

# Back-Running: Seeking and Hiding Fundamental Information in Order Flows\*

Liyan Yang<sup>†</sup>      Haoxiang Zhu<sup>‡</sup>

October 2015

## Abstract

We study the strategic interaction between fundamental informed trading and order-flow informed trading. In a standard two-period Kyle (1985) model, we add a “back-runner” who observes, *ex post* and potentially with noise, the order flow of the fundamental informed investor in the first period. Learning from order-flow information, the back-runner competes with the fundamental investor in the second period. If order-flow information is accurate, the fundamental investor hides her information by endogenously adding noise into her first-period trade, resulting in a mixed strategy equilibrium. As order-flow information becomes less precise, the equilibrium switches to a pure strategy one. The presence of back-running harms price discovery in the first period but improves it in the second period; the impact of back-running on market liquidity is mixed. A lower cost of acquiring order-flow information reduces the fundamental investor’s incentive to acquire fundamental information.

JEL code: G14, G18

Keywords: order flow, order anticipation, high-frequency trading, price discovery, market liquidity

---

\*For helpful discussions and comments, we thank Henry Cao, Eric Hughson (discussant), Pete Kyle, Doron Levit, Wei Li, Albert Menkveld, Christine Parlour, Ioanid Rosu (discussant), Duane Seppi, Matthew Spiegel, Yajun Wang, Mao Ye, and Bart Yueshen, as well as seminar and conference participants at Finance Theory Group Summer School, Northern Finance Association annual meeting, the 11th Central Bank Workshop on the Microstructure of Financial Markets, Gothenburg University, Lund University, and Stockholm School of Economics. The most updated version of this paper is posted at <http://ssrn.com/abstract=2583915>.

<sup>†</sup>Rotman School of Management, University of Toronto; liyan.yang@rotman.utoronto.ca.

<sup>‡</sup>MIT Sloan School of Management; zhuh@mit.edu.

# 1 Introduction

This paper studies the strategic interaction between fundamental informed trading and order-flow informed trading, as well as its implications for market equilibrium outcomes. By order-flow informed trading, we refer to strategies that begin with no innate trading motives—be it fundamental information or liquidity needs—but instead first learn about other investors’ order flows and then act accordingly.

A primary example of order-flow informed trading is “order anticipation” strategies. According to the Securities and Exchange Commission (2010, p. 54–55), order anticipation “involves any means to ascertain the existence of a large buyer (seller) that does not involve violation of a duty, misappropriation of information, or other misconduct. Examples include the employment of sophisticated pattern recognition software to ascertain from *publicly available information* the existence of a large buyer (seller), or the sophisticated use of orders to ‘ping’ different market centers in an attempt to locate and trade in front of large buyers and sellers [emphasis added].”

Always been controversial,<sup>1</sup> order anticipation strategies have recently attracted intense attention and generated heated debates in the context of high-frequency trading (HFT). In a colorful account of today’s U.S. equity market, Lewis (2014) argues that high-frequency traders observe part of investors’ orders on one exchange and “front-run” the remaining orders before they reach other exchanges.<sup>2</sup> Although most (reluctantly) agree that such strategies are legal in today’s regulatory framework, many investors and regulators have expressed severe concerns that they could harm market quality and long-term investors.<sup>3</sup> For example, in its influential Concept Release on Equity Market Structure, Securities and Exchange Commission (2010, p. 56) asks: “Do commenters believe that order anticipation significantly detracts from market quality and harms institutional investors?”

To address important policy questions like this, we need to first address the more fun-

---

<sup>1</sup>For example, Harris (2003) writes “Order anticipators are *parasitic traders*. They profit only when they can prey on other traders [emphasis in original].”

<sup>2</sup>In its original sense, front-running refers to the illegal practice that a broker executes orders on his own account before executing a customer order. In recent discussions of market structure, this term is often used more broadly to refer to any type of trading strategy that takes advantage of order-flow information, including some academic papers that we will discuss shortly. When discussing these papers we use “front-running” to denote the broader meaning, as the original authors do.

<sup>3</sup>It should be noted that order anticipation strategies are not restricted to HFT; they also apply to other market participants such as broker-dealers who have such a technology. “The successful implementation of this strategy (order anticipation) depends less on low-latency communications than on high-quality pattern-recognition algorithms,” remarks Harris (2013). “The order anticipation problem is thus not really an HFT problem.”

damental questions of market equilibrium. For example, how do order anticipators take advantage of their superior order-flow information of fundamental investors (such as mutual funds and hedge funds)? How do these fundamental investors, in turn, respond to potential information leakage? How does the strategic interaction between these two types of traders affect market equilibrium and associated market quality?

In this paper, we take up this task. Our analysis builds on a simple theoretical model of strategic trading. We start from a standard two-period Kyle (1985) model, which has a “fundamental investor” who is informed of the true asset value, noise traders, and a competitive market maker. The novel part of our model is that we add a “back-runner” who begins with no fundamental information nor liquidity needs, but who receives a signal of the fundamental investor’s order flow after that order is executed by the market maker. We emphasize that it is the fundamental investor’s order flow, not her information,<sup>4</sup> that is partially observed by the back-runner (otherwise, the back-runner would be conceptually indistinguishable from a fundamental investor); and this order-flow information is observed ex post, not ex ante. This important feature is directly motivated by the “publicly available information” part in the SEC’s definition above. In this sense, back-running should be distinguished from “front-running,” which is trading on information about *future* order flows.

The trading mechanism of this market is the same as the standard Kyle (1985) model. In the first period, only the fundamental investor and noise traders submit market orders, which are filled by the market maker at the market-clearing price. Only after the period-1 market clears does the back-runner observe the fundamental investor’s period-1 order flow. In the second period, the fundamental investor, the back-runner, and noise traders all submit market orders, which are filled at a new market-clearing price.

It may be tempting to conjecture that the order flow of the fundamental investor would completely reveal her private information to the back-runner. We show that the fundamental investor can do better than this. In particular, we identify a mixed strategy equilibrium in which the fundamental investor injects an endogenous, normally-distributed noise into her period-1 order flow to hide her information. This garbled order flow, in turn, confuses the back-runner in his inference of the asset fundamental value. This mixed strategy equilibrium obtains if the back-runner’s order-flow signal is sufficiently precise, in which case playing a pure strategy is suboptimal as it costs the fundamental investor too much of her information advantage. As a result, the mixed strategy equilibrium is the unique one among linear

---

<sup>4</sup>Throughout this paper, we will use “her”/“she” to refer to the fundamental investor and use “his”/“he” to refer to the back-runner.

equilibria. In other words, if order-flow information is precise, then randomization is the best defense. This result echoes nicely Stiglitz (2014, p. 8)’s remark on high-frequency trading: “[T]he informed, knowing that there are those who are trying to extract information from observing (directly or indirectly) their actions, will go to great lengths to make it difficult for others to extract such information.”

In contrast, if the back-runner’s order-flow information is sufficiently noisy, he has a hard time inferring the fundamental investor’s information anyway. In this case, the fundamental investor does not need to inject additional noise; she simply plays a pure strategy.

Our analysis points out a new channel—i.e., the amount of noise in the back-runner’s signal—that determines whether a mixed strategy equilibrium or a pure strategy one should prevail in a Kyle-type auction game. In particular, if there is less *exogenous* noise in the back-runner’s signal, the fundamental investor *endogenously* injects more noise into her own period-1 order flow. As a result, as the amount of noise in the back-runner’s signal increases from 0 to  $\infty$ , the unique linear equilibrium switches from a mixed strategy equilibrium, which involves randomization by the fundamental investor in period 1, to a pure strategy equilibrium. Characterizing the endogenous switch between a mixed strategy equilibrium and a pure strategy one is the first, theoretical contribution of our paper.

The second, applied contribution of our paper is to investigate the implications of back-running for market quality and traders’ welfare. The natural benchmark is a standard two-period Kyle model without the back-runner. Our results reveal that back-running has ambiguous impacts on price discovery and market liquidity. In the presence of back-running, the fundamental investor trades less aggressively in the first period, harming price discovery. Price discovery is improved in the second period, however, since the back-runner also has value-relevant information and trades with the fundamental investor.

Market liquidity, measured by the inverse of Kyle’s lambda, is also affected differently in the two periods by back-running. The first-period liquidity is generally improved because the more cautious trading of the fundamental investor weakens the market maker’s adverse selection problem. However, the presence of back-running can either improve or harm the second-period market liquidity. It will harm liquidity if the back-runner’s order-flow signal is sufficiently precise, which means that his trading will inject more private information into the second-period order flow, aggravating the market maker’s adverse selection problem.

Unsurprisingly, taking the two periods together, the fundamental investor suffers from the presence of back-running, but noise traders benefit from it. Because institutional investors like mutual funds, pension funds, and ETFs employ a wide variety of investment strategies,

they may act as either fundamental investors or liquidity (noise) traders, depending on the context. Since the back-runner makes a positive expected profit, the net result is that the other two trader types suffer collectively. We thus confirm the suspicion by regulators that order-flow informed trading tends to harm institutional investors on average.<sup>5</sup>

We consider endogenous information acquisition in an extension of the baseline model. The fundamental investor chooses to acquire a certain amount of fundamental information, and the back-runner simultaneously chooses the precision of order-flow information that he acquires. Our prior results are robust to endogenous information acquisition. We also find that a lower cost of acquiring order-flow information reduces the fundamental investor’s incentive to acquire fundamental information. Moreover, many equilibrium outcomes are invariant to the cost of order-flow information if that cost is within a certain range, which is a novel and unique feature of our framework. This stickiness comes from the property that the back-runner’s profit, as a function of his information precision, may have a kink; it is precisely at the kink that the fundamental investor switches from a pure strategy to a mixed strategy. As is usually the case with corner solutions, a small change in the underlying parameter, in this case the cost of order-flow information, does not affect the optimal precision of order-flow information, hence the stickiness.

The theoretical results of this paper are consistent with recent studies that link the execution cost of institutional investors to the activities of (certain) high-frequency traders (HFTs). In the Swedish stock market, van Kervel and Menkveld (2015) find that as large institutional orders are executed over time, HFTs first “lean against the wind” and provide liquidity, but eventually “go with the wind” and accumulate inventory in the same direction as the institutional order. They further find that implementation shortfall (a measure of execution cost) is substantially higher if HFTs trade in the same direction as the institution, controlling for standard covariates. In the U.S. equity market, Tong (2015) finds that an increase in HFT activities increases the implementation shortfall cost of institutions. Also using U.S. data, Hirschey (2013) documents that aggressive HFT orders predict non-HFT orders in the immediate future, and he interprets this result as consistent with the hypothesis that HFTs make money partly “by identifying patterns in trade and order data that allow them to anticipate and trade ahead of other investors’ order flow”—which is precisely back-running. In the Canadian equity market, Korajczyk and Murphy (2014) find that implementation shortfalls are higher if HFTs take more liquidity, controlling for the level of

---

<sup>5</sup>We caution against interpreting noise traders as retail orders. In today’s U.S. equity market, retail market orders rarely make their way to stocks exchanges; instead, they are filled predominantly by broker-dealers acting as agents or principals.

activities of HFTs and designated market makers. Overall, evidence from these studies supports both the premise of our theory (certain traders like HFTs can learn the direction of informed orders over time) and the result (back-running harms institutional investors). In turn, our model can serve as a theoretical framework for interpreting the evidence.

While our model is made as simple and parsimonious as possible, a number of extensions could be entertained. First, one could allow multiple informed investors and multiple back-runners. We expect that the additional informed investors create a free-riding problem and weaken the incentives to add noise to their period-1 strategies, but the additional back-runners increase the risk of information leakage for informed investors and encourage them to add more noise in the first period. A second possible extension is to combine back-running with information acquisition. Because back-runners harm fundamental investors, we expect fundamental investors to invest less in information acquisition, which would further reduce the informativeness of prices. A third possible extension is to write a dynamic back-running model with more than two periods. A challenge of this extension is history-dependence, that is, strategies in period  $t$  can potentially depend on variables in periods  $1, 2, \dots, t - 1$ . These extensions, while potentially interesting, are unlikely to change our main results, and we leave them for future research.

## 1.1 Related Literature

Our paper contributes to two branches of literature: mixed strategies in trading models and order-flow informed trading.

**Mixed strategies in trading models.** The model of our paper is closest to that of Huddart, Hughes, and Levine (2001), also an extension of the Kyle (1985) model. Motivated by the mandatory disclosure of trades by firm insiders, they assume that the insider's orders are disclosed *publicly and perfectly* after being filled. They show that the only equilibrium in their setting is a mixed strategy one, for otherwise the market maker would always perfectly infer the asset value, preventing any further trading profits of the insider. In their model the mandatory public disclosure unambiguously improves price discovery and market liquidity in each period.

Buffa (2013) studies disclosure of insider trades when the insider is risk-averse. His equilibrium with disclosure also features mixed strategies. In contrast to Huddart, Hughes, and Levine (2001), however, he shows that disclosing insider trades can make the market informationally less efficient. The reason is that without disclosure, the risk-averse insider trades very aggressively in early periods to minimize the uncertainty in profits in the future; but

trade disclosure forces the insider to slow down trading, which reduces the informativeness of prices.

Our results differ from those of Huddart, Hughes, and Levine (2001) and Buffa (2013) in at least two important aspects. First, while their models apply to public disclosure of insider trades, our model is much more suitable to analyze the *private* learning of order-flow information by proprietary firms such as HFTs. As we have shown, private learning of order-flow information has mixed effects on price discovery and market liquidity. Second, a theoretical contribution of our analysis is that we identify the endogenous switching between the mixed strategy equilibrium and the pure strategy one, depending on the precision of the order-flow information. To the best of our knowledge, ours is the first model that presents a pure-mix strategy equilibrium switch among many extensions of Kyle (1985).

Other studies that identify a mixed strategy equilibrium in microstructure models include Back and Baruch (2004) and Baruch and Glosten (2013). In a continuous-time extension of Glosten and Milgrom (1985) model, Back and Baruch (2004) show that there is a mixed strategy equilibrium in which the informed trader’s strategy is a point process with stochastic intensity. Baruch and Glosten (2013) show that “flickering quotes” and “fleeting orders” can arise from a mixed strategy equilibrium in which quote providers repeatedly undercut each other. Neither paper explores the question of trading on order-flow information, and more importantly, neither identifies a switch between pure and mixed strategy equilibria.

**Order-flow information.** Among papers studying order-flow information, the one closest to ours is Madrigal (1996), who also considers a two-period Kyle (1985) model with an insider and a “(non-fundamental) speculator.” The speculator observes part of the period-1 noise trading. Although Madrigal’s model and ours have many similarities, he only considers pure strategy equilibria and does not verify the second-order condition. Consequently, his result misses the mixed strategy equilibrium entirely, and the second-order condition for his pure strategy equilibrium turns out to be violated when the speculator observes a precise signal of noise trading (hence infers the insider’s trade accurately). For this interesting region of parameters, if one were to apply Madrigal’s pure strategy equilibrium, one would conclude incorrectly that the presence of the speculator would improve price discovery in the first period relative to the standard Kyle model. To the contrary, we show that when the back-runner’s information of past order flows is accurate, only the mixed strategy equilibrium exists, and price discovery in the first period becomes *worse* than the standard Kyle model because the informed investor injects noise to her orders. Thus, characterizing the mixed strategy equilibrium matters a great deal for market quality implications.

Li (2014) models high-frequency trading “front-running,” whereby multiple HFTs with various speeds observe the aggregate order flow *ex ante* with noise and front-run it before it reaches the market maker. In his model the informed trader has one trading opportunity and does not counter information leakage by adding noise.

A few other earlier models explore information of liquidity-driven order flows. For example, Cao, Evans, and Lyons (2006) analyze a type of asymmetric information—inventory information—that is unrelated to asset cash flows but still forecasts future prices by forecasting future discount factors. In the two-period model of Bernhardt and Taub (2008), a single informed speculator observes liquidity trades *ex ante* in both periods. In period 1, the speculator front-runs the period-2 liquidity trades and later reverses the position. Attari, Mello, and Ruckes (2005) study a setting in which a strategic investor observes the initial position of an arbitrageur who faces capital constraint. They find that the strategic investor can benefit from the predictable price deviations caused by the arbitrageur’s trades. Brunnermeier and Pedersen (2005) model predatory trading whereby some arbitrageurs take advantage of others that are subject to liquidity shock. Predators first trade in the same direction as distressed investors and later reverse the position. Carlin, Lobo, and Viswanathan (2007) consider repeated interaction among traders and characterize conditions under which they predate each other or provide liquidity to each other. Neither Brunnermeier and Pedersen (2005) nor Carlin, Lobo, and Viswanathan (2007) addresses information asymmetry or price discovery.

Relative to the prior literature on order-flow information, our model differs in two important aspects. First, in most existing models strategic traders explore information about incoming order flows (“front-running”). By contrast, our back-runner learns from *past* order flow, which is much more realistic in today’s market than observing incoming orders.<sup>6</sup> Second, to hide her information, the fundamental investor in our model optimally injects noise into her orders as camouflage. This endogenous response is absent in other studies.

---

<sup>6</sup>A typical example that involves information leakage of incoming order flows is the in-person observation on the trading floor of exchanges. Harris (2003) gives an example of “legal front-running,” in which an “observant trader” on an options trading floor notices the subtle behavior difference of another trader working on the same floor and front-runs him. In today’s electronic markets, gaining information by physical proximity is no longer possible.



## 2 A Model of Back-Running

This section provides a model of back-running, based on the standard Kyle (1985) model. For ease of reference, main model variables are tabulated and explained in Appendix A. All proofs are in Appendix B.

### 2.1 Setup

There are two trading periods,  $t = 1$  and  $t = 2$ . The timeline of the economy is described by Figure 1. A risky asset pays a liquidation value  $v \sim N(p_0, \Sigma_0)$  at the end of period 2, where  $p_0 \in \mathbb{R}$  and  $\Sigma_0 > 0$ . A single “fundamental investor” learns  $v$  at the start of the first period and places market orders  $x_1$  and  $x_2$  at the start of periods 1 and 2, respectively. Noise traders’ net demands in the two periods are  $u_1$  and  $u_2$ , both distributed  $N(0, \sigma_u^2)$ , with  $\sigma_u > 0$ . Random variables  $v$ ,  $u_1$  and  $u_2$  are mutually independent. Asset prices  $p_1$  and  $p_2$  are set by a competitive market maker who observes the total order flow at each period,  $y_1$  and  $y_2$ , and sets the price equal to the posterior expectation of  $v$  given public information.

