|\rho|$. As before, let $R(\alpha_L)$ denote the right hand side common to (A.8) and (A.9). Note that R is continuous on $(\hat{\alpha}, \infty)$, and it has the properties $\lim_{\alpha_L \to +\infty} R(\alpha_L) = +\infty$, $\lim_{\alpha_L \downarrow \hat{\alpha}} R(\alpha_L) = -\infty$, and $R(\alpha^K) = 0$. The left hand side of (A.8) is strictly positive and bounded, so by the intermediate value theorem (IVT), there exists a solution $\alpha_L^* \in (\alpha^K, \infty)$ to (A.8). Similarly, the left hand side of (A.9) is strictly negative and bounded, so by the IVT, there exists a solution $\alpha_L^\flat \in (\hat{\alpha}, \alpha^K)$ to (A.9).

Define $\alpha_F^* := \alpha_{F,1}(\alpha_L^*)$ and define $\alpha_F^\flat = \alpha_{F,2}(\alpha_L^\flat)$. By definition, both (α_L^*, α_F^*) and $(\alpha_L^\flat, \alpha_F^\flat)$ are candidate signaling pairs.

To assess other candidate signaling pairs, we derive a polynomial equation such that (α_L, α_F) is a candidate signaling pair only if α_L is a root of this equation. By squaring either (A.8) or (A.9), we obtain a necessary condition

$$\frac{\sigma^4 + \alpha_L^2\sigma^2\phi}{\sigma^2\phi + \alpha_L^2(-(\rho)^2 + (\phi)^2)} = \left(\frac{(\rho + \phi + \phi\alpha_L)(\alpha_L^2\phi - \sigma^2)}{\rho[\sigma^2 - \alpha_L(1 + \alpha_L)\phi]}\right)^2, \tag{A.18}$$

and by cross multiplying, an eighth-degree polynomial equation

$$0 = Q(\alpha_L;\rho) = \sum_{i=0}^{8} A_i \alpha_L^i, \qquad \text{where} \tag{A.19}$$

$$A_8 = -\phi^4(\phi^2-\rho^2), \qquad A_7 = -2(\phi-\rho)\phi^3(\rho+\phi)^2,$$
$$A_6 = \phi^2(\rho^2-\phi^2)[\rho^2+2\rho\phi+\phi(-\sigma^2+\phi)], \qquad A_5 = 2\sigma^2\phi^2[-2\rho^3-\rho^2\phi+\rho\phi^2+\phi^3],$$
$$A_4 = \sigma^2\phi[-2\rho^4-4\rho^3\phi+2\rho\phi^3+\phi^3(\sigma^2+\phi)], \qquad A_3 = 2\sigma^4\phi[\rho^3+\rho^2\phi+\rho\phi^2+\phi^3],$$
$$A_2 = \sigma^4[\rho^4+2\rho^3\phi+2\rho\phi^3+\phi^3(-\sigma^2+\phi)+\rho^2\phi(-\sigma^2+3\phi)], \qquad A_1 = -2\sigma^6\phi[\rho^2+\phi\rho+\phi^2],$$
$$A_0 = \sigma^6[\rho^2(\sigma^2-\phi)-2\rho\phi^2-\phi^3].$$

Being an eighth-degree polynomial, $Q(\cdot;\rho)$ has exactly eight complex roots, counting multiplicity; two of these are α_L^* and $\alpha_L^\flat$.

We now show that of all candidate signaling pairs, when $|\rho|$ is sufficiently small, only (α_L^*, α_F^*) satisfies both activists' SOCs. To that end, it is useful to approximate all of the roots of (A.19) for small $|\rho|$. We will make use of a standard result on the continuous dependence of the (complex) roots of a polynomial on its coefficients:

Lemma A.1 (Uherka and Sergott (1977))**.** *Let $p(x) = x^n + \sum_{k=1}^n a_i x^{n-k}$ and $p^*(x) = x^n + \sum_{k=1}^n a_i^* x^{n-k}$ be two nth degree polynomials. Suppose λ^* is a root of p^* with multiplicity m and $\epsilon > 0$. Then for $|a_i - a_i^*|$ sufficiently small ($i = 1, \ldots, n$), p has at least m roots within ϵ of λ^*.*

For a proof, see Uherka and Sergott (1977) or the references therein.

We apply this lemma to the polynomial Q indexed by ρ. (While Lemma A.1 assumes a leading coefficient of 1, we can divide through our polynomial $Q(\cdot;\rho)$ in (A.19) by A_8, which is bounded away from 0 provided that $|\rho| < |\phi|$, allowing us to apply the lemma.) In the limit as $\rho \to 0$,

$$Q(\alpha_L;0) = -(1+\alpha_L)^2\phi^3(\sigma^2-\alpha_L^2\phi)^2(\sigma^2+\alpha_L^2\phi). \tag{A.20}$$

By inspection, $Q(\cdot;0)$ is nonpositive and has double roots at -1 and $\pm\alpha^K$, and it has complex roots at $\pm\alpha^K i$.

Lemma A.1 then has two important implications about candidate signaling pairs. We state the first one as a corollary.

Corollary A.1. *As $\rho \to 0$, $\alpha_L^* \to \alpha^K$ and $\alpha_L^\flat \to \alpha^K$.*

This simply follows from the fact that $\alpha_L^*, \alpha_L^\flat \geq 0$, so they can only converge to α^K. The second implication of Lemma A.1 is that for any $\epsilon > 0$, there exists $\overline{\rho} > 0$ such that for all

ρ with $0 < |\rho| < \overline{\rho}$ all of the other six roots of $Q(\cdot;\rho)$ lie within ϵ of -1, $-\alpha^K$, or $\pm\alpha^K i$. Hence, for such ρ, α_L^* and $\alpha_L^\flat$ are roots with multiplicity 1, and they are uniquely defined.

We can now check SOCs: for the leader in Lemma A.2 and the follower in Lemma A.3.

Lemma A.2. *For $|\rho| > 0$ sufficiently small, the candidate signaling pairs (α_L^*, α_F^*) and $(\alpha_L^\flat, \alpha_F^\flat)$ satisfy (A.6) and are the only candidate signaling pairs that do.*

Proof. First, we show that (α_L^*, α_F^*) satisfy (A.6) for sufficiently small $|\rho| > 0$. As $\rho \to 0$, the left hand side of (A.6) tends to

$$\sigma^2 - (\alpha^K)^2\phi - 2\alpha^K\phi = -2\sigma\sqrt{\phi} < 0, \tag{A.21}$$

where we have used that $\alpha_L^* \to \alpha^K$ by Corollary A.1. A nearly identical calculation shows $(\alpha_L^\flat, \alpha_F^\flat)$ also satisfy (A.6) for sufficiently small $|\rho| > 0$.

The remaining candidates for equilibria are associated with the real roots of (A.19) other than $\alpha_L^*, \alpha_L^\flat$. By Lemma A.1, as $\rho \to 0$, these roots must converge to the other roots of $Q(\cdot;0)$, namely -1, $-\alpha^K$, or $\pm\alpha^K i$. Any root of $Q(\cdot;\rho)$ that is in a sufficiently small neighborhood of $\pm\alpha^K i$ has a nonzero complex component, and is not an equilibrium candidate. Therefore, we need only consider candidates in neighborhoods of -1 or $-\alpha^K$. In the first case, for any $\alpha_F \in \{\alpha_{F,1}, \alpha_{F,2}\}$, the left hand side of (A.6) converges to

$$\sigma^2 - (-1)^2\phi - 2(-1)\phi = \sigma^2 + \phi > 0. \tag{A.22}$$

In the second case, for any $\alpha_F \in \{\alpha_{F,1}, \alpha_{F,2}\}$, the left hand side of (A.6) converges to

$$\sigma^2 - \left(-\alpha^K\right)^2\phi - 2\left(-\alpha^K\right)\phi = 2\sigma\sqrt{\phi} > 0. \tag{A.23}$$

Thus, for $|\rho| > 0$ sufficiently small, all roots of $Q(\cdot;\rho)$ other than α_L^* and $\alpha_L^\flat$ violate the leader's SOC. □

Lemma A.3. *For $|\rho| > 0$ sufficiently small, the candidate signaling pair (α_L^*, α_F^*) satisfies (A.7), while the pair $(\alpha_L^\flat, \alpha_F^\flat)$ does not.*

Proof. For the pair (α_L^*, α_F^*), the left hand side of (A.7) tends to $-[(\alpha^K)^2\phi^2 + \sigma^2\phi] < 0$ as $\rho \to 0$. For the pair $(\alpha_L^\flat, \alpha_F^\flat)$, however, it tends to $(\alpha^K)^2\phi^2 + \sigma^2\phi > 0$, violating (A.7). □

From Lemmas A.2 and A.3, we conclude that for $|\rho| > 0$ sufficiently small, (α_L^*, α_F^*) is the unique candidate signaling pair satisfying both (A.6) and (A.7). Hence, in any linear equilibrium, (α_L, α_F) must equal (α_L^*, α_F^*).

