### Preliminary Program for the Stern Microstructure Meeting, Friday, May 10, 2013

Supporting funding is provided by NASDAQ OMX through a grant to the Salomon Center at Stern.

**Organizer**  
Joel Hasbrouck, Stern School

**Program Committee**  
Tarun Chordia, Goizueta School, Emory University  
Amit Goyal, HEC Lausanne  
Gideon Saar, Johnson School, Cornell University

The Stern Microstructure Conference is open to everyone with an interest in market microstructure research. The sessions will be held at the Management Education Center, 44 W. 4th St., NYC (near the southeast corner of Washington Square Park). For more complete directions see [http://www.stern.nyu.edu/AboutStern/VisitStern/index.htm](http://www.stern.nyu.edu/AboutStern/VisitStern/index.htm). There will be a registration desk in the lobby.

Registration Instructions: E-mail salomon@stern.nyu.edu with “SMC2013” in the subject line. Please indicate if you will be joining us for the dinner the night before. Other inquiries: jhasbrou@stern.nyu.edu.

Please note: Hard copies of the papers will not be available at the conference. A single pdf containing the schedule and all papers is available [here](http://www.stern.nyu.edu/AboutStern/VisitStern/index.htm).

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thursday, May 9</td>
<td></td>
</tr>
<tr>
<td>6:30 pm</td>
<td>Dinner (open to all registered conference attendees)</td>
</tr>
<tr>
<td>Friday, May 10</td>
<td></td>
</tr>
<tr>
<td>8:30 am - 9:00</td>
<td>Continental Breakfast</td>
</tr>
</tbody>
</table>
| 9:00 - 10:00 | **Informed Trading Before Unscheduled Corporate Announcements**  
Shmuel Baruch, University of Utah  
Marios Panayides, University of Pittsburgh  
Kumar Venkataraman, Southern Methodist University  
Discussant: Ingrid Werner, Fisher College, Ohio State University |
| 10:00 - 11:00 | **Adverse Selection and Intermediation Chains**  
(formerly: Intermediating Adverse Selection)  
Vincent Glode, Wharton  
Christian Opp, Wharton  
Discussant TBA |

---

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:00 - 11:15</td>
<td>Break</td>
</tr>
</tbody>
</table>
| 11:15 - 12:15| **Exploratory Trading**  
Adam Clark-Joseph, University of Illinois  
Discussant TBA |
| 12:15-1:15   | Lunch  
David Mechner, Pragma Trading                                         |
| 1:15-2:15    | **Competition of High-Frequency Market Makers and Market Quality**  
Johannes Breckenfelder, Stockholm School of Economics  
Discussant: Charles Jones, Columbia University |
| 2:15-3:15    | **Asset Pricing Frictions in Fragmented Markets**  
Emilian Pagnotta, NYU Stern  
Discussant TBA |
| 3:15-3:30    | Break                                                                |
| 3:30-4:30    | **Fleeting Orders and the Competitive Equilibrium**  
Shmuel Baruch, University of Utah  
Larry Glosten, Columbia  
Discussant: Ioanid Rosu, HEC |
| 4:30         | Adjourn                                                               |
Informed Trading before Unscheduled Corporate Events: Theory and Evidence

Abstract

Despite widespread evidence that informed agents are active before corporate events, there is little work describing how informed agents accumulate positions and what explains their trading strategies. We use the prisoners' dilemma to model the execution risk that informed traders impose on each other and explain why they forgo the price benefit of limit orders and use instead market orders. However the efficient limit-orders outcome is obtained if there is sufficient uncertainty about the presence of informed traders. We link the level of uncertainty to costly short selling and test theoretical predictions using detailed order level data from Euronext Paris. We find strong empirical support for the prediction that informed traders use limit orders when the news is negative, especially when (a) the investor base is not broad, (b) security borrowing costs are high, and (c) the magnitude of the event is small so potential profits cannot justify the cost of borrowing. When the news is positive, we show that informed buyers face more competition and use market orders. These results help explain the buy-sell asymmetry in price impact of trades and provide a framework for surveillance systems that are designed to detect insider trading.

This Draft: March 2013

Keywords: Price impact; Limit versus market orders; Buy-sell asymmetry; Insider Trading; Dark Pools.
1. Introduction

Insider trading has been a focus of regulatory efforts in recent years. A large academic finance literature has examined the open market stock trades of registered insiders (Seyhun (1992), Meulbroek (1992), Agrawal and Jaffe (1995), among others). These studies find that stock price tends to increase following insider purchase and decrease following insider sale and that insider trades are more profitable before corporate events, such as earnings announcements, SEOs, and earnings restatements. Related empirical work shows that institutional investors build profitable positions before corporate events.¹ Despite the widespread evidence suggesting that informed traders are active before corporate events, there is little work describing how informed agents accumulate positions and what explains their order submission strategies. In this study, we present a theoretical model based on the prisoners dilemma framework that builds on the idea that competition among informed agents impose execution risk on each other. We examine detailed order level data from Euronext Paris and document informed traders’ order submission strategies that are consistent with theoretical predictions.

The model is based on the following intuition. Informed traders face a tradeoff between transacting with certainty at a current market price by placing a market order versus risking non-execution in an attempt to get a better price by placing a limit order. In addition to worse price, market orders might tip off market participants about the presence of informed traders and result in high market impact cost. On the other hand, the non-execution risk of limit order strategy is particularly high when other informed agents use market orders and cause an adverse price move away from the limit price. In a multiple traders’ game, the joint decision of order submission of informed agents fit the prisoners’ dilemma framework, i.e., despite the price benefit of limit orders, informed traders use market orders. We posit the market order outcome when the nature of private information conveys an increase in stock price.

Informed traders face less competition when the nature of private information is unfavorable. This is because informed agents are less likely to sell stocks with unfavorable information if they do not already own the stock (see Saar (2001)). Therefore, when the nature of private information conveys a decline in stock price, informed sellers who are not “trade constrained” will face less competition from other traders. We show that informed traders use limit orders when there is sufficient uncertainty about the presence of other informed traders in the market. To model the uncertainty about the presence of other informed traders, we extend our model and assume that the informed agent can be one of two types; the first type already owns the stock while the second type does not. The probability that a trader is of the first type increases with the broadness of the investors’ base. We find that when the investors’ base is narrow, the costs of borrowing shares are sufficiently large, or the event is small so potential gains cannot justify the borrowing costs, a limit order equilibrium emerges in which the first type uses sell limit orders and the second type abstains from trade. On the other hand, when borrowing costs are sufficiently low or the event is sufficiently large, then the second type borrows the shares, and both types trade. Because of the execution risk they impose on each other, both types use market orders.

Surprisingly, despite the obvious importance of understanding how informed traders build positions, there is little systematic empirical evidence on their trading strategies, mainly because empirical studies need to overcome the lack of useful data. Some data on insider trades are available from regulatory filings, such as Form 4 filed with U.S. Securities and Exchange Commission (SEC); however, the data is not sufficiently detailed to study the order submission strategies of insiders. Detailed order level data is available from some markets, such as Euronext-Paris, but there is no information on trader identity. We employ a methodology similar in spirit to Chae (2005), Graham, Koski and Loewenstein (2006), and Sarkar and Schwartz (2009), who study information flow surrounding corporate events.

---

2 Studies on insider trading, such as Marin and Olivier (2008), observe that corporate insiders face more portfolio constraints when they trade on bad news than on good news. For example, insiders in many markets are prohibited from selling short their own stock, or corporate managers may be unable to sell stock holdings that are part of a compensation contract below a certain threshold. While insiders face constraints when they are in possession of bad news, they do not face threshold constraints on purchases when they are in possession of good news.

3 The prisoners’ dilemma analogy of this scenario occurs when there is a high likelihood that the accomplice has been released from custody for lack of evidence; i.e., the interrogator is bluffing.
announcements. These studies document that volume and liquidity decline before scheduled events, such as earnings announcements, because liquidity traders optimize the timing of their trades to minimize adverse selection risk. In contrast, because timing information is unavailable, liquidity traders do not change behavior before unscheduled events. Thus the abnormal activity observed before an unscheduled event, benchmarked against a non-event window for the firm, can be attributed to informed traders.

Using detailed Euronext data, we are able to uncover the order submission strategies of informed traders before unscheduled corporate announcements. We identify a sample of 95 French stocks and 101 unscheduled mergers, acquisitions, Seasoned Equity Offerings (SEOs), repurchases, and dividend initiations and terminations in the year 2003. These announcements convey a considerable amount of new information to the market – the absolute value of event day return for our sample exceeds 4.5% - but the timing of these announcements is not public information. Following our model’s predictions, we identify positive and negative events based on the announcement returns. We examine several observable attributes of the traders’ order submission strategy including price aggressiveness and the decision to hide order size. The event study framework implemented by the study holds each firm as its own control and avoids concerns associated with omitted cross-sectional determinants of trading strategies. Nonetheless, in our main specifications, we control for order characteristics and market conditions, including the state of the limit order book, following Bessembinder, Panayides and Venkataraman (BPV hereafter, 2009).

The key testable prediction from the model is that informed traders employ more aggressive strategies preceding positive events and less aggressive strategies preceding negative events. We find strong empirical support for these predictions. Specifically, preceding positive events, we document an increase in aggressively priced buy orders, and further an increase in the magnitude of limit order size that is not hidden. Both using aggressively priced orders and exposing order size lower the probability of non-execution and reduces the time-to execution; however, using aggressively priced orders signal the

---

4 During our sample period, the vast majority of trading in sample stocks occurred on Euronext-Paris. The consolidated market structure allows for a clearer examination of order submission strategies since there is no need to explicitly model the choice of trading in alternative trading venues. Similar to U.S. equity markets, trading in the Euronext-listed stocks has become highly fragmented in more recent years.
presence of informed agents with positive private information and increase the opportunity cost of non-execution. In contrast, preceding negative events, we observe a *decrease* in aggressively priced sell orders, suggesting that informed sellers use passive limit orders to build positions. We find that limit sell orders before negative events are less likely to be hidden and expose more order size. The exposure strategy increases the execution probability and lowers the time-to-execution by attracting counterparties. The latter evidence is consistent with theoretical framework in Moinas (2006), where informed traders supply liquidity using limit orders in an attempt to mimic the behavior of uninformed liquidity suppliers.

We develop further cross-sections tests of the model based on the idea that informed sellers face more competition before negative events if the stock is easy to borrow. Prior research has shown institutional ownership is higher in stocks with index membership. We therefore classify stocks with SBF120 Index membership as being associated with broad investor base and low borrowing costs (D’Avolio (2002), Nagel (2005)). For stocks with listed options, informed agents can avoid costly security borrowing by taking positions in the options market. We therefore classify stocks with listed option as being less expensive to short. Informed sellers in stocks with index membership or listed options will thus face more competition from other traders. The model predicts that, under these conditions, both informed buyers and sellers will implement aggressive strategies. For these stocks, we document an increase in price aggressiveness *symmetrically* for both buy orders before positive events and sell orders before negative events.

For stocks with no index membership or listed options, we observe *an asymmetry* in order submission strategies. Informed buyers implement more price aggressive strategies while informed sellers implement less price aggressive strategies, consistent with model predictions. We show that, although strategies differ, informed agents reduce the limit order time-to-execution for all sub-samples. That informed agents achieve similar outcomes while implementing different trading strategies provides further support for the theoretical model. The asymmetry in order submission strategies before positive and negative events is larger when the magnitude of announcement return is small as compared to events.
when announcement return is large. This is supportive of model predictions that informed sellers face less competition before small events since the potential gains cannot justify the security borrowing costs.

Since informed agents can choose among many potential trading strategies, our model predicts that the choice of a specific strategy reflects an optimization across many dimensions of execution quality, including non-execution risk, time-to-execution, and implementation costs. We further investigate to what extent informed trader strategies affect the price patterns observed in the cross-section of stocks. Focusing on opportunity cost of non-execution, we find that informed buyers experience an adverse drift in stock price before positive events while informed sellers do not before negative events. However, for index stocks or those with listed options, both informed buyers and informed sellers experience an adverse move in stock price. In contrast, for non-index stocks or those without listed options, we observe an adverse price move for buy orders before positive events but not for sell orders before negative events. These patterns in cross-sectional variation suggest that the uncertainty in trader competition influences the drift in stock price. Specifically, informed agents use “passive” trading strategies when trader competition is low and aggressive strategies when trader competition is high. These trading strategies tip off market participants to varying degree about the presence of informed agents and these variations in information leakage influence price adjustments before corporate events.

Our paper contributes to the literature on the price impact asymmetry. Past studies on block trading indicate that security prices reverse after block sale but not after block purchase, suggesting that block purchase conveys more information than block sale. As in Diamond and Verrecchia (1987), our theory uses costly short selling to match the price impact asymmetry observed in the data. However, in our model the asymmetry is not only because some informed traders decide to abstain, but also because informed agents become liquidity providers; i.e. they use limit orders based on the uncertainty about the presence of other informed traders. An important related paper by Saar (2001) takes a different approach.

---

than ours on the buy-sell asymmetry. In his model, informed traders face capital constraints rather than costly short selling. Thus, to finance investment in undervalued securities, informed traders sell securities that are priced correctly; i.e. informed traders may sell for liquidity reasons, which is a source of the price impact buy-sell asymmetry.