The main difference from a standard Kyle model is that there is a “back-runner” who can extract private information from public order flows and trades on this private information in period 2. We call this trader a back-runner instead of a “front-runner” to highlight the fact that he observes past, not future, order flow information. Specifically, after seeing the aggregate period-1 order flow  $y_1$ , which is public information in period 2, the back-runner observes a signal about the fundamental investor’s period-1 trades  $x_1$  as follows:

$$s = x_1 + \varepsilon, \tag{1}$$

where  $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ , where  $\sigma_\varepsilon \in [0, \infty]$ , is independent of all other random variables ( $v$ ,  $u_1$  and  $u_2$ ). (Since  $y_1$  is public, observing a signal of  $u_1$  would be equivalent to observing a signal of  $x_1$ .) Parameter  $\sigma_\varepsilon$  controls the information quality of the signal  $s$ —a larger  $\sigma_\varepsilon$  means less accurate information about  $x_1$ . In particular, we deliberately allow  $\sigma_\varepsilon$  to take values of 0 and  $\infty$ , which respectively corresponds to the case in which  $s$  perfectly reveals  $x_1$  and the case in which  $s$  reveals nothing about  $x_1$ .

After receiving the signal  $s$ , the back-runner places a market order  $d_2$  in period 2. As a result, the market maker receives an aggregate order flow

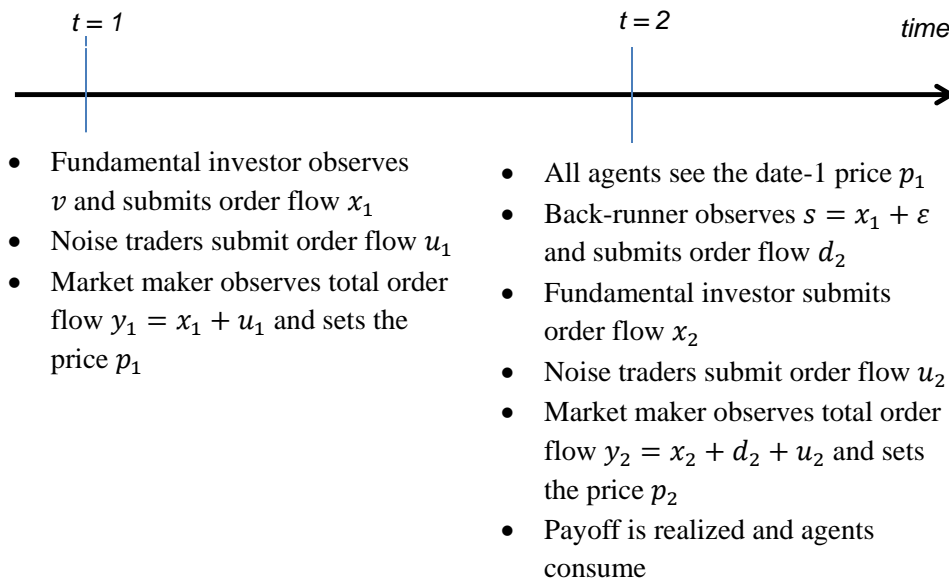
$$y_2 = x_2 + d_2 + u_2. \tag{2}$$

Of course, in period 1, the aggregate order flow is

$$y_1 = x_1 + u_1, \tag{3}$$

since during period 1 the back-runner has no private information and does not send any

Figure 1: Model Timeline



order. The weak-form-efficiency pricing rule of the market maker implies

$$p_1 = E(v|y_1) \text{ and } p_2 = E(v|y_1, y_2). \quad (4)$$

At the end of period 2, all agents receive their payoffs and consume, and the economy ends.

As discussed in the introduction, a practical interpretation of this back-runner is that he uses advanced technology and processes public information better than the general public, represented by the market maker in the model. Such superior ability of processing public information has long been recognized in the literature. For example, Kim and Verrecchia (1994) argue that savvy market participants, such as asset managers and analysts, can process information better than the market by converting a firm's noisy public signals (e.g., earnings announcements) into more accurate information. Engelberg, Reed, and Ringgenberg (2012) show that a significant portion of short sellers profitability actually comes from their skills in analyzing public information. Our objective is to explore how the presence of the back-runner—a trader who has superior skills in processing public trading data to extract the patterns of trades—affects the trading strategies of the fundamental investor (such as pension funds, mutual funds, or hedge funds) as well as the resulting market equilibrium outcomes.

## 2.2 Equilibrium Definitions

A *perfect Bayesian equilibrium* of the trading game is given by a strategy profile

$$\{x_1^*(v), x_2^*(v, p_1, x_1), d_2^*(s, p_1), p_1^*(y_1), p_2^*(y_1, y_2)\},$$

that satisfies:

1. Profit maximization:

$$x_2^* \in \arg \max_{x_2} E[x_2(v - p_2) | v, p_1, x_1],$$

$$d_2^* \in \arg \max_{d_2} E[d_2(v - p_2) | s, p_1],$$

$$\text{and } x_1^* \in \arg \max_{x_1} E[x_1(v - p_1) + x_2^*(v - p_2) | v].$$

2. Market efficiency:  $p_1$  and  $p_2$  are determined according to equation (4).

We consider linear equilibria, i.e., the trading strategies and pricing functions are linear. Specifically, a *linear equilibrium* is defined as a perfect Bayesian equilibrium in which there exist constants

$$(\beta_{v,1}, \beta_{v,2}, \beta_{x_1}, \beta_{y_1}, \delta_s, \delta_{y_1}, \lambda_1, \lambda_2) \in \mathbb{R}^8 \quad \text{and} \quad \sigma_z \geq 0,$$

such that

$$x_1 = \beta_{v,1}(v - p_0) + z \text{ with } z \sim N(0, \sigma_z^2), \quad (5)$$

$$x_2 = \beta_{v,2}(v - p_1) - \beta_{x_1}x_1 + \beta_{y_1}y_1, \quad (6)$$

$$d_2 = \delta_s s - \delta_{y_1}y_1, \quad (7)$$

$$p_1 = p_0 + \lambda_1 y_1 \text{ with } y_1 = x_1 + u_1, \quad (8)$$

$$p_2 = p_1 + \lambda_2 y_2 \text{ with } y_2 = x_2 + d_2 + u_2, \quad (9)$$

where  $z$  is independent of all other random variables  $(v, u_1, u_2, \varepsilon)$ .

Equations (5)–(9) are intuitive. Equations (5)–(7) simply say that the fundamental investor and the back-runner trade on their information advantage. Importantly, our specification (5) allows the fundamental investor to play a mixed strategy in period 1. We have followed Huddart, Hughes, and Levine (2001) and restricted attention to normally distributed  $z$  in order to maintain tractability. If  $\sigma_z = 0$ , the fundamental investor plays a pure strategy in period 1, and we refer to the resulting linear equilibrium as a *pure strategy equilibrium*. If  $\sigma_z > 0$ , the fundamental investor plays a mixed strategy in period 1, and we refer to the resulting linear equilibrium as a *mixed strategy equilibrium*. As we show shortly, by adding noise into her orders, the fundamental investor limits the back-runner's ability to infer  $x_1$  and hence  $v$ . To an outside observer, the endogenously added noise  $z$  may look like exogenous noise trading.

Although in principle the fundamental investor and the back-runner can play mixed strategies in period 2, we show later that using mixed strategies in period 2 is suboptimal in equilibrium. Thus, the linear period-2 trading strategies specified in equations (6) and (7) are without loss of generality. They are also the most general linear form, as each equation spans the information set of the relevant trader in the relevant period. Note that at this stage we do not require that  $\beta_{x_1}$ ,  $\beta_{y_1}$ ,  $\delta_s$  or  $\delta_{y_1}$  be positive, although in equilibrium they will be positive.

Equation (6) has three terms. The first term  $\beta_{v,2}(v - p_1)$  captures how aggressively the fundamental investor trades on her information advantage about  $v$ . The other two terms  $-\beta_{x_1}x_1$  and  $\beta_{y_1}y_1$  say that the fundamental investor potentially adjusts her period-2 market order by using lagged information  $x_1$  and  $y_1$ . Because the back-runner generally uses  $y_1$  and his signal  $s$  about  $x_1$  to form his period-2 order (see equation (7)), the fundamental investor takes advantage of this predictive pattern by using  $x_1$  and  $y_1$  in her period-2 order as well.

In equilibrium characterized later, the conjectured strategy in equation (7) can also be written alternatively as:

$$d_2 = \alpha [E(v|s, y_1) - E(v|y_1)], \quad (10)$$

for some constant  $\alpha > 0$  (see Appendix B.1 for a proof). That is, the back-runner's order is proportional to his information advantage relative to the market maker's. By the joint normality of  $s$  and  $y_1$ , this alternative form implies that  $d_2$  is linear in  $s$  and  $y_1$ . We nonetheless start with (7) because it is the most general and does not impose any structure as (10) does. We start with equation (6) for a similar reason.

The pricing equations (8) and (9) state that the price in each period is equal to the expected value of  $v$  before trading, adjusted by the information carried by the new order flow. Although the conjectured  $p_2$  may in principle depend on  $y_1$ , in equilibrium  $p_1$  already incorporates all information of  $y_1$ .<sup>7</sup> Thus, we can start with (9).

## 2.3 Equilibrium Derivation

We now derive by backward induction all possible linear equilibria. Along the derivations, we will see that the distinction between pure strategy and mixed strategy equilibria lies only in the conditions characterizing the fundamental investor's period-1 decision. Explicit statements of the equilibria and their properties are presented in the next subsection.

---

<sup>7</sup>Strictly speaking the most general form is  $p_2 = p_1 + \lambda_2 [y_2 - E(y_2|y_1)]$ . But in equilibrium we can show that  $E(y_2|y_1) = 0$ , so the more general form reduces to (9).

**Fundamental investor's date-2 problem.** In period 2, the fundamental investor has information  $\{v, p_1, x_1\}$ . Given  $\lambda_1 \neq 0$ , which holds in equilibrium, the fundamental investor can infer  $y_1$  from  $p_1$  by equation (8). Using equations (7) and (9), we can compute

$$E[x_2(v - p_2) | v, p_1, x_1] = -\lambda_2 x_2^2 + [v - p_1 - \lambda_2(\delta_s x_1 - \delta_{y_1} y_1)] x_2. \quad (11)$$

Taking the first-order-condition (FOC) results in the solution as follows:

$$x_2 = \frac{v - p_1}{2\lambda_2} - \frac{\delta_s}{2} x_1 + \frac{\delta_{y_1}}{2} y_1. \quad (12)$$

The second-order-condition (SOC) is<sup>8</sup>

$$\lambda_2 > 0. \quad (13)$$

Equation (12) also implies that the fundamental investor optimally chooses to play a pure strategy in equilibrium, which verifies our conjectured pure strategy specification (6).

Comparing equation (12) with the conjectured strategy (6), we have

$$\beta_{v,2} = \frac{1}{2\lambda_2}, \beta_{x_1} = \frac{\delta_s}{2} \text{ and } \beta_{y_1} = \frac{\delta_{y_1}}{2}. \quad (14)$$

Let  $\pi_{F,2} = x_2(v - p_2)$  denote the fundamental investor's profit that is directly attributable to her period-2 trade. Inserting (12) into (11) yields

$$E(\pi_{F,2} | v, p_1, x_1) = \frac{[v - p_1 - \lambda_2(\delta_s x_1 - \delta_{y_1} y_1)]^2}{4\lambda_2}. \quad (15)$$

**Back-runner's date-2 problem.** In period 2, the back-runner chooses  $d_2$  to maximize  $E(\pi_{B,2} | s, p_1)$ , where

$$\pi_{B,2} = d_2(v - p_2). \quad (16)$$

Using (6) and (9), we can compute the FOC, which delivers

$$d_2 = \frac{(1 - \lambda_2 \beta_{v,2}) E(v - p_1 | s, y_1) - \lambda_2 \beta_{y_1} y_1 + \lambda_2 \beta_{x_1} E(x_1 | s, y_1)}{2\lambda_2}. \quad (17)$$

The SOC is still  $\lambda_2 > 0$ , as given by (13) in the fundamental investor's problem. Again, equation (17) means that the back-runner optimally chooses to play a pure strategy in a linear equilibrium.

We then employ the projection theorem and equations (1), (3), and (5) to find out the expressions of  $E(v - p_1 | s, y_1)$  and  $E(x_1 | s, y_1)$ , which are in turn inserted into (17) to express  $d_2$  as a linear function of  $s$  and  $y_1$ . Finally, we compare this expression with the conjectured

---

<sup>8</sup>The SOC cannot be  $\lambda_2 = 0$ , because otherwise, we have  $p_2 = p_1 = p_0 + \lambda_1 y_1 = p_0 + \lambda_1(x_1 + u_1)$ , and thus  $E(p_2 | v) = p_0 + \lambda_1 x_1$ , which means that the fundamental investor can choose  $x_1$  and  $x_2$  to make infinite profit in period 2. Thus, in any linear equilibrium, we must have  $\lambda_2 > 0$ .

strategy (7) to arrive at the following two equations:

$$\begin{aligned}\delta_s &= \frac{\left[ (1 - \lambda_2 \beta_{v,2}) \frac{\beta_{v,1} \Sigma_0}{\beta_{v,1}^2 \Sigma_0 + \sigma_z^2} + \lambda_2 \beta_{x_1} \right] \frac{\sigma_\varepsilon^{-2}}{(\beta_{v,1}^2 \Sigma_0 + \sigma_z^2)^{-1} + \sigma_\varepsilon^{-2} + \sigma_u^{-2}}}{2\lambda_2}, \\ \delta_{y_1} &= -\delta_s \frac{\sigma_u^{-2}}{\sigma_\varepsilon^{-2}} + \frac{\lambda_1 (1 - \lambda_2 \beta_{v,2}) + \lambda_2 \beta_{y_1}}{2\lambda_2}.\end{aligned}$$

Using (14), we can further simplify the above two equations as follows:

$$\delta_s = \frac{\frac{\sigma_\varepsilon^{-2}}{(\beta_{v,1}^2 \Sigma_0 + \sigma_z^2)^{-1} + \sigma_\varepsilon^{-2} + \sigma_u^{-2}}}{4 - \frac{\sigma_\varepsilon^{-2}}{(\beta_{v,1}^2 \Sigma_0 + \sigma_z^2)^{-1} + \sigma_\varepsilon^{-2} + \sigma_u^{-2}}} \frac{\beta_{v,1} \Sigma_0}{\lambda_2 (\beta_{v,1}^2 \Sigma_0 + \sigma_z^2)}, \quad (18)$$

$$\delta_{y_1} = \frac{\lambda_1}{3\lambda_2} - \delta_s \frac{4\sigma_\varepsilon^2}{3\sigma_u^2}. \quad (19)$$

**Market maker's decisions.** In period 1, the market maker sees the aggregate order flow  $y_1$  and sets  $p_1 = E(v|y_1)$ . Accordingly, we have  $\lambda_1 = \frac{Cov(v, y_1)}{Var(y_1)}$ . By equation (5) and the projection theorem, we can compute

$$\lambda_1 = \frac{Cov(v, y_1)}{Var(y_1)} = \frac{\beta_{v,1} \Sigma_0}{\beta_{v,1}^2 \Sigma_0 + \sigma_z^2 + \sigma_u^2}. \quad (20)$$

Similarly, in period 2, the market maker sees  $\{y_1, y_2\}$  and sets  $p_2 = E(v|y_1, y_2)$ . By equations (6), (7), and (14) and applying the projection theorem, we have

$$\begin{aligned}\lambda_2 &= \frac{Cov(v, y_2|y_1)}{Var(y_2|y_1)} \\ &= \frac{\left( \frac{1}{2\lambda_2} + \frac{\delta_s}{2} \beta_{v,1} \right) \Sigma_0 - \frac{\beta_{v,1} \Sigma_0 \left[ \left( \frac{1}{2\lambda_2} + \frac{\delta_s}{2} \beta_{v,1} \right) \beta_{v,1} \Sigma_0 + \frac{\delta_s}{2} \sigma_z^2 \right]}{\beta_{v,1}^2 \Sigma_0 + \sigma_z^2 + \sigma_u^2}}{\left( \frac{1}{2\lambda_2} + \frac{\delta_s}{2} \beta_{v,1} \right)^2 \Sigma_0 + \frac{\delta_s^2}{4} \sigma_z^2 + \delta_s \sigma_\varepsilon^2 + \sigma_u^2 - \frac{\left[ \left( \frac{1}{2\lambda_2} + \frac{\delta_s}{2} \beta_{v,1} \right) \beta_{v,1} \Sigma_0 + \frac{\delta_s}{2} \sigma_z^2 \right]^2}{\beta_{v,1}^2 \Sigma_0 + \sigma_z^2 + \sigma_u^2}}.\end{aligned} \quad (21)$$

**Fundamental investor's date-1 problem.** We denote by  $\pi_{F,1} = x_1 (v - p_1)$  the fundamental investor's profit that comes from her period-1 trade. In period 1, the fundamental investor chooses  $x_1$  to maximize

$$E(\pi_{F,1} + \pi_{F,2}|v) = x_1 E(v - p_1|v) + E \left[ \frac{[v - p_1 - \lambda_2 (\delta_s x_1 - \delta_{y_1} y_1)]^2}{4\lambda_2} \middle| v \right],$$

where the equality follows from equation (15). Using (8), we can further express  $E(\pi_{F,1} + \pi_{F,2}|v)$  as follows:

$$\begin{aligned}
E(\pi_{F,1} + \pi_{F,2}|v) &= - \left[ \lambda_1 - \frac{(\lambda_1 + \lambda_2 \delta_s - \lambda_2 \delta_{y_1})^2}{4\lambda_2} \right] x_1^2 \\
&\quad + \left[ 1 - \frac{\lambda_1 + \lambda_2 \delta_s - \lambda_2 \delta_{y_1}}{2\lambda_2} \right] (v - p_0) x_1 \\
&\quad + \frac{(v - p_0)^2 + \sigma_u^2 (\lambda_1 - \lambda_2 \delta_{y_1})^2}{4\lambda_2}. \tag{22}
\end{aligned}$$