To conclude, observe that as $\rho \to 0$, $\phi(1+\alpha_L^*)+\rho \to \phi(1+\alpha^K) > 0$, allowing us to apply the "converse" part of Proposition A.1 when $|\rho|$ is sufficiently small, giving us existence. Since we have already shown that $0 < \alpha_L^* < \alpha^K$ if $\rho > 0$, (A.5) implies $-\delta_L > \alpha^K$ in this case, and likewise when $\rho < 0$, we have $\alpha_L^* > \alpha^K$ which implies $0 < -\delta_L < \alpha^K$. □

By the results above, a unique PBS equilibrium exists if ρ is (i) positive or (ii) sufficiently close to zero. Thus, $\underline{\rho} := \inf\{\rho' \in [-\phi, \phi] : \text{a PBS equilibrium exists for all } \rho \in [\rho', \phi]\} < 0$ and $\rho_0 := \inf\{\rho' \in [-\phi, \phi] : \text{a unique PBS equilibrium exists for all } \rho \in [\rho', \phi]\} < 0$, where $\rho_0 \geq \underline{\rho}$ is obvious. To show that $\underline{\rho} > -\phi$, we invoke the following result.

Proposition A.4. *Fix $\sigma, \phi > 0$. There exists $\hat{\rho} \in (-\phi, 0)$ such that if $\rho < \hat{\rho}$, there is no PBS equilibrium.*

Proof. The proof is based on the following two lemmas.

Lemma A.4. *There is no $[-\phi, \phi]$-valued sequence $(\rho_n)_{n\in\mathbb{N}}$ that converges to $-\phi$ and has the property that there is an associated sequence of PBS equilibria such that $(\alpha_{F,n})_{n\in\mathbb{N}}$ is bounded.*

Proof. Suppose by way of contradiction that there exists such a sequence with associated PBS equilibria indexed by n. We claim that $(\alpha_{L,n})_{n\in\mathbb{N}}$ is bounded. To see this, take n sufficiently large that $\rho_n \neq 0$, and note that the right hand side of (A.8) must be bounded, since it equals $\alpha_{F,n}$ which we have supposed is bounded. Since the numerator on the right hand side is cubic while the denominator is quadratic, it must be that $(\alpha_{L,n})_{n\in\mathbb{N}}$ is bounded.

Given that $(\alpha_{F,n})_{n\in\mathbb{N}}$ and $(\alpha_{L,n})_{n\in\mathbb{N}}$ are both bounded, we can pass to a subsequence such $\alpha_{F,n} \to \overline{\alpha}_F \geq 0$ and $\alpha_{L,n} \to \overline{\alpha}_L \geq 0$, where the inequalities follow from $\alpha_{F,n}, \alpha_{L,n} \geq 0$ in PBS equilibria by definition. Then taking limits in (A.8), we have

$$\overline{\alpha}_F = \sqrt{\frac{\sigma^2}{\phi} + \overline{\alpha}_L^2} > \overline{\alpha}_L. \tag{A.24}$$

The right hand side of (A.6) then has limit

$$\sigma^2 + \overline{\alpha}_L^2\phi - 2\overline{\alpha}_L[-\phi(1+\overline{\alpha}_F) + \phi(1+\overline{\alpha}_L)] = \sigma^2 + \overline{\alpha}_L^2\phi + 2\overline{\alpha}_L\phi(\overline{\alpha}_F - \overline{\alpha}_L) > 0, \tag{A.25}$$

where $\overline{\alpha}_F - \overline{\alpha}_L > 0$ by (A.24). But since (A.6) is satisfied for all n, this limit must be nonpositive, a contradiction. □

Lemma A.5. *There is no $[-\phi, \phi]$-valued sequence $(\rho_n)_{n\in\mathbb{N}}$ that converges to $-\phi$ and has the property that there is an associated sequence of PBS equilibria such that $(\alpha_{F,n}) \to +\infty$.*

Proof. Suppose by way of contradiction that there were such a sequence. From the expression for $\alpha_{F,n}$ in (A.8), it must be that $\alpha_{L,n} \to +\infty$. We claim that $\frac{\alpha_{F_n}}{\alpha_{L,n}} \to 1$. To obtain this, divide through (A.8) by $\alpha_{L,n}$ to get

$$\frac{\alpha_{F_n}}{\alpha_{L,n}} = \frac{(\rho_n + \phi + \phi\alpha_{L,n})(\alpha_L^2\phi - \sigma^2)}{\rho_n\alpha_{L,n}\left[\sigma^2 - \alpha_{L,n}(1+\alpha_{L,n})\phi\right]} \to 1. \tag{A.26}$$

We now show that (A.6) eventually fails. The right hand side of (A.6) is

$$\sigma^2 + \alpha_{L,n}^2\phi - 2\alpha_{L,n}[\phi + \rho_n + \alpha_{L,n}(\rho_n\alpha_{F,n}/\alpha_{L,n} + \phi)]. \tag{A.27}$$

Since $\phi + \rho_n \to 0$ and $\frac{\alpha_{F,n}}{\alpha_{L,n}} \to 1$, for any $\epsilon > 0$, the expression in square brackets in (A.27) is less than $\epsilon\alpha_{L,n}$ for sufficiently large n. Hence, (A.27) is eventually greater than $\sigma^2 + \alpha_{L,n}^2\phi - 2\epsilon\alpha_{L,n}^2$, which is positive for $\epsilon < \phi/2$, violating (A.6), contradicting equilibrium. □

The existence of $\hat{\rho} > -\phi$ then follows immediately from Lemmas A.4 and A.5, since if there is no such $\hat{\rho}$ there would exist a sequence $(\rho_n)_{n\in\mathbb{N}}$ with $\rho_n \to -\phi$ and an associated sequence of PBS equilibria such that either (i) $\alpha_{F,n} \to +\infty$ along some subsequence (which is ruled out by Lemma A.5) or (ii) $(\alpha_{F,n})_{n\in\mathbb{N}}$ is bounded (ruled out by Lemma A.4). Since Proposition A.3 shows that a PBS equilibrium exists for some $\rho < 0$, we have $\hat{\rho} < 0$. □

Now, given any $\hat{\rho}$ as in Proposition A.4, we have $\underline{\rho} \geq \hat{\rho}$, so $\underline{\rho} \geq \hat{\rho} > -\phi$. We conclude the proof of Proposition 2 by showing that α_L is decreasing in ρ. First suppose $\rho > 0$. The right hand side of (A.8) crosses the left hand side from above at α_L. Moreover, when $\rho > 0$, the right hand side is (positive and) decreasing in ρ at α_L while the left hand side is increasing in ρ. Hence, α_L is decreasing in ρ. In turn, when $\rho < 0$, the right hand side of (A.8) crosses the left hand side from below; the left hand side is decreasing in ρ; and the right hand side is increasing in ρ at α_L. Hence, again, α_L is unambiguously decreasing in ρ. The result then follows since α_L is continuous in ρ at $\rho = 0$ by Corollary A.1.

A.4 Proof of Proposition 3

For part (i), the expected first-period order flow is $\mathbb{E}[\Psi_1] = \mu(\alpha_L + \delta_L)$, which by Proposition 2 is negative if and only if $\rho > 0$. For the second period, note that $\mathbb{E}[\Psi_2|\mathcal{F}_1^\Psi] = \mathbb{E}[\theta^F|\mathcal{F}_1^\Psi] + \mathbb{E}[Z_2|\mathcal{F}_1^\Psi] = \mathbb{E}[\theta^F|\mathcal{F}_1^\Psi] = 0$ by Proposition 1. Similarly, $\mathbb{E}[\Psi_2] = \mathbb{E}[\theta^F] + \mathbb{E}[Z_2] = \mathbb{E}[\theta^F] = 0$.

For part (ii.1), ex ante expected firm value is

$$\mathbb{E}[W^L + W^F] = \mathbb{E}[X_0^L + \theta^L + X_0^F + \theta^F] = \mu + (\alpha_L + \delta_L)\mu + \mu,$$

where we have used that terminal positions coincide with terminal efforts, that $\mathbb{E}[\theta^F] = 0$ from the proof of part (i), and that θ^L is given in (5). The last statement of the proposition then follows from the fact that, again, $\alpha_L + \delta_L < 0$ is negative if and only if $\rho > 0$.

For part (ii.2), we show that $\alpha_L + \delta_L > -1$. Using (A.5), we have $\alpha_L + \delta_L = \alpha_L - \frac{\sigma^2}{\phi\alpha_L} =: h(\alpha_L)$. Now h is increasing in α_L for $\alpha_L > 0$, and from the proof of Proposition 2, $\alpha_L > \hat{\alpha}$. By direct calculation, $h(\hat{\alpha}) = -1$, so we are done. Further, since h is increasing in α_L and α_L is decreasing in ρ by Proposition 2, firm value is decreasing in ρ.

For part (iii), first choose ρ sufficiently small that there exists a unique linear equilibrium by Proposition A.3. By writing (A.8) in the form $G(\alpha_L, \rho) = 0$, we have $G(\alpha^K, 0) = 0$ and $\frac{\partial G(\alpha^K,0)}{\partial \alpha_L} \neq 0$, so by the implicit function theorem, there exists a C^1 function $\alpha_L(\rho)$ satisfying (A.8) on a neighborhood of 0, and this coincides with our unique linear equilibrium α_L from the proof of Proposition A.3. In turn, by substituting the characterization of β_F via (A.2) into (10), we obtain Λ_1 as a C^1 function of ρ:

$$\Lambda_1 = \frac{\alpha_L[\rho + \phi(1 + \alpha_L)]}{\sigma^2 + \alpha_L^2 \phi},$$

suppressing dependence of α_L on ρ. Therefore, to prove part (ii) it suffices to show that $\frac{d\Lambda_1}{d\rho} > 0$ at $\rho = 0$. Differentiating wrt ρ and evaluating at $\rho = 0$ yields

$$\frac{d\Lambda_1}{d\rho}|_{\rho=0} = \frac{1 + \phi\alpha_L'(0)}{2\sigma\sqrt{\phi}}. \tag{A.28}$$

By the implicit function theorem, $\alpha_L'(0) = -\frac{\sigma}{2\phi(\sigma+\sqrt{\phi})}$. Plugging this into (A.28) yields $\frac{d\Lambda_1}{d\rho}|_{\rho=0} = \frac{2-\frac{\sigma}{\sigma+\sqrt{\phi}}}{4\sigma\sqrt{\phi}}$, which is strictly positive for all $\sigma > 0, \phi > 0$ by inspection.