Our study extends the growing theoretical literature on informed trader strategies. Most theoretical work posits that informed agents trade aggressively in order to exploit their information advantage. On the other hand, in Kumar and Seppi (1994), Chakravarty and Holden (1995), Kaniel and Liu (2006), Goettler, Parlour and Rajan (2009), and Boulatov and George (2013), informed traders do find it optimal, under certain conditions, to submit limit orders. An important insight from our study is that no single trading pattern sufficiently describes the optimal strategy of informed agents. We suggest that designers of surveillance systems, including exchanges, broker-dealers, and regulators should first specify the ex-ante optimal strategies of insiders based on our model’s framework and then examine trading patterns in the data that are consistent with the strategy.

Recent years have witnessed a proliferation in electronic markets that provide traders with the option to hide order size. Market venues range from stock exchanges that allow “iceberg” orders with some displayed size to the opaque dark pool venues that do not display any information. Despite the widespread use of hidden orders, it is still unclear whether informed traders choose to hide order size. BPV (2009) examine Euronext-Paris data and conclude that hidden orders are primarily used by uninformed traders. However, since most market participants are likely to be uninformed traders, the result that the typical hidden order user is uninformed is not surprising. Our analysis of order exposure strategies before unscheduled events permits a careful examination of the informed agent behavior. We

---

6 For studies of strategic trading in a dealer setting, see Kyle (1985) and Glosten and Milgrom (1985); For studies of strategic trading in a limit order book setting, see Glosten (1994), Rock (1996), Seppi (1997) and Back and Baruch (2013).
7 There is also recent experimental and empirical findings that suggest that informed traders do use limit orders (see Barclay, Hendershott, and McCormick (2003); Bloomfield, O’Hara, and Saar (2005); Anand, Chakravarty and Martell (2005), Hautsch and Huang (2012), among others).
8 Buti and Rindi (2012) present a theoretical model describing the trader’s decision to hide a portion of the order. Recent empirical work on dark pools include Nimalendran and Ray (2011) and Buti, Rindi and Werner (2011). Industry report from the Tabb Group estimates that dark pools account for 8-9% of trading volume in U.S. equities.
show that informed traders prefer to expose order size when they submit limit orders, which lowers time-to-execution, the risk of non-execution, and opportunity costs. These findings suggest that dark pool venues are more likely to attract uninformed investors who hide order size to control order exposure risk and while ‘lit’ markets are more likely to attract informed investors who care about opportunity cost of non-execution. Our evidence should inform recent regulatory initiatives on dark pools and the impact of dark pools on the informational efficiency of security prices.

The rest of the paper is organized as follows. Section II presents a theoretical model on the competition among informed traders and identifies testable predictions on order submission strategies preceding positive and negative events. Section III describes the institutional details of the Euronext-Paris, the data sources and sample selection. Section IV presents the informed trading strategies on the price aggressiveness attribute and Section V presents the results on the order exposure attribute. Section VI examines the impact of trading strategies on the likelihood of achieving full execution, the time-to-execution, and execution costs. Section VII summarizes the main results and presents the implications of the study.

2. The Model

2.1. The prisoner’s dilemma game

When choosing between limit orders and market orders, a trader weighs the price benefit of a limit order against the risk of non-execution. To predict the behavior of informed traders, we employ a static model with two informed traders. The informed traders learn the realization of a signal, \( v \), after which they perceive the asset as either overvalued or undervalued. We denote the expected value of the signal by \( \bar{v} \). These informed traders face the choice between using market orders or limit orders to build their position before the information becomes widely known. We model the symmetric payoff, which depends on their joint decision to use market or limit order, as a prisoners’ dilemma. Trader one’s payoff is given by the following payoff matrix.
Trader One’s payoff

<table>
<thead>
<tr>
<th></th>
<th>Market Orders</th>
<th>Limit Orders</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trader One</td>
<td>$(v-v)\gamma^2 b$</td>
<td>$(v-v)\gamma^2 a$</td>
</tr>
<tr>
<td>Trader Two</td>
<td>$(v-v)\gamma^2 d$</td>
<td>$(v-v)\gamma^2 c$</td>
</tr>
</tbody>
</table>

where $0 < a < b < c < d$. The payoff rankings are based on the idea that both traders observe the same signal and trade in the same direction. If trader two uses market orders, then these market orders impose severe execution risk on trader one’s limit orders. Therefore, using limit orders when trader two uses market orders should be suboptimal; i.e. $a < b$. By the same token, if trader one uses market orders, these orders should be cheaper to execute if trader two does not compete for available liquidity by employing market orders; i.e. $b < d$. These discussion lead to the following: $a < b < d$.

The counter parties to informed traders one and two are uninformed traders. Though we don’t model these traders explicitly, we can divide them into two subgroups, patient and impatient uninformed traders. Patient traders use limit orders while impatient trades buy liquidity (i.e. use market orders). If the informed traders could collude, they would prefer trading with the impatient uninformed traders rather than the patient ones; i.e. $2b < 2c$. Even better, informed traders would prefer trading with both subgroups, the combined payoff is then $d + a$. Having established that $a < b < c$, we conclude that $a < b < c < d$. In the appendix we present a simple dynamic model that gives rise to the order $0 < a < b < c < d$ we postulate.

To sum our discussion thus far: we use the prisoners’ dilemma to model the interaction between the informed traders. The outcome of the game is that despite the price benefit of limit orders, informed traders competing on similar private information use market orders.\(^9\)

If the assumption is not true, then together with $a < b < d$, the equilibrium outcome would be to use limit orders. This is inconsistent with the empirical literature that private information mostly arrives via

\(^9\) Holden and Subrahmanyam (1992) also show that competition among informed traders leads to an inefficient outcome. In their model, informed traders can only trade using market orders but have a choice between trading on their information gradually or rapidly. Holden and Subrahmanyam (1992) show that if traders could collude, they would choose to trade gradually, however the equilibrium outcome is to trade fast.
market orders. On the other hand, the assumption that \( b < c \) may or may not be reasonable, depending on market conditions. The assumption can be stated as follows: if informed traders could cooperate, they would prefer to use limit orders (i.e. \( 2b < 2c \)). This assumption is violated when there is severe execution risk, e.g. when private information is short-lived. We present in the appendix a dynamic model that gives rise to the order \( 0 < a < b < c < d \) we postulate.

To sum our discussion thus far: we use the prisoners’ dilemma to model the interaction between the informed traders. The outcome of the game is that despite the price benefit of limit orders, informed traders competing on similar private information use market orders.

### 2.2. Trader competition and buy-sell asymmetry

Up until now we assumed that the direction of the information was irrelevant. However, when shares are hard to borrow (either short selling is expensive, prohibited, or the shares are simply hard to locate), then informed sellers might be forced to abstain from trade. In extreme situations, where it is virtually impossible to short, informed traders can sell only if prior to learning that the stock is overvalued, they were long on the stock.

To model the difficulty in short selling, we extend the payoff matrix:

<table>
<thead>
<tr>
<th>Trader Two</th>
<th>Market Orders</th>
<th>Limit Orders</th>
<th>Abstain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trader One</td>
<td>((v - \bar{v})^2 b)</td>
<td>((v - \bar{v})^2 a)</td>
<td>0</td>
</tr>
<tr>
<td>Market Orders</td>
<td>((v - \bar{v})^2 d)</td>
<td>((v - \bar{v})^2 c)</td>
<td>0</td>
</tr>
<tr>
<td>Limit Orders</td>
<td>((v - \bar{v})^2 e)</td>
<td>((v - \bar{v})^2 f)</td>
<td>0</td>
</tr>
<tr>
<td>Abstain</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Based on the above discussions, numerical values are assumed to correspond to alphabetical order: \( 0 < a < b < c < d \leq e < f \). The assumption that \( d \leq e \) is natural since the competitor submits a limit order in the former while being absent in the latter. The assumption that \( e < f \) is similar to our assumption that absent strategic consideration, informed traders prefer to employ limit orders to market orders.
If the event is negative and short selling is prohibited, traders can sell only if the stock is in their portfolios. Further each trader perceives the probability that the stock is in the other's portfolio to be \( p \). Then, if \( p \) is sufficiently small, informed traders use limit orders. Indeed, let us conjecture trader two, when owns the stock, uses limit orders. Then the expected payoff for trader one is \((v - \bar{v})^2 (pd + (1-p)e)\) when using market orders and \((v - \bar{v})^2 (pc + (1-p)f)\) when using limit orders. Thus, when

\[
p < \frac{f - e}{f - e + d - c}
\]

then the expected payoff when using limit orders is greater than the expected payoff when using market orders. Because the game is symmetric, we conclude that if inequality (1) holds and short selling is prohibited, then it is an equilibrium strategy to use limit orders.

We extend the above result to a world where traders can short the stock at a cost. We assume each of the informed traders is one of two types; each type faces a different cost of selling. One type corresponds to an informed trader that already located the shares, perhaps because the shares were in the trader's portfolio to begin with. The second type has yet to borrow the shares. We denote the borrowing costs by \( C > 0 \), with the convention that if the shares are impossible to short or locate then \( C \) is infinity. We use the cost of locating the shares, zero or \( C \), to denote the type of the trader. Consistent with the previous discussion, we let \( p \) be the probability that a trader is of type zero, and \( 1-p \) the probability that a trader is of type \( C \). The payoff of type \( C \) is

<table>
<thead>
<tr>
<th>Trader One's payoff (Type C)</th>
<th>Trader One</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Market Orders</td>
</tr>
<tr>
<td>Trader Two</td>
<td></td>
</tr>
<tr>
<td>Market Orders</td>
<td>((v - \bar{v})^2 b-C)</td>
</tr>
<tr>
<td>Limit Orders</td>
<td>((v - \bar{v})^2 d-C)</td>
</tr>
<tr>
<td>Abstain</td>
<td>((v - \bar{v})^2 e-C)</td>
</tr>
</tbody>
</table>

**Theorem 1** (Limit Order vs. Market Order Equilibrium): Assume the event is negative, \( 0 < a < b < c < d \leq e < f \), and \( C > 0 \).
(I) If inequality (1) holds, and in addition

\[ C > (v - \bar{v})^2 (pc + (1 - p)f) \]  \hspace{1cm} (2)

then there is a limit order equilibrium in which type 0 uses limit orders and type C abstains from trade.

(II) If

\[ C < (v - \bar{v})^2 b \]  \hspace{1cm} (3)

then, whether or not inequality (1) holds, there is a market order equilibrium in which type C borrows the shares and both types use market orders.

The proof of the theorem goes as follows. Consider the limit order equilibrium and assume that inequalities (1) and (2) hold, and trader two follows the equilibrium strategy; i.e. trader two uses limit orders when two’s type is zero, and otherwise abstains from trade. If trader one chooses to trade, then the expected payoff when using market orders is \((v - \bar{v})^2 (pd + (1 - p)e)\), while the expected payoff when using limit orders is \((v - \bar{v})^2 (pc + (1 - p)f)\). Inequality (1) ensures that the latter is optimal. Thus, if the type of trader one is zero, it is optimal to use limit orders. Inequality (2) ensures that the cost of borrowing is greater than the expected payoff, and hence trader one abstains when one’s type is \(C\). We therefore verified the limit-order equilibrium.

When the cost of borrowing is sufficiently low then type \(C\) borrows the shares and the market order equilibrium emerges. To verify that this is indeed an equilibrium, we only need to check that the cost of borrowing is lower than \((v - \bar{v})^2 b\), which the payoff when both traders use market orders. This is (3) in the theorem. This concludes the proof of the theorem.

To sum up, we have shown in the previous section that when the news is positive the outcome is a market order equilibrium. When the news is negative, depending on parameters, a limit order equilibrium may emerge. This leads to the following testable predictions.

**Hypothesis I:** Informed traders use more aggressive (market) orders before positive unscheduled events and less aggressive (limit) orders before negative unscheduled events.
On the other hand, if the event is sufficiently large; i.e. \((v - \bar{v})^2\) is large, then inequality (3) holds even if borrowing costs are high.

**Hypothesis II:** *The symmetry in the buy-sell order submission strategies is restored when the event is large.*

In addition, when \(p\), the probability that the informed traders owns the stock prior to learning the negative information, is sufficiently high then inequality (1) is violated, and the limit order equilibrium breaks down. It is conceivable that \(p\) is high for stocks that belong to an index and low for stocks that are not part of an index.

**Hypothesis III:** *Symmetry in buy-sell order submission strategies is restored for stocks that belong to an index.*

If the cost of borrowing is sufficiently small, then (3) holds, and the informed traders use market orders.