Depending on whether the fundamental investor plays a mixed or a pure strategy (i.e., whether  $\sigma_z$  is equal to 0), we have two cases:

Case 1. Mixed Strategy ( $\sigma_z > 0$ )

For a mixed strategy to sustain in equilibrium, the fundamental investor has to be indifferent between any realized pure strategy. This in turn means that coefficients on  $x_1^2$  and  $x_1$  in (22) are equal to zero, that is,

$$\lambda_1 - \frac{(\lambda_1 + \lambda_2 \delta_s - \lambda_2 \delta_{y_1})^2}{4\lambda_2} = 0 \text{ and } 1 - \frac{\lambda_1 + \lambda_2 \delta_s - \lambda_2 \delta_{y_1}}{2\lambda_2} = 0.$$

These two equations, together with equation (19), imply

$$\lambda_1 = \lambda_2 \text{ and } \delta_s = \frac{\frac{4}{3}}{1 + \frac{4\sigma_z^2}{3\sigma_u^2}}. \tag{23}$$

Case 2. Pure Strategy ( $\sigma_z = 0$ )

When the fundamental investor plays a pure strategy,  $z = 0$  (and  $\sigma_z = 0$ ) in the conjectured strategy, and thus (5) degenerates to  $x_1 = \beta_{v,1}(v - p_0)$ . The FOC of (22) yields

$$x_1 = \frac{\left(1 - \frac{\lambda_1 + \lambda_2 \delta_s - \lambda_2 \delta_{y_1}}{2\lambda_2}\right)}{2 \left[ \lambda_1 - \frac{(\lambda_1 + \lambda_2 \delta_s - \lambda_2 \delta_{y_1})^2}{4\lambda_2} \right]} (v - p_0),$$

which, compared with the conjectured pure strategy  $x_1 = \beta_{v,1}(v - p_0)$ , implies

$$\beta_{v,1} = \frac{1 - \frac{\lambda_1 + \lambda_2 \delta_s - \lambda_2 \delta_{y_1}}{2\lambda_2}}{2 \left[ \lambda_1 - \frac{(\lambda_1 + \lambda_2 \delta_s - \lambda_2 \delta_{y_1})^2}{4\lambda_2} \right]}. \tag{24}$$

The SOC is

$$\lambda_1 - \frac{(\lambda_1 + \lambda_2 \delta_s - \lambda_2 \delta_{y_1})^2}{4\lambda_2} > 0. \tag{25}$$

## 2.4 Equilibrium Characterization and Properties

A mixed strategy equilibrium is characterized by equations (14), (18), (19), (20), (21), and (23), together with one SOC,  $\lambda_2 > 0$  (given by (13)). These conditions jointly define a system that determine nine unknowns,  $\sigma_z, \beta_{v,1}, \beta_{v,2}, \beta_{x_1}, \beta_{y_1}, \delta_s, \delta_{y_1}, \lambda_1$ , and  $\lambda_2$ . The following proposition formally characterizes a linear mixed strategy equilibrium.

**Proposition 1** (Mixed Strategy Equilibrium). *Let  $\gamma \equiv \frac{\sigma_\varepsilon}{\sigma_u}$ . If and only if  $\gamma < \frac{\sqrt{\sqrt{17}-4}}{2} \approx 0.175$ , there exists a linear mixed strategy equilibrium, and it is specified by equations (5)–(9), where*

$$\begin{aligned}\sigma_z &= \sigma_u \sqrt{\frac{(1+4\gamma^2)(1-32\gamma^2-16\gamma^4)}{(3+4\gamma^2)(13+40\gamma^2+16\gamma^4)}}, \\ \beta_{v,1} &= \frac{\sigma_u}{\sqrt{\Sigma_0}} \sqrt{\frac{1-4\gamma^2-(3+4\gamma^2)\frac{\sigma_z^2}{\sigma_u^2}}{3+4\gamma^2}}, \\ \lambda_1 &= \lambda_2 = \frac{\beta_{v,1}\Sigma_0}{\beta_{v,1}^2\Sigma_0 + \sigma_z^2 + \sigma_u^2} > 0, \\ \beta_{v,2} &= \frac{1}{2\lambda_2}, \delta_s = \frac{4}{3+4\gamma^2}, \delta_{y_1} = \frac{1-4\gamma^2\delta_s}{3}, \\ \beta_{x_1} &= \frac{\delta_s}{2} \text{ and } \beta_{y_1} = \frac{\delta_{y_1}}{2}.\end{aligned}$$

*When it exists, this equilibrium is the unique linear mixed strategy equilibrium.*

To illustrate the intuition of the equilibrium strategies, it is useful to explicitly decompose  $d_2$  as follows:

$$\begin{aligned}d_2 &= \delta_s(x_1 + \varepsilon) - \delta_{y_1}(x_1 + u_1) = (\delta_s - \delta_{y_1})x_1 + \delta_s\varepsilon - \delta_{y_1}u_1 \\ &= x_1 + \delta_s\varepsilon - \delta_{y_1}u_1 = \beta_{v,1}(v - p_0) + (\delta_s\varepsilon + z) - \delta_{y_1}u_1,\end{aligned}\tag{26}$$

where we have used the fact that  $\delta_s - \delta_{y_1} = 1$  in equilibrium. Equation (26) says that the back-runner's order  $d_2$  consists of three parts. The first part is the fundamental investor's order  $x_1$  in period 1. The second part,  $\delta_s\varepsilon + z$ , reflects the imprecision of his signal, caused by both the exogenous noise  $\varepsilon$  in his signal-processing technology and the endogenous noise  $z$  added by the fundamental investor. The third part,  $-\delta_{y_1}u_1$ , says that the back-runner trades against the period-1 noise demand  $u_1$ , which is profitable in expectation because the back-runner can tell  $x_1$  from  $u_1$  better than the market maker does. Note that equation (26) should be read as purely as a decomposition but not the strategy used by the back-runner, as  $v, \varepsilon, z$ , and  $u_1$  are not separately observable to him.

In the mixed strategy equilibrium,  $x_1, x_2, v - p_0$ , and  $v - p_1$  need not always have the same sign. For example, if  $v > p_0$  but  $z$  is sufficiently negative, the fundamental investor



ends up selling in period 1 (with  $x_1 < 0$ ), before purchasing in period 2 ( $x_2 > 0$ ). While such a pattern in the data may raise red flags of potential “manipulation” (trading in the opposite direction of the true intention), it could simply be part of an optimal execution strategy that involves randomizing.

Proposition 1 reveals that a mixed strategy equilibrium exists if and only if the size  $\sigma_\varepsilon$  of the noise in the back-runner’s signal is sufficiently small relative to  $\sigma_u$ . This result is natural and intuitive. A small  $\sigma_\varepsilon$  implies that the back-runner can observe  $x_1$  relatively accurately. The back-runner will in turn compete aggressively with the fundamental investor in period 2, which reduces the fundamental investor’s profit substantially. Worried about information leakage, the fundamental investor optimally plays a mixed strategy in period 1 by injecting an endogenous noise  $z$  into her order  $x_1$ , with  $\sigma_z$  uniquely determined in equilibrium. This garbled  $x_1$  limits the back-runner’s ability to learn about  $v$ . In other words, if the back-runner’s order-parsing technology is accurate enough, randomization is the fundamental investor’s best camouflage.

Conversely, if  $\sigma_\varepsilon$  is sufficiently large already, the fundamental investor retains much of her information advantage, and further obscuring  $x_1$  is unnecessary. In this case a linear pure strategy equilibrium, characterized shortly, would be more natural.

Looked another way, all else equal, the mixed strategy equilibrium obtains if and only if  $\sigma_u$  is sufficiently *large*. Traditional Kyle-type models would not generate this result, as noise trading provides camouflage for the informed investor. In our model, however, a large  $\sigma_u$  confuses only the market maker, not the back-runner. Thus, more noise trading implies a higher profit for the fundamental investor and hence a stronger incentive to retain her proprietary information by adding noise. A natural implication of this observation is that the exogenous noise  $\sigma_u$  reinforces the endogenous noise  $\sigma_z$ .

The threshold value for the existence of the mixed strategy equilibrium is  $\sigma_\varepsilon/\sigma_u \approx 17.5\%$ ; whether it is large or small is an empirical question. We believe that in a setting with more than two periods or a two-period setting with  $Var(u_1) < Var(u_2)$ ,<sup>9</sup> this threshold value is likely to increase. The intuition is that in these extended settings, the fundamental investor has more time or a larger market to trade on her information and hence has a stronger incentive to prevent information leakage.

Now we turn to pure strategy equilibria. In a pure strategy equilibrium, we have  $\sigma_z = 0$ . This type of equilibrium is characterized by equations (14), (18), (19), (20), (21), and (24),

---

<sup>9</sup>The setting with  $Var(u_1) < Var(u_2)$  is natural if the second period corresponds to a longer trading interval (e.g., one hour versus one minute) or a larger market (multiple exchanges versus a single exchange).

together with two SOC's, (13) and (25). These conditions jointly define a system that determine eight unknowns,  $\beta_{v,1}, \beta_{v,2}, \beta_{x_1}, \beta_{y_1}, \delta_s, \delta_{y_1}, \lambda_1$ , and  $\lambda_2$ . The following proposition formally characterizes a linear pure strategy equilibrium.

**Proposition 2** (Pure Strategy Equilibrium). *A linear pure strategy equilibrium is characterized by equations (5)–(9) with  $\sigma_z = 0$  as well as the following two conditions on  $\beta_{v,1} \in \left(0, \frac{\sigma_u}{\sqrt{\Sigma_0}}\right]$ :*

(1)  $\beta_{v,1}^2$  solves the 7<sup>th</sup> order polynomial:

$$f(\beta_{v,1}^2) = A_7\beta_{v,1}^{14} + A_6\beta_{v,1}^{12} + A_5\beta_{v,1}^{10} + A_4\beta_{v,1}^8 + A_3\beta_{v,1}^6 + A_2\beta_{v,1}^4 + A_1\beta_{v,1}^2 + A_0 = 0,$$

where the coefficients  $A$ 's are given by equations (B14)–(B21) in Appendix B; and

(2) The following SOC (i.e., (25)) is satisfied:

$$\lambda_1 - \frac{(\lambda_1 + \lambda_2\delta_s - \lambda_2\delta_{y_1})^2}{4\lambda_2} > 0,$$

where  $\lambda_1, \lambda_2, \delta_s$ , and  $\delta_{y_1}$  are expressed as functions of  $\beta_{v,1}$  as follows:

$$\begin{aligned} \lambda_1 &= \frac{\beta_{v,1}\Sigma_0}{\beta_{v,1}^2\Sigma_0 + \sigma_u^2}, \\ \lambda_2 &= \sqrt{\frac{\Sigma_0(2\sigma_u^4 + 4\sigma_\varepsilon^4 + 5\sigma_u^2\sigma_\varepsilon^2)\Sigma_0^2\beta_{v,1}^4 + (8\sigma_u^2\sigma_\varepsilon^4 + 5\sigma_u^4\sigma_\varepsilon^2)\Sigma_0\beta_{v,1}^2 + 4\sigma_u^4\sigma_\varepsilon^4}{(\beta_{v,1}^2\Sigma_0 + \sigma_u^2)(3\sigma_u^2\Sigma_0\beta_{v,1}^2 + 4\sigma_\varepsilon^2\Sigma_0\beta_{v,1}^2 + 4\sigma_u^2\sigma_\varepsilon^2)^2}}, \\ \delta_s &= \frac{\beta_{v,1}\Sigma_0\sigma_u^2}{\lambda_2[(3\sigma_u^2 + 4\sigma_\varepsilon^2)\Sigma_0\beta_{v,1}^2 + 4\sigma_u^2\sigma_\varepsilon^2]}, \\ \delta_{y_1} &= \frac{\lambda_1}{3\lambda_2} - \frac{4\sigma_\varepsilon^2}{3\sigma_u^2}\delta_s. \end{aligned}$$

Propositions 1 and 2 respectively characterize mixed strategy and pure strategy equilibria. The following proposition provides sufficient conditions under which either equilibrium prevails as the unique one among linear equilibria.

**Proposition 3** (Mixed vs. Pure Strategy Equilibria). *If the back-runner has a sufficiently precise signal about  $x_1$  (i.e.,  $\sigma_\varepsilon^2$  is sufficiently small), there is no pure strategy equilibrium, and the unique linear strategy equilibrium is the mixed strategy equilibrium characterized by Proposition 1. If the back-runner has a sufficiently noisy signal about  $x_1$  (i.e.,  $\sigma_\varepsilon^2$  is sufficiently large), there is no mixed strategy equilibrium, and there is a unique pure strategy equilibrium characterized by Proposition 2.*

Given Proposition 1 and the discussion of its properties, the mixed strategy part of Proposition 3 is relatively straightforward. The existence of a pure strategy equilibrium for a sufficiently large  $\sigma_\varepsilon$  is also natural, as in this case the back-runner's signal has little

information and does not deter the fundamental investor from using a pure strategy. In fact, as  $\sigma_\varepsilon \uparrow \infty$  our setting degenerates to a standard two-period Kyle (1985) setting, and the unique linear equilibrium in our model indeed converges to the pure strategy equilibrium of Kyle (1985). This result is shown in the following corollary.

**Corollary 1.** *As  $\sigma_\varepsilon \rightarrow \infty$ , the linear equilibrium in the two-period economy with a back-runner converges to the linear equilibrium in the standard two-period Kyle model.*

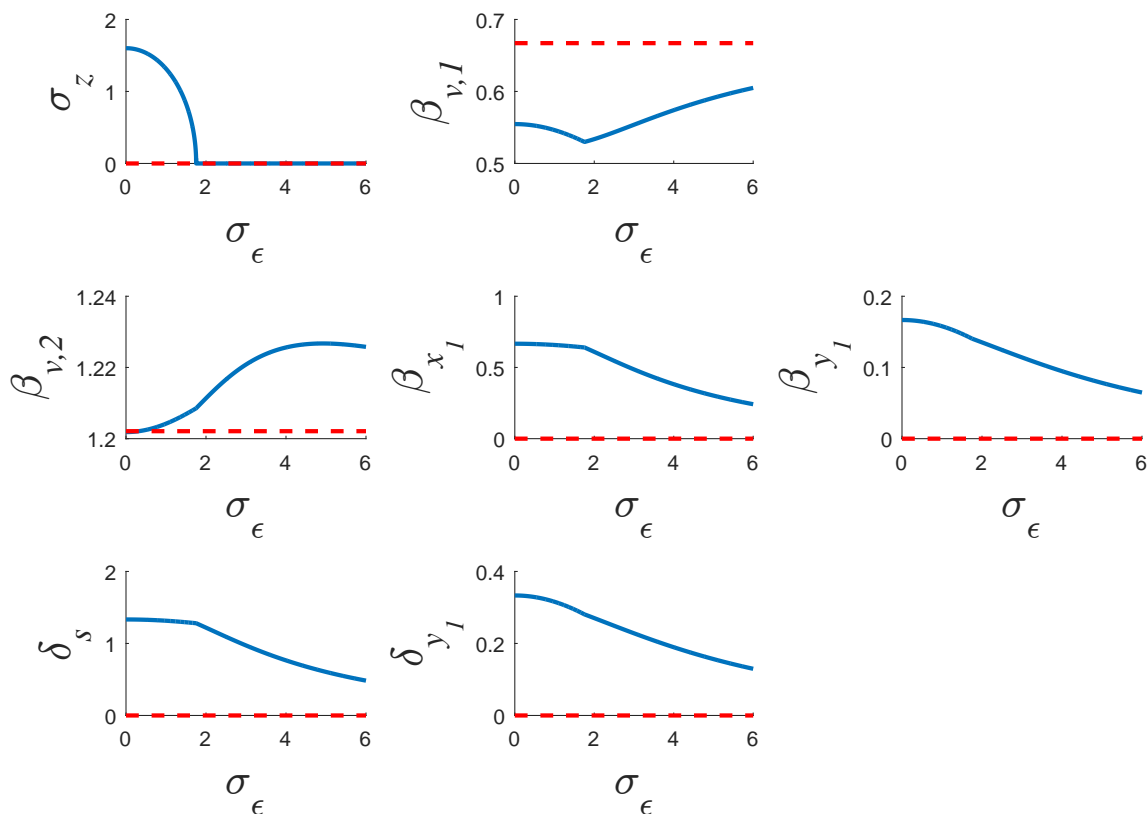
Proposition 3 analytically proves the uniqueness of a linear equilibrium only for sufficiently small or sufficiently large values of  $\sigma_\varepsilon^2$ . It would be desirable to generalize this uniqueness result to any value of  $\sigma_\varepsilon$ , but we have not managed to do so due to the complexity of the 7<sup>th</sup> order polynomial characterizing a pure strategy equilibrium in Proposition 2. In particular, given Proposition 1, a reasonable conjecture is that the boundary between pure and mix strategy equilibria is at  $\frac{\sigma_\varepsilon}{\sigma_u} = \frac{\sqrt{\sqrt{17}-4}}{2}$ . This conjecture, albeit not formally proven, seems to hold numerically. That is, if  $\frac{\sigma_\varepsilon}{\sigma_u} < \frac{\sqrt{\sqrt{17}-4}}{2}$ , only a mixed strategy linear equilibrium exists, and if  $\frac{\sigma_\varepsilon}{\sigma_u} \geq \frac{\sqrt{\sqrt{17}-4}}{2}$ , only a pure strategy linear equilibrium exists. Either way, the linear equilibrium seems unique for all parameter values.

Propositions 1 and 2 suggest the following three-step algorithm to compute all possible linear equilibria:

- Step 1: Compute all the positive root of the polynomial  $f(\beta_{v,1}^2) = 0$  in Proposition 2. Retain the values of  $\beta_{v,1} \in \left(0, \frac{\sigma_u}{\sqrt{\Sigma_0}}\right]$  to serve as candidates for a pure strategy equilibrium.
- Step 2: For each  $\beta_{v,1}$  retained in Step 1, check whether the SOC in Proposition 2 is satisfied. If yes, then it is a pure strategy equilibrium; otherwise, it is not.
- Step 3: If  $\frac{\sigma_\varepsilon}{\sigma_u} < \frac{\sqrt{\sqrt{17}-4}}{2}$ , employ Proposition 1 to compute a mixed strategy equilibrium.