A.5 Proof of Proposition 4

For part (i), from the proof of Proposition 8, in the PBS equilibrium, α_L/σ converges to a positive constant as $\sigma \to 0$, so it follows that $\lim_{\sigma\to 0} \alpha_L = 0$. By (A.5), $\delta_L/\sigma = -1/(\phi\alpha_L/\sigma)$ converges to a negative constant, and thus $\lim_{\sigma\to 0} \delta_L = 0$.

To establish the results for $\sigma \to +\infty$, recall from the proof of Proposition A.2 that $\hat{\alpha} < \alpha_L < \alpha^K$. But $\lim_{\sigma\to+\infty} \frac{\hat{\alpha}}{\sigma} = 1/\sqrt{\phi} = \lim_{\sigma\to+\infty} \frac{\alpha^K}{\sigma}$, so $\lim_{\sigma\to+\infty} \frac{\alpha_L}{\sigma} = 1/\sqrt{\phi}$. Then by A.5, $\lim_{\sigma\to+\infty} \frac{\delta_L}{\sigma} = -1/\sqrt{\phi}$. These limits imply $\lim_{\sigma\to+\infty} \alpha_L = +\infty$ and $\lim_{\sigma\to+\infty} \delta_L = -\infty$.

For part (ii), $\lim_{\sigma\to 0} |\alpha_L - \alpha^K| = |0 - 0| = 0$ by inspection. Next, let $x_L = \alpha_L/\sigma$ and $x_F = \alpha_F/\sigma$, and recall from above that x_L converges to a positive constant as $\sigma \to \infty$. Using the expression for α_F in (A.8), it is easy to see that x_F also converges to a positive constant as $\sigma \to \infty$. Now rearrange (A.4) using the fact that $\alpha^K := \sqrt{\sigma^2/\phi}$ to write

$$\frac{(\alpha^K)^2}{\alpha_L} - \alpha_L = \frac{\rho\alpha_F}{\phi(1+\alpha_L) + \rho(1+\alpha_F)}$$
$$\implies \alpha^K - \alpha_L = \frac{x_L}{1/\sqrt{\phi} + x_L} \frac{\rho x_F}{(\rho+\phi)/\sigma + \phi x_L + \rho x_F},$$

which converges to a positive constant since both x_L and x_F converge to positive constants as $\sigma \to \infty$ and $\rho + \phi \geq 0$. Hence $\lim\limits_{\sigma\to\infty} |\alpha_L - \alpha^K| > 0$.

References

Admati, A. R. and P. Pfleiderer (2009): "The "Wall Street Walk" and shareholder activism: Exit as a form of voice," *The Review of Financial Studies*, 22, 2645–2685.

Ahnert, T., C. Machado, and A. Elisa Pereira (2020): "Trading for bailouts," Tech. rep., Bank of Canada.

Attari, M., S. Banerjee, and T. H. Noe (2006): "Crushed by a rational stampede: Strategic share dumping and shareholder insurrections," *Journal of Financial Economics*, 79, 181–222.

Back, K., P. Collin-Dufresne, V. Fos, T. Li, and A. Ljungqvist (2018): "Activism, strategic trading, and liquidity," *Econometrica*, 86, 1431–1463.

Bebchuk, L. A., A. Brav, R. J. Jackson Jr, and W. Jiang (2013): "Pre-disclosure accumulations by activist investors: Evidence and policy," *J. Corp. L.*, 39.

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

Berle, A. A. and G. C. Means (1932): *The modern corporation and private property*, Macmillan.

Bliss, B. A., P. Molk, and F. Partnoy (2019): "Negative activism," *Wash. UL Rev.*, 97, 1333.

Boleslavsky, R., D. L. Kelly, and C. R. Taylor (2017): "Selloffs, bailouts, and feedback: Can asset markets inform policy?" *Journal of Economic Theory*, 169, 294–343.

Bolton, P. and C. Harris (1999): "Strategic experimentation," *Econometrica*, 67, 349–374.

Bonatti, A. and G. Cisternas (2020): "Consumer scores and price discrimination," *The Review of Economic Studies*, 87, 750–791.

Bonatti, A., G. Cisternas, and J. Toikka (2017): "Dynamic oligopoly with incomplete information," *The Review of Economic Studies*, 84, 503–546.

Brav, A., A. Dasgupta, and R. Mathews (2021a): "Wolf pack activism," *Management Science*.

Brav, A., W. Jiang, and R. Li (2021b): "Governance by persuasion: hedge fund activism and market-based shareholder influence," *European Corporate Governance Institute–Finance Working Paper*.

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

Briggs, T. W. (2007): "Corporate Governance and the New Hedge Fund Activism: An Empirical Analysis," *Journal of Corporation Law*, 32, 681.

Brunnermeier, M. K. (2005): "Information leakage and market efficiency," *The Review of Financial Studies*, 18, 417–457.

Cetemen, D. (2020): "Efficiency in Repeated Partnerships," Tech. rep.

Cetemen, D., I. Hwang, and A. Kaya (2019): "Uncertainty-driven cooperation," *Theoretical Economics*.

Chakraborty, A. and B. Yilmaz (2004): "Informed manipulation," *Journal of Economic theory*, 114, 132–152.

Cisternas, G. (2018): "Two-sided learning and the ratchet principle," *The Review of Economic Studies*, 85, 307–351.

Coffee Jr, J. C. and D. Palia (2016): "The wolf at the door: The impact of hedge fund activism on corporate governance," *Annals of Corporate Governance*, 1, 1–94.

Collin-Dufresne, P. and V. Fos (2015): "Shareholder activism, informed trading, and stock prices," *Swiss Finance Institute Research Paper*.

Cookson, J. A., V. Fos, and M. Niessner (2022): "Does disagreement facilitate informed trading? Evidence from activist investors," *working paper*.

Di Maggio, M., F. Franzoni, A. Kermani, and C. Sommavilla (2019): "The relevance of broker networks for information diffusion in the stock market," *Journal of Financial Economics*, 134, 419–446.

Doidge, C., A. Dyck, and L. Yang (2021): "Collective Activism," .

Edmans, A. (2009): "Blockholder trading, market efficiency, and managerial myopia," *The Journal of Finance*, 64, 2481–2513.

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

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

Ekmekci, M., L. Gorno, L. Maestri, J. Sun, and D. Wei (2020): "Learning from Manipulable Signals," *arXiv preprint arXiv:2007.08762*.

Foster, F. D. and S. Viswanathan (1996): "Strategic trading when agents forecast the forecasts of others," *The Journal of Finance*, 51, 1437–1478.

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

Goldstein, I. and A. Guembel (2008): "Manipulation and the allocational role of prices," *The Review of Economic Studies*, 75, 133–164.

Holmström, B. (1999): "Managerial incentive problems: A dynamic perspective," *The review of Economic studies*, 66, 169–182.

Huddart, S., J. S. Hughes, and C. B. Levine (2001): "Public disclosure and dissimulation of insider trades," *Econometrica*, 69, 665–681.

Kahn, C. and A. Winton (1998): "Ownership structure, speculation, and shareholder intervention," *The Journal of Finance*, 53, 99–129.

Khanna, N. and R. Mathews (2012): "Doing battle with short sellers: The conflicted role of blockholders in bear raids," *Journal of financial economics*, 106, 229–246.

Kyle, A. S. (1985): "Continuous auctions and insider trading," *Econometrica: Journal of the Econometric Society*, 1315–1335.

Li, T., P. Saffi, and D. Yang (2022): "Power Grab: Activists, Short Sellers, and Disagreement," *Columbia University working paper*.

Maug, E. (1998): "Large shareholders as monitors: Is there a trade-off between liquidity and control?" *The journal of finance*, 53, 65–98.

Ordonez-Calafi, G. and D. Bernhardt (2022): "Blockholder disclosure thresholds and hedge fund activism," *Journal of Financial and Quantitative Analysis*, 1–48.

Shleifer, A. and R. W. Vishny (1986): "Large shareholders and corporate control," *Journal of political economy*, 94, 461–488.

Uherka, D. and A. M. Sergott (1977): "On the continuous dependence of the roots of a polynomial on its coefficients," *The American mathematical monthly*, 84, 368–370.

Williams, B. and A. Skrzypacz (2020): "Spoofing in Equilibrium," *Available at SSRN: 374327*.

Wong, Y. T. F. (2020): "Wolves at the door: A closer look at hedge fund activism," *Management Science*, 66, 2347–2371.

Yang, L. and H. Zhu (2021): "Strategic Trading When Central Bank Intervention Is Predictable," *The Review of Asset Pricing Studies*, 11, 735–761.