**Hypothesis IV:** *Symmetry in buy-sell order submission strategies is restored for stocks that are inexpensive to short.*

3. **Data and Methodology**

3.1. **Sample and data**

To understand the trading strategies of informed traders, we examine the Euronext-Paris, Base de Donnees de Marche (BDM) database for the year 2003. The BDM database contains information on the characteristics of all orders submitted for all stocks listed in the Euronext-Paris market. This includes the firm symbol; the date and time of order submission; a buy or sell indicator; the total size of the order (in shares); the displayed size (in shares); an order type indicator for identifying market, open or limit orders; a limit price in the case of a limit order; and instructions on when the order will expire. In addition, each order contains fields that allow tracking of any modifications made to the order prior to the expiration, with the exception of cancellations.\(^{10}\)

\(^{10}\) The database contains fields that allow us to track any modifications made to the order (typically order size and limit price) with complete accuracy. Cancelled orders can be identified as of the end of the day with complete accuracy, but cannot always be identified intraday. We are able in many instances to infer the exact time when an order has been cancelled, based on quote updates that do not reflect completed trades or order modifications, as in
We examine the 2003 sample period because more recent order-level data purchased from the Euronext market have important inaccuracies. In particular, orders that never get executed, or orders with a hidden component that are partly executed, do not get reported to the database. The omission affects the accuracy of the reconstructed limit order book, the analysis of order submission strategies, and control variables used in some specifications (e.g., displayed depth on the same price in the book, or book order imbalance). In addition, an important advantage is that the Euronext market in 2003 is highly consolidated with the vast majority of orders being submitted and executed in the main exchange. In more recent periods, the European equity markets have become highly fragmented due to the growth in alternative trading venues, including dark pools.

To test the predictions of our model, we focus the analysis on unanticipated corporate events where it is easier to identify informed trading. Following prior literature, “unanticipated” events are those whose timing is not predictable in advance, while “anticipated” events are those whose timing is known in advance of the event. As shown by Chae (2005), Graham, Koski and Loewenstein (2006) and Sarkar and Schwartz (2009), although market participants do not know in advance the information contained in anticipated events, those traders with some discretion on timing of trades tend to alter behavior before anticipated events in order to lower adverse selection risk. Lee, Mucklow and Ready (1993), for example, show that market makers widen the bid-ask spread and lower the inside depth before earnings announcements. Anticipated corporate events – such as earnings announcements or macro news announcements - are characterized by two-sided markets due to trading motivated by differential information and/or heterogeneous beliefs. Unanticipated events on the other hand are characterized by one-sided markets since order flow reflects trading that is motivated by private information (Sarkar and Schwartz (2009)).

We identify unanticipated events in the year 2003 using both the Global SDC database compiled by Thomson Financial Securities Data and the AMADEUS database provided by Bureau van Dijk. We

Bessembinder and Venkataraman (2004). Since the database identifies the cancellation date, any minor errors in the reconstructed limit order book attributable to undetected order cancellations do not accumulate across trading days.
focus on five types of unscheduled corporate announcements: acquisitions, targets, seasoned equity offerings (SEOs), repurchases, dividend initiations and dividend termination. We capture the exact date of these announcements using Bloomberg and Factiva search engines and identifying the date of the first news story about the event. We focus on companies that are publicly traded on Euronext-Paris market. We eliminate stocks that switched from continuous trading to call auctions (or vice-versa) or were delisted from the exchange during a 40-day period surrounding the event (30 days before and 10 days after). Our final sample consists of 101 unscheduled corporate events for 95 unique stocks.

Table 1 shows the number of the different types of unanticipated corporate events in our sample. We separate the events into positive and negative based on the two day (Days [0,1]) cumulative abnormal returns (CAR). We use the CAC40 daily index return as a benchmark. We report mean and median measures of returns. Overall, the events in our sample have similar magnitude of positive and negative returns. In particular, we have 58 positive events with mean (median) CAR of 4.84% (2.74%) and 43 negative events with mean (median) CAR of -4.53% (-2.57%). For acquisitions, targets, seos, repurchases and dividend initiation announcements we have both positive and negative event days (day 0 and day 1) CAR while only one stock in our sample announced a dividend termination. The most significant announcement effect is observed for Targets followed by Acquisition and SEO events.

For each unscheduled corporate event we look at order level submission activity around the announcement day 0. We define as control period (base period) the period of 20 days preceding the announcement day, starting on day -30 and ending on day -10. We focus on identifying informed activity during Days [-5,-1] preceding the announcement day. Table 2 reports descriptive statistics (mean and medians) of order usage characteristics during the control and sample periods for positive and negative events. In particular, the table shows the daily number of orders submitted and average order size of a) all orders submitted b) market/marketable submitted and c) limit orders submitted. We also report mean and median numbers of the average daily ratio of marketable/market to limit orders submitted and the average

---

11 We eliminate mergers and acquisitions in which the deal value relative to the market value of the acquirer is less than 5%.
daily hidden order usage. A higher value of the ratio is consistent with aggressive order submission.

We find that the number of buy orders increases before positive events and the number of sell orders increases before negative events relative to the control period suggesting that informed traders are active before unanticipated events. We also observe similar increases in the average order size of orders submitted. This is particular true for the less aggressively priced limit orders, where buy (sell) order size increased from 1,095 (1,909) shares in control period to 1,562 (2,168) shares in the sample period before positive (negative) events. Although large orders increase the risk of front-running by ‘parasitic’ traders, exposing size attracts the more patient counterparties who react to displayed orders. Consistent with this idea, we observe no increase in the option to hide order size. Importantly, in preliminary support of hypothesis I, we find a marginal increase in proportion of aggressively priced buy orders before positive events and a marginal decrease in proportion of aggressively priced sell orders before negative events.

3.2 Cross-sectional aggregation

We estimate all of our subsequent multivariate analyses on an event-by-event basis. In the interest of parsimony, we present results that are aggregated across events. Harris and Piwowar (2006) emphasize the desirability of assigning larger weights in cross-sectional aggregation to those securities whose parameters are estimated more precisely. To do so, we assess statistical significance by relying on a Bayesian framework attributable to DuMouchel (1994) and also implemented by BVP (2009). The method assumes that, for each estimated firm i coefficient, $\beta_i$:

$$\hat{\beta}_i \mid \beta_i \sim i.i.d.N(\beta_i, s_i^2)$$

and

$$\beta_i \sim i.i.d.N(\beta, \sigma^2)$$

where $N$ is the Gaussian distribution. The standard errors, $s_i$, are estimated by use of the Newey-West method to correct for autocorrelation and heteroskedasticity, and $\sigma^2$ is estimated by maximum
likelihood. The aggregated $\beta$ estimate is obtained from the $N$ individual firm estimates as

$$
\hat{\beta} = \frac{\sum_{i=1}^{N} \hat{\beta}_i}{\sum_{i=1}^{N} \frac{1}{s_i^2 + \hat{\sigma}_{m.i.e}^2}}
$$

(6)

Assuming independence across firms, the variance of the aggregate estimate is:

$$
Var(\hat{\beta}) = \frac{1}{\sum_{i=1}^{N} \frac{1}{s_i^2 + \hat{\sigma}_{m.i.e}^2}}
$$

(7)

where $\hat{\sigma}_{m.i.e}^2$ is the maximum likelihood estimator of $\sigma^2$. The aggregate $t$-statistic is based on the aggregated coefficient estimate relative to the standard error of the aggregate estimate. This method allows for variation across stocks in the true $\beta_i$ and also for cross-sectional differences in the precision with which $\beta_i$ is estimated. The key feature of the aggregation method is that it places more weight on those coefficients that are estimated more precisely.\(^\text{13}\)

In all of our multivariate specifications of order submission strategies we include daily dummy variables of the sample period (Days [-5,-1]) to identify abnormal order activity –using the control period as our benchmark. We report estimated daily dummy coefficients and test statistics in all of our regression output. In addition, since any abnormal activity that is observed can happen at any time during the five days before the event day 0, we also report abnormal activity obtained by aggregating the individual day dummy coefficients during the sample period. Our objective is to capture the cumulative abnormal activity across the sample period without any econometric constraint on each of the days -5 to -1. Throughout the paper, we focus on both the individual day dummy coefficients and the cumulative effects and the corresponding $t$-statistics in our interpretation of the results.

\(^{12}\) To estimate the Newey-West corrected standard errors, we use the Generalized Method of Moments (GMM) with a Bartlett kernel and a maximum lag length of 10.

\(^{13}\) The method does not control for dependence of estimation errors across events. We believe that this dependence should be small since the events are not clustered in time.
4. Price Aggressiveness of Informed Traders

4.1. Patterns before unscheduled corporate events

The theoretical model assumes that informed investors weigh the price benefit of a limit order against the risk of non-execution when selecting the trading strategy. In particular, the model predicts an asymmetry in trading strategies of informed buyers and sellers motivated by the asymmetry in competition among informed agents. In this section we test the model predictions by examining the price aggressiveness of orders submitted before unanticipated corporate events.

The regression specification controls for market conditions at the time of order submission and accounts for all orders submitted during the control period of Days [-30,-10] before the event and the event Days [-5,2], as follows:

\[
\text{PriceAggressive}_{it} = \gamma_0 + \gamma_1 \text{DayMinus5} + \gamma_2 \text{DayMinus4} + \gamma_3 \text{DayMinus3} + \gamma_4 \text{DayMinus2} + \gamma_5 \text{DayMinus1} + \gamma_6 \text{Day0&Plus1} + \gamma_7 \text{DayPlus2} \quad + \gamma_8 \text{OrderExposure}_{it} + \gamma_9 \text{PriceAggressive}_{i,t-1} + \gamma_{10} \text{HiddenOppSide}_{it} + \gamma_{11} \text{DisplayedSize}_{i,t-1} + \gamma_{12} \text{OrderSize}_{it} + \gamma_{13} \text{Spread}_{it} + \gamma_{14} \text{DepthSame}_{it} + \gamma_{15} \text{DepthOpp}_{it} + \gamma_{16} \text{Volatility}_{it} + \gamma_{17} \text{WaitTime}_{it} + \gamma_{18} \text{TradeFreqHour}_{it} + \gamma_{19} \text{BookOrderImbalance}_{it} + \gamma_{20} \text{TradeSize}_{i,t-1} + \gamma_{21} \text{MktVolatility}_{i,t-1} + \gamma_{22} \text{Ind.Volatility}_{i,t-1}
\]

where the subscript \(i,t\) refers to the time \(t\) order for event \(i\). \text{PriceAggressive} is an ordinal variable that takes the value of 1 for the most aggressive order and 7 for the least aggressive, following the approach in BVP (2009) and Biais, Hillion and Spatt (1995).\(^{14}\) \text{DayMinus5} to \text{DayPlus2} are the dummy variables that equal one for each of the event days, and equals zero otherwise.

The control variables are defined as follows. \text{OrderExposure} is an indicator variable that equals

---

\(^{14}\) The first four categories represent orders that demand liquidity from the book and the last three categories represent orders that supply liquidity to the book. The most aggressive orders (category 1) represents buy (sell) orders with order size greater than those displayed in the inside ask (bid) and with instructions to walk up (down) the book until the order is fully executed. Category 2 represents buy (sell) orders with order size greater than those displayed in the inside ask (bid) and with instructions to walk up (down) the book, but the order specifies a limit price such that the order is not expected to execute fully based on displayed book. Category 3 represents buy (sell) orders with the limit price equal to the inside ask (bid) and with order sizes greater than those displayed in the inside ask (bid). Orders in categories 2 and 3 may execute fully due to hidden liquidity but may also clear the book and convert to a standing limit order. Category 4 represents buy (sell) orders with the limit price equal to the inside ask (bid) and with order size less than those displayed in the inside ask (bid). These orders are expected to immediately execute in full. Category 5 represents orders with limit prices that lie within the inside bid and ask prices. Category 6 represents buy (sell) orders with limit price equal to the inside bid (ask). Finally, Category 7 represents buy (sell) orders with limit price less (greater) than the inside bid (ask).
one if the order has a hidden size and equals zero otherwise. \textit{DisplayedOrderSize} is exposed size of the order divided by average daily trading volume. \textit{TotalOrderSize} is total (displayed plus hidden) size of the order divided by average daily trading volume. \textit{Spread} is the percentage bid-ask spread. \textit{DepthSame} is the displayed depth at the best bid (ask) for a buy (sell) order divided by the monthly median. \textit{DepthOpp} is the displayed depth at the best ask (bid) for a buy (sell) order divided by the monthly median. \textit{Volatility} is the standard deviation of quote midpoint returns over the preceding hour. \textit{WaitTime} is the average elapsed time between the prior three order arrivals on the same side, refreshing the time clock each day. \textit{TradeFreqHour} is the number of transactions in the last hour. \textit{HiddenSameSide} is the number of hidden shares at the best quote on the same side (i.e., at the bid side for a buy order) revealed by the most recent transaction. \textit{HiddenOppSide} is the number of hidden shares at the best quote on the opposite side revealed by the most recent transaction. \textit{TradesSize} is the size of the most recent transaction divided by the average daily trading volume. \textit{SamePriceDisplayedDepth}\textsubscript{\textit{it}} is the current depth at the price level of the limit order that is submitted divided by the average daily trading volume. \textit{BookOrderImbalance} is the percentage difference between the displayed liquidity in the best five prices on the buy and sell side of the book, suitable signed (i.e., the variable is positive when same size liquidity exceeds opposite side liquidity). \textit{Ind.Volatility} is the return volatility of a portfolio of stocks in the same industry during the prior hour. \textit{Mkt.Volatility} is the return volatility of the CAC40 Index during the prior hour. \textit{Last Hour} is an indicator variable that equals one for orders submitted in the last hour of the trading day and is zero otherwise.