Figure 2 plots in solid lines the equilibrium trading strategies of the fundamental investor and the back-runner as functions of  $\sigma_\varepsilon$ , where we set  $\sigma_u = 10$  and  $\Sigma_0 = 100$ . As a comparison, the dashed lines show corresponding strategies in the standard two-period Kyle model without the back-runner. The first panel confirms that  $\sigma_z > 0$  if and only if  $\sigma_\varepsilon < 0.175\sigma_u = 1.75$ . Also, when  $\sigma_\varepsilon < 1.75$ , the equilibrium value of  $\sigma_z$  decreases with  $\sigma_\varepsilon$ . That is, when there is more *exogenous* noise in the back-runner's signal, the fundamental investor *endogenously* injects less noise into her own period-1 orders. This result points out a new channel—i.e., the amount of noise in the back-runner's signal—that determines whether a mixed strategy equilibrium or a pure strategy one should prevail in a Kyle-type auction game.

Figure 2: Implications for Trading Strategies



This figure plots the implications of back-running for trading strategies of the fundamental investor and the back-runner. In each panel, the blue solid line plots the value in the equilibrium of this paper, and the dashed red line plots the value in a standard Kyle economy (i.e.,  $\sigma_\epsilon = \infty$ ). The horizontal axis in each panel is the standard deviation  $\sigma_\epsilon$  of the noise in the back-runner's private signal about the fundamental investor's past order. The other parameters are:  $\sigma_u = 10$  and  $\Sigma_0 = 100$ .

The other panels in Figure 2 are also intuitive. For instance,  $\beta_{v,1}$  decreases with  $\sigma_\epsilon$  in the mixed strategy regime, but increases with  $\sigma_\epsilon$  in the pure strategy regime. This is because in the mixed strategy regime, as  $\sigma_\epsilon$  increases, the fundamental investor adds less noise  $z$  to her order; to avoid revealing too much information to the back-runner, she trades less aggressively on  $v$  in period 1. In contrast, in the pure strategy equilibrium, as  $\sigma_\epsilon$  increases, the fundamental investor knows that the back-runner will learn less from her order due to the increased exogenous noise  $\epsilon$ , and so she can afford to trade more aggressively in period 1. The intensity  $\beta_{v,1}$  with order-flow information is smaller than its counterpart without order-flow information in a standard Kyle model.

An interesting observation is that  $\beta_{v,2}$  is hump-shaped in  $\sigma_\varepsilon$ , but the peak obtains when  $\sigma_\varepsilon$  is substantially above  $\sigma_u\sqrt{\sqrt{17}-4/2}$ . This is a combination of two effects. First,  $\beta_{v,2}$  should have a negative relation with  $\beta_{v,1}$ , as the fundamental investor smoothes her trades across the two periods. Thus, the U-shaped  $\beta_{v,1}$  leads to a hump-shaped  $\beta_{v,2}$ . Second, the fundamental investor also faces competition from the back-runner in the second period, and as  $\sigma_\varepsilon$  increases, this competition is less intense, so that the fundamental investor can afford to trade more aggressively on her private information. The second competition effect, adding to the first smoothing effect, implies that the hump-shaped  $\beta_{v,2}$  achieves its peak above  $\sigma_u\sqrt{\sqrt{17}-4/2}$ .

It is straightforward to understand that  $\delta_s$  decreases with  $\sigma_\varepsilon$ : A higher value of  $\sigma_\varepsilon$  means that the back-runner’s private information  $s$  is less precise, and so he trades less aggressively on this information.

### 3 Implications of Back-Running for Market Quality and Welfare

In this section we discuss the positive and normative implications of back-running, including price discovery, market liquidity, and the trading profits (or losses) of various trader types. Because these measures are proxies for market quality and welfare, our analysis generates important policy implications regarding the use of order-flow informed trading strategies.

We first examine the behavior of positive variables that represent market quality. In the microstructure literature, two leading positive variables are price discovery and market liquidity.<sup>10</sup> Price discovery measures how much information about the asset value  $v$  is revealed in prices  $p_1$  and  $p_2$ . Given price functions (8) and (9), prices are linear transformations of aggregate order flows  $y_1$  and  $y_2$ , and thus the literature has measured price discovery by the market maker’s posterior variances of  $v$  in periods 1 and 2:

$$\Sigma_1 \equiv \text{Var}(v|y_1) \text{ and } \Sigma_2 \equiv \text{Var}(v|y_1, y_2).$$

A lower  $\Sigma_t$  implies a more informative period- $t$  price about  $v$ , for  $t \in \{1, 2\}$ . Price discovery is important because it helps allocation efficiency by conveying information that is useful for real decisions (see, for example, O’Hara (2003) and Bond, Edmans, and Goldstein (2012)).

In Kyle-type models (including ours), market liquidity is measured by the inverse of Kyle’s lambda ( $\lambda_1$  and  $\lambda_2$ ), which are price impacts of trading. A lower  $\lambda_t$  means that the period- $t$

---

<sup>10</sup>For example, O’Hara (2003) states that “Markets have two important functions—liquidity and price discovery—and these functions are important for asset pricing.”

market is deeper and more liquid. One important reason to care about market liquidity is that it is related to the welfare of noise traders, who can be interpreted as investors trading for non-informational, liquidity or hedging reasons that are decided outside the financial markets. In general, noise traders are better off in a more liquid market, because their expected trading loss is  $(\lambda_1 + \lambda_2) \sigma_u^2$  in our economy.

Next, the normative variables are the payoffs of each group of players in the economy, that is, the expected profit  $E(\pi_{F,1} + \pi_{F,2})$  of the fundamental investor, the expected profit  $E(\pi_{B,2})$  of the back-runner, and the expected loss  $(\lambda_1 + \lambda_2) \sigma_u^2$  of noise traders. This approach allows us to discuss who wins and who loses as a result of a particular policy. In practice, investors' trading motives range from fundamental analysis to liquidity shocks (e.g., client withdrawal from mutual funds or hedge funds). Our fundamental investor can be viewed as investors trading for informational reasons, and noise traders as those trading for liquidity reasons. The back-runner is more in line with broker-dealers or HFTs who employ sophisticated trading technology and may possess superior order-flow information. If the regulator wishes to protect liquidity-driven traders, the welfare of noise traders would be the relevant measure. If the regulator wishes to protect investors who acquire fundamental information, then the informed profit  $E(\pi_{F,1} + \pi_{F,2})$  would be a relevant measure.

The following proposition gives a comparison between two “extreme” economies: the economy with  $\sigma_\varepsilon = 0$  and the one with  $\sigma_\varepsilon = \infty$  (i.e. the standard Kyle setting). For instance, the first economy corresponds to one in which back-runners are able to extract very precise information about the past orders submitted by large institutions. The second economy may represent one in which institutional investors manage to hide order-flow information almost completely (or an economy in which back-runners do not participate in the market, due to high technological costs or strict regulations). In the proposition, we have used superscripts “0” and “Kyle” to indicate these two economies.

**Proposition 4** (Perfect Order-Flow Information vs. Standard Kyle). *In the two-period setting, the following orderings apply:*

$$\begin{aligned} \Sigma_1^0 &> \Sigma_1^{Kyle}, \Sigma_2^0 < \Sigma_2^{Kyle}, \\ \lambda_1^0 &< \lambda_1^{Kyle}, \lambda_2^0 > \lambda_2^{Kyle}, \\ E(\pi_{F,1}^0) &< E(\pi_{F,1}^{Kyle}), E(\pi_{F,2}^0) < E(\pi_{F,2}^{Kyle}) \text{ and} \\ (\lambda_1^0 + \lambda_2^0) \sigma_u^2 &< (\lambda_1^{Kyle} + \lambda_2^{Kyle}) \sigma_u^2. \end{aligned}$$

The positive implications in Proposition 4 are in sharp contrast to those presented by Huddart, Hughes, and Levine (2001), although both studies consider a comparison between

an economy featuring a mixed strategy equilibrium and a standard Kyle economy. In Huddart, Hughes, and Levine (2001), the market maker perfectly observes the past order placed by an informed trader. They find that market liquidity and price discovery unambiguously improve in both periods of their economy relative to a standard Kyle setting (i.e., their  $\lambda_1$ ,  $\lambda_2$ ,  $\Sigma_1$ , and  $\Sigma_2$  are all smaller than the Kyle setting counterparts). In contrast, in our setting, the market maker does not observe the informed fundamental investor's past trade  $x_1$ ; it is the back-runner who does, with some noise. As a result of the endogenous noise  $z$  placed by the fundamental investor in the mixed strategy and her more cautious trading on fundamental information (i.e., a smaller  $\beta_{v,1}$ ), the first-period price discovery is harmed by back-running in our setting (i.e.,  $\Sigma_1^0 > \Sigma_1^{Kyle}$ ). The presence of *perfect* information about past order flows also worsens the second period market liquidity relative to the standard Kyle setting (i.e.,  $\lambda_2^0 > \lambda_2^{Kyle}$ ). This is again opposite to the effect of publicly revealing the informed orders in period 1, as in Huddart, Hughes, and Levine (2001).

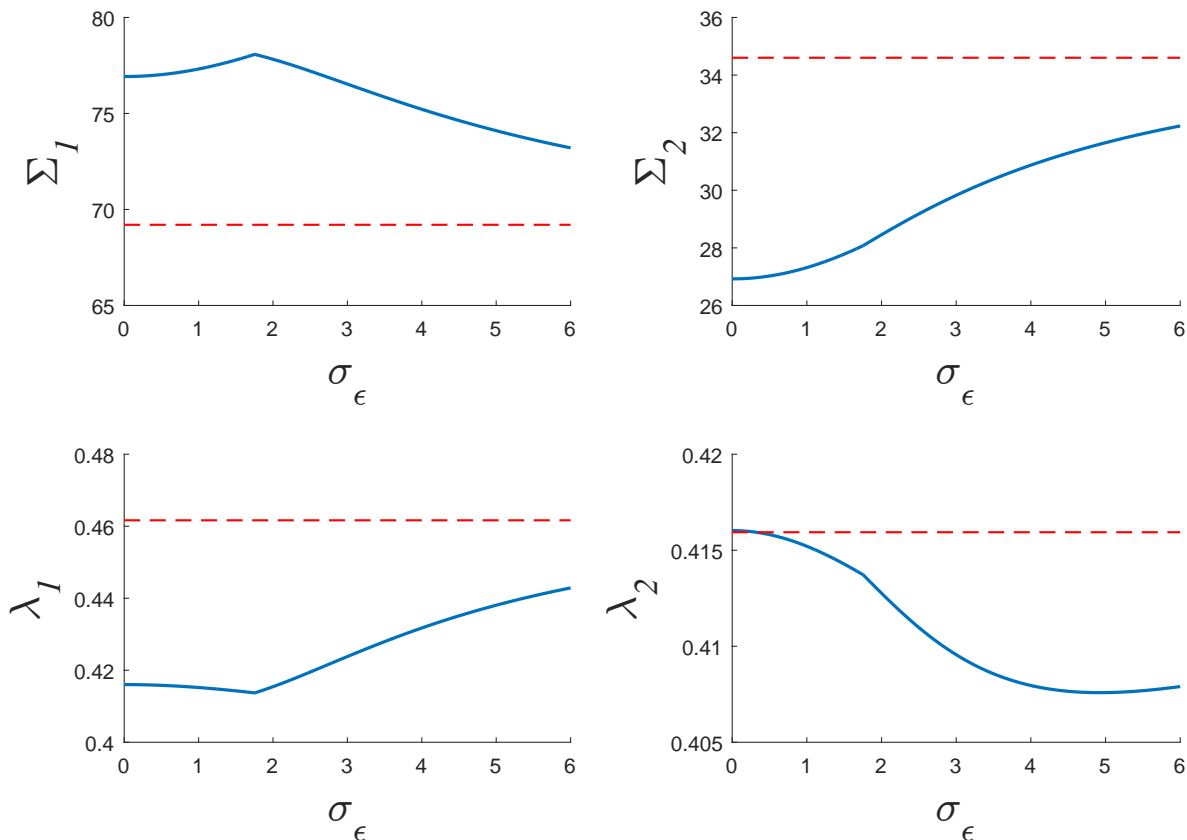
Figures 3 and 4 respectively plot in solid lines the positive and normative implications as we continuously increase  $\sigma_\varepsilon$  from 0 to  $\infty$ . The other two exogenous parameters are the same as those in Figure 2 ( $\sigma_u = 10$  and  $\Sigma_0 = 100$ ). The dashed lines plot the corresponding variables in a standard two-period Kyle model without the back-runner.

In Figure 3, we see that  $\Sigma_1$  is hump-shaped in  $\sigma_\varepsilon$ , with the peak at the cutoff  $\sigma_\varepsilon = \sigma_u \sqrt{\sqrt{17} - 4}/2$ . The intuition is as follows. In the first period, only the fundamental investor's trade brings information about  $v$  into the market. Since her trading sensitivity  $\beta_{v,1}$  on fundamental information is U-shaped in  $\sigma_\varepsilon$  (see Figure 2),  $\Sigma_1$  should have the opposite pattern, i.e., hump-shaped. By contrast,  $\Sigma_2$  monotonically increases with  $\sigma_\varepsilon$  in Figure 3. This is because in period 2, both the fundamental investor and the back-runner trade on value-relevant information, and as  $\sigma_\varepsilon$  increases, the back-runner's order brings less information about  $v$  into the price. Comparing the solid lines to dashed lines, we see that adding the back-runner harms price discovery in period 1 but improves price discovery in period 2.

The illiquidity measures in both periods,  $\lambda_1$  and  $\lambda_2$ , first decrease and then increase with  $\sigma_\varepsilon$ . Since adverse selection from the fundamental investor is the sole source of price impact in period 1, it is rather intuitive that  $\lambda_1$  has a similar U-shape as  $\beta_{v,1}$  (see equation (20)). The period-2 illiquidity measure  $\lambda_2$  is also U-shaped and opposite to the humped-shaped  $\beta_{v,2}$ , by the first-order condition in period 2 (i.e.,  $\lambda_2 = \frac{1}{2\beta_{v,2}}$  by (14)).

Comparing the solid lines to dashed lines, we find that back-running generally improves the first-period market liquidity because the fundamental investor trades less aggressively on her private information, but its impact on the second-period market liquidity is ambiguous.

Figure 3: Implications for Positive Variables

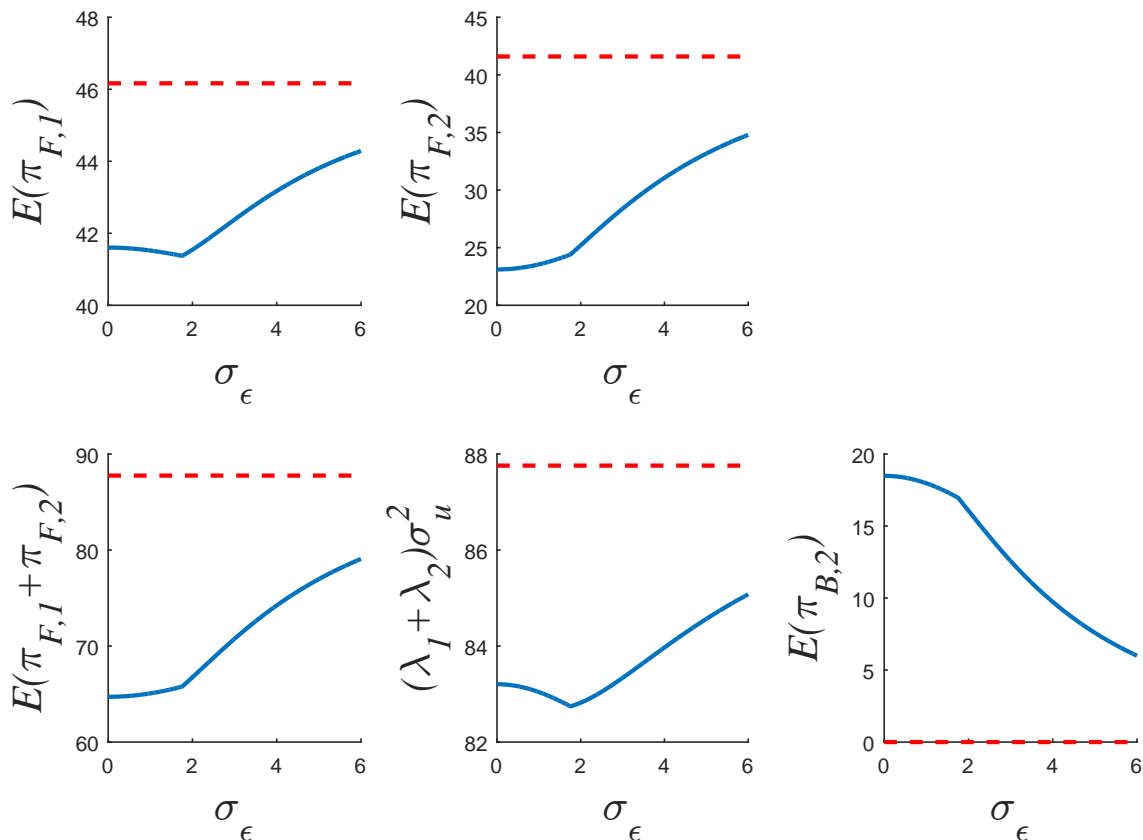


This figure plots the market quality implications of back-running. In each panel, the blue solid line plots the value in the equilibrium of this paper, and the dashed red line plots the value in a standard Kyle economy (i.e.,  $\sigma_\varepsilon = \infty$ ). The horizontal axis in each panel is the standard deviation  $\sigma_\varepsilon$  of the noise in the back-runner's private signal about the fundamental investor's past order. The other parameters are:  $\sigma_u = 10$  and  $\Sigma_0 = 100$ .