Online Appendix for "Activist Trading Dynamics"

B Omitted proofs, supporting details, and additional results

B.1 Supporting details for learning and pricing

Using the fact that $P_0 = \mathbb{E}[P_1]$, the quoted price P_0 satisfies

$$P_0 = \mathbb{E}[(1+\alpha_L)X_0^L + \delta_L\mu + (1+\alpha_F)X_0^F + \beta_F P_0 + \delta_F\mu].$$

Solving for P_0 yields [37]

$$P_0 = \frac{\mu(2+\alpha_L+\alpha_F+\delta_L+\delta_F)}{1-\beta_F}. \tag{B.1}$$

After period 1, the posterior covariance matrix of the market maker's beliefs about (X_T^L, X_0^F) is $\Gamma_1 = \begin{pmatrix} \gamma_1^L & \rho_1 \\ \rho_1 & \gamma_1^F \end{pmatrix}$, where

$$\gamma_1^L = \frac{\phi\sigma^2(1+\alpha_L)^2}{\alpha_L^2\phi + \sigma^2}, \tag{B.2}$$

$$\gamma_1^F = \frac{\alpha_L^2[\phi^2-\rho^2] + \phi\sigma^2}{\alpha_L^2\phi + \sigma^2}, \tag{B.3}$$

$$\rho_1 = \frac{\rho\sigma^2(1+\alpha_L)}{\alpha_L^2\phi + \sigma^2}. \tag{B.4}$$

The expressions for γ_1^L, γ_1^F, and ρ_1 can be obtained using the law of total variance and law of total covariance.[38]

The market maker's updated beliefs about (X_T^L, X_T^F) after the second-period order flow is observed are given by

$$\begin{aligned} M_T^F &:= \mathbb{E}[X_T^F|\mathcal{F}_2^\Psi] \\ &= (1+\alpha_F)M_1^F + \beta_F P_1 + \delta_F\mu + \frac{\alpha_F\gamma_1^F(1+\alpha_F)}{\alpha_F^2\gamma_1^F + \sigma^2}[\Psi_2 - \alpha_F M_1^F - \beta_F P_1 - \delta_F\mu] \end{aligned} \tag{B.5}$$

[37] The leader's SOC requires $\beta_F \neq 1$, and thus in any equilibrium, the denominator in (B.1) is nonzero.

[38] For instance, $(1+\alpha_L)\rho = \mathrm{Cov}(X_T^L, X_0^F) = \mathbb{E}[\mathrm{Cov}(X_T^L, X_0^F|\mathcal{F}_1^\Psi)] + \mathrm{Cov}(M_1^L, M_1^F) = \rho_1 + \frac{\alpha_L^2(1+\alpha_L)\phi\rho}{(\alpha_L^2\phi+\sigma^2)^2}(\alpha_L^2\phi+\sigma^2) = \rho_1 + \frac{\alpha_L^2(1+\alpha_L)\phi\rho}{\alpha_L^2\phi+\sigma^2}$, so $\rho_1 = (1+\alpha_L)\rho - \frac{\alpha_L^2(1+\alpha_L)\phi\rho}{\alpha_L^2\phi+\sigma^2}$.

$$M_T^L := \mathbb{E}[X_T^L|\mathcal{F}_2^\Psi] = M_1^L + \frac{\alpha_F\rho_1}{\alpha_F^2\gamma_1^F + \sigma^2}[\Psi_2 - \alpha_F M_1^F - \beta_F P_1 - \delta_F\mu]. \tag{B.6}$$

B.2 Proof of Proposition 5

We consider symmetric linear strategies of the form

$$\theta^i = \alpha X_0^i + \beta\mu. \tag{B.7}$$

We begin by characterizing belief updating and pricing, and then we use these to set up the best-response problem of either trader.

After observing the total order flow, the market maker updates her beliefs about the activists' positions. Given the form of strategies and symmetry, it is sufficient for the market maker to only estimate the sum of initial positions. By standard Gaussian filtering,

$$\begin{aligned}\mathbb{E}[X_0^i + X_0^j|\mathcal{F}_1^\Psi] &= 2\mu + \frac{\operatorname{Cov}(X_0^i + X_0^j, \Psi_1)}{\operatorname{Var}(\Psi_1)}\left\{\Psi_1 - \underbrace{[2\alpha\mu + 2\beta\mu]}_{=\mathbb{E}[\theta^i+\theta^j]}\right\} \\ &= 2\mu + \frac{2\alpha(\phi+\rho)}{2\alpha^2(\phi+\rho)+\sigma^2}\left\{\Psi_1 - 2\mu(\alpha+\beta)]\right\}.\end{aligned}$$

Hence, P_1 is equal to

$$\begin{aligned}P_1 = \mathbb{E}[W|\mathcal{F}_1^\Psi] = \mathbb{E}[X_T^i + X_T^j|\mathcal{F}_1^\Psi] &= (1+\alpha)\mathbb{E}[X_0^i + X_0^j|\mathcal{F}_1^\Psi] + 2\mu\beta && \text{(B.8)}\\ = P_0^S + \Lambda_1^S\left\{\Psi_1 - 2\mu(\alpha+\beta)]\right\}, && && \text{(B.9)}\end{aligned}$$

where $P_0^S := 2\mu(1+\alpha+\beta)$ is the ex ante expected firm value and $\Lambda_1^S := (1+\alpha)\frac{2\alpha(\phi+\rho)}{2\alpha^2(\phi+\rho)+\sigma^2}$ is Kyle's lambda.

Each activist then maximizes

$$\sup_{\theta^i}\mathbb{E}\left[\frac{(X_0^i+\theta^i)^2 + 2X_T^{-i}(X_0^i+\theta^i)}{2} - P_1\theta^i|X_0^i,\theta^i\right]. \tag{B.10}$$

The FOC is

$$\frac{2(X_0^i+\theta^i) + 2\mathbb{E}[X_T^{-i}|X_0^i]}{2} - \theta^i\frac{\partial P_1}{\partial\theta^i} - P_1 = 0. \tag{B.11}$$

Plugging in the expression for Λ_1^S, evaluating at the conjectured strategy (B.7), and setting the coefficient on X_0^i to 0 yields an equation for α with the following three roots:

$$\alpha = \frac{\sigma}{\sqrt{2\phi}}, \quad -\frac{\sigma}{\sqrt{2\phi}}, \quad -1. \tag{B.12}$$

Similarly, setting the coefficient on μ to 0, we can pin down β from α via the following equation

$$\beta = \frac{\sigma^2}{2\sigma^2 - 4\alpha(1+\alpha)\phi}. \tag{B.13}$$

Since the second and third roots are negative, we have a unique candidate for a symmetric PBS equilibrium.

For existence, we must check the SOC

$$1 - 2\Lambda_1^S \leq 0. \tag{B.14}$$

Plugging in $\alpha = \frac{\sigma}{\sqrt{2\phi}}$, this condition is equivalent to the inequality

$$\sigma^2 - 2\alpha(2+\alpha)(\rho+\phi) = \sigma^2 - 2\frac{\sigma}{\sqrt{2\phi}}\left(2 + \frac{\sigma}{\sqrt{2\phi}}\right)(\phi+\rho) \leq 0.$$

The left hand side is decreasing and continuous in ρ, and it is strictly negative when $\rho = 0$, so there exists $\rho_0^{\text{sim}} \in (-\phi, 0)$ such that the inequality is satisfied, and in turn a unique PBS equilibrium exists, whenever $\rho \in [\rho_0^{\text{sim}}, \phi]$.

To compare payoffs to the sequential-move game, choose $|\rho|$ sufficiently small such that by the first part of the current result and by Proposition 2, there is a unique PBS equilibrium of the both the simultaneous-move and sequential-move games. The leader's expected payoff in either case is continuous in ρ and in α_L (which is in turn continuous in ρ at $\rho = 0$ by Corollary A.1).[39] Hence, it suffices to prove the result for $\rho = 0$.

Recall that when $\rho = 0$, the equilibrium is characterized in Proposition 1, and $\alpha_L = \alpha_F = \sqrt{\frac{\sigma^2}{\phi}}$. The coefficient in the simultaneous-move game is $\alpha_S := \sqrt{\frac{\sigma^2}{2\phi}}$ (see (B.12)), where $\alpha_L = \alpha_F > \alpha_S$.

To calculate the leader's expected payoff in the sequential case, plug the equilibrium strategies into (4) to obtain

$$\mathbb{E}\Bigg[\tfrac{1}{2}\left(X_0^L\left(1+\sqrt{\tfrac{\sigma^2}{\phi}}\right) - \sqrt{\tfrac{\sigma^2}{\phi}}\mu\right)^2 + \left(X_0^F + \sqrt{\tfrac{\sigma^2}{\phi}}(X_0^F - \mu)\right)\left(X_0^L + \sqrt{\tfrac{\sigma^2}{\phi}}(X_0^L - \mu)\right)$$
$$- \left(P_0 + \Lambda_1\left(\sqrt{\tfrac{\sigma^2}{\phi}}(X_0^L - \mu) + \sigma Z_1\right)\right)\sqrt{\tfrac{\sigma^2}{\phi}}(X_0^L - \mu)\Bigg].$$

[39]Full expressions for general ρ are available from the authors upon request.