Table 3 reports regression coefficients along with corresponding $t$-statistics, estimated on an event-by-event basis and aggregated across firms using the approach described in Subsection 5.2. The coefficients on the control variables are consistent with those reported in the prior studies. Specifically, traders submit less aggressively priced orders (i.e., prefer limit orders over market orders) when the inside bid-ask spread is wide; depth on same side of limit order book is small indicating less competition from same side for liquidity provision; depth on opposite side of the limit order book is large or revealed the presence of hidden orders which indicates the presence of counterparties eager to trade; when volatility is high which is consistent with a volatility capture strategy as described in Handa and Schwartz (1996);
book imbalance indicates that there is less competition on the same side of the book relative to the opposite side of the book; and the previous order is less aggressive which likely captures market conditions that are missing from the model. Regression coefficients capturing the time series variation in market and industry-level volatility are not statistically significant which suggests that own-stock volatility plays a more important role than market-wide volatility.

The main tests of the theoretical model are based on coefficient estimates on Dummy variables, DayMinus5 to DayMinus1. For buy orders submitted in the days preceding positive events (column (1)), we estimate that all the five coefficients corresponding to Days [-5,-1] are negative and DayMinus3 and DayMinus1 coefficients have t-statistics below -2.0. Since the day dummies can be positive or negative on an event day purely by chance, we estimate the cumulative effect across the event days. The negative coefficient on cumulative effect (t-statistic=-2.96) suggests that informed buyers submit more aggressively priced orders in the days preceding a positive event. In contrast, for sell orders submitted in the days preceding negative events (column (2)), we estimate that four of the five coefficients corresponding to Days [-5,-1] are positive and the coefficient on DayMinus5 is highly statistically significant. The positive coefficient on cumulative effect (t-statistic=1.70) suggests that informed sellers submit less aggressively priced orders in the days preceding a negative event. These patterns are consistent with Hypothesis I and support the idea that buy-sell asymmetry can be explained, at least in part, by the asymmetry in the order submission strategies of informed agents.

4.2. Cross-sectional Patterns in Price Aggressiveness

We develop further cross-sectional tests of the model by examining patterns in buy-sell asymmetry along two dimensions. First, we examine corporate events that are characterized by large versus small announcement period returns. Second, we examine sub-samples of firms characterized by the ease of locating shares and the cost of borrowing shares to implement a short sale.

4.2.1. Announcement return and buy-sell asymmetry

Announcement period return influences buy-sell asymmetry in our framework because informed agents weigh the benefits of take a short position against the cost of borrowing shares. If the information
content is large, informed sellers have more incentives to locate shares that are difficult or costly to borrow. However, if the information content is small, the benefits of short selling might not out-weight the cost, thus leading informed agents to abstain from trading if they do not already own the stock. The model predicts that buy-sell asymmetry in price aggressiveness before corporate events is smaller for large event announcement relative to small event announcement returns.

In columns (3) to (6) of Table 3, we present regression coefficients that are conditional on the magnitude of announcement period returns. We use absolute return of 5% to identify large and small announcement events. For small announcements, the cumulative effect of informed traders in Days[-5,-1] before positive events suggests an increase in price aggressiveness (coefficient=-0.35 with t-statistic =-2.09) while those before negative events indicates a decrease in price aggressiveness (coefficient=0.31 with t-statistic=1.95). For large announcements, we observe an increase in price aggressiveness before positive events (coefficient=-1.60 with t-statistic=-2.40) but no change in price aggressiveness before negative events (coefficient=0.17 with t-statistic=0.45). The latter result are supportive that informed sellers use more aggressive strategies when the announcement effect is economically large.

4.2.2. Short sale constraints and buy-sell asymmetry

In Table 4, Index membership influences short sale constraints because index constituent stocks are owned by index tracking funds who are active participants in the security lending market. Prior work has shown that it is easier to borrow stocks belonging to major indices (D’Avolio (2002), Nagel (2005)). We therefore classify the sample firms based on whether the stock belongs to the SBF120 index. The model predicts that informed sellers are more likely to trade a stock with unfavorable information if the stock is easy to borrow. Informed sellers will face more competition from other traders in index constituent stocks. For these stocks, both informed buyers and sellers implement aggressive strategies. For stocks not belonging to SBF Index, we expect informed buyers to trade aggressively while informed sellers to trade more passively.

In columns (1) - (4) of Table 4, we report the regression coefficients for subsamples of firms based on SBF120 index membership. For buy orders preceding positive events (columns (1) and (3)), the
cumulative effect coefficients indicate that informed traders increase price aggressiveness in both stocks included the SBF120 Index and those not included in the index. However, for sell orders preceding negative events (columns (2) and (4)), we observe a strikingly different pattern in informed trading based on index membership. For index stocks, we observe an increase in price aggressiveness by informed sellers before the event while for non-index stocks, we observe a decrease in price aggressiveness by informed sellers before the event. These results provide direct empirical support for the central prediction of the model that informed agent strategy is influenced by extent of trader competition and contributes to buy-sell asymmetry. Specifically, there is no buy-sell asymmetry in order submission strategies when competition among informed agents is high. Thus the evidence of buy-sell asymmetry in the data can be explained, at least in part, by short sale constraints on informed traders.

We build supporting evidence using another well-know proxy for short sale constraints – the presence of listed option in the stock. For stocks with listed option, informed sellers have the ability to build a position using the option market. As option market makers hedge their position in the underlying security, the informed order flow in the option market will be transmitted via the market maker trades to the stock market. Therefore the model predicts that stocks with listed options will not exhibit buy-sell asymmetry in price aggressiveness because both informed buyers and sellers face competition and therefore implement aggressive strategies. In contrast, for stocks without listed options, the model predicts that informed buyers will implement aggressive strategies while informed sellers will implement passive strategies.

In columns (5) - (8) of Table 4, we report the regression coefficients for firms with and without listed options. The empirical evidence provides strong support for the theoretical predictions. For firms with listed options (column (5) and (6)), we observe that the cumulative effect coefficient is negative for both buy orders before positive events and sell orders before negative events, suggesting that both

---

We note that, although many index constituents stocks have listed options, the fit is far from perfect. In particular, for positive events we have 25 stocks in our sample that belong to the SBF120 index where 14 have listed options. Similarly, for negative events we have 17 stocks that belong to the Index and 14 that have listed options. The overlap in both positive and negative events subsamples is less than 70%.
informed buyers and sellers face more competition from other traders and use price aggressive orders. In contrast, for stocks without listed option, the coefficient on buy order before positive events is negative, which indicates more price aggressive orders, but the coefficient on sell orders before negative events is positive, which indicates less price aggressive orders. These results provide support for the linkage between buy-sell asymmetry and informed trader behavior.

5. Order exposure strategies of informed investors

The option to partially or fully hide order size is a widely available feature in many electronic markets. In fact, one important category of trading venues called Dark Pool only accept orders that are fully hidden. Despite the wide-spread usage of hidden orders, only a handful of academic papers have studied empirically the determinants of order exposure.\textsuperscript{16} BVP (2009) examine Euronext-Paris data and show that hidden orders are associated with smaller opportunity costs of non-execution, suggesting that hidden orders are primarily used by traders with no information regarding future price movements. However, since the majority of market participants are likely to be either noise or liquidity traders, the result that the typical hidden order user is not informed is not surprising.

But do informed traders hide order size? The line of enquiry can guide the recent regulatory debate on dark pools venues and the extent to which hidden orders, which reduce pre-trade transparency, impact the information efficiency of prices. Specifically, a result that informed traders choose to expose order size implies that dark pools venues are predominantly used by uninformed traders who are looking to control order exposure risk. In this section, we use unscheduled corporate events as a setting to investigate the informed traders’ order exposure strategies. Harris (1996) argues that exposing size will attract interest from “reactive” traders, who have not revealed their orders but wait for other traders to post orders at favorable prices. Informed traders might prefer to exposure size when they submit passive orders to attract counterparties. However, exposing order size reveals the presence of informed agents on one side of the market. Some counterparties might react by withdrawing trading interest, or exploit the

\textsuperscript{16} For theoretical studies, see Buti and Rindi (2008), Bloomfield, O’Hara and Saar (2012) and Boulatov and George (2013).
information content of the order by implementing front-running strategies (Harris (1997)). Building on this reasoning, Moinas (2006) presents a theoretical model where exposing the size of a large limit order lowers the probability of execution.

5.1. The decision to hide an order.

In table 5, we report regression coefficients, along with corresponding t-statistics, of a logistical model on the decision to hide order size, following the approach implemented by BPV (2009). The dependent variable is an indicator variable that equals one for orders that contain hidden size and zero for orders that do not. The model is estimated using limit orders that will stand on the book (those in categories 5, 6 and 7) because the option to hide size is more relevant for such orders. The Day Dummies are defined similar to those described in Section 4 and capture the strategies of informed traders. Consistent with BPV (2009) and De Winne and D’Hondt (2007), we find that market conditions and order attributes at the time of order submission influence the exposure decision. For example, hidden orders are more likely when the depth at the best quote on the same side is greater, or when previous trades have revealed hidden depth on the same side. We also find that traders who submit larger orders are more likely to hide size. Focusing on individual day cumulative effect coefficients, Panel B of Table 5 results indicate that informed buyers are more likely to use the option to hide size before positive events and informed sellers are less likely to use the option to hide size before negative events.

We obtain additional insights by examining the variation across firms in informed trader strategies. Specifically, in Panels C-F of Table 5, we present the individual day cumulative effect coefficients for various subsamples. The following pattern is noteworthy. For buy orders before positive events, the cumulative effect coefficient is positive in all sub-samples; however the effect is marginally significant only for stocks in SBF120 Index (t-statistic=1.87) and those with listed options (t-statistic=1.85). When we examine sub-samples of negative events where short sale constraints are less binding; i.e., those stocks in SBF120 Index or those with listed option, the cumulative effect coefficient for sell orders before negative events is negative but not statistically significant. Note that all sub-samples discussed thus far are associated with informed traders selecting more aggressively priced orders, based
Collectively, when informed traders face more competition, we find that informed traders increase the use of aggressive (market) orders, but conditional on submitting a standing limit order, they are more likely to use the option to hide size.

In contrast, for sub-samples where short sale constraints are more binding; i.e., those stocks not in SBF120 Index (Panel D) or those without listed options (Panel F), the cumulative effect coefficient for sell orders before negative events is negative and statistically significant. Collectively, when informed agents face less competition, the evidence suggests that informed traders increase the use of passive (limit) orders, but conditional on doing so, they are less likely to use the option to hide size.

5.2. The magnitude of hidden order size

In table 6, we report regression coefficients, along with corresponding t-statistics, of a tobit analysis, focusing on the quantity of shares that are hidden. The analysis builds on the logistical analysis above that focuses on whether the incoming order has any hidden shares. Similar to Table 5, the empirical specification includes variables that control for the state of the limit order book, the market conditions such as recent volatility and trades, the order attributes such as price aggressiveness and total order size, and Day Dummies that capture informed trader exposure strategies before positive and negative events. The model is estimated using limit orders that will stand on the book (those in categories 5, 6 and 7). Consistent with prior work, we find that the number of hidden shares is positively associated with the total size of the incoming order.

Controlling for market conditions and order attributes, in Panel B of Table 6 we find that the individual day cumulative effect coefficient is negative and statistically significant for both buy orders before positive events and sell orders before negative events. In Panels C-F of Table 6, we examine the number of hidden shares for sub-sample of firms. For sub-samples with more competition among informed traders – index constituent stocks and those with listed options – the individual day cumulative effect coefficient is negative but not statistically significant both before positive and negative events. For stocks not in the index or those without listed options, the cumulative effect coefficient is negative and statistically significant. The finding that the cumulative effect is negative in all sub-samples suggests that
when informed traders submit limit orders, they choose to expose order size. These results are consistent with Harris (1996) prediction that limit order traders who face a large opportunity cost of non-execution will choose to expose order size in order to attract the reactive traders.

In summary, the results thus far point to two types of informed trader strategies, which we broadly classify as “aggressive” and “passive”. Scenarios with more competition among informed traders are characterized by (a) the use of market orders over limit orders, and (b) the exercise more often of the option to hide limit order size but exposing a larger quantity of shares to the market. Scenarios with less competition among informed trades are characterized by (a) the use of limit orders over market orders, and (b) the exercise less often of the option to hide size and exposing a larger quantity of shares to the market. With passive strategies, informed traders face a higher risk of non-execution as the price drifts away due to leakage effects. Empirical evidence in BPV (2009) suggests that exposing order size lowers the time to execution and increases the probability of full execution of a limit order. The increased usage of non-hidden limit orders within passive strategies controls the opportunity cost of non-execution.

6. Order submission strategies of informed investors and execution costs

Informed traders have the ability to choose among many possible trading strategies. The analysis thus far identifies two types of strategies based on the degree of trader competition in the stock. Our model predicts that the choice of a specific strategy reflects an optimization across many dimensions of execution quality, including non-execution risk, time-to-execution and implementation costs. In this section, we examine the realized outcome on execution quality for the two types of strategies. A rational selection of strategies implies that different strategies achieve favorable outcomes.