Consistent with Proposition 4, back-running worsens the second-period liquidity relative to the standard Kyle setting, if and only if the back-runner's order-flow information is sufficiently precise. (In the neighborhood of  $\sigma_\varepsilon = 0$ , the solid line is strictly above the dashed line in the plot for  $\lambda_2$ .) This is due to a combination of two effects. First, adding the back-runner introduces competition, which makes the period-2 aggregate order flow reflect more of the fundamental than noise trading. This tends to reduce  $\lambda_2$ . Second, back-running also increases the amount of *private* information, which makes the adverse selection problem faced by the market maker more severe. This generally tends to increase  $\lambda_2$ . When  $\sigma_\varepsilon$  is



Figure 4: Implications for Normative Variables



This figure plots the profits of various groups of traders. In each panel, the blue solid line plots the value in the equilibrium of this paper, and the dashed red line plots the value in a standard Kyle economy (i.e.,  $\sigma_\epsilon = \infty$ ). The horizontal axis in each panel is the standard deviation  $\sigma_\epsilon$  of the noise in the back-runner's private signal about the fundamental investor's past order. The other parameters are:  $\sigma_u = 10$  and  $\Sigma_0 = 100$ .

small, the back-runner has very precise private information and the second effect dominates, so that  $\lambda_2$  is higher than its counterpart in a standard Kyle setting.

The top two panels of Figure 4 plot the fundamental investor's expected profits in the two periods,  $E(\pi_{F,1})$  and  $E(\pi_{F,2})$ . We observe that  $E(\pi_{F,2})$  monotonically increases with  $\sigma_\epsilon$ . This result is intuitive: A higher  $\sigma_\epsilon$  means that the fundamental investor faces a less competitive back-runner in period 2, so her period-2 profit is higher on average. The period-1 profit  $E(\pi_{F,1})$  first decreases with  $\sigma_\epsilon$  (in the mixed strategy regime) and then increases with  $\sigma_\epsilon$  (in the pure strategy regime). This U-shaped profit pattern is natural given the U-shaped  $\beta_{v,1}$  pattern in Figure 2. Comparing the solid lines to dashed lines, we clearly see

that back-running lowers the profit of the fundamental investor.

The bottom three panels of Figure 4 present the total profit  $E(\pi_{F,1} + \pi_{F,2})$  of the fundamental investor, the total loss  $(\lambda_1 + \lambda_2)\sigma_u^2$  of noise traders, and the expected profit  $E(\pi_{B,2})$  of the back-runner. All the results are as expected. As  $\sigma_\varepsilon$  increases, the back-runner has less precise private information, and thus  $E(\pi_{B,2})$  decreases. Meanwhile, a higher  $\sigma_\varepsilon$  also implies that the fundamental investor faces less competition from the back-runner, and  $E(\pi_{F,1} + \pi_{F,2})$  increases. The U-shaped total loss  $(\lambda_1 + \lambda_2)\sigma_u^2$  of noise traders is a direct result of the U-shaped  $\lambda_1$  and  $\lambda_2$  in Figure 3. In general, back-running reduces the loss of noise traders (the entire solid line of  $(\lambda_1 + \lambda_2)\sigma_u^2$  lies below the dashed line).

## 4 Information Acquisition

So far, we have taken the information of the fundamental investor and the back-runner as given. In this section, we explicitly model information acquisition. Besides showing the robustness of our earlier results, this additional step sheds light on questions like “Does back-running discourage acquisition of fundamental information?”

### 4.1 Setup

We add one period,  $t = 0$ , before the two-period economy considered in previous sections. At  $t = 0$ , the fundamental investor decides the amount of fundamental information she acquires, and the back-runner decides the precision of order-flow information he acquires. Specifically, for the fundamental trader, we follow Admati and Pfleiderer (1989), Madrigal (1996), and Bond, Goldstein, and Prescott (2010) and assume that the fundamental investor can pay a cost  $C_F(\phi)$  upfront to observe the fundamental value  $v$  with probability  $\phi \in (0, 1)$ . For the back-runner, we follow Verrecchia (1982) and Vives (2008) and assume that the back-runner can pay a cost  $C_B\left(\frac{1}{\sigma_\varepsilon^2}\right)$  upfront to observe a signal  $s$  of  $x_1$  with precision  $\frac{1}{\sigma_\varepsilon^2}$ . These information-acquisition decisions are simultaneous. After time 0, the choices of  $\phi$  and  $\sigma_\varepsilon$  become public information. In reality, investment in fundamental research, such as hiring analysts, and investment in advanced trading technology, such as high-speed connections to exchanges, are usually observable.

To ensure interior solutions of  $\sigma_\varepsilon^2$  and  $\phi$ , we make the standard technical assumptions: (i)  $C_B(\cdot)$  and  $C_F(\cdot)$  are increasing and convex; and (ii)  $C_B(0) = C'_B(0) = 0, C_B(\infty) = C'_B(\infty) = \infty, C_F(0) = C'_F(0) = 0$ , and  $C_F(1) = C'_F(1) = \infty$ .

For simplicity, we assume that at the beginning of period 1, it becomes public knowledge whether the fundamental investor has successfully observed  $v$ . It is a standard assumption in Kyle-type models whether such an (fundamentally) informed investor exists. Then, the subsequent game has two possible outcomes:

1. If the fundamental investor observes  $v$ , then the economy is the one that we analyzed in the previous two sections.
2. If the fundamental investor does not observe  $v$ , then as an uninformed investor she will not trade in either period. As a result, the back-runner will not trade in period 2, either, despite receiving the signal of the fundamental investor's (zero) order flow. In this case, only noise traders submit orders, and so the price is  $p_1 = p_2 = E(v) = p_0$ .

## 4.2 Analysis and Results

Our objective is to find the equilibrium levels of  $\phi$  and  $\sigma_\varepsilon$ . They are determined jointly by the period-0 maximization problems of the fundamental investor and the back-runner.

Recall that  $\pi_{F,1}$  and  $\pi_{F,2}$  denote the realized profits of the fundamental investor in dates 1 and 2, respectively. The fundamental investor's period-0 expected net profit is:

$$\Pi_{F,0} \equiv \phi E(\pi_{F,1} + \pi_{F,2}) - C_F(\phi),$$

and her problem is to choose  $\phi$  to maximize  $\Pi_{F,0}$ , taking her conjectured equilibrium value of  $\sigma_\varepsilon$  as given. Because  $E(\pi_{F,1} + \pi_{F,2})$  does not depend on  $\phi$ , and given the technical assumption on  $C_F(\phi)$ , we know that the solution to the fundamental investor's problem is characterized by the first-order condition:

$$E(\pi_{F,1} + \pi_{F,2}) = C'_F(\phi).$$

Now we consider the back-runner's information acquisition problem. Recall that

$$\pi_{B,2} \equiv (v - p_2) d_2$$

is the back-runner's realized period-2 profit. So, his period-0 expected net profit of acquiring order-flow information is:

$$\Pi_{B,0} \equiv \phi E(\pi_{B,2}) - C_B\left(\frac{1}{\sigma_\varepsilon^2}\right).$$

The back-runner takes the equilibrium value  $\phi$  as given and chooses  $\sigma_\varepsilon$  to maximize  $\Pi_0^B$ .

The back-runner's choice of  $\sigma_\varepsilon$  affects  $E(\pi_{B,2})$  through its effect on the equilibrium strategies,  $\sigma_z, \beta_{v,1}, \beta_{v,2}, \beta_{x_1}, \beta_{y_1}, \delta_s, \delta_{y_1}, \lambda_1$  and  $\lambda_2$ . Specifically, we can compute

$$E(\pi_{B,2}) = \lambda_2 [(\delta_s - \delta_{y_1})^2 \beta_{v,1}^2 \Sigma_0 + (\delta_s - \delta_{y_1})^2 \sigma_z^2 + \delta_s^2 \sigma_\varepsilon^2 + \delta_{y_1}^2 \sigma_u^2],$$

and hence

$$\Pi_{B,0} = \phi \lambda_2 [(\delta_s - \delta_{y_1})^2 \beta_{v,1}^2 \Sigma_0 + (\delta_s - \delta_{y_1})^2 \sigma_z^2 + \delta_s^2 \sigma_\varepsilon^2 + \delta_{y_1}^2 \sigma_u^2] - C_B \left( \frac{1}{\sigma_\varepsilon^2} \right).$$

There is an important complication in the solution to the back-runner's information-acquisition problem. Although this problem has an interior solution, as ensured by the cost function  $C_B(\cdot)$ , the optimal choice of  $\sigma_\varepsilon$  cannot in general be guaranteed by setting the first-order derivative to zero. This is because whether the equilibrium has a mixed strategy or a pure strategy (used by the fundamental investor) depends on  $\sigma_\varepsilon$ . As  $\sigma_\varepsilon$  decreases and drops below the threshold value of  $(\sqrt{\sqrt{17}-4}/2)\sigma_u$ , the equilibrium switches from pure strategy to mixed strategy, giving rise to a kink in  $E(\pi_{B,2})$ . If the optimal value of  $\sigma_\varepsilon$  occurs at the kink, the first-order condition is characterized by two inequalities rather than an equality. (This complication does not apply to the fundamental investor's problem.)

To solve the equilibrium explicitly and numerically, we need explicit functional forms of  $C_B$  and  $C_F$ . Following Vives (2008), we choose the following parametrization:

$$\begin{aligned} C_B \left( \frac{1}{\sigma_\varepsilon^2} \right) &= k_B \left( \frac{1}{\sigma_\varepsilon^2} \right)^{h_B} = k_B \sigma_\varepsilon^{-2h_B}, \\ C_F(\phi) &= k_F \left( \frac{\phi}{1-\phi} \right)^{h_F}, \end{aligned}$$

where

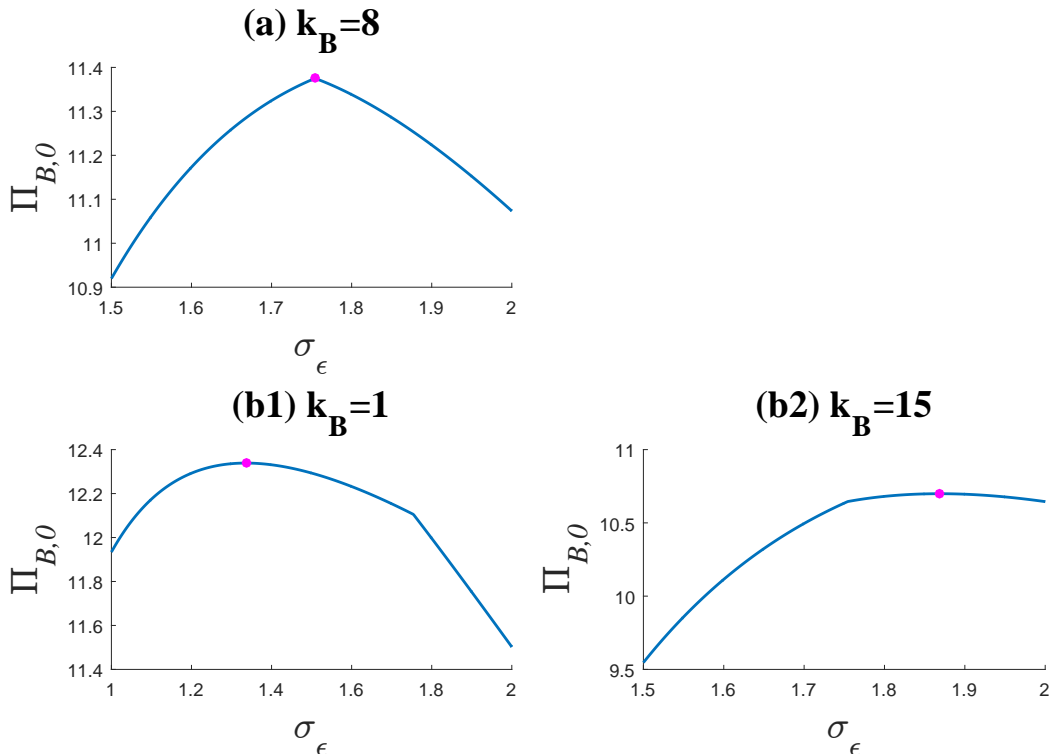
$$k_B > 0, k_F > 0, h_B > 1 \text{ and } h_F > 1.$$

We will conduct comparative statics with respect to parameter  $k_B$ , which is taken as a proxy for the cost of acquiring order-flow information. A larger  $k_B$  means a higher cost.

In order to gain better intuition of the comparative statics, it is useful to first illustrate the kink in  $\Pi_{B,0}$ . Figure 5 plots the profit function  $\Pi_{B,0}$  against  $\sigma_\varepsilon$ , for  $k_B \in \{1, 8, 15\}$ , in the three panels. In each panel,  $\phi$  is set to its equilibrium value corresponding to the particular  $k_B$  and does not vary with  $\sigma_\varepsilon$ , since this value of  $\phi$  is the belief of the back-runner at the information-acquisition stage. But for each  $\sigma_\varepsilon$ , other equilibrium variables in periods 1 and 2 are determined according to Propositions 1 and 2 for this particular  $\sigma_\varepsilon$  (and the fixed equilibrium value of  $\phi$ ), because at the information-acquisition stage, the back-runner takes into account how the fundamental investor and the market maker react in future periods. As in earlier figures, we set  $\sigma_u = 10$  and  $\Sigma_0 = 100$ . We also set  $k_F = 1$  and  $h_F = h_B = 2$ .

In Panel (a), where  $k_B = 8$ , the optimal  $\sigma_\varepsilon$  occurs exactly at the kink. In Panels (b1) and (b2), where  $k_B = 1$  and  $k_B = 15$  respectively, the optimal values of  $\sigma_\varepsilon$  are found in the smooth regions. Intuitively, if the information-acquisition cost  $k_B$  is very high or very low, the unconstrained optimal  $\sigma_\varepsilon$ —the solution without considering the equilibrium switch—is

Figure 5: Illustration of Possible Kink in  $\Pi_{B,0}$



This figure plots  $\Pi_{B,0}$  against  $\sigma_\epsilon$  for three values of  $k_B$ . In each panel,  $\phi$  is set to its equilibrium value corresponding to the particular  $k_B$  and does not vary with  $\sigma_\epsilon$ . For each  $\sigma_\epsilon$ , other equilibrium variables in periods 1 and 2 are optimized to this particular  $\sigma_\epsilon$ . The red dot is the global maximum. Other parameters:  $\sigma_u = 10$ ,  $\Sigma_0 = 100$ ,  $k_F = 1$ , and  $h_F = h_B = 2$ .

sufficiently far away from the threshold  $(\sqrt{\sqrt{17}-4}/2)\sigma_u$ , so the switch in equilibrium does not bind, as in Panels (b1) and (b2). If, however,  $k_B$  takes an intermediary value, the nature of equilibrium depends heavily on  $\sigma_\epsilon$ . In the mixed strategy region of Panel (a), i.e. if  $\sigma_\epsilon < (\sqrt{\sqrt{17}-4}/2)\sigma_u$ , the back-runner prefers to acquire less precise information because the fundamental investor injects noise anyway; but in the pure strategy region of Panel (a), i.e. if  $\sigma_\epsilon \geq (\sqrt{\sqrt{17}-4}/2)\sigma_u$ , the back-runner prefers more precise information because the fundamental investor does not inject any noise. The result is that the unique maximum of  $\Pi_{B,0}$  is obtained when  $\sigma_\epsilon$  is exactly at the threshold  $(\sqrt{\sqrt{17}-4}/2)\sigma_u$ . As we see shortly, this corner solution leads to the stickiness in the responses of equilibrium outcomes to changes in  $k_B$ .

Now we proceed with describing the comparative statics. The variables of interest

include:

- Equilibrium values of  $\phi, \sigma_\varepsilon, \sigma_z, \beta_{v,1}, \beta_{v,2}, \beta_{x_1}, \beta_{y_1}, \delta_s, \delta_{y_1}, \lambda_1$  and  $\lambda_2$ ;
- Equilibrium profits:  $\Pi_{F,0}$  and  $\Pi_{B,0}$ , and the expected cost of noise traders  $\Pi_{F,0} + \Pi_{B,0}$ ;
- Price discovery:  $\phi\Sigma_1 + (1 - \phi)\Sigma_0$  for period 1 and  $\phi\Sigma_2 + (1 - \phi)\Sigma_0$  for period 2;
- Illiquidity:  $\phi\lambda_1 + (1 - \phi)0 = \phi\lambda_1$  for period 1 and  $\phi\lambda_2 + (1 - \phi)0 = \phi\lambda_2$  for period 2.

Here, we follow the literature and measure average price discovery and illiquidity, where  $\lambda_1$  and  $\lambda_2$  are defined in Propositions 1 and 2, and  $\Sigma_1$  and  $\Sigma_2$  are defined at the beginning of Section 3.

Figure 6 plots the implications of changes in information acquisition cost  $k_B$  for information-acquisition decisions, profits of various groups of traders, price discovery, and market illiquidity.

An interesting and salient pattern is that all but one of these variables are entirely irresponsive to changes in  $k_B$  when  $k_B$  is in an intermediate range. As discussed earlier, in this range, the optimal  $\sigma_\varepsilon$  is always equal to  $(\sqrt{\sqrt{17}-4}/2)\sigma_u$  regardless of  $k_B$ . As a result, the equilibrium has zero sensitivity to  $k_B$ , leading to the flat parts of equilibrium variables.