Opening up the expectation and simplifying we can write the first line as

$$\frac{1}{2}\left(\mu^2+(\sigma+\sqrt{\phi})^2\right)+\mu^2,$$

and second line simplifies to

$$-\frac{\sigma(\sigma+\sqrt{\phi})}{2}.$$

Hence, the leader's total expected payoff is

$$\frac{1}{2}\left[3\mu^2+\phi+\sigma\sqrt{\phi}\right]. \tag{B.15}$$

Following similar steps for the simultaneous case, we can write the equilibrium payoff of player i $(i=1,2)$ as

$$\mathbb{E}\Bigg[\tfrac{1}{2}\left(X_0^i\left(1+\sqrt{\tfrac{\sigma^2}{2\phi}}\right)-\sqrt{\tfrac{\sigma^2}{2\phi}}\mu\right)^2+2\left(X_0^j+\sqrt{\tfrac{\sigma^2}{2\phi}}(X_0^j-\mu)\right)\left(X_0^i+\sqrt{\tfrac{\sigma^2}{2\phi}}(X_0^i-\mu)\right)$$
$$-\left(P_0^S+\Lambda_1^S\left(\sqrt{\tfrac{\sigma^2}{2\phi}}(X_0^i-\mu)+\epsilon_i\right)\right)\sqrt{\tfrac{\sigma^2}{2\phi}}(X_0^i-\mu)\Bigg].$$

Opening up the expectation, the first line simplifies to

$$\frac{1}{2}\left(\mu^2+\frac{(\sigma+\sqrt{2\phi})^2}{2}\right)+\mu^2,$$

while the second line simplifies to

$$-\frac{\sigma(\sigma+\sqrt{2\phi})}{4},$$

for a total expected payoff of

$$\frac{1}{2}\left[3\mu^2+\phi+\frac{\sigma\sqrt{2\phi}}{2}\right]. \tag{B.16}$$

Subtracting (B.16) from (B.15) yields $\frac{1}{2}\left(1-\frac{\sqrt{2}}{2}\right)\sigma\sqrt{\phi}$, which is strictly positive. Therefore, the leader unambiguously prefers the sequential-move game when $\rho=0$.

B.3 Proofs for Section 5.3

Proof of Proposition 6. Fix μ,σ,ϕ,ρ. Let μs_μ denote the prior mean for each follower, ϕs_ϕ the variance, and ρs_ρ the covariance between the leader and each follower, where s_μ, s_ϕ, s_ρ

will vary with N. The setup described in Section 5.3 is captured by $s_\mu = 1/N$, $s_\phi = 1/N^2$, and $s_\rho = 1/N$.

Define $\gamma_1^{\text{sum}} = N^2\gamma_1^F$, the market maker's posterior variance of the sum of all followers' positions. In any PBS equilibrium, the followers play gap strategies and their FOC yields $\alpha_F = \sqrt{\frac{N\sigma^2}{\gamma_1^{\text{sum}}}} = \sqrt{\frac{\sigma^2}{N\gamma_1^F}}$. By adapting the proof of Proposition 2, the leader's FOC yields the following equation for α_L:

$$\frac{(N\rho s_\rho + \phi + \alpha_L\phi)(\sigma^2 - \alpha_L^2\phi)}{N\rho s_\rho[\alpha_L(1+\alpha_L)\phi - \sigma^2]} = \sqrt{\frac{\sigma^4 + \sigma^2\alpha_L^2\phi}{N[\phi s_\phi\sigma^2 + \alpha_L^2(\phi^2 s_\phi - (\rho s_\rho)^2)]}}. \tag{B.17}$$

Familiar arguments show that for $\rho > 0$, (B.17) has a solution α_L in $(\hat{\alpha}, \alpha^K)$, there is no other solution for $\alpha_L \geq 0$, and SOCs are satisfied. The FOC also implies that the coefficient on μ is $\delta_L = -\frac{\sigma^2}{\phi\alpha_L}$. Hence, we have characterized the unique PBS equilibrium.

We now turn to comparative statics wrt N. After plugging in our values for (s_μ, s_ϕ, s_ρ), (B.17) reduces to

$$\frac{(\rho + \phi + \alpha_L\phi)(\sigma^2 - \alpha_L^2\phi)}{\rho[\alpha_L(1+\alpha_L)\phi - \sigma^2]} = \sqrt{\frac{N(\sigma^4 + \sigma^2\alpha_L^2\phi)}{\phi\sigma^2 + \alpha_L^2(\phi^2 - \rho^2)}}. \tag{B.18}$$

When these intersect at $\alpha_L \in (\hat{\alpha}, \alpha^K)$, the left hand side crosses the right hand side from above. Then since the right hand side is increasing in N, the equilibrium value of α_L is decreasing in N. It is also straightforward to show that the left side of (B.18) is decreasing in α_L on $(\hat{\alpha}, \infty)$, so each side of (B.18) is increasing in N. Since the right hand side is precisely α_F, this establishes that α_F is increasing in N.

Since the followers play gap strategies, ex ante firm value is still $(2 + \alpha_L + \delta_L)\mu = (2 + \alpha_L - \sigma^2/(\phi\alpha_L))\mu$ for all N. Since α_L is decreasing in N, ex ante firm value is decreasing in N.

For later use, we show that $\lim_{N\to\infty} \alpha_L = \hat{\alpha} > 0$, where $\hat{\alpha}$ was defined earlier as the positive root of $\alpha_L(1+\alpha_L)\phi - \sigma^2$. As $N \to \infty$, the right hand side of (B.18) explodes as the rest of the expression in the square root is bounded. Thus, the left hand side must also explode, which requires its denominator to vanish. Given that $\alpha_L > 0$, this implies that α_L converges to $\hat{\alpha}$.

We now turn to the asymptotic result. The leader's expected payoff is

$$\mathbb{E}\left[-P_1\theta_L + \frac{(X_0^L + \theta_L)^2}{2} + (X_0^L + \theta_L)N(X_0^F + \alpha_F(X_0^F - M_1^F))\right]. \tag{B.19}$$

We simplify (B.19) one term at a time. The first term equals

$$
\begin{aligned}
&-\mathbb{E}[(P_0 + \Lambda_1[\Psi_1 - (\alpha_L + \delta_L)\mu])\theta_L] \\
&= -\mathbb{E}[P_0(\alpha_L X_0^L + \delta_L\mu) + \Lambda_1\alpha_L(X_0^L - \mu)(\alpha_L X_0^L + \delta_L\mu)] \\
&= -\mathbb{E}[(2 + \alpha_L + \delta_L)\mu(\alpha_L X_0^L + \delta_L\mu) + \Lambda_1\alpha_L(X_0^L - \mu)(\alpha_L X_0^L + \delta_L\mu)] \\
&= -\mathbb{E}[(2 + \alpha_L + \delta_L)\mu(\alpha_L X_0^L + \delta_L\mu) + \Lambda_1\alpha_L(X_0^L - \mu)\alpha_L X_0^L] \\
&= -[(2 + \alpha_L + \delta_L)(\alpha_L + \delta_L)\mu^2 + \Lambda_1\alpha_L^2\phi] =: S_1, \qquad \text{(B.20)}
\end{aligned}
$$

where the third equality uses that $\mathbb{E}[X_0^L - \mu] = 0$. Since α_L and δ_L have finite limits as $N \to \infty$, and $\Lambda_1 = \frac{\alpha_L(\rho+\phi(1+\alpha_L)}{\sigma^2+\alpha_L^2\phi}$ also has a finite limit, this term overall is therefore uniformly bounded in N.

The expectation of the second term in (B.19) equals

$$
S_2 := \frac{1}{2}\mathbb{E}\left[(X_0^L(1+\alpha_L) + \delta_L\mu)^2\right] = \frac{1}{2}[(1 + \alpha_L + \delta_L)^2\mu^2 + \phi(1+\alpha_L)^2], \qquad \text{(B.21)}
$$

which is also uniformly bounded in N.

Using that $\mathbb{E}[X_0^F - M_1^F] = 0$ by the law of iterated expectations, the third term in (B.19) simplifies as:

$$
\begin{aligned}
&\mathbb{E}[(X_0^L(1+\alpha_L) + \delta_L\mu)N(X_0^F + \alpha_F(X_0^F - M_1^F))] \\
&= (1+\alpha_L)(1+\alpha_F)N\mathbb{E}[X_0^L X_0^F] + \delta_L N\mu^2 s_\mu - \mathbb{E}[X_0^L(1+\alpha_L)N\alpha_F M_1^F] \\
&= (1+\alpha_L)(1+\alpha_F)N(\mu^2 s_\mu + \rho s_\rho) + \delta_L N\mu^2 s_\mu - \mathbb{E}[X_0^L(1+\alpha_L)N\alpha_F M_1^F] \\
&= (1+\alpha_L)(1+\alpha_F)N(\mu^2 s_\mu + \rho s_\rho) + \delta_L N\mu^2 s_\mu \\
&\quad - \mathbb{E}[X_0^L(1+\alpha_L)N\alpha_F\left\{\mu s_\mu + \frac{\alpha_L\rho s_\rho}{\alpha_L^2\phi + \sigma^2}[\alpha_L X_0^L + \delta_L\mu - (\alpha_L + \delta_L)\mu]\right\}. \qquad \text{(B.22)}
\end{aligned}
$$