6.1. Trader strategies and execution time

In Table 7, we reports results of an econometric model of limit order time-to-execution using survival analysis, as described in Lo, Mackinlay, and Zhang (2002). The model describes an accelerated failure time specification of limit order execution under the generalized gamma distribution. The control variables, which are identical to Lo et. al. (2002), capture the state of the book, market conditions and order attributes. We include a dummy variable that equals one for a hidden order and equals zero
otherwise because BPV (2009) find that exposing an order reduces execution time. The individual day dummy variable coefficients capture the change in time-to-execution before unanticipated events. Similar to earlier tables, coefficients are estimated for each event and aggregated across events using the Bayesian framework of DuMouchel (1994) and Panayides (2007). We find that more aggressively priced orders and orders that are fully exposed are associated with shorter execution time. Consistent with prior work, we also find that execution time increases when there is more competition on the same side of the market, and vice-versa.

More importantly for our investigation, for both buy orders before positive events and sell orders before negative events, we find that individual day cumulative effect coefficient is negative and statistically significant. Thus the abnormal limit order flow, which we attribute to informed traders, are associated with shorter execution time. In Panel C, we report the time-to-execution individual day cumulative effect coefficients for sub-samples of stocks. We find that cumulative effect coefficient is negative for all subsamples and statistically significant in seven of the eight sub-samples. Since execution time is an empirical proxy for price risk associated with a delayed trade, the evidence in Table 7 suggests that informed agents implement trading strategies that control for delayed execution.

6.2. Trader strategies and implementation shortfall

In this section, we investigate how informed trader strategies are influenced by execution costs. To measure execution costs, we rely on the implementation shortfall framework proposed by Perold (1988) and implemented by Harris and Hasbrouck (1996), Griffiths, Smith, Turnbull, and White (2000), and BPV (2009). We estimate two components of the implementation shortfall for each order: (a) effective spread cost is the appropriately signed difference between the fill price and the quote mid-point at the time of order submission, and (b) the opportunity cost is the appropriately signed difference between the closing price on the order expiration or cancellation date and the quote midpoint at the time of order submission. For a limit order that goes unfilled, the effective spread cost is zero. For an order that is fully executed, the opportunity cost is zero. For orders that are not fully executed, the opportunity cost is positive if the stock price rises for buy orders and falls for sell orders after order submission. The model
predicts, when informed agents face more competition from other traders, they employ more aggressive strategies which cause an adverse price move. More competition among informed agents causes high opportunity costs while less competition causes low opportunity costs. The implementation shortfall cost for an order is the weighted sum of effective spread cost and opportunity cost, where the weights are the proportion of the order size that is filled and unfilled, respectively.

Table 8 presents the coefficient estimates of regressions of execution costs on market conditions and order characteristics. Consistent with Harris and Hasbrouck (1996), we find that price aggressiveness is positively associated with effective spread cost. The negative coefficients on hidden order dummy in the opportunity cost regression suggests that hidden orders are associated with smaller opportunity costs. These findings are consistent with BPV (2009) who conclude that hidden orders are primarily used by uninformed traders who use the option to hide order size to control order exposure risk.

To test model predictions, we focus on Day Dummy coefficients in opportunity cost regressions. These coefficients capture the abnormal execution costs before information events relative to the non-event benchmark costs. For buy orders before positive events, the individual day cumulative effect variable is positive and statistically significant (t-statistic=2.17), suggesting that the stock price on average tends to drift upwards before a positive event. For sell orders before negative events, the cumulative effect variable is negative and marginally significant (t-statistic=-1.77), suggesting that the stock price does not drift downwards before negative events. The asymmetry in opportunity cost is consistent with evidence in Table 4 that informed buyers use aggressive strategies before positive events because they face more competition from other traders. The aggressive strategies are associated with positive drift in stock price and therefore the positive opportunity cost before positive events.

We next examine the patterns in opportunity cost for sub-samples based on short sale constraints. For stocks included in the SBF120 index, the opportunity cost coefficient based on individual day cumulative effects is positive for both buy orders before positive events and sell orders before negative events. In other words, the stock price tends to drift upwards relative to pre-event benchmarks before positive news and tends to drift downwards before negative news. The drift in stock price is consistent
with both informed buyers and informed sellers implementing aggressive trading strategies for these stocks. Since both informed buyer and sellers implement aggressive strategies, other participants detect the presence of informed agents in order flow and the price impact of buy and sell orders are similar. The positive opportunity costs results in higher implementation shortfall costs for informed traders relative to a non-event benchmark.

For stocks not included in SBF 120, the opportunity cost coefficient based on cumulative effects is positive (t-statistic=2.21) for buy orders before positive events and negative (t-statistic=-1.72) for sell orders before negative events. The asymmetry in opportunity cost (or the adverse drift in stock price) is consistent with the asymmetry in order submission strategies for this sub-sample documented in Table 4. Collectively, the results suggest that aggressive strategies implemented by informed buyers before positive events causes an upward drift in price. In contrast, the stock price does not drift downwards before negative events because informed sellers implement passive trading strategies. The point estimates of cumulative effects before positive and negative events are broadly similar for the sub-samples of stocks based on options listing but the statistical significance of the results is smaller.

7. Conclusion

In this study, we present a theoretical model that describes the order submission strategies of informed agents. All else the same, an informed agent would prefer to place a limit order in an attempt to get a better price. However the opportunity cost of non-execution is high when other informed traders use market orders and cause an adverse price move. When opportunity costs are present, informed traders have a strong incentive to use market orders before prices can fully adjust to the new information. Thus when the competition among informed traders is high, the equilibrium strategy in the game is to use market orders. In contrast a monopolist informed trader prefers to use limit orders in order to manage order exposure risk and contain the market impact costs. We posit that costly short selling causes an asymmetry in informed trader strategies before positive and negative events. Specifically, when the event
is negative, informed agents choose to abstain if (i) the shares are not in their portfolio, (ii) borrowing costs are high, (iii) the event is small so potential profits cannot justify the costs.

Using detailed order level data from Euronext-Paris, we examine whether the optimal response of informed agents is influenced by uncertainty about the presence of other informed agents in the market. Trading activity attributable to informed agents is identified by focusing on unscheduled corporate events where timing information is not available. Our findings are strongly supportive of theoretical predictions. We show that informed agents employ more aggressive strategies preceding positive events and less aggressive strategies preceding negative events. Examining sub-samples, we find that the buy-sell asymmetry in order submission strategies exists only for sub-samples where security borrowing is difficult or too expensive or the announcement returns are small. The model predicts that informed agents will abstain from seller in these sub-samples if they do not already own the stock. For sub-samples where announcement results are large or short sale constraints are less binding, we observe an increase in price aggressiveness symmetrically for both buy orders before positive events and sell orders before negative events. This is supportive of model predictions that both informed buyers and sellers face competition and therefore employ more aggressive trading strategies.

The study contributes to a better understanding of the well-documented asymmetry in the price effects surrounding block purchases and sales. Specifically, prior work has shown that block purchases contain more information than block sales. Our results that informed agents with positive private information use aggressive strategies while those with negative information use passive strategies suggest that the information content is driven by the extent to which the trading strategies tip off market participants about the presence of informed traders. We verify that the asymmetry in execution costs exist for sub-samples where short sale constraints are bindings but not for sub-samples where stocks are easy to borrow. Thus the competition among informed agents signals their presence to other market participants who adjust the security price to reflect the higher adverse selection risk. These price adjustments are captured as higher opportunity costs for buy orders as compared with sell orders, particularly for sub-samples with a narrow investor base and short selling constraints.
The study provides a framework detecting unusual trading activity that can be attributed to inside information. An important insight from our model is that informed agents order submission strategies are influenced by the breadth of the investor base, the cost of borrowing shares and the nature of private information. For regulators, broker-dealers, and exchanges who are interested in detecting patterns attributable to inside information, our study provides guidance on ex-ante optimal strategies of insiders and how patterns might vary in the cross-section of stocks. Further, we show that informed traders prefer to expose limit order size in order to increase execution probability and lower non-execution risk, suggesting that informed traders are unlikely to be attracted to dark pool venues. Our evidence should inform the debate on the impact of dark pools on informational efficiency of prices.
References


Harris, L., 1996. Does a minimum price variation encourage order exposure? Unpublished working paper, University of Southern California, Los Angeles, CA.


Jegadeesh, N., Tang, Y., 2010. Institutional trades around takeover announcements: skill vs. inside information. Unpublished working paper, Emory University, Atlanta, GA.


Table 1: Unanticipated Corporate Events: abnormal returns

The table reports number of the different types of unanticipated corporate events in our sample. We look at 5 different types of events: acquisitions, targets, season equity offerings, repurchases, dividend initiations and dividend termination. We separate the events into positive and negative based on the two day (day 0 and day plus 1) cumulative abnormal returns. These are calculated by subtracting the CAC40 daily index returns which is used as a benchmark. We report mean and median measures of the returns.

<table>
<thead>
<tr>
<th>Type of Events</th>
<th># of Events</th>
<th>Abnormal Returns</th>
<th>Abnormal Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Positive</td>
<td>Negative Events</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>Overall</td>
<td>101</td>
<td>4.84%</td>
<td>2.74%</td>
</tr>
<tr>
<td>Acquisitions</td>
<td>35</td>
<td>3.61%</td>
<td>2.56%</td>
</tr>
<tr>
<td>Targets</td>
<td>25</td>
<td>9.52%</td>
<td>7.12%</td>
</tr>
<tr>
<td>SEOs</td>
<td>22</td>
<td>3.41%</td>
<td>3.42%</td>
</tr>
<tr>
<td>Repurchases</td>
<td>14</td>
<td>2.88%</td>
<td>1.59%</td>
</tr>
<tr>
<td>Dividend Initiations</td>
<td>4</td>
<td>2.32%</td>
<td>2.32%</td>
</tr>
<tr>
<td>Dividend Terminations</td>
<td>1</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 2: Descriptive Statistics of Order Usage characteristics

The table reports descriptive statistics of order usage characteristics for all 101 unanticipated corporate events in our sample. The relevant characteristics are calculated for each firm-event and the table reports the (cross-sectional) statistics across all firm-events. We report mean and median statistics of order activity and order size of 1) all orders submitted, 2) market/marketable orders submitted and, 3) limit orders submitted. We also report mean and median percentage numbers of the ratio of market/marketable orders to limit orders, and hidden order usage. In Panel A (control period) we report statistics during our control period of 10 days to 30 days before the corporate announcements. Panel B (sample period) reports similar statistics during day minus 5 to day minus 1 before the event announcement.

<table>
<thead>
<tr>
<th>Daily Descriptive statistics</th>
<th>Positive</th>
<th>Negative Events</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>Daily Number of Orders</td>
<td>638.5</td>
<td>50.7</td>
</tr>
<tr>
<td>Average Order Size</td>
<td>1,004.0</td>
<td>436.6</td>
</tr>
<tr>
<td>Daily Number of Marketable/Market Orders</td>
<td>147.5</td>
<td>16.0</td>
</tr>
<tr>
<td>Average Order Size of Marketable/Market Orders</td>
<td>731.0</td>
<td>298.1</td>
</tr>
<tr>
<td>Daily Number of Limit Orders</td>
<td>432.1</td>
<td>45.6</td>
</tr>
<tr>
<td>Average Order Size of Limit Orders</td>
<td>1,095.0</td>
<td>481.2</td>
</tr>
<tr>
<td>Average Percentage Marketable/Market Orders to Limit</td>
<td>45.1</td>
<td>44.3</td>
</tr>
<tr>
<td>Average Hidden Orders Usage</td>
<td>18.4</td>
<td>15.3</td>
</tr>
</tbody>
</table>

Panel A: Control Period

<table>
<thead>
<tr>
<th>Daily Descriptive statistics</th>
<th>Positive</th>
<th>Negative Events</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>Daily Number of Orders</td>
<td>808.8</td>
<td>63.9</td>
</tr>
<tr>
<td>Average Order Size</td>
<td>1,358.0</td>
<td>461.7</td>
</tr>
<tr>
<td>Daily Number of Marketable/Market Orders</td>
<td>158.4</td>
<td>15.4</td>
</tr>
<tr>
<td>Average Order Size of Marketable/Market Orders</td>
<td>703.2</td>
<td>345.0</td>
</tr>
<tr>
<td>Daily Number of Limit Orders</td>
<td>444.5</td>
<td>41.2</td>
</tr>
<tr>
<td>Average Order Size of Limit Orders</td>
<td>1,562.0</td>
<td>492.5</td>
</tr>
<tr>
<td>Average Percentage Marketable/Market Orders to Limit</td>
<td>48.1</td>
<td>46.1</td>
</tr>
<tr>
<td>Average Hidden Orders Usage</td>
<td>18.8</td>
<td>14.8</td>
</tr>
</tbody>
</table>

Panel B: Sample Period
Table 3: Regressions of Price Aggressiveness and Event Period Returns

The table shows regression coefficients that report changes in price aggressiveness around unanticipated corporate events (5 days before the event to 1 day after the event) controlling for order attributes and market conditions. Detailed definitions of the explanatory variables are provided in Section 4. Our sample is a set of 95 Euronext-Paris stocks around 101 unanticipated corporate events. We investigate buy (sell) orders around positive (negative) events separately and for subsamples of large and small event returns (less or more than 5%). Panel A reports individual day dummy effects and Panel B reports cumulative coefficient effects of the 5 days dummies before the event (day minus 5 to day minus 1). The time series coefficients are estimated on an event-by-event basis. Reported results are aggregated across events using the Bayesian framework of DuMouchel (1994).