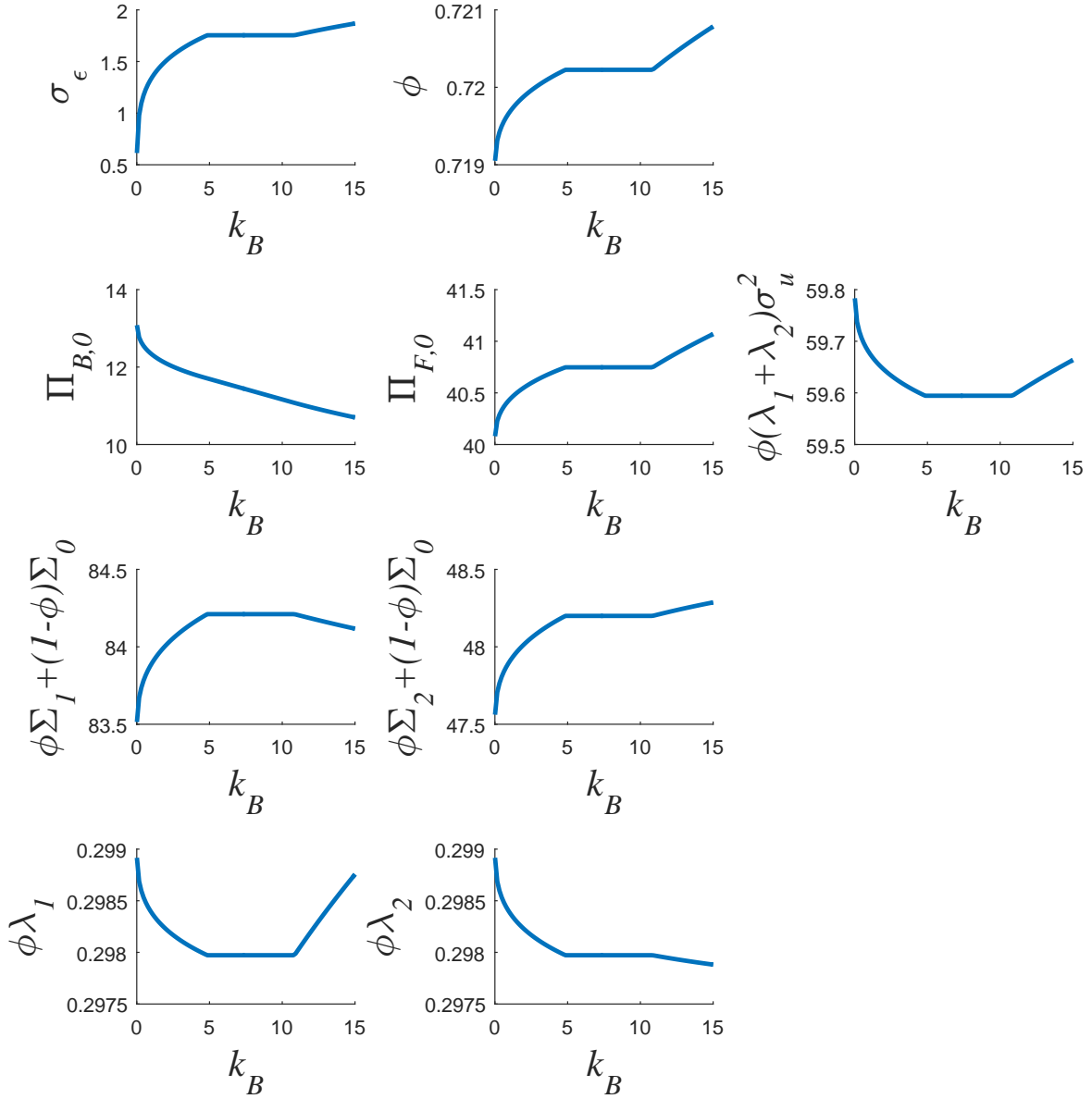
Moreover, we observe that a lower  $k_B$  weakly reduces  $\phi$ , which implies that technology improvement in processing order-flow information reduces investment in fundamental information (top row of Figure 6). A lower cost of acquiring order-flow information leads to a higher profit of the back-runner but a lower profit of the fundamental investor. The loss of noise traders, the period-1 price discovery, and the period-1 market liquidity are all non-monotone in  $k_B$ . This last result mirrors the patterns in Figures 3 and 4 that these variables are also non-monotone in  $\sigma_\varepsilon$ .

Overall, results of the previous sections are robust to information acquisition. A unique and novel prediction with information acquisition is that equilibrium outcomes can be insensitive to the cost of order-flow information. This insensitivity is the consequence of the switch between a pure strategy equilibrium and a mixed strategy one.

## 5 Conclusion

Order-flow informed trading is a salient part of modern financial markets. This type of trading strategies, such as order anticipation, often starts with no innate trading motive, but instead seeks and exploits information from other investors' past order flows. We refer

Figure 6: Equilibrium Strategies and Implications of Information Acquisition



This figure plots the equilibrium levels of information acquisition, expected profits of various parties, price discovery, and market illiquidity, as functions of  $k_B$ . Other parameters:  $\sigma_u = 10, \Sigma_0 = 100, k_F = 1$ , and  $h_F = h_B = 2$ .

to such strategies as back-running. While back-running has long existed in financial markets, its latest incarnation in certain high-frequency trading strategies caused renewed and severe concerns among investors and regulators.

In this paper we study the strategic interaction between back-runners and fundamental informed investors. In our two-period model, which is based on Kyle (1985), a back-runner observes, *ex post* and potentially with noise, the executed trades of the informed investor in period 1. The informed order flow thus provides a signal to the back-runner regarding the asset fundamental value. Using this information, the back-runner competes with the informed investor in period 2. While simple, this model structure parsimoniously captures the key idea of back-running.

If the back-runner's signal is sufficiently precise, the fundamental investor hides her information by endogenously adding noise into her period-1 order flow, leading to a mixed strategy equilibrium. The more precise is the order-flow signal, the more volatile is the added noise. As the back-runner's signal becomes sufficiently imprecise, the equilibrium switches to a pure strategy one, in which the fundamental investor adds no endogenous noise in her order flows. We prove uniqueness of equilibrium under natural conditions. The characterization of the equilibria, in particular the endogenous switch between a mixed strategy equilibrium and a pure strategy one, is the first main contribution of this paper.

Our second main contribution is to identify the effects of back-running on market quality and welfare. Because the fundamental investor trades more cautiously and potentially adds noise into her period-1 orders, the presence of the back-runner harms price discovery in the first period. In the second period, however, price discovery is improved because of competition. Effects on market liquidity, measured by the inverse of Kyle's lambda, are mixed: Liquidity improves in the first period but can either improve or worsen in the second period. Overall, back-running harms the fundamental investor but benefits noise traders.

Our main results are robust to endogenous information acquisition. A lower cost of acquiring order-flow information reduces the fundamental investor's incentive to acquire fundamental information. The equilibrium with information acquisition also has the novel property that many equilibrium outcomes are invariant to changes, within a certain range, in the cost of order-flow information. Such stickiness is driven by the endogenous switch between pure strategy equilibrium and mixed strategy equilibrium.



# Appendix

## A List of Model Variables

Variables	Description
<u>Random Variables</u>	
$v$	Asset liquidation value at the end of period 2, $N(p_0, \Sigma_0)$
$x_1, x_2$	Orders placed by the fundamental investor in periods 1 and 2
$z$	Noise component in the period-1 order $x_1$ of the fundamental investor
$d_2$	Order placed by the back-runner in period 2
$s, \varepsilon$	Signal observed by the back-runner, and its noise component
$u_1, u_2$	Noise trading in periods 1 and 2
$y_1, y_2$	Aggregate order flows in periods 1 and 2
$p_1, p_2$	Asset prices in periods 1 and 2
$\pi_{F,1}, \pi_{F,2}$	Fundamental investor's profits attributable to trades in periods 1 and 2
$\pi_{B,2}$	Back-runner's profit in period 2
$\sigma_\varepsilon$	Noise in back-runner's signal of $x_1$
$\phi$	Fundamental investor's probability of observing $v$ (only in Section 4)
$\Pi_{F,0}$	$\phi E(\pi_{F,1} + \pi_{F,2}) - C_F(\phi)$ (only in Section 4)
$\Pi_{B,0}$	$\phi E(\pi_{B,2}) - C_B\left(\frac{1}{\sigma_\varepsilon^2}\right)$ (only in Section 4)
<u>Deterministic Variables</u>	
$p_0, \Sigma_0$	Prior mean and variance of the asset value
$\sigma_u^2$	Variance of noise trading in periods 1 and 2
$\sigma_z$	Standard deviation of the noise component $z$ in the period-1 order $x_1$ placed by the fundamental investor
$\Sigma_1, \Sigma_2$	Posterior variance of the asset value in periods 1 and 2 (i.e., $Var(v y_1)$ and $Var(v y_1, y_2)$ )
$C_F(\phi)$	Fundamental investor's cost to observe $v$ with probability $\phi$
$C_B(1/\sigma_\varepsilon^2)$	Back-runner's cost of observing a signal of $x_1$ with precision $1/\sigma_\varepsilon^2$
<u>Strategy Summary</u>	
$\beta_{v,1}$	$x_1 = \beta_{v,1}(v - p) + z$
$\beta_{v,2}, \beta_{x_1}, \beta_{y_1}$	$x_2 = \beta_{v,2}(v - p) - \beta_{x_1}x_1 + \beta_{y_1}y_1$
$\delta_s, \delta_{y_1}$	$d_2 = \delta_s s - \delta_{y_1}y_1$
$\lambda_1$	$p_1 = p_0 + \lambda_1 y_1$ , with $y_1 = x_1 + u_1$
$\lambda_2$	$p_2 = p_1 + \lambda_2 y_2$ , with $y_2 = x_2 + d_2 + u_2$

## B Proofs

### B.1 Proof of Equation (10)

Define  $\sigma_x^2 \equiv Var(x_1) = \beta_{v,1}^2 \Sigma_0 + \sigma_z^2$ . Direct computation shows

$$E(v|s, y_1) - E(v|y_1) = \frac{\beta_{v,1} \Sigma_0 \sigma_\varepsilon^{-2}}{\sigma_x^2 (\sigma_x^{-2} + \sigma_\varepsilon^{-2} + \sigma_u^{-2})} \left( s - \frac{\sigma_x^2}{\sigma_u^2 + \sigma_x^2} y_1 \right).$$

Thus, it suffices to show that

$$\frac{\delta_{y_1}}{\delta_s} = \frac{\sigma_x^2}{\sigma_u^2 + \sigma_x^2} \quad (\text{B1})$$

holds in equilibrium, in order for  $d_2$  in equation (7) to admit a form given by equation (10).

By equation (18), we have:

$$\frac{\beta_{v,1} \Sigma_0}{\delta_s \lambda_2} = \frac{\sigma_x^2 [4(\sigma_x^{-2} + \sigma_\varepsilon^{-2} + \sigma_u^{-2}) - \sigma_\varepsilon^{-2}]}{\sigma_\varepsilon^{-2}}. \quad (\text{B2})$$

Plugging the expression of  $\lambda_1 = \frac{\beta_{v,1} \Sigma_0}{\sigma_z^2 + \sigma_u^2}$  (i.e. equation (20)) into equation (19) yields

$$\frac{\delta_{y_1}}{\delta_s} = \frac{\beta_{v,1} \Sigma_0}{\delta_s \lambda_2} \frac{1}{3(\sigma_x^2 + \sigma_u^2)} - \frac{4\sigma_\varepsilon^2}{3\sigma_u^2}. \quad (\text{B3})$$

Inserting equation (B2) into (B3) and simplifying, we have equation (B1).

### B.2 Proof of Proposition 1

A mixed strategy equilibrium is characterized by nine parameters,  $\sigma_z$ ,  $\beta_{v,1}$ ,  $\beta_{v,2}$ ,  $\beta_{y_1}$ ,  $\beta_{x_1}$ ,  $\delta_{y_1}$ ,  $\delta_s$ ,  $\lambda_1$ , and  $\lambda_2$ . These parameters are jointly determined by a system consisting of nine equations (given by (14), (18), (19), (20), (21), and (23)) as well as one SOC ( $\lambda_2 > 0$  given by (13)). Note that by equation (23),  $\delta_s$  is already known, and also  $\lambda_1 = \lambda_2$  degenerates to one parameter, denoted by  $\lambda$ . So, the system characterizing a mixed strategy equilibrium essentially has six unknowns. To solve this system, we first simplify it to a 3-equation system in terms of  $(\lambda, \beta_{v,1}, \sigma_z)$  and then solve this new system analytically.

Given that  $\delta_s$  is known, parameter  $\delta_{y_1}$  is also known by (19). Also, once  $\lambda$  is solved, the three equations in (14) will yield solutions of  $\beta_{v,2}$ ,  $\beta_{x_1}$ , and  $\beta_{y_1}$ . Thus, the three equations left to compute  $(\lambda, \beta_{v,1}, \sigma_z)$  are given by equations (18), (20) and (21). To solve this 3-equation system, we first express  $\beta_{v,1}$  and  $\lambda$  as functions of  $\sigma_z$ , and then solve the single equation of  $\sigma_z$ .

By (18) and noting that  $\lambda \equiv \lambda_1 = \lambda_2$ , we have

$$\lambda = \frac{1}{\delta_s} \frac{\frac{\sigma_\varepsilon^{-2}}{(\beta_{v,1}^2 \Sigma_0 + \sigma_z^2)^{-1} + \sigma_\varepsilon^{-2} + \sigma_u^{-2}}}{4 - \frac{\sigma_\varepsilon^{-2}}{(\beta_{v,1}^2 \Sigma_0 + \sigma_z^2)^{-1} + \sigma_\varepsilon^{-2} + \sigma_u^{-2}}} \frac{\beta_{v,1} \Sigma_0}{\beta_{v,1}^2 \Sigma_0 + \sigma_z^2}.$$

Combining the above equation with (20) and the expression  $\delta_s = \frac{\frac{4}{3}}{1 + \frac{4\sigma_\varepsilon^2}{3\sigma_u^2}}$ , we can compute

$$\beta_{v,1}^2 = \frac{\sigma_u^4 - 3\sigma_u^2\sigma_z^2 - 4\sigma_u^2\sigma_\varepsilon^2 - 4\sigma_z^2\sigma_\varepsilon^2}{3\Sigma_0\sigma_u^2 + 4\Sigma_0\sigma_\varepsilon^2}. \quad (\text{B4})$$

Equation (B4) puts an restriction on the endogenous value of  $\sigma_z$ , i.e.,  $\sigma_u^4 - 3\sigma_u^2\sigma_z^2 - 4\sigma_u^2\sigma_\varepsilon^2 - 4\sigma_z^2\sigma_\varepsilon^2 > 0$ , which can be shown to hold in equilibrium.

By (20) and (B4), we can express  $\lambda^2$  as a function of  $\sigma_z^2$  as follows:

$$\lambda^2 = \frac{\frac{\sigma_u^4 - 3\sigma_u^2\sigma_z^2 - 4\sigma_u^2\sigma_\varepsilon^2 - 4\sigma_z^2\sigma_\varepsilon^2}{3\Sigma_0\sigma_u^2 + 4\Sigma_0\sigma_\varepsilon^2} \Sigma_0^2}{\left( \frac{\sigma_u^4 - 3\sigma_u^2\sigma_z^2 - 4\sigma_u^2\sigma_\varepsilon^2 - 4\sigma_z^2\sigma_\varepsilon^2}{3\Sigma_0\sigma_u^2 + 4\Sigma_0\sigma_\varepsilon^2} \Sigma_0 + \sigma_z^2 + \sigma_u^2 \right)^2}. \quad (\text{B5})$$

Inserting  $\delta_s = \frac{\frac{4}{3}}{1 + \frac{4\sigma_\varepsilon^2}{3\sigma_u^2}}$  into equation (21) and further simplification yield

$$\left( \begin{aligned} & (52\Sigma_0\sigma_u^6 + 160\Sigma_0\sigma_u^4\sigma_\varepsilon^2 + 64\Sigma_0\sigma_u^2\sigma_\varepsilon^4) \beta_{v,1}^2 \\ & + (36\sigma_u^8 + 52\sigma_u^6\sigma_z^2 + 160\sigma_u^6\sigma_\varepsilon^2 + 160\sigma_u^4\sigma_z^2\sigma_\varepsilon^2 + 64\sigma_u^4\sigma_\varepsilon^4 + 64\sigma_u^2\sigma_z^2\sigma_\varepsilon^4) \end{aligned} \right) \lambda^2 \\ = (9\Sigma_0\sigma_u^6 + 9\Sigma_0\sigma_u^4\sigma_z^2 + 24\Sigma_0\sigma_u^4\sigma_\varepsilon^2 + 24\Sigma_0\sigma_u^2\sigma_z^2\sigma_\varepsilon^2 + 16\Sigma_0\sigma_u^2\sigma_\varepsilon^4 + 16\Sigma_0\sigma_z^2\sigma_\varepsilon^4).$$

Inserting equations (B4) and (B5) into the above equation, we can compute

$$\sigma_z^2 = \frac{\sigma_u^2 (\sigma_u^2 + 4\sigma_\varepsilon^2) (\sigma_u^4 - 16\sigma_\varepsilon^4 - 32\sigma_u^2\sigma_\varepsilon^2)}{(3\sigma_u^2 + 4\sigma_\varepsilon^2) (13\sigma_u^4 + 16\sigma_\varepsilon^4 + 40\sigma_u^2\sigma_\varepsilon^2)}, \quad (\text{B6})$$

which gives the expression of  $\sigma_z$  in Proposition 1.

In order for equation (B6) to indeed construct a mixed strategy equilibrium, we need

$$\sigma_z^2 = \frac{\sigma_u^2 (\sigma_u^2 + 4\sigma_\varepsilon^2) (\sigma_u^4 - 16\sigma_\varepsilon^4 - 32\sigma_u^2\sigma_\varepsilon^2)}{(3\sigma_u^2 + 4\sigma_\varepsilon^2) (13\sigma_u^4 + 16\sigma_\varepsilon^4 + 40\sigma_u^2\sigma_\varepsilon^2)} > 0 \Leftrightarrow \frac{\sigma_\varepsilon^2}{\sigma_u^2} < \frac{\sqrt{17}}{4} - 1.$$

Also, inserting equation (B6) into equation (B4), we see that (B4) is always positive. Finally, by equation (20) and  $\lambda_2 = \lambda_1$ , we know  $\lambda_2 > 0$ , i.e., the SOC is satisfied. Thus, when  $\frac{\sigma_\varepsilon^2}{\sigma_u^2} < \frac{\sqrt{17}}{4} - 1$ , the expression of  $\sigma_z^2$  in equation (B6) indeed constructs a mixed strategy equilibrium.

Clearly, if  $\frac{\sigma_\varepsilon^2}{\sigma_u^2} \geq \frac{\sqrt{17}}{4} - 1$ , then the solved  $\sigma_z^2$  would be non-positive in (B6), which implies the non-existence of a linear mixed strategy equilibrium.