We now simplify the last term in (B.22):

$$
\begin{aligned}
&\mathbb{E}\left[X_0^L(1+\alpha_L)N\alpha_F\left\{\mu s_\mu + \frac{\alpha_L\rho s_\rho}{\alpha_L^2\phi + \sigma^2}[\alpha_L X_0^L + \delta_L\mu - (\alpha_L + \delta_L)\mu]\right\}\right] \\
&= \mathbb{E}\left[X_0^L(1+\alpha_L)N\alpha_F\left\{\mu s_\mu + \frac{\alpha_L\rho s_\rho}{\alpha_L^2\phi + \sigma^2}\alpha_L(X_0^L - \mu)\right\}\right] \\
&= (1+\alpha_L)N\alpha_F\mu^2 s_\mu + (1+\alpha_L)N\alpha_F\frac{\alpha_L\rho s_\rho}{\alpha_L^2\phi + \sigma^2}\alpha_L\mathbb{E}[X_0^L(X_0^L - \mu)] \\
&= (1+\alpha_L)N\alpha_F\mu^2 s_\mu + (1+\alpha_L)N\alpha_F\frac{\alpha_L\rho s_\rho}{\alpha_L^2\phi + \sigma^2}\alpha_L\mathrm{Var}(X_0^L) \\
&= (1+\alpha_L)\alpha_F\mu^2 + (1+\alpha_L)\alpha_F\frac{\alpha_L\rho}{\alpha_L^2\phi + \sigma^2}\alpha_L\phi.
\end{aligned}
$$

Incorporating this in (B.22), the third term of (B.19) equals

$$S_3 := (1+\alpha_L)(1+\alpha_F)(\mu^2+\rho) + \delta_L\mu^2 - \left[(1+\alpha_L)\alpha_F\mu^2 + (1+\alpha_L)\alpha_F\frac{\alpha_L^2\rho\phi}{\alpha_L^2\phi+\sigma^2}\right]$$
$$= (1+\alpha_L)(\mu^2+\rho) + \delta_L\mu^2 + \alpha_F\rho(1+\alpha_L)\frac{\sigma^2}{\alpha_L^2\phi+\sigma^2}, \quad \text{(B.23)}$$

where we have canceled N with $1/N$ in s_μ and s_ρ. Again, $(1+\alpha_L)(\mu^2+\rho)+\delta_L\mu^2$ is uniformly bounded in N, so S_3 has the form $C(N)+\alpha_F\rho(1+\alpha_L)\frac{\sigma^2}{\alpha_L^2\phi+\sigma^2}$ as noted in Section 5.3.

The leader's payoff is the sum of (B.20), (B.21), and (B.23):

$$\Pi_L = S_1 + S_2 + S_3. \quad \text{(B.24)}$$

To show that the rate of growth is $\sqrt{N}$, we calculate

$$\begin{aligned}\lim_{N\to\infty}\frac{\Pi_L}{\sqrt{N}} &= \lim_{N\to\infty}\frac{S_1}{\sqrt{N}} + \lim_{N\to\infty}\frac{S_2}{\sqrt{N}} + \lim_{N\to\infty}\frac{S_3}{\sqrt{N}} \\ &= 0 + 0 + \lim_{N\to\infty}\frac{S_3}{\sqrt{N}} \\ &= \lim_{N\to\infty}\frac{\alpha_F}{\sqrt{N}}(1+\alpha_L)\rho\frac{\sigma^2}{\alpha_L^2\phi+\sigma^2} \\ &= \left(\lim_{N\to\infty}\frac{\alpha_F}{\sqrt{N}}\right)\left(\lim_{N\to\infty}(1+\alpha_L)\rho\frac{\sigma^2}{\alpha_L^2\phi+\sigma^2}\right). \quad \text{(B.25)}\end{aligned}$$

To take limits in the last line, we use the fact that for $\rho \in (0,\phi]$, $\lim_{N\to\infty}\alpha_L = \hat{\alpha} > 0$, as shown earlier in the proof. We have

$$\lim_{N\to\infty}\frac{\alpha_F}{\sqrt{N}} = \lim_{N\to\infty}\sqrt{\frac{(\sigma^4+\sigma^2\alpha_L^2\phi)}{\phi\sigma^2+\alpha_L^2(\phi^2-\rho^2)}} = \sqrt{\frac{(\sigma^4+\sigma^2\hat{\alpha}^2\phi)}{\phi\sigma^2+\hat{\alpha}^2(\phi^2-\rho^2)}} \quad \text{(B.26)}$$
$$\lim_{N\to\infty}(1+\alpha_L)\rho\frac{\sigma^2}{\alpha_L^2\phi+\sigma^2} = (1+\hat{\alpha})\rho\frac{\sigma^2}{\hat{\alpha}^2\phi+\sigma^2}. \quad \text{(B.27)}$$

Since these limits are positive and finite, so is their product, and we conclude that Π_L grows asymptotically at rate $\sqrt{N}$. □

The following result was referred to in Section 5.3.

Lemma B.1. *Assume $\rho = \phi$, and let Π_L^{seq} and Π_L^{sim} denote the leader's payoff in the sequential- and simultaneous-move games, respectively. When N is sufficiently large, the leader's payoff advantage from going first is increasing in N. Specifically, Π_L^{seq} and Π_L^{sim} grow at rate $\sqrt{N}$ asymptotically, and $\lim_{N\to\infty}\frac{\Pi_L^{seq}-\Pi_L^{sim}}{\sqrt{N}} > 0$.*

Proof. Proposition 6 characterizes the asymptotics of Π_L^{seq}, so consider the simultaneous-move game. The FOCs lead to the following system of equations:

$$\alpha_L = \frac{1 - \frac{\rho}{\phi}\Lambda\alpha_F + \frac{\rho}{\phi}(1+\alpha_F)}{2\Lambda - 1}, \tag{B.28}$$

$$\alpha_F = \frac{N(1 - \frac{\rho}{\phi}\Lambda\alpha_L + \frac{\rho}{\phi}(1+\alpha_L))}{(N+1)\Lambda - N}, \tag{B.29}$$

where $\Lambda = \frac{(1+\alpha_L)(\phi\alpha_L+\rho\alpha_F)+(1+\alpha_F)(\phi\alpha_F+\rho\alpha_L)}{\phi(\alpha_L^2+\alpha_F^2)+2\alpha_L\alpha_F\rho+\sigma^2}$.

For the case $\rho = \phi$, we obtain $(\alpha_L, \alpha_F) = \left(\frac{\sigma}{\sqrt{(N+1)\phi}}, \frac{N\sigma}{\sqrt{(N+1)\phi}}\right)$. The leader's payoff is again of the order $\sqrt{N}$, with coefficient $\lim_{N\to\infty} \frac{\alpha_F}{\sqrt{N}}(1+\alpha_L)\text{Cov}(X_0^L, X_0^F) = \lim_{N\to\infty} \frac{\alpha_F}{\sqrt{N}}(1+\alpha_L)\phi = \sigma\sqrt{\phi}$. To complete the proof, we show that this is strictly less than the corresponding coefficient in the sequential-move game, namely $\sqrt{\frac{(\sigma^4+\sigma^2\hat{\alpha}^2\phi)}{\phi\sigma^2}}(1+\hat{\alpha})\phi\frac{\sigma^2}{\hat{\alpha}^2\phi+\sigma^2}$. By routine simplifications,

$$\begin{aligned}
\sigma\sqrt{\phi} &\leq \sqrt{\frac{(\sigma^4+\sigma^2\hat{\alpha}^2\phi)}{\phi\sigma^2}}(1+\hat{\alpha})\phi\frac{\sigma^2}{\hat{\alpha}^2\phi+\sigma^2} \\
\iff 1 &\leq \sqrt{\sigma^2+\hat{\alpha}^2\phi}(1+\hat{\alpha})\frac{\sigma}{\hat{\alpha}^2\phi+\sigma^2} \\
\iff \sqrt{\sigma^2+\hat{\alpha}^2\phi} &\leq (1+\hat{\alpha})\sigma \\
\iff \sigma^2+\hat{\alpha}^2\phi &\leq (1+\hat{\alpha})^2\sigma^2 \qquad \text{(since both sides are positive)} \\
\iff 0 &\leq \hat{\alpha}[\hat{\alpha}(\sigma^2-\phi)+2\sigma^2].
\end{aligned}$$

Since $\hat{\alpha}$ solves $\sigma^2 - \hat{\alpha}(1+\hat{\alpha})\phi = 0$, the right hand side is

$$\hat{\alpha}[\hat{\alpha}(\sigma^2-\phi)+2\sigma^2] = \hat{\alpha}[\hat{\alpha}\sigma^2+\hat{\alpha}^2\phi-\sigma^2+2\sigma^2] = \hat{\alpha}[\hat{\alpha}\sigma^2+\hat{\alpha}^2\phi+\sigma^2] \geq 0,$$

establishing the inequality. □

B.4 Proof of Proposition 7

For part (i), we prove that for sufficiently large σ, there is a solution to (A.9) with $\alpha_L < 0$. We then check the conditions (A.6), (A.7), and $\phi(1+\alpha_L)+\rho \neq 0$ and apply the "converse" part of Proposition A.1.