### Table 3: Regressions of Price Aggressiveness and Event Period Returns

<table>
<thead>
<tr>
<th>Variable</th>
<th>Buy orders, Positive Events</th>
<th>Sell orders, Negative Events</th>
<th>Buy orders, Positive Events</th>
<th>Sell orders, Negative Events</th>
<th>Buy orders, Positive Events</th>
<th>Sell orders, Negative Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Events</td>
<td>Coefficient (1)</td>
<td>Coefficient (2)</td>
<td>Coefficient (3)</td>
<td>Coefficient (4)</td>
<td>Coefficient (5)</td>
<td>Coefficient (6)</td>
</tr>
<tr>
<td></td>
<td>(75.78)</td>
<td>(23.50)</td>
<td>(70.47)</td>
<td>(60.92)</td>
<td>(34.17)</td>
<td>(48.62)</td>
</tr>
<tr>
<td>Day Minus 5 (dummy)</td>
<td>-0.1802</td>
<td>0.0996</td>
<td>-0.0234</td>
<td>0.1563</td>
<td>-0.5875</td>
<td>0.0289</td>
</tr>
<tr>
<td></td>
<td>(-1.50)</td>
<td>(2.76)</td>
<td>(-0.26)</td>
<td>(2.45)</td>
<td>(-1.89)</td>
<td>(0.25)</td>
</tr>
<tr>
<td>Day Minus 4 (dummy)</td>
<td>-0.0362</td>
<td>-0.0066</td>
<td>-0.0412</td>
<td>0.0605</td>
<td>-0.2910</td>
<td>0.0105</td>
</tr>
<tr>
<td></td>
<td>(-1.07)</td>
<td>(-1.04)</td>
<td>(-0.56)</td>
<td>(0.89)</td>
<td>(-1.21)</td>
<td>(1.59)</td>
</tr>
<tr>
<td>Day Minus 3 (dummy)</td>
<td>-0.1069</td>
<td>0.0480</td>
<td>-0.0593</td>
<td>0.1071</td>
<td>-0.2080</td>
<td>-0.3191</td>
</tr>
<tr>
<td></td>
<td>(-2.08)</td>
<td>(1.30)</td>
<td>(-1.00)</td>
<td>(2.18)</td>
<td>(-2.22)</td>
<td>(-1.38)</td>
</tr>
<tr>
<td>Day Minus 2 (dummy)</td>
<td>-0.1524</td>
<td>0.0198</td>
<td>-0.0461</td>
<td>0.0111</td>
<td>-0.4219</td>
<td>0.2013</td>
</tr>
<tr>
<td></td>
<td>(-1.85)</td>
<td>(0.58)</td>
<td>(-1.18)</td>
<td>(0.21)</td>
<td>(-1.69)</td>
<td>(1.53)</td>
</tr>
<tr>
<td>Day Minus 1 (dummy)</td>
<td>-0.1147</td>
<td>0.0335</td>
<td>-0.1188</td>
<td>0.1069</td>
<td>-0.1077</td>
<td>0.0167</td>
</tr>
<tr>
<td></td>
<td>(-2.96)</td>
<td>(0.60)</td>
<td>(-2.43)</td>
<td>(0.86)</td>
<td>(-1.23)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>Day 0 &amp; Plus 1 (dummy)</td>
<td>-0.0179</td>
<td>0.0312</td>
<td>-0.0520</td>
<td>0.0432</td>
<td>0.0611</td>
<td>0.1278</td>
</tr>
<tr>
<td></td>
<td>(-0.48)</td>
<td>(0.82)</td>
<td>(-1.30)</td>
<td>(0.69)</td>
<td>(2.40)</td>
<td>(1.18)</td>
</tr>
<tr>
<td>Order exposure</td>
<td>0.7033</td>
<td>0.3215</td>
<td>0.6509</td>
<td>0.6803</td>
<td>0.8131</td>
<td>0.5921</td>
</tr>
<tr>
<td></td>
<td>(14.79)</td>
<td>(4.95)</td>
<td>(10.80)</td>
<td>(8.21)</td>
<td>(11.52)</td>
<td>(7.14)</td>
</tr>
<tr>
<td>Total order size (norm)</td>
<td>-0.0132</td>
<td>-0.0513</td>
<td>-0.0116</td>
<td>-0.0354</td>
<td>-0.0321</td>
<td>-0.0061</td>
</tr>
<tr>
<td></td>
<td>(-0.76)</td>
<td>(-0.78)</td>
<td>(-0.62)</td>
<td>(-0.71)</td>
<td>(-0.84)</td>
<td>(-0.50)</td>
</tr>
<tr>
<td>Bid-ask spread (norm)</td>
<td>26.4288</td>
<td>3.1271</td>
<td>31.9928</td>
<td>30.9306</td>
<td>15.0717</td>
<td>27.5898</td>
</tr>
<tr>
<td></td>
<td>(3.12)</td>
<td>(0.41)</td>
<td>(2.81)</td>
<td>(3.32)</td>
<td>(1.37)</td>
<td>(1.73)</td>
</tr>
<tr>
<td>Depth - same side (norm)</td>
<td>-1.3418</td>
<td>-4.8590</td>
<td>-9.2097</td>
<td>-2.3023</td>
<td>-0.1543</td>
<td>-17.0091</td>
</tr>
<tr>
<td></td>
<td>(-3.63)</td>
<td>(-2.56)</td>
<td>(-3.05)</td>
<td>(-2.69)</td>
<td>(-2.07)</td>
<td>(-2.90)</td>
</tr>
<tr>
<td>Depth - opposite side (norm)</td>
<td>0.5511</td>
<td>2.4014</td>
<td>0.7286</td>
<td>0.3726</td>
<td>0.1437</td>
<td>4.0519</td>
</tr>
<tr>
<td></td>
<td>(2.60)</td>
<td>(1.58)</td>
<td>(1.90)</td>
<td>(0.80)</td>
<td>(1.65)</td>
<td>(1.66)</td>
</tr>
<tr>
<td></td>
<td>(0.57)</td>
<td>(-0.69)</td>
<td>(-0.44)</td>
<td>(0.82)</td>
<td>(1.45)</td>
<td>(-0.98)</td>
</tr>
<tr>
<td>Waiting time</td>
<td>0.0010</td>
<td>0.0007</td>
<td>0.0012</td>
<td>0.0014</td>
<td>0.0004</td>
<td>0.0036</td>
</tr>
<tr>
<td></td>
<td>(2.23)</td>
<td>(1.32)</td>
<td>(2.34)</td>
<td>(2.04)</td>
<td>(0.54)</td>
<td>(1.89)</td>
</tr>
<tr>
<td>Trade frequency</td>
<td>-0.0014</td>
<td>-0.0008</td>
<td>-0.0015</td>
<td>-0.0004</td>
<td>-0.0013</td>
<td>-0.0009</td>
</tr>
<tr>
<td></td>
<td>(-2.00)</td>
<td>(-1.12)</td>
<td>(-1.72)</td>
<td>(-1.01)</td>
<td>(-1.02)</td>
<td>(-0.67)</td>
</tr>
<tr>
<td></td>
<td>(-2.85)</td>
<td>(-2.92)</td>
<td>(-3.78)</td>
<td>(-4.29)</td>
<td>(-3.12)</td>
<td>(-2.62)</td>
</tr>
<tr>
<td>Book order imbalance (norm)</td>
<td>-0.0351</td>
<td>-0.0673</td>
<td>-0.0425</td>
<td>-0.0782</td>
<td>-0.1328</td>
<td>0.0034</td>
</tr>
<tr>
<td></td>
<td>(-2.52)</td>
<td>(-2.41)</td>
<td>(-3.09)</td>
<td>(-4.71)</td>
<td>(-0.75)</td>
<td>(0.05)</td>
</tr>
<tr>
<td></td>
<td>(-9.20)</td>
<td>(-4.65)</td>
<td>(-8.03)</td>
<td>(-6.27)</td>
<td>(-5.08)</td>
<td>(-2.58)</td>
</tr>
<tr>
<td>Lag (displayed order size)</td>
<td>-0.4658</td>
<td>3.7116</td>
<td>-0.1595</td>
<td>1.5042</td>
<td>-0.1126</td>
<td>20.5119</td>
</tr>
<tr>
<td></td>
<td>(-2.24)</td>
<td>(1.47)</td>
<td>(-0.32)</td>
<td>(1.59)</td>
<td>(-2.49)</td>
<td>(2.40)</td>
</tr>
<tr>
<td>Last trade size (norm)</td>
<td>-0.0209</td>
<td>0.4518</td>
<td>1.0078</td>
<td>2.7113</td>
<td>-0.0112</td>
<td>-0.1815</td>
</tr>
<tr>
<td></td>
<td>(-1.64)</td>
<td>(0.29)</td>
<td>(2.99)</td>
<td>(2.25)</td>
<td>(-2.22)</td>
<td>(-1.32)</td>
</tr>
<tr>
<td>Market volatility</td>
<td>0.1982</td>
<td>-0.0653</td>
<td>-0.1107</td>
<td>0.0057</td>
<td>0.9537</td>
<td>-0.1916</td>
</tr>
<tr>
<td></td>
<td>(0.82)</td>
<td>(-0.79)</td>
<td>(-1.08)</td>
<td>(0.03)</td>
<td>(1.59)</td>
<td>(-1.15)</td>
</tr>
<tr>
<td>Industry volatility</td>
<td>0.0454</td>
<td>-0.0111</td>
<td>0.0204</td>
<td>-0.0209</td>
<td>0.0949</td>
<td>0.0043</td>
</tr>
<tr>
<td></td>
<td>(2.21)</td>
<td>(-1.07)</td>
<td>(1.68)</td>
<td>(-1.17)</td>
<td>(1.63)</td>
<td>(0.88)</td>
</tr>
</tbody>
</table>
### Panel B: Cumulative Effect of Day Minus 5 to Day Minus 1 Coefficients

<table>
<thead>
<tr>
<th>Variable</th>
<th>All Events</th>
<th>Event Period Absolute Return &lt;5%</th>
<th>Event Period Absolute Return &gt;5%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Buy orders, Positive Events</td>
<td>Sell orders, Negative Events</td>
<td>Buy orders, Positive Events</td>
</tr>
<tr>
<td></td>
<td>Coefficient</td>
<td>Coefficient</td>
<td>Coefficient</td>
</tr>
<tr>
<td>Cumulative Effect</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Day Minus 5 to Day Minus 1 (t-statistic)</td>
<td>-0.6045</td>
<td>0.1702</td>
<td>-0.3462</td>
</tr>
</tbody>
</table>
Table 4: Regressions of Price Aggressiveness - Inclusion in the SBF120 Index or Presence of Options Market