### B.3 Proof of Proposition 2

For a pure strategy equilibrium, we have  $\sigma_z = 0$  and need to compute eight parameters,  $\beta_{v,1}, \beta_{v,2}, \beta_{y_1}, \beta_{x_1}, \delta_s, \delta_{y_1}, \lambda_1$ , and  $\lambda_2$ . These parameters are determined by equations (14), (18), (19), (20), (21), and (24), together with two SOC's, (13) and (25). In particular, after setting  $\sigma_z = 0$ , we can simplify equations (18), (20), and (21) as follows:

$$\delta_s = \frac{\frac{\sigma_\varepsilon^{-2}}{(\beta_{v,1}^2 \Sigma_0)^{-1} + \sigma_\varepsilon^{-2} + \sigma_u^{-2}}}{4 - \frac{\sigma_\varepsilon^{-2}}{(\beta_{v,1}^2 \Sigma_0)^{-1} + \sigma_\varepsilon^{-2} + \sigma_u^{-2}}} \frac{1}{\lambda_2 \beta_{v,1}}, \quad (\text{B7})$$

$$\lambda_1 = \frac{\beta_{v,1} \Sigma_0}{\beta_{v,1}^2 \Sigma_0 + \sigma_u^2}, \quad (\text{B8})$$

$$\lambda_2 = \frac{\left( \frac{1}{2\lambda_2} + \frac{\delta_s}{2} \beta_{v,1} \right) \frac{1}{\Sigma_0^{-1} + \beta_{v,1}^2 \sigma_u^{-2}}}{\left( \frac{1}{2\lambda_2} + \frac{\delta_s}{2} \beta_{v,1} \right)^2 \frac{1}{\Sigma_0^{-1} + \beta_{v,1}^2 \sigma_u^{-2}} + \delta_s^2 \sigma_\varepsilon^2 + \sigma_u^2}. \quad (\text{B9})$$

Note that equation (B8) is the expression of  $\lambda_1$  in Proposition 2.

The idea to compute the system characterizing a pure strategy equilibrium is to simplify it to a system in terms of  $(\lambda_1, \lambda_2, \beta_{v,1}, \delta_s)$  and then characterize this simplified system as a single equation of  $\beta_{v,1}$ .

If we know  $(\lambda_1, \lambda_2, \delta_s)$ , then  $\delta_{y_1}$  is known by equation (19), and  $\beta_{y_1}, \beta_{v,2}$ , and  $\beta_{x_1}$  are known by equation (14). Thus, the four unknowns  $(\lambda_1, \lambda_2, \beta_{v,1}, \delta_s)$  are determined by the remaining four equations, (24) and (B7)–(B9), and the two SOC's, (13) and (25).

Now, we simplify this four-equation system as a single equation of  $\beta_{v,1}$ . The idea is to express  $\lambda_1, \lambda_2 \delta_s$  and  $\lambda_2$  as functions of  $\beta_{v,1}$ , and then insert these expressions into equation (24). By (B7),

$$\lambda_2 \delta_s = \frac{\beta_{v,1} \sigma_u^2 \Sigma_0}{4\sigma_u^2 \sigma_\varepsilon^2 + 3\beta_{v,1}^2 \Sigma_0 \sigma_u^2 + 4\beta_{v,1}^2 \Sigma_0 \sigma_\varepsilon^2}. \quad (\text{B10})$$

By (B9),

$$\lambda_2 = \sigma_u^{-1} \sqrt{\left( \frac{1}{2} + \frac{\lambda_2 \delta_s}{2} \beta_{v,1} \right) \frac{1}{\Sigma_0^{-1} + \beta_{v,1}^2 \sigma_u^{-2}} - \left( \frac{1}{2} + \frac{\lambda_2 \delta_s}{2} \beta_{v,1} \right)^2 \frac{1}{\Sigma_0^{-1} + \beta_{v,1}^2 \sigma_u^{-2}} - (\lambda_2 \delta_s)^2 \sigma_\varepsilon^2}.$$

Inserting (B10) into the above expression, we obtain

$$\lambda_2^2 = \Sigma_0 \frac{(2\sigma_u^4 + 4\sigma_\varepsilon^4 + 5\sigma_u^2 \sigma_\varepsilon^2) \Sigma_0^2 \beta_{v,1}^4 + (8\sigma_u^2 \sigma_\varepsilon^4 + 5\sigma_u^4 \sigma_\varepsilon^2) \Sigma_0 \beta_{v,1}^2 + 4\sigma_u^4 \sigma_\varepsilon^4}{(\beta_{v,1}^2 \Sigma_0 + \sigma_u^2) (3\sigma_u^2 \Sigma_0 \beta_{v,1}^2 + 4\sigma_\varepsilon^2 \Sigma_0 \beta_{v,1}^2 + 4\sigma_u^2 \sigma_\varepsilon^2)^2}, \quad (\text{B11})$$

which gives the expression of  $\lambda_2$  in Proposition 2.

We can rewrite equation (24) as

$$2\lambda_2(2\beta_{v,1}\lambda_1 - 1) = \left[ \beta_{v,1} \left( \frac{2}{3}\lambda_1 + \lambda_2\delta_s \left( 1 + \frac{4\sigma_\varepsilon^2}{3\sigma_u^2} \right) \right) - 1 \right] \times \left[ \frac{2}{3}\lambda_1 + \lambda_2\delta_s \left( 1 + \frac{4\sigma_\varepsilon^2}{3\sigma_u^2} \right) \right]. \quad (\text{B12})$$

We then want to take square on both sides of (B12) in order to use (B11) to substitute  $\lambda_2^2$ . Doing this requires that the terms  $2\beta_{v,1}\lambda_1 - 1$  and  $\beta_{v,1} \left( \frac{2}{3}\lambda_1 + \lambda_2\delta_s \left( 1 + \frac{4\sigma_\varepsilon^2}{3\sigma_u^2} \right) \right) - 1$  have the same sign, that is,

$$(2\beta_{v,1}\lambda_1 - 1) \left[ \beta_{v,1} \left( \frac{2}{3}\lambda_1 + \lambda_2\delta_s \left( 1 + \frac{4\sigma_\varepsilon^2}{3\sigma_u^2} \right) \right) - 1 \right] \geq 0.$$

Inserting the expression of  $\lambda_1$  and  $\lambda_2\delta_s$  in (B8) and (B10) into the above condition, we find that the above inequality is equivalent to requiring

$$\beta_{v,1} \leq \frac{\sigma_u}{\sqrt{\Sigma_0}}.$$

Thus, given  $\beta_{v,1} \leq \frac{\sigma_u}{\sqrt{\Sigma_0}}$ , we can take square of (B12), and set

$$4\lambda_2^2(2\beta_{v,1}\lambda_1 - 1)^2 - \left[ \beta_{v,1} \left( \frac{2}{3}\lambda_1 + \lambda_2\delta_s \left( 1 + \frac{4\sigma_\varepsilon^2}{3\sigma_u^2} \right) \right) - 1 \right]^2 \times \left[ \frac{2}{3}\lambda_1 + \lambda_2\delta_s \left( 1 + \frac{4\sigma_\varepsilon^2}{3\sigma_u^2} \right) \right]^2 = 0.$$

Inserting the expression of  $\lambda_1$ ,  $\lambda_2\delta_s$  and  $\lambda_2^2$  in (B8), (B10), and (B11) into the above equation, we have the 7<sup>th</sup> order polynomial of  $\beta_{v,1}^2$  as follows:

$$f(\beta_{v,1}^2) = A_7\beta_{v,1}^{14} + A_6\beta_{v,1}^{12} + A_5\beta_{v,1}^{10} + A_4\beta_{v,1}^8 + A_3\beta_{v,1}^6 + A_2\beta_{v,1}^4 + A_1\beta_{v,1}^2 + A_0 = 0, \quad (\text{B13})$$

where

$$A_7 = \Sigma_0^7 (2\sigma_u^4 + 4\sigma_\varepsilon^4 + 5\sigma_u^2\sigma_\varepsilon^2) (3\sigma_u^2 + 4\sigma_\varepsilon^2)^2, \quad (\text{B14})$$

$$A_6 = 2\Sigma_0^6\sigma_u^2 (4\sigma_\varepsilon^2 - \sigma_u^2) (3\sigma_u^2 + 4\sigma_\varepsilon^2) (3\sigma_u^4 + 6\sigma_\varepsilon^4 + 8\sigma_u^2\sigma_\varepsilon^2), \quad (\text{B15})$$

$$A_5 = -\Sigma_0^5\sigma_u^6 (27\sigma_u^6 + 336\sigma_\varepsilon^6 + 524\sigma_u^2\sigma_\varepsilon^4 + 246\sigma_u^4\sigma_\varepsilon^2), \quad (\text{B16})$$

$$A_4 = 4\Sigma_0^4\sigma_u^6 (3\sigma_u^8 - 144\sigma_\varepsilon^8 - 304\sigma_u^2\sigma_\varepsilon^6 - 182\sigma_u^4\sigma_\varepsilon^4 - 23\sigma_u^6\sigma_\varepsilon^2), \quad (\text{B17})$$

$$A_3 = -\Sigma_0^3\sigma_u^8 (\sigma_u^8 + 704\sigma_\varepsilon^8 + 752\sigma_u^2\sigma_\varepsilon^6 + 76\sigma_u^4\sigma_\varepsilon^4 - 57\sigma_u^6\sigma_\varepsilon^2), \quad (\text{B18})$$

$$A_2 = -4\Sigma_0^2\sigma_u^{10}\sigma_\varepsilon^2 (\sigma_u^6 + 48\sigma_\varepsilon^6 - 24\sigma_u^2\sigma_\varepsilon^4 - 31\sigma_u^4\sigma_\varepsilon^2), \quad (\text{B19})$$

$$A_1 = -4\Sigma_0\sigma_u^{12}\sigma_\varepsilon^4 (\sigma_u^4 - 32\sigma_\varepsilon^4 - 36\sigma_u^2\sigma_\varepsilon^2), \quad (\text{B20})$$

$$A_0 = 64\sigma_u^{14}\sigma_\varepsilon^8. \quad (\text{B21})$$

The final requirement is to ensure that a root to the polynomial also satisfies the two SOC's, (13) and (25). Given the expression of  $\lambda_2$  in Proposition 2, (13) is redundant. Also, (25) implies  $\beta_{v,1} > 0$ , because (25) implies  $\lambda_1 > 0$ , which by (B8), in turn implies  $\beta_{v,1} > 0$ . So, the final constraint on  $\beta_{v,1}$  is  $0 < \beta_{v,1} \leq \frac{\sigma_u}{\sqrt{\Sigma_0}}$  and condition (25).

## B.4 Proof of Proposition 3

**When  $\sigma_\varepsilon$  is small:** By Proposition 1, when  $\sigma_\varepsilon$  is small, there is a mixed strategy equilibrium. The task is to show that there is no pure strategy equilibrium. By (24) and the fact  $\beta_{v,1} > 0$  in a pure strategy equilibrium, we have

$$\beta_{v,1} = \frac{1 - \frac{\lambda_1 + \lambda_2 \delta_s - \lambda_2 \delta_{y_1}}{2\lambda_2}}{2 \left[ \lambda_1 - \frac{(\lambda_1 + \lambda_2 \delta_s - \lambda_2 \delta_{y_1})^2}{4\lambda_2} \right]} > 0. \quad (\text{B22})$$

Note that the denominator is the SOC in (25), which is positive. So, we must have

$$1 - \frac{\lambda_1 + \lambda_2 \delta_s - \lambda_2 \delta_{y_1}}{2\lambda_2} > 0 \Rightarrow 4\lambda_2^2 - (\lambda_1 + \lambda_2 \delta_s - \lambda_2 \delta_{y_1})^2 > 0.$$

Using (19) we can rewrite the above inequality as follows:

$$4\lambda_2^2 - \left( \frac{2}{3}\lambda_1 + \lambda_2 \delta_s \left( 1 + \frac{4\sigma_\varepsilon^2}{3\sigma_u^2} \right) \right)^2 > 0. \quad (\text{B23})$$

Plugging the expression of  $\lambda_1$ ,  $\lambda_2 \delta_s$ , and  $\lambda_2^2$  in (B8), (B10), and (B11) into the left-hand-side (LHS) of (B23), we find that (B23) is equivalent to

$$\begin{aligned} & (16\beta_{v,1}^4 \Sigma_0^2 + 32\beta_{v,1}^2 \Sigma_0 \sigma_u^2 + 16\sigma_u^4) \sigma_\varepsilon^4 + (8\beta_{v,1}^4 \Sigma_0^2 \sigma_u^2 - 4\beta_{v,1}^6 \Sigma_0^3 + 12\beta_{v,1}^2 \Sigma_0 \sigma_u^4) \sigma_\varepsilon^2 \\ & - \beta_{v,1}^2 \Sigma_0 \sigma_u^2 (\beta_{v,1}^2 \Sigma_0 - \sigma_u^2)^2 > 0. \end{aligned} \quad (\text{B24})$$

We prove that the above condition is not satisfied in a pure strategy equilibrium, as  $\sigma_\varepsilon \rightarrow 0$ . Proposition 2 implies that in a pure strategy equilibrium,  $\beta_{v,1}^2 \in \left( 0, \frac{\sigma_u^2}{\Sigma_0} \right]$ . So, as  $\sigma_\varepsilon \rightarrow 0$ , the first two terms of the LHS of (B24) go to 0. Thus, if as  $\sigma_\varepsilon \rightarrow 0$ ,  $\beta_{v,1}^2$  does not go to 0 or  $\frac{\sigma_u^2}{\Sigma_0}$  in a pure strategy equilibrium, then the third term of the LHS of (B24) is strictly negative, which proves our statement.

Now we consider the two cases that  $\beta_{v,1}^2$  converges to 0 or to  $\frac{\sigma_u^2}{\Sigma_0}$  as  $\sigma_\varepsilon \rightarrow 0$ . We will show that both lead to contradictions to a pure strategy equilibrium.

Note that if  $\sigma_\varepsilon = 0$ , the polynomial (B13) is negative at  $\frac{\sigma_u^2}{\Sigma_0}$ ; that is,  $f\left(\frac{\sigma_u^2}{\Sigma_0}\right) = -16\sigma_u^{22} < 0$  if  $\sigma_\varepsilon = 0$ . Thus, if for any sequence of  $\sigma_\varepsilon \rightarrow 0$ , we have  $\beta_{v,1}^2 \rightarrow \frac{\sigma_u^2}{\Sigma_0}$  in a pure strategy equilibrium, then we must have  $f(\beta_{v,1}^2) \rightarrow -16\sigma_u^{22} < 0$ , which contradicts with Proposition 2 which says that  $f(\beta_{v,1}^2) \equiv 0$  in a pure strategy equilibrium. Thus,  $\beta_{v,1}^2 \not\rightarrow \frac{\sigma_u^2}{\Sigma_0}$  as  $\sigma_\varepsilon \rightarrow 0$ .

Suppose  $\beta_{v,1}^2 \rightarrow 0$  in a pure strategy equilibrium for some sequence of  $\sigma_\varepsilon^2 \rightarrow 0$ . By (24), we have

$$\left( \frac{2}{3}\lambda_1 + \lambda_2 \delta_s \left( 1 + \frac{4\sigma_\varepsilon^2}{3\sigma_u^2} \right) \right)^2 > \left( 1 - 2 \frac{\beta_1 \Sigma_0}{\beta_1^2 \Sigma_0 + \sigma_u^2} \beta_1 \right)^2 4\lambda_2^2.$$

Combining the above condition with condition (B23), we know  $\left(\frac{2}{3}\lambda_1 + \lambda_2\delta_s \left(1 + \frac{4\sigma_\varepsilon^2}{3\sigma_u^2}\right)\right)^2$  has the same order as  $\lambda_2^2$ :

$$O\left(\left(\frac{2}{3}\lambda_1 + \lambda_2\delta_s \left(1 + \frac{4\sigma_\varepsilon^2}{3\sigma_u^2}\right)\right)^2\right) = O(\lambda_2^2).$$

Substituting into the above equation the expression of  $\lambda_1$ ,  $\lambda_2\delta_s$ , and  $\lambda_2^2$  from (B8), (B10), and (B11) and matching the highest-order terms, we can show that  $\beta_{v,1}^2$  has the same order as  $\sigma_\varepsilon^4$ . As a result, by (B8),  $\lambda_1 \rightarrow 0$ ; by (B10),  $\lambda_2\delta_s$  goes to a positive finite number; and by (B11),  $\lambda_2$  goes to a positive finite number. This in turn implies the SOC (25) is violated. Specifically, by (19), the SOC is equivalent to

$$\lambda_1 - \frac{\left(\frac{2}{3}\lambda_1 + \lambda_2\delta_s \left(1 + \frac{4\sigma_\varepsilon^2}{3\sigma_u^2}\right)\right)^2}{4\lambda_2} > 0. \quad (\text{B25})$$

However, as  $\sigma_\varepsilon^2 \rightarrow 0$ , we have  $\lambda_1 - \frac{\left(\frac{2}{3}\lambda_1 + \lambda_2\delta_s \left(1 + \frac{4\sigma_\varepsilon^2}{3\sigma_u^2}\right)\right)^2}{4\lambda_2} \rightarrow -\frac{(\lambda_2\delta_s)^2}{4\lambda_2} < 0$ , a contradiction.

**When  $\sigma_\varepsilon$  is large:** By Proposition 1, when  $\sigma_\varepsilon$  is sufficiently large, there is no mixed strategy equilibrium. The task is to show that a linear pure strategy equilibrium exists and is unique.

By equations (B14)–(B21), we have  $A_7 > 0$ ,  $A_6 > 0$ ,  $A_5 < 0$ ,  $A_4 < 0$ ,  $A_3 < 0$ ,  $A_2 < 0$ ,  $A_1 > 0$ , and  $A_0 > 0$ , when  $\sigma_\varepsilon^2$  is sufficiently large. Thus, by Descartes' Rule of Signs, there are at most two positive roots of  $\beta_{v,1}^2$ .

By equation (B13), we have

$$\begin{aligned} f(0) &= 64\sigma_u^{14}\sigma_\varepsilon^8 > 0, \\ \lim_{\beta_{v,1}^2 \rightarrow \infty} f(\beta_{v,1}^2) &\propto \Sigma_0^7 (2\sigma_u^4 + 4\sigma_\varepsilon^4 + 5\sigma_u^2\sigma_\varepsilon^2) (3\sigma_u^2 + 4\sigma_\varepsilon^2)^2 \times \infty > 0. \end{aligned}$$

In addition, as  $\sigma_\varepsilon^2 \rightarrow \infty$ ,  $f\left(\frac{\sigma_u^2}{\Sigma_0}\right) \propto -1024\sigma_u^{14}\sigma_\varepsilon^8 < 0$ . So, there is exactly one root of  $\beta_{v,1}^2$  in the range of  $\left(0, \frac{\sigma_u^2}{\Sigma_0}\right)$  and one root in the range of  $\left(\frac{\sigma_u^2}{\Sigma_0}, \infty\right)$ . Given that in a pure strategy equilibrium, we require  $0 < \beta_{v,1}^2 \leq \frac{\sigma_u}{\Sigma_0}$  by Proposition 2, only the small root is a possible equilibrium candidate (which is indeed an equilibrium if the SOC is also satisfied).