After a change of variables $x = \alpha_L/\sigma$ in (A.9), we obtain

$$-\sqrt{\frac{1+x^2\phi}{\phi+x^2(\phi^2-\rho^2)}} = \frac{\left(\frac{\rho+\phi}{\sigma}+\phi x\right)(x^2\phi-1)}{\rho[1-x\phi/\sigma-x^2\phi]}. \tag{B.30}$$

When $x = -1/\sqrt{\phi}$, the right hand side vanishes, while the left hand side is strictly negative. Now choose σ sufficiently large that $\left(\frac{\rho+\phi}{\sigma}+\phi x\right) < 0$ for all $x \le -1/\sqrt{\phi}$. Define $\alpha^\dagger$ to be the negative root of $\alpha_L(1+\alpha_L)\phi - \sigma^2$, and define $x^\dagger = \alpha^\dagger/\sigma < -1/\sqrt{\phi}$ to be the unique negative root of the denominator of (B.30), where $x^\dagger \uparrow -1/\sqrt{\phi}$ as $\sigma \uparrow \infty$. The right hand side of (B.30) is well-defined and continuous on $(x^\dagger, -1/\sqrt{\phi}]$ and moreover, it has limit $-\infty$ as $x \downarrow x^\dagger$. Thus, by the intermediate value theorem, there exists a solution x_L to (B.30) in $(x^\dagger, -1/\sqrt{\phi}]$, and by the squeeze theorem, $\lim_{\sigma\uparrow\infty} x_L = -1/\sqrt{\phi}$. (By reversing the change of variables, one can recover α_L solving the leader's FOC.) Note that as $\sigma \uparrow \infty$, $x_F := \alpha_F/\sigma = -\sqrt{\frac{1+x^2\phi}{\phi+x^2(\phi^2-\rho^2)}} \to -\sqrt{\frac{2}{2\phi-\rho^2/\phi}} =: x_F^\infty$

To verify (A.6), note that this is equivalent to the condition $1-x_L^2\phi-2x_L\left(\frac{\rho+\phi}{\sigma}+\rho x_F\right) \le 0$. As $\sigma \uparrow +\infty$, the left hand side has limit $1-1-2(-1/\sqrt{\phi})\rho x_F^\infty = 2\rho x_F^\infty/\sqrt{\phi} < 0$, so (A.6) is satisfied for sufficiently large σ.

As for (A.7), using that $\alpha_{F,2} < 0$, it suffices to show that

$$\sigma^2[x_L^2(\phi^2-\rho^2) + x_L\sigma\rho + (\phi+\rho)] \le 0.$$

Recall that x_L has finite limit as $\sigma \to +\infty$, so the dominating term is $\sigma^3 x_L \rho < 0$. We conclude that (A.7) is satisfied for sufficiently large σ.

Finally, observe that since the left side of (B.30) is nonzero, at our solution the right side is also nonzero, and thus $\frac{\rho+\phi}{\sigma}+\phi x_L = \frac{1}{\sigma}[\phi(1+\alpha_L)+\rho] \neq 0$. Hence Proposition A.1 applies, giving us existence for large σ.

For part (ii), we begin with the observation that for $\rho = -\phi$, (A.7) becomes

$$\sigma^2\phi\alpha_F\alpha_L \le 0. \tag{B.31}$$

Hence, there is no equilibrium in which α_F and α_L are both strictly positive or both strictly negative, and (17)-(18) imply $\alpha_L \neq 0$ and $\alpha_F \neq 0$.

We now establish the existence of an equilibrium with $\alpha_L < 0 < \alpha_F$. Note that for $\rho = -\phi$, as long as $\alpha_L \neq 0$ (which must hold in any equilibrium), the condition $\phi(1+\alpha_L)+\rho \neq 0$ is satisfied. When $\rho = -\phi$ and $\alpha_F = \alpha_{F,1}$, (A.8) simplifies to

$$\sqrt{\sigma^2/\phi+\alpha_L^2} = \alpha_L\frac{\alpha_L^2\phi-\sigma^2}{\alpha_L(1+\alpha_L)\phi-\sigma^2}. \tag{B.32}$$

In particular, an equilibrium with $\alpha_F = \alpha_{F,1}$ exists if and only if there exists α_L satisfying (B.32) such that both SOCs are satisfied. Now the left hand side of (B.32) is positive, while the right hand side vanishes at $\alpha_L = -\sigma/\sqrt{\phi}$, has limit $+\infty$ as $\alpha_L \downarrow \alpha^\dagger$, and is continuous on $(\alpha^\dagger, -\sigma/\sqrt{\phi})$, where $\alpha^\dagger$ was previously defined as the negative root of $\alpha_L(1+\alpha_L)\phi - \sigma^2$, and recall that $\hat{\alpha}$ is the positive root. Thus, (B.32) has a solution in this interval. We finally check (A.6), which is now $\sigma^2 - \alpha_L^2\phi + 2\alpha_L\phi\alpha_F \leq 0$. This is satisfied since $\alpha_L < -\sigma/\sqrt{\phi}$ implies $\sigma^2 - \alpha_L^2\phi < 0$, and clearly $2\alpha_L\phi\alpha_F < 0$. Since α_F and α_L have opposite signs, (A.7) is satisfied. Hence, existence follows from Proposition A.1.

B.5 Proof of Proposition 8

Since Proposition 1 establishes existence and uniqueness for all $\sigma > 0$ when $\rho = 0$, assume $\rho \neq 0$. We will show that for sufficiently small $\sigma > 0$, there is a unique pair (α_L, α_F) satisfying (A.1), (A.14), (A.6), and (A.7). Further, we will show that $\phi(1+\alpha_L) + \rho \neq 0$, so existence follows from Proposition A.1.

In any equilibrium, (α_L, α_F) must solve (A.14). By squaring both sides of this equation, using (A.1), and multiplying through by the nonzero denominator, we get (A.19). Now as $\sigma \to 0$, the coefficients of the polynomial Q converge to those of

$$Q^{\sigma=0}(\alpha_L) := -\alpha_L^6\phi^2[\rho + \phi + \alpha_L\phi]^2(\phi^2 - \rho^2), \tag{B.33}$$

which has a root of multiplicity 6 at 0 and of multiplicity 2 at $-\frac{\rho+\phi}{\phi}$.

By Lemma A.1, for any $\epsilon > 0$, there exists $\delta > 0$ such that if $\sigma \in (0, \delta)$, Q has 6 complex roots within distance ϵ of 0 and 2 complex roots within ϵ of $-\frac{\rho+\phi}{\phi}$. For ϵ sufficiently small that these neighborhoods do not intersect, and δ chosen accordingly, let $\alpha_1, \ldots, \alpha_6$ denote the 6 roots near 0, and let α_7 and α_8 denote the roots near $-\frac{\rho+\phi}{\phi}$. We maintain these assumptions on ϵ and δ throughout the proof.

The following lemma rules out α_7 and α_8 from being part of an equilibrium.

Lemma B.2. *For sufficiently small $\sigma > 0$, each of α_7 and α_8 is either complex or otherwise fails (A.6).*

Proof. The left side of (A.6) is continuous in (σ, α_L) at $\left(0, -\frac{\rho+\phi}{\phi}\right)$, where it evaluates to $(\phi+\rho)^2/\phi > 0$. Hence, choosing $\epsilon > 0$ sufficiently small, and $\delta > 0$ sufficiently small as described before the lemma, if either α_7 or α_8 is real, it fails (A.6). $\square$

Remark 2. *Having ruled out α_7 and α_8, note that if σ is sufficiently small, then for any real $\alpha_L \in \{\alpha_1, \ldots, \alpha_6\}$, $\rho + \phi + \alpha_L\phi \neq 0$. This fact is useful two fold: (i) this criterion appears in* *the sufficiency part of Proposition A.1, and (ii) due to (A.14), using that $\rho \neq 0$ and $\alpha_{F,1} \neq 0$ and $\alpha_{F,2} \neq 0$ for α_L real, we have $\sigma^2 - \alpha_L(1+\alpha_L) \neq 0$ for sufficiently small σ for α_L real. Thus, any real solution to (A.19) solves (A.18).*

We can now rule out equilibria in which $\alpha_F = \alpha_{F,2}$, as these fail the follower's second order condition when σ is sufficiently small. To do so, we use asymptotic properties of the roots of (A.19) as $\sigma \to 0$.

It is useful to define a change of variables $z = \alpha_L/\sigma$ in (A.19) and divide through the resulting equation by σ^6, obtaining an equivalent equation

$$0 = \tilde{Q}(z,\sigma) := \sigma H(z) + F(z), \tag{B.34}$$

where $H(z)$ is a polynomial of degree 8 and where $F(z)$ is a polynomial independent of σ that has the form $c_6 z^6 + c_4 z^4 + c_2 z^2 + c_0$.[40] For each $i \in \{1, 2, \ldots, 6\}$, define $z_i = \alpha_i/6$.

Lemma B.3. *F has 6 distinct roots, denoted $\hat{z}_1, \ldots, \hat{z}_6$, of which exactly two are positive, two are negative, and two are complex. As $\sigma \to 0$, $z_1, \ldots, z_6$ converge to $\hat{z}_1, \ldots, \hat{z}_6$.*

Proof. We first characterize the roots of F. Consider the cubic polynomial $G(y) = c_6 y^3 + c_4 y^2 + c_2 y + c_0$, where $F(y) = G(y^2)$. We have $G(0) < 0$ and $\lim_{y\to-\infty} G(y) = +\infty$, so G has a negative root. Also, we have $\lim_{y\to+\infty} G(y) = -\infty$ and $G(1/\phi) = 2\rho^2\phi > 0$, so G has two distinct positive roots: one in $(0, 1/\phi)$ and one in $(1/\phi, +\infty)$. Since G is cubic, there are no other roots (real or complex). Now the negative root of G corresponds to two distinct complex roots of F, and the positive roots of G each correspond to both one positive and one negative root of F, all distinct.