The table shows regression coefficients that report changes in price aggressiveness around unanticipated corporate events (5 days before the event to 1 day after the event) controlling for order attributes and market conditions. Detailed definitions of the explanatory variables are provided in Section 4. Our sample is a set of 95 Euronext-Paris stocks around 101 unanticipated corporate events. We investigate buy (sell) orders around positive (negative) events for subsamples of companies based on 1) whether or not they belong to the SBF 120 index and 2) have an active options marker. Panel A reports individual day dummy effects and Panel B reports cumulative coefficient effects of the 5 days dummies before the event (day minus 5 to day minus 1). The time series coefficients are estimated on an event-by-event basis. Reported results are aggregated across events using the Bayesian framework of DuMouchel (1994).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Events From Companies in SBF120</th>
<th>Events From Companies not in SBF120</th>
<th>Events From Companies with Options</th>
<th>Events From Companies without Options</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Buy orders, Positive Events</td>
<td>Sell orders, Positive Events</td>
<td>Buy orders, Negative Events</td>
<td>Sell orders, Negative Events</td>
</tr>
<tr>
<td></td>
<td>Coefficient (1)</td>
<td>Coefficient (2)</td>
<td>Coefficient (3)</td>
<td>Coefficient (4)</td>
</tr>
<tr>
<td>Day Minus 5 (dummy)</td>
<td>-0.0687</td>
<td>-0.0488</td>
<td>0.1799</td>
<td>-0.0713</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-1.56)</td>
<td>(-0.68)</td>
<td>(2.21)</td>
<td>(-2.01)</td>
</tr>
<tr>
<td>Day Minus 4 (dummy)</td>
<td>-0.1212</td>
<td>-0.1221</td>
<td>0.0981</td>
<td>-0.0249</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-2.27)</td>
<td>(-1.71)</td>
<td>(1.20)</td>
<td>(-0.53)</td>
</tr>
<tr>
<td>Day Minus 3 (dummy)</td>
<td>-0.0892</td>
<td>-0.0575</td>
<td>-0.1001</td>
<td>-0.0417</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-1.67)</td>
<td>(-0.91)</td>
<td>(-1.61)</td>
<td>(-0.83)</td>
</tr>
<tr>
<td>Day Minus 2 (dummy)</td>
<td>-0.0877</td>
<td>-0.1014</td>
<td>0.1499</td>
<td>-0.0490</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-2.50)</td>
<td>(-2.42)</td>
<td>(1.50)</td>
<td>(-1.51)</td>
</tr>
<tr>
<td>Day Minus 1 (dummy)</td>
<td>-0.0762</td>
<td>-0.0463</td>
<td>-0.1932</td>
<td>-0.0449</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-1.94)</td>
<td>(-0.87)</td>
<td>(-2.75)</td>
<td>(-1.21)</td>
</tr>
<tr>
<td>Day 0 &amp; Plus 1 (dummy)</td>
<td>-0.0656</td>
<td>-0.0196</td>
<td>0.0601</td>
<td>-0.0529</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-2.20)</td>
<td>(-0.32)</td>
<td>(1.37)</td>
<td>(-2.66)</td>
</tr>
<tr>
<td>Day Plus 2 (dummy)</td>
<td>-0.0681</td>
<td>-0.0828</td>
<td>0.1760</td>
<td>-0.0488</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-1.60)</td>
<td>(-1.16)</td>
<td>(1.56)</td>
<td>(-1.02)</td>
</tr>
<tr>
<td>Order exposure</td>
<td>0.6328</td>
<td>0.5981</td>
<td>0.7256</td>
<td>0.6143</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(8.21)</td>
<td>(8.62)</td>
<td>(10.51)</td>
<td>(6.45)</td>
</tr>
<tr>
<td>Total order size (norm)</td>
<td>-4.7856</td>
<td>0.1762</td>
<td>0.0016</td>
<td>-0.0230</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-2.82)</td>
<td>(0.39)</td>
<td>(0.10)</td>
<td>(-0.64)</td>
</tr>
<tr>
<td>Control variables</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Cumulative Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day Minus 5 to Day Minus 1</td>
<td>-0.3781</td>
<td>-0.3521</td>
<td>-0.8574</td>
<td>0.4246</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-2.31)</td>
<td>(-1.74)</td>
<td>(-2.14)</td>
<td>(2.17)</td>
</tr>
</tbody>
</table>

Panel A: Individual Day Minus 5 to Day Minus 1 Day Dummies

Panel B: Cumulative Effect of Day Minus 5 to Day Minus 1 Coefficients
Table 5: Logistic Regressions of Decision to Hide

The table shows logistic regression coefficients that report changes in the decision to hide for standing limit orders submitted around unanticipated corporate events (5 days before the event to 1 day after the event) controlling for order attributes and market conditions. Detailed definitions of the explanatory variables are provided in Section 4. Our sample is a set of 95 Euronext-Paris stocks around 101 unanticipated corporate events. We investigate buy (sell) orders around positive (negative) events separately. Panel A reports individual day dummy effects and Panel B reports cumulative coefficient effects of the 5 days dummies before the event (day minus 5 to day minus 1). Panels C and D report cumulative coefficient effects of the 5-day dummies before the event (day minus 5 to day minus 1) for subsamples of companies based on whether (or not) they belong to the SBF 120 index. Panels E and F report similar cumulative coefficient effects for subsample of companies based on whether or not they have an active options marker. The time series coefficients are estimated on a event-by-event basis. Reported results are aggregated across events using the Bayesian framework of DuMouchel (1994).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient (Buy orders, Positive Events)</th>
<th>Coefficient (Sell orders, Negative Events)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-2.8920 (-20.08)</td>
<td>-2.8394 (-13.12)</td>
</tr>
<tr>
<td>Price aggressiveness</td>
<td>-0.6754 (-1.24)</td>
<td>1.0847 (1.27)</td>
</tr>
<tr>
<td>Total order size (norm)</td>
<td>284.7097 (4.23)</td>
<td>187.4860 (4.69)</td>
</tr>
<tr>
<td>Bid-ask spread (norm)</td>
<td>-14.1272 (-1.20)</td>
<td>-21.1568 (-1.41)</td>
</tr>
<tr>
<td>Depth - same side (norm)</td>
<td>-20.3182 (-4.90)</td>
<td>-18.5887 (-2.88)</td>
</tr>
<tr>
<td>Depth - opposite side (norm)</td>
<td>-1.1419 (1.07)</td>
<td>-5.4541 (-1.62)</td>
</tr>
<tr>
<td>Volatility</td>
<td>-121.7095 (-3.34)</td>
<td>-38.1052 (-1.84)</td>
</tr>
<tr>
<td>Trade frequency</td>
<td>-6.94E-06 (-0.01)</td>
<td>0.0003 (0.52)</td>
</tr>
<tr>
<td>HiddenSameSide (norm)</td>
<td>4.4443 (2.25)</td>
<td>6.6096 (2.09)</td>
</tr>
<tr>
<td>Same price book displayed depth (norm)</td>
<td>0.4473 (1.60)</td>
<td>-1.2861 (-2.26)</td>
</tr>
<tr>
<td>Book order imbalance (norm)</td>
<td>-0.0449 (-1.19)</td>
<td>-0.4683 (-1.69)</td>
</tr>
<tr>
<td>Last trade size (norm)</td>
<td>-2.8053 (-4.01)</td>
<td>-1.3999 (-0.78)</td>
</tr>
<tr>
<td>Industry volatility</td>
<td>0.2611 (0.32)</td>
<td>0.0035 (0.10)</td>
</tr>
<tr>
<td>Variable</td>
<td>Buy orders, Positive Events</td>
<td>Sell orders, Negative Events</td>
</tr>
<tr>
<td>----------</td>
<td>----------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td></td>
<td>Coefficient (1)</td>
<td>Coefficient (2)</td>
</tr>
<tr>
<td>Panel B: Cumulative Effect of Day Dummies - Overall</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumulative Effect: Day Minus 5 to Day Minus 1</td>
<td>1.4193</td>
<td>-1.7257</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(2.01)</td>
<td>(-2.34)</td>
</tr>
<tr>
<td>Panel C: Cumulative Effect of Day Dummies - Companies in the SBF120 Index</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumulative Effect: Day Minus 5 to Day Minus 1</td>
<td>1.1066</td>
<td>-1.9511</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(1.87)</td>
<td>(-1.55)</td>
</tr>
<tr>
<td>Panel D: Cumulative Effect of Day Dummies - Companies not in the SBF120 Index</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumulative Effect: Day Minus 5 to Day Minus 1</td>
<td>1.0482</td>
<td>-2.3973</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(0.85)</td>
<td>(-2.35)</td>
</tr>
<tr>
<td>Panel E: Cumulative Effect of Day Dummies - Companies with Options</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumulative Effect: Day Minus 5 to Day Minus 1</td>
<td>1.7273</td>
<td>-1.6068</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(1.85)</td>
<td>(-1.22)</td>
</tr>
<tr>
<td>Panel F: Cumulative Effect of Day Dummies - Companies without Options</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumulative Effect: Day Minus 5 to Day Minus 1</td>
<td>0.8468</td>
<td>-1.8679</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(1.02)</td>
<td>(-2.46)</td>
</tr>
</tbody>
</table>
Table 6: Tobit Regressions of Magnitude of Hidden Size

The table shows tobit regression coefficients that report changes of the magnitude of hidden size for standing limit orders submitted around unanticipated corporate events (5 days before the event to 1 day after the event) controlling for order attributes and market conditions. Detailed definitions of the explanatory variables are provided in Section 4. Our sample is a set of 95 Euronext-Paris stocks around 101 unanticipated corporate events. We investigate buy (sell) orders around positive (negative) events separately. Panel A reports individual day dummy effects and Panel B reports cumulative coefficient effects of the 5 days dummies before the event (day minus 5 to day minus 1). Panels C and D report cumulative coefficient effects of the 5-day dummies before the event (day minus 5 to day minus 1) for subsamples of companies based on whether (or not) they belong to the SBF 120 index. Panels E and F report similar cumulative coefficient effects for subsample of companies based on whether or not they have an active options marker. The time series coefficients are estimated on a event-by-event basis. Reported results are aggregated across events using the Bayesian framework of DuMouchel (1994).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Buy orders, Positive Events</th>
<th>Sell orders, Negative Events</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>Coefficient</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.1027</td>
<td>-0.2880</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-0.67)</td>
<td>(1.79)</td>
</tr>
<tr>
<td>Day Minus 5 (dummy)</td>
<td>-0.2054</td>
<td>-0.2451</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-1.85)</td>
<td>(-1.41)</td>
</tr>
<tr>
<td>Day Minus 4 (dummy)</td>
<td>-0.1493</td>
<td>-0.0001</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-1.15)</td>
<td>(-0.00)</td>
</tr>
<tr>
<td>Day Minus 3 (dummy)</td>
<td>-0.0442</td>
<td>-0.1186</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-0.57)</td>
<td>(-2.14)</td>
</tr>
<tr>
<td>Day Minus 2 (dummy)</td>
<td>0.0791</td>
<td>0.0140</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(0.61)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>Day Minus 1 (dummy)</td>
<td>0.0390</td>
<td>0.0765</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(0.40)</td>
<td>(0.73)</td>
</tr>
<tr>
<td>Day 0 &amp; Plus 1 (dummy)</td>
<td>0.0580</td>
<td>-0.0824</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(0.92)</td>
<td>(-0.42)</td>
</tr>
<tr>
<td>Day Plus 2 (dummy)</td>
<td>0.0076</td>
<td>-0.3375</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(0.05)</td>
<td>(-1.83)</td>
</tr>
<tr>
<td>Price aggressiveness</td>
<td>-2.3127</td>
<td>-2.1466</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-1.68)</td>
<td>(-1.29)</td>
</tr>
<tr>
<td>Total order size (norm)</td>
<td>0.3034</td>
<td>1.3982</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(4.25)</td>
<td>(3.52)</td>
</tr>
<tr>
<td>Bid-ask spread (norm)</td>
<td>-0.1073</td>
<td>-0.7237</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-0.08)</td>
<td>(-0.35)</td>
</tr>
<tr>
<td>Depth - same side (norm)</td>
<td>-0.1544</td>
<td>-0.2238</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-0.91)</td>
<td>(-0.44)</td>
</tr>
<tr>
<td>Depth - opposite side (norm)</td>
<td>0.0523</td>
<td>-0.7361</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(0.33)</td>
<td>(-1.12)</td>
</tr>
<tr>
<td>Volatility</td>
<td>-17.4197</td>
<td>5.7855</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-1.57)</td>
<td>(0.56)</td>
</tr>
<tr>
<td>Waiting time</td>
<td>0.0002</td>
<td>0.0000</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(0.33)</td>
<td>(0.21)</td>
</tr>
<tr>
<td>Trade frequency</td>
<td>0.0050</td>
<td>0.0026</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(1.25)</td>
<td>(1.18)</td>
</tr>
<tr>
<td>HiddenSameSide (norm)</td>
<td>0.3363</td>
<td>1.6736</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(0.60)</td>
<td>(1.05)</td>
</tr>
<tr>
<td>Same price book displayed depth (norm)</td>
<td>-0.0382</td>
<td>-0.5880</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-0.56)</td>
<td>(-0.99)</td>
</tr>
<tr>
<td>Book order imbalance (norm)</td>
<td>-0.0688</td>
<td>0.0097</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-1.69)</td>
<td>(0.16)</td>
</tr>
<tr>
<td>Last trade size (norm)</td>
<td>-0.4486</td>
<td>1.1302</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-1.29)</td>
<td>(1.68)</td>
</tr>
<tr>
<td>Market volatility</td>
<td>-3.3918</td>
<td>0.2542</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-2.48)</td>
<td>(1.00)</td>
</tr>
<tr>
<td>Industry volatility</td>
<td>0.0944</td>
<td>0.0287</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(0.68)</td>
<td>(0.27)</td>
</tr>
<tr>
<td>Variable</td>
<td>Magnitude of hidden order size</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>--------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Buy orders, Positive Events</td>
<td>Sell orders, Negative Events</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel B: Cumulative Effect of Day Dummies - Overall</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumulative Effect: Day Minus 5 to Day Minus 1</td>
<td>-1.7683</td>
<td>-1.2516</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-3.33)</td>
<td>(-2.44)</td>
</tr>
<tr>
<td>Panel C: Cumulative Effect of Day Dummies - Companies in the SBF120 Index</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumulative Effect: Day Minus 5 to Day Minus 1</td>
<td>-1.2805</td>
<td>-0.5382</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-1.48)</td>
<td>(-1.16)</td>
</tr>
<tr>
<td>Panel D: Cumulative Effect of Day Dummies - Companies not in the SBF120 Index</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumulative Effect: Day Minus 5 to Day Minus 1</td>
<td>-2.1779</td>
<td>-1.7885</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-2.99)</td>
<td>(-2.24)</td>
</tr>
<tr>
<td>Panel E: Cumulative Effect of Day Dummies - Companies with Options</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumulative Effect: Day Minus 5 to Day Minus 1</td>
<td>-0.3717</td>
<td>-0.7375</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-0.61)</td>
<td>(-1.32)</td>
</tr>
<tr>
<td>Panel F: Cumulative Effect of Day Dummies - Companies without Options</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cumulative Effect: Day Minus 5 to Day Minus 1</td>
<td>-2.3479</td>
<td>-1.7502</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-3.42)</td>
<td>(-2.38)</td>
</tr>
</tbody>
</table>
Table 7: Order submission strategies and limit order time-to-execution