Finally, we can show that the small root of  $\beta_{v,1}^2 \in \left(0, \frac{\sigma_u^2}{\Sigma_0}\right)$  satisfies the SOC as  $\sigma_\varepsilon^2 \rightarrow \infty$ . Specifically, by (B25), the SOC is

$$\lambda_1 - \frac{\left(\frac{2}{3}\lambda_1 + \lambda_2\delta_s \left(1 + \frac{4\sigma_\varepsilon^2}{3\sigma_u^2}\right)\right)^2}{4\lambda_2} > 0 \Leftrightarrow$$

$$16\lambda_2^2\lambda_1^2 - \left( \frac{2}{3}\lambda_1 + \lambda_2\delta_s \left( 1 + \frac{4\sigma_\varepsilon^2}{3\sigma_u^2} \right) \right)^4 > 0.$$

Plugging the expression of  $\lambda_1$ ,  $\lambda_2\delta_s$ , and  $\lambda_2^2$  in (B8), (B10) and (B11) into the LHS of the above condition, we can show that the above condition holds if and only if

$$B_4\sigma_\varepsilon^8 + B_3\sigma_\varepsilon^6 + B_2\sigma_\varepsilon^4 + B_1\sigma_\varepsilon^2 + B_0 > 0, \quad (\text{B26})$$

where,

$$\begin{aligned} B_4 &= 768\beta_{v,1}^{10}\Sigma_0^5 + 4096\beta_{v,1}^8\Sigma_0^4\sigma_u^2 + 8704\beta_{v,1}^6\Sigma_0^3\sigma_u^4 + 9216\beta_{v,1}^4\Sigma_0^2\sigma_u^6 + 4864\beta_{v,1}^2\Sigma_0\sigma_u^8 + 1024\sigma_u^{10}, \\ B_3 &= 2048\beta_{v,1}^{10}\Sigma_0^5\sigma_u^2 + 8704\beta_{v,1}^8\Sigma_0^4\sigma_u^4 + 13824\beta_{v,1}^6\Sigma_0^3\sigma_u^6 + 9728\beta_{v,1}^4\Sigma_0^2\sigma_u^8 + 2560\beta_{v,1}^2\Sigma_0\sigma_u^{10}, \\ B_2 &= 2144\beta_{v,1}^{10}\Sigma_0^5\sigma_u^4 + 6720\beta_{v,1}^8\Sigma_0^4\sigma_u^6 + 6912\beta_{v,1}^6\Sigma_0^3\sigma_u^8 + 2240\beta_{v,1}^4\Sigma_0^2\sigma_u^{10} - 96\beta_{v,1}^2\Sigma_0\sigma_u^{12}, \\ B_1 &= 1056\beta_{v,1}^{10}\Sigma_0^5\sigma_u^6 + 2112\beta_{v,1}^8\Sigma_0^4\sigma_u^8 + 912\beta_{v,1}^6\Sigma_0^3\sigma_u^{10} - 160\beta_{v,1}^4\Sigma_0^2\sigma_u^{12} - 16\beta_{v,1}^2\Sigma_0\sigma_u^{14}, \\ B_0 &= 207\beta_{v,1}^{10}\Sigma_0^5\sigma_u^8 + 180\beta_{v,1}^8\Sigma_0^4\sigma_u^{10} - 54\beta_{v,1}^6\Sigma_0^3\sigma_u^{12} - 12\beta_{v,1}^4\Sigma_0^2\sigma_u^{14} - \beta_{v,1}^2\Sigma_0\sigma_u^{16}. \end{aligned}$$

Given that  $\beta_{v,1}$  is bounded, we have that as  $\sigma_\varepsilon^2$  is large, the LHS of condition (B26) is determined by  $B_4\sigma_\varepsilon^8$ , which is always positive:  $B_4\sigma_\varepsilon^8 > 1024\sigma_u^{10}\sigma_\varepsilon^8 > 0$ .

## B.5 Proof of Corollary 1

Now suppose  $\sigma_\varepsilon \rightarrow \infty$ . By Proposition 3, as  $\sigma_\varepsilon$  is large, there is a unique linear equilibrium, which is a pure strategy equilibrium. In a pure strategy equilibrium, we always have  $f(\beta_{v,1}^2) = 0$ . If we rewrite the polynomial  $f$  as a polynomial in terms of  $\sigma_\varepsilon$ , we must have that as  $\sigma_\varepsilon \rightarrow \infty$ , the coefficients on the highest order of  $\sigma_\varepsilon$  goes to 0. This exercise yields the following condition that as  $\sigma_\varepsilon \rightarrow \infty$ , we have

$$64\Sigma_0^7\beta_{v,1}^{14} + 192\Sigma_0^6\sigma_u^2\beta_{v,1}^{12} - 576\Sigma_0^4\sigma_u^6\beta_{v,1}^8 - 704\Sigma_0^3\sigma_u^8\beta_{v,1}^6 - 192\Sigma_0^2\sigma_u^{10}\beta_{v,1}^4 + 128\Sigma_0\sigma_u^{12}\beta_{v,1}^2 + 64\sigma_u^{14} \rightarrow 0. \quad (\text{B27})$$

Define  $x \equiv \beta_{v,1}^2 \frac{\Sigma_0}{\sigma_u^2} \in [0, 1]$  in a pure strategy equilibrium. Condition (B27) becomes

$$-2x - x^2 + x^3 + 1 \rightarrow 0, \text{ as } \sigma_\varepsilon \rightarrow \infty. \quad (\text{B28})$$

That is, as  $\sigma_\varepsilon \rightarrow \infty$ , we must have that (B28) holds.

In a standard Kyle setting, the unique equilibrium is defined by

$$-2x^* - x^{*2} + x^{*3} + 1 = 0. \quad (\text{B29})$$

Specifically, Proposition 1 of Huddart, Hughes, and Levine (2001) characterizes the equilibrium in a two-period Kyle model by a cubic in terms of  $K$ ,

$$8K^3 - 4K^2 - 4K + 1 = 0. \quad (\text{B30})$$



By the expressions of  $\beta_1 = \frac{2K-1}{4K-1} \frac{1}{\lambda_1}$  and  $\lambda_1 = \frac{\sqrt{2K(2K-1)}}{4K-1} \frac{\sqrt{\Sigma_0}}{\sigma_u}$  in Proposition 1 of Huddart, Hughes, and Levine (2001), we have  $K = \frac{1}{2(1-x^*)}$ , where  $x^* = \beta_1^2 \frac{\Sigma_0}{\sigma_u^2}$ . Then, equation (B30) is equivalent to equation (B29). Given that  $-2x - x^2 + x^3 + 1$  is monotone and continuous in the range of  $[0, 1]$ , we have  $x \rightarrow x^*$  as  $\sigma_\varepsilon \rightarrow \infty$ , by conditions (B28) and (B29).

## B.6 Proof of Proposition 4

We here give the expression of the variables in the proposition. The comparison follows from setting  $\sigma_\varepsilon = 0$  and  $\sigma_\varepsilon = \infty$  in these expressions and from straightforward computations.

Setting  $\sigma_\varepsilon = 0$  in Proposition 1 yields  $\sigma_z = \sqrt{\frac{1}{39}} \sigma_u$  and  $\beta_{v,1} = \frac{2}{\sqrt{13}} \frac{\sigma_u}{\sqrt{\Sigma_0}}$ . Plugging these two expressions of  $\sigma_z$  and  $\beta_{v,1}$  into the expressions of  $\lambda_1$  and  $\lambda_2$  in Proposition 1 gives  $\lambda_1^0$  and  $\lambda_2^0$ . In a pure strategy equilibrium,  $\lambda_1$  and  $\lambda_2$  are given by Proposition 2. Setting  $\sigma_\varepsilon = \infty$  yields the expressions of  $\lambda_1^{Kyle}$  and  $\lambda_2^{Kyle}$ .

Direct computation shows that in a mixed strategy equilibrium, the price discovery variables are given by

$$\begin{aligned} \Sigma_1^{mixed} &= \frac{(\sigma_z^2 + \sigma_u^2) \Sigma_0}{\beta_{v,1}^2 \Sigma_0 + \sigma_z^2 + \sigma_u^2}, \\ \Sigma_2^{mixed} &= \frac{(4\sigma_u^4 + 4\sigma_u^2 \sigma_z^2 + \sigma_u^2 \delta_s^2 \sigma_z^2 + 4\sigma_u^2 \delta_s^2 \sigma_\varepsilon^2 + 4\delta_s^2 \sigma_z^2 \sigma_\varepsilon^2) \lambda_2^2 \Sigma_0}{\left( (\Sigma_0 \lambda_2^2 \sigma_u^2 \delta_s^2 + 4\Sigma_0 \lambda_2^2 \sigma_u^2 + 4\Sigma_0 \lambda_2^2 \delta_s^2 \sigma_\varepsilon^2) \beta_{v,1}^2 + 2\lambda_2 \Sigma_0 \sigma_u^2 \delta_s \beta_{v,1} \right. \\ &\quad \left. + \lambda_2^2 (4\sigma_u^4 + 4\sigma_u^2 \sigma_z^2 + \sigma_u^2 \delta_s^2 \sigma_z^2 + 4\sigma_u^2 \delta_s^2 \sigma_\varepsilon^2 + 4\delta_s^2 \sigma_z^2 \sigma_\varepsilon^2) + \Sigma_0 (\sigma_u^2 + \sigma_z^2) \right)}. \end{aligned}$$

Plugging  $\sigma_\varepsilon = 0$ ,  $\sigma_z = \sqrt{\frac{1}{39}} \sigma_u$ , and  $\beta_{v,1} = \frac{2}{\sqrt{13}} \frac{\sigma_u}{\sqrt{\Sigma_0}}$  into the above expressions yields  $\Sigma_1^0$  and  $\Sigma_2^0$ . In a pure strategy equilibrium, we can compute

$$\begin{aligned} \Sigma_1^{pure} &= \frac{\sigma_u^2 \Sigma_0}{\sigma_u^2 + \beta_{v,1}^2 \Sigma_0}, \\ \Sigma_2^{pure} &= \frac{1}{\Sigma_0^{-1} + \beta_{v,1}^2 \sigma_u^{-2} + \left( \frac{1}{2\lambda_2} + \frac{\delta_s}{2} \beta_{v,1} \right)^2 (\delta_s^2 \sigma_\varepsilon^2 + \sigma_u^2)^{-1}}. \end{aligned}$$

Setting  $\sigma_\varepsilon = \infty$  in Proposition 2, computing  $\beta_{v,1}$ ,  $\lambda_2$  and  $\delta_s$ , and inserting these expressions into the above expressions, we have the expressions of  $\Sigma_1^{Kyle}$  and  $\Sigma_2^{Kyle}$ .

Finally, we present the profit expressions. By (15), the fundamental investor's *ex-ante* expected period-2 profit is

$$E(\pi_{F,2}) = \frac{E[v - p_1 - \lambda_2(-\delta_y y_1 + \delta_s x_1)]^2}{4\lambda_2}.$$

Using equations (5) and (8), we can show

$$E(\pi_{F,2}) = \frac{[1 - (\lambda_1 - \lambda_2\delta_y + \lambda_2\delta_s)\beta_{v,1}]^2 \Sigma_0 + (\lambda_1 - \lambda_2\delta_y)^2 \sigma_u^2 + (\lambda_1 - \lambda_2\delta_y + \lambda_2\delta_s)^2 \sigma_z^2}{4\lambda_2}.$$

For  $E(\pi_{F,2}^0)$ , we set  $\sigma_\varepsilon = 0$ , compute  $\sigma_z, \beta_{v,1}, \lambda_1, \lambda_2, \delta_s$ , and  $\delta_y$  in Proposition 1, and insert these expressions into the above equation. For  $E(\pi_{F,2}^{Kyle})$ , we set  $\sigma_\varepsilon = \infty$  in Proposition 2 and compute the relevant parameters accordingly.

To give the expression of  $E(\pi_{F,1})$ , we first compute  $E(\pi_{F,1} + \pi_{F,2})$ , and then use the above computed  $E(\pi_{F,2})$  to compute  $E(\pi_{F,1}) = E(\pi_{F,1} + \pi_{F,2}) - E(\pi_{F,2})$ . Using equation (22), we can show that in a mixed strategy equilibrium,

$$E(\pi_{F,1}^{mixed} + \pi_{F,2}^{mixed}) = \frac{\Sigma_0 + \sigma_u^2 \lambda^2 (1 - \delta_y)^2}{4\lambda}$$

while in a pure strategy equilibrium

$$E(\pi_{F,1}^{pure} + \pi_{F,2}^{pure}) = \frac{\left(1 - \frac{\lambda_1 + \lambda_2\delta_s - \lambda_2\delta_y}{2\lambda_2}\right)^2}{4\left(\lambda_1 - \frac{(\lambda_1 + \lambda_2\delta_s - \lambda_2\delta_y)^2}{4\lambda_2}\right)} \Sigma_0 + \frac{\Sigma_0 + \sigma_u^2 (\lambda_1 - \lambda_2\delta_y)^2}{4\lambda_2}.$$

Then, setting  $\sigma_\varepsilon = 0$  and  $\sigma_\varepsilon = \infty$  in the above two expressions gives  $E(\pi_{F,1}^0 + \pi_{F,2}^0)$  and  $E(\pi_{F,1}^{Kyle} + \pi_{F,2}^{Kyle})$ , respectively.

## References

- ADMATI, A. AND P. PFLEIDERER (1989): “Divide and conquer: A theory of intraday and day-of-the-week mean effects.” *Review of Financial Studies*, 2, 189–223.
- ATTARI, M., A. S. MELLO, AND M. E. RUCKES (2005): “Arbitraging Arbitrageurs,” *Journal of Finance*, 60, 2471–2511.
- BACK, K. AND S. BARUCH (2004): “Information in Securities Markets: Kyle Meets Glosten and Milgrom,” *Econometrica*, 72, 433–465.
- BARUCH, S. AND L. R. GLOSTEN (2013): “Fleeting Orders,” Working paper.
- BERNHARDT, D. AND B. TAUB (2008): “Front-Running Dynamics,” *Journal of Economic Theory*, 138, 288–296.
- BOND, P., A. EDMANS, AND I. GOLDSTEIN (2012): “The Real Effects of Financial Markets,” *Annual Reviews of Financial Economics*, 4, 339–360.
- BOND, P., I. GOLDSTEIN, AND E. S. PRESCOTT (2010): “Market-Based Corrective Actions,” *Review of Financial Studies*, 23, 781–820.
- BRUNNERMEIER, M. K. AND L. H. PEDERSEN (2005): “Predatory Trading,” *Journal of Finance*, 60, 1825–1863.
- BUFFA, A. (2013): “Insider Trade Disclosure, Market Efficiency, and Liquidity,” Working paper.
- CAO, H. H., M. D. EVANS, AND R. K. LYONS (2006): “Inventory Information,” *Journal of Business*, 79, 325–364.
- CARLIN, B. I., M. S. LOBO, AND S. VISWANATHAN (2007): “Episodic Liquidity Crisis: Cooperative and Predatory Trading,” *Journal of Finance*, 62, 2235–2274.
- ENGELBERG, J. E., A. V. REED, AND M. C. RINGGENBERG (2012): “How are Shorts Informed? Short Sellers, News, and Information Processing,” *Journal of Financial Economics*, 105, 260–278.
- GLOSTEN, L. R. AND P. R. MILGROM (1985): “Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders,” *Journal of Financial Economics*, 14, 71–100.
- HARRIS, L. (2003): *Trading and Exchanges: Market Microstructure for Practitioners*, Oxford University Press.
- (2013): “What to Do about High-Frequency Trading,” *Financial Analysts Journal*, 69, 6–9.

- HIRSCHEY, N. (2013): “Do High-Frequency Traders Anticipate Buying and Selling Pressure?” Working paper.
- HUDDART, S., J. S. HUGHES, AND C. B. LEVINE (2001): “Public Disclosure and Dissimulation of Insider Trades,” *Econometrica*, 69, 665–681.
- KIM, O. AND R. E. VERRECCHIA (1994): “Market Liquidity and Volume around Earnings Announcements,” *Journal of Accounting and Economics*, 17, 41–67.
- KORAJCZYK, R. AND D. MURPHY (2014): “High Frequency Market Making to Large Institutional Trades,” Working paper.
- KYLE, A. S. (1985): “Continuous Auctions and Insider Trading,” *Econometrica*, 53, 1315–1335.
- LEWIS, M. (2014): *Flash Boys*, Norton.
- LI, W. (2014): “High Frequency Trading with Speed Hierarchies,” Working paper.
- MADRIGAL, V. (1996): “Non-Fundamental Speculation,” *Journal of Finance*, 51, 553–578.
- O’HARA, M. (2003): “Presidential Address: Liquidity and Price Discovery,” *Journal of Finance*, 58, 1335–1354.
- SECURITIES AND EXCHANGE COMMISSION (2010): “Concept Release on Equity Market Structure,” Technical report.
- STIGLITZ, J. (2014): “Tapping the Brakes: Are Less Active Markets Safer and Better for the Economy?” Speech at the 2014 Financial Markets Conference at the Federal Reserve Bank of Atlanta.
- TONG, L. (2015): “A Blessing or a Curse? The Impact of High Frequency Trading on Institutional Investors,” Working paper.
- VAN KERVEL, V. AND A. MENKVELD (2015): “High-Frequency Trading around Large Institutional Orders,” Working paper.
- VERRECCHIA, R. E. (1982): “Information Acquisition in a Noisy Rational Expectations Economy,” *Econometrica*, 50, 1415–1430.
- VIVES, X. (2008): *Information and Learning in Markets: The Impact of Market Microstructure*, Princeton University Press, Princeton and Oxford.