We now turn to the convergence claim in the lemma. Next, set $K = 1 + \max_{i\in\{1,\ldots,6\}} |\hat{z}_i|$, and define a compact set $\mathcal{K} = \{z \in \mathbb{C} : |z| \leq K\}$. By definition, all roots of F lie in $\mathcal{K}$. Further, note that on $\mathcal{K}$, for any sequence $(\sigma_n)_{n\in\mathbb{N}}$ with $\sigma_n \downarrow 0$, the sequence $(\tilde{Q}(\cdot, \sigma_n))_{n\in\mathbb{N}}$ of functions defined on $\mathcal{K}$ is equicontinuous and converges pointwise to F since $\sigma H(z)$ vanishes; thus, by the Arzela-Ascoli theorem, the sequence converges uniformly to F on $\mathcal{K}$.

Choose $\overline{\eta} > 0$ less than 1 and less than the minimum distance between any $\hat{z}_i$ and $\hat{z}_j$, where $i, j \in \{1, \ldots, 6\}$ and $i \neq j$. Then for all $\eta \in (0, \overline{\eta})$, for each $i \in \{1, \ldots, 6\}$, 0 is the unique value of $t \in (1-\eta, 1+\eta)$ such that $0 = F(t\hat{z}_i)$. Further, $F(t\hat{z}_i)$ takes opposite signs at $t = 1+\eta$ and $t = 1-\eta$. By uniform convergence, for each such η, it holds that for all sufficiently small $\sigma > 0$, and for all $i \in \{1, \ldots, 6\}$, $\tilde{Q}((1+\eta)\hat{z}_i, \sigma)$ and $\tilde{Q}((1-\eta)\hat{z}_i, \sigma)$ have the same signs as $F((1+\eta)\hat{z}_i)$ and $F((1-\eta)\hat{z}_i)$, respectively; thus, for all sufficiently small $\sigma > 0$, there exists $t_i(\sigma)$ in $(1-\eta, 1+\eta)$ such that $\tilde{Q}(t_i(\sigma)\hat{z}_i, \sigma) = 0$, and therefore,

[40] In particular, $F(z) = -z^6(\phi-\rho)\phi^2(\phi+\rho)^3 + z^4\phi[-2\rho^4 - 4\rho^3\phi + 2\rho\phi^3 + \phi^4] + z^2(\rho^2+\rho\phi+\phi^2)^2 - \phi(\rho+\phi)^2$.

$\{z_1, \ldots, z_6\} = \{t_1(\sigma), \ldots, t_6(\sigma)\}$. Relabelling so that $z_i = t_i(\sigma)$, we have $z_i \to \hat{z}_i$ for each $i \in \{1, \ldots, 6\}$. □

We now analyze the follower's SOC.

Lemma B.4. *If $\sigma > 0$ is sufficiently small, then (i) there is no equilibrium in which $\alpha_F = \alpha_{F,2}$, and (ii) for $\alpha_F = \alpha_{F,1}$, (A.7) is satisfied for all real roots of Q among $a_1, \ldots, a_6$.*

Proof. Having ruled out equilibria in which $\alpha_L \in \{\alpha_7, \alpha_8\}$ (when $\sigma > 0$ is small), we show that for $\alpha_F = \alpha_{F,2}$ and for sufficiently small $\sigma > 0$, (A.7) fails for all real roots among $\alpha_1, \ldots, \alpha_6$. By Lemma B.3, each α_i/σ, $i \in \{1, \ldots, 6\}$, converges to a finite nonzero limit $\hat{z}_i$. Hence, for sufficiently small $\sigma > 0$, if $\alpha_L = \alpha_i$, for some $i \in \{1, \ldots, 6\}$ is real, the factor in square brackets in (A.7) is bounded below by

$$\begin{aligned} \alpha_i^2(\phi^2 - \rho^2) + \sigma^2(\phi + \rho) - |\alpha_i \rho|\sigma^2 &\geq \alpha_i^2(\phi^2 - \rho^2) + \sigma^2(\phi + \rho) - |\rho z_i|\sigma^3 \\ &= \sigma^2(z_i^2(\phi^2 - \rho^2) + \phi + \rho - |\rho z_i|\sigma), \end{aligned}$$

where $z_i^2(\phi^2 - \rho^2) + \phi + \rho - |\rho z_i|\sigma \to \hat{z}_i(\rho^2 - \rho^2) + \phi + \rho > 0$. Since $-\alpha_{F,2} > 0$, this implies that (A.7) fails.

For $\alpha_F = \alpha_{F,1}$, the same bound above holds, but since $-\alpha_{F,1} < 0$, (A.7) is satisfied. □

From the proof of Proposition A.3, any equilibrium value of α_L must solve (A.8) (with $\alpha_F = \alpha_{F,1}$) or (A.9) (with $\alpha_F = \alpha_{F,2}$). By Lemma B.4 part (i), α_L must solve (A.8).

We now turn to the leader's SOC.

Lemma B.5. *If $\sigma > 0$ is sufficiently small, then (i) there is no equilibrium in which $\alpha_L \leq 0$, and (ii) if $\alpha_L > 0$ is a real root of (A.19) and $\alpha_F = \alpha_{F,1}$, then (A.6) is satisfied.*

Proof. For part (i), we only need to consider the roots $\alpha_1, \ldots, \alpha_6$, since for sufficiently small σ α_7 and α_8 cannot be part of an equilibrium by Lemma B.2. By Lemma B.4, we further only need to consider $\alpha_F = \alpha_{F,1}$, for which (A.6) becomes

$$\sigma^2 - \alpha_L^2\phi - 2\alpha_L\left(\rho + \phi + \rho\sigma\sqrt{\frac{\sigma^2 + (\alpha_L/\sigma)^2\sigma^2\phi}{\phi + (\alpha_L/\sigma)^2(-(\rho)^2 + (\phi)^2)}}\right) \leq 0. \tag{B.35}$$

Clearly, this is violated if $\alpha_L = 0$. And since $\alpha_L \to 0$ in proportion to σ by Lemma B.3, for small σ, the dominating term is $-2\alpha_L(\rho + \phi)$, which is positive (violating (B.35)) if $\alpha_L < 0$.

For part (ii), we again only need to consider the roots $\alpha_1, \ldots, \alpha_6$, since for sufficiently small σ, α_7 and α_8 are not positive real numbers as they converge to $-\frac{\rho+\phi}{\phi}$. Following the same calculation above, for sufficiently small σ, the left hand side of (A.6) has the same sign as $-2\alpha_L(\rho + \phi)$, which is negative for $\alpha_L > 0$, satisfying (A.6). □

In light of Lemma B.5, we use Lemma B.3 to show that for sufficiently small $\sigma > 0$, there is exactly one positive solution to (A.8), and thus one equilibrium candidate. We establish this in the following lemma:

Lemma B.6. *For sufficiently small $\sigma > 0$, equation* (A.19) *has exactly two positive roots, one solving* (A.8) *and the other solving* (A.9)*.*

Proof. Any (positive) solution to (A.8) or (A.9) must be a (positive) root of (A.19). From the proof of Proposition A.3, (A.19) has *at least* two positive roots, one for each equation (A.8) and (A.9), so it suffices to show that these are the only two positive roots of (A.19). Using the change of variables $z = \alpha_L/\sigma$, $\tilde{Q}(\cdot, \sigma)$ has at least two positive real roots for all sufficiently small σ. But $\tilde{Q}(\cdot, \sigma)$ cannot have more than two positive roots for all sufficiently small σ. To see this, recall that for small σ, α_7 and α_8 are complex or negative, so any positive roots must be among $\alpha_1, \ldots, \alpha_6$. And if there were more than two such positive roots, then by Lemma B.3, F would have more than two nonnegative roots, a contradiction. Mapping back to $\alpha_L = z\sigma$, this implies that (A.19) has exactly two roots for sufficiently small σ, (A.8) and (A.9) each have exactly one. □

From Lemmas B.4, B.5, and B.6, for sufficiently small $\sigma > 0$, there is exactly one pair (α_L, α_F) solving (A.1), (A.4), (A.7), and (A.6), and thus at most one equilibrium. By Remark 2, we can invoke the "converse" part of Proposition A.1, establishing existence.

C Market valuations and shares outstanding

We access the publicly available holdings of the iShares Core S&P Total U.S. Stock Market ETF as of January 2023.[41] The ETF intends to track the whole U.S. stock market by holding a broad set of about 3,380 companies in the U.S. The holdings data includes the market capitalization and stock price of every company. We use this information to calculate shares outstanding.

To examine the relationship between firm size and number of shares outstanding, we regress log shares outstanding on log market value. The results are shown in Table 1. Both shares outstanding and market value are expressed in thousands. The estimate of 0.502 implies that a one percent change in market value yields a half a percent change in shares outstanding; the estimate is significant at a 1% level.

Table 1

	Dependent variable:
	Log Shares Outstanding
Log Market Value	0.502^{***}
	(0.007)
Constant	3.903^{***}
	(0.092)
Observations	3,375
R^2	0.635
Adjusted R^2	0.635
Residual Std. Error	1.113 (df = 3373)
F Statistic	$5{,}876.948^{***}$ (df = 1; 3373)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

In Figure C.1, we segment the log market value variable into 50 bins, take the average log shares outstanding within each bin, generate a scatter plot of the log shares outstanding against log market value, and then plot a regression line. Note that since the regression line shown in the figure is on the binned data, it is different from the regression estimated for Table 1.

[41] https://www.ishares.com/us/products/239724/ishares-core-sp-total-us-stock-market-etf.

Figure C.1