The table reports parameter estimates of an econometric model of limit order time-to-execution using survival analysis, following Lo, Mackinlay, and Zhang (2002). The model describes an accelerated failure time specification of limit order execution times under the generalized gamma distribution. Our sample is a set of 95 Euronext-Paris stocks around 101 unanticipated corporate events. We investigate buy (sell) orders around positive (negative) events. We report changes in time-to-execution around unanticipated corporate events (5 days before the event to 1 day after the event). Our control variables are: the distance in basis points of the order’s limit price from the quote midpoint (midquote - limit price); an indicator variable that equals one if the prior trade is buyer-initiated and equals zero otherwise (last trade buy indicator); the displayed depth at the best bid (ask) for a buy (sell) order (same side depth); the square of the previous measure to account for non-linearity (same side depth squared); the displayed depth at the best ask (bid) for a buy (sell) order (opposite side depth); the total (exposed plus hidden) size of the order (order Size); the number of trades in the last hour (trade frequency); an indicator valuable that equals one if the order has hidden size and equals zero otherwise (hidden order). Panel A reports individual day dummy effects and Panel B reports cumulative coefficient effects of the 5 days dummies before the event (day minus 5 to day minus 1). Panels C and D report cumulative coefficient effects of the 5-day dummies before the event (day minus 5 to day minus 1) for subsamples of companies based on whether (or not) they belong to the SBF 120 index. Panels E and F report similar cumulative coefficient effects for subsample of companies based on whether or not they have an active options marker. Reported results are aggregated across events using the Bayesian framework of DuMouchel (1994).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Event regressions</th>
<th>Buy orders, Positive Coefficient (1)</th>
<th>Sell orders, Negative Coefficient (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Coefficient (1)</td>
<td>Coefficient (2)</td>
</tr>
<tr>
<td>Intercept</td>
<td></td>
<td>9.1890 (13.06)</td>
<td>14.7985 (13.63)</td>
</tr>
<tr>
<td>Day Minus 5 (dummy)</td>
<td></td>
<td>-0.7588 (-1.87)</td>
<td>-0.8058 (-2.35)</td>
</tr>
<tr>
<td>Day Minus 4 (dummy)</td>
<td></td>
<td>-0.7109 (-2.19)</td>
<td>-0.2320 (-0.75)</td>
</tr>
<tr>
<td>Day Minus 3 (dummy)</td>
<td></td>
<td>-0.1890 (-0.73)</td>
<td>-0.3581 (-1.09)</td>
</tr>
<tr>
<td>Day Minus 2 (dummy)</td>
<td></td>
<td>-0.3274 (-1.01)</td>
<td>-0.2397 (-1.23)</td>
</tr>
<tr>
<td>Day Minus 1 (dummy)</td>
<td></td>
<td>0.1344 (0.50)</td>
<td>0.0637 (0.25)</td>
</tr>
<tr>
<td>Day 0 &amp; Plus 1 (dummy)</td>
<td></td>
<td>-0.1417 (-0.43)</td>
<td>-0.1877 (-1.22)</td>
</tr>
<tr>
<td>Day Plus 2 (dummy)</td>
<td></td>
<td>-0.4730 (-2.85)</td>
<td>-0.2739 (-0.69)</td>
</tr>
<tr>
<td>Midquote - limit price</td>
<td></td>
<td>3.6879 (3.02)</td>
<td>-6.6568 (-2.88)</td>
</tr>
<tr>
<td>Last trade buy indicator</td>
<td></td>
<td>-0.1330 (-2.52)</td>
<td>-0.3285 (-1.72)</td>
</tr>
<tr>
<td>Same side depth (norm)</td>
<td></td>
<td>0.0872 (3.49)</td>
<td>0.0478 (2.64)</td>
</tr>
<tr>
<td>Same side depth squared</td>
<td></td>
<td>0.0029 (0.03)</td>
<td>0.0156 (0.36)</td>
</tr>
<tr>
<td>Variable</td>
<td>Coefficient (1)</td>
<td>Coefficient (2)</td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td>Opposite side depth (norm)</td>
<td>-0.2199</td>
<td>-0.3902</td>
<td></td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-7.18)</td>
<td>(-5.85)</td>
<td></td>
</tr>
<tr>
<td>Order Size</td>
<td>0.1884</td>
<td>0.1091</td>
<td></td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(4.81)</td>
<td>(2.83)</td>
<td></td>
</tr>
<tr>
<td>Trade frequency</td>
<td>-0.0060</td>
<td>-0.0045</td>
<td></td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-4.27)</td>
<td>(-4.40)</td>
<td></td>
</tr>
<tr>
<td>Hidden order indicator</td>
<td>0.9269</td>
<td>1.2866</td>
<td></td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(5.53)</td>
<td>(6.45)</td>
<td></td>
</tr>
<tr>
<td>Scale (fitted distribution)</td>
<td>3.0459</td>
<td>2.0031</td>
<td></td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(8.46)</td>
<td>(5.80)</td>
<td></td>
</tr>
<tr>
<td>Shape (fitted distribution)</td>
<td>0.3185</td>
<td>3.6109</td>
<td></td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(0.52)</td>
<td>(3.78)</td>
<td></td>
</tr>
</tbody>
</table>

**Panel B: Cumulative Effect of Day Dummies - Overall**

| Cumulative Effect: Day Minus 5 to Day Minus 1 | -1.7089         | -1.2612         |
| (t-statistic)                                | (-3.17)         | (-3.04)         |

**Panel C: Cumulative Effect of Day Dummies - Companies in the SBF120 Index**

| Cumulative Effect: Day Minus 5 to Day Minus 1 | -1.9295         | -1.1736         |
| (t-statistic)                                | (-2.10)         | (-2.88)         |

**Panel D: Cumulative Effect of Day Dummies - Companies not in the SBF120 Index**

| Cumulative Effect: Day Minus 5 to Day Minus 1 | -1.6231         | -1.7499         |
| (t-statistic)                                | (-2.52)         | (-1.89)         |

**Panel E: Cumulative Effect of Day Dummies - Companies with Options**

| Cumulative Effect: Day Minus 5 to Day Minus 1 | -2.1491         | -1.3092         |
| (t-statistic)                                | (-1.43)         | (-3.27)         |

**Panel F: Cumulative Effect of Day Dummies - Companies without Options**

| Cumulative Effect: Day Minus 5 to Day Minus 1 | -2.0108         | -1.2612         |
| (t-statistic)                                | (-3.34)         | (-3.04)         |
Table 8: Regressions of implementation shortfall, effective spreads trading costs, and opportunity cost

The table shows regression coefficients that report changes in execution costs around unanticipated corporate events (5 days before the event to 1 day after the event) controlling for order attributes and market conditions. Our sample is a set of 95 Euronext-Paris stocks around 101 unanticipated corporate events. We investigate buy (sell) orders around positive (negative) events. Execution costs are based on the implementation shortfall approach proposed by Perold (1988), defined as follows. For a buy order, effective spread cost is defined as the difference between the filled price of each submitted order and the mid-quote price at the time of order submission. Opportunity cost is defined as the difference between the closing price on the day of order cancellation or expiration and the quote midpoint at the time of order submission. Implementation shortfall is the summation of the two costs. We control for three variables that represent order attributes (price aggressiveness, order size, and hidden order indicator) and two variables that represent market conditions during the trading hour prior to order submission (trading frequency and return volatility). For effective spread cost, we report regression results conditional on partial execution (effective spread cost ≠ 0, Columns 3 and 4). For opportunity cost, we report regression results conditional on partial non-execution (opportunity cost ≠ 0, columns 5 and 6). Panel A reports individual day dummy effects and Panel B reports cumulative coefficient effects of the 5-day dummies before the event (day minus 5 to day minus 1). Panels C and D report cumulative coefficient effects of the 5-day dummies before the event (day minus 5 to day minus 1) for subsamples of companies based on whether (or not) they belong to the SBF 120 index. Panels E and F report similar cumulative coefficient effects for subsample of companies based on whether or not they have an active options marker. Reported results are aggregated across events using the Bayesian framework of DuMouchel (1994).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Implementation Shortfall</th>
<th>Effective Spread cost: fill rate &gt;0%</th>
<th>Opportunity cost: fill rate &lt;100%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Buy orders, Positive Events</td>
<td>Sell orders, Negative Events</td>
<td>Buy orders, Positive Events</td>
</tr>
<tr>
<td></td>
<td>Coefficient (1)</td>
<td>Coefficient (2)</td>
<td>Coefficient (3)</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.0541</td>
<td>0.0560</td>
<td>0.0651</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(3.03)</td>
<td>(2.81)</td>
<td>(5.77)</td>
</tr>
<tr>
<td>Day Minus 5 (dummy)</td>
<td>0.0275</td>
<td>-0.0553</td>
<td>0.0044</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(0.88)</td>
<td>(-1.42)</td>
<td>(0.39)</td>
</tr>
<tr>
<td>Day Minus 4 (dummy)</td>
<td>0.0037</td>
<td>-0.0115</td>
<td>-0.0055</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(0.12)</td>
<td>(-0.42)</td>
<td>(-1.23)</td>
</tr>
<tr>
<td>Day Minus 3 (dummy)</td>
<td>0.0846</td>
<td>-0.0350</td>
<td>-0.0006</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(2.67)</td>
<td>(-2.59)</td>
<td>(-0.10)</td>
</tr>
<tr>
<td>Day Minus 2 (dummy)</td>
<td>0.0161</td>
<td>0.0114</td>
<td>0.0029</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(0.36)</td>
<td>(0.33)</td>
<td>(0.21)</td>
</tr>
<tr>
<td>Day Minus 1 (dummy)</td>
<td>0.1018</td>
<td>-0.0142</td>
<td>0.0327</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(2.61)</td>
<td>(-0.48)</td>
<td>(2.14)</td>
</tr>
<tr>
<td>Day 0 &amp; Plus 1 (dummy)</td>
<td>0.0830</td>
<td>-0.0181</td>
<td>0.0005</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(3.51)</td>
<td>(-0.78)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>Day Plus 2 (dummy)</td>
<td>0.0011</td>
<td>-0.0381</td>
<td>0.0043</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(0.05)</td>
<td>(-1.67)</td>
<td>(0.74)</td>
</tr>
<tr>
<td>Price aggressiveness</td>
<td>-0.1933</td>
<td>0.0005</td>
<td>22.4275</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-0.80)</td>
<td>(0.43)</td>
<td>(6.05)</td>
</tr>
<tr>
<td>Order size (million shares)</td>
<td>0.0584</td>
<td>-5.3622</td>
<td>-5.9165</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(0.83)</td>
<td>(-4.11)</td>
<td>(-1.33)</td>
</tr>
<tr>
<td>Hidden order (dummy)</td>
<td>-0.0175</td>
<td>-0.0034</td>
<td>-0.0236</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-2.95)</td>
<td>(-6.69)</td>
<td>(-6.15)</td>
</tr>
<tr>
<td>Trading frequency</td>
<td>-0.0001</td>
<td>0.0008</td>
<td>0.0000</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-1.77)</td>
<td>(1.36)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Volatility</td>
<td>-0.8071</td>
<td>-3.6028</td>
<td>8.1933</td>
</tr>
<tr>
<td>(t-statistic)</td>
<td>(-0.25)</td>
<td>(-2.45)</td>
<td>(1.63)</td>
</tr>
</tbody>
</table>

Panel A: Individual Day Minus 5 to Day Minus 1 Day Dummies
<table>
<thead>
<tr>
<th>Variable</th>
<th>Implementation Shortfall</th>
<th>Effective Spread cost: fill rate &gt;0%</th>
<th>Opportunity cost: fill rate &lt;100%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Buy orders, Positive Events</td>
<td>Coefficient</td>
<td>(1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(t-statistic)</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>Sell orders, Negative Events</td>
<td>Coefficient</td>
<td>(3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(t-statistic)</td>
<td>(4)</td>
</tr>
<tr>
<td></td>
<td>Buy orders, Positive Events</td>
<td>Coefficient</td>
<td>(5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(t-statistic)</td>
<td>(6)</td>
</tr>
<tr>
<td></td>
<td>Sell orders, Negative Events</td>
<td>Coefficient</td>
<td>(7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(t-statistic)</td>
<td>(8)</td>
</tr>
</tbody>
</table>

**Panel B: Cumulative Effect of Day Dummies - Overall**

**Cumulative Effect:**
Day Minus 5 to Day Minus 1

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>(t-statistic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2422</td>
<td>(2.43)</td>
</tr>
<tr>
<td>-0.0771</td>
<td>(-1.10)</td>
</tr>
<tr>
<td>0.0107</td>
<td>(0.50)</td>
</tr>
<tr>
<td>0.0020</td>
<td>(0.19)</td>
</tr>
<tr>
<td>0.3234</td>
<td>(2.17)</td>
</tr>
<tr>
<td>-0.1849</td>
<td>(-1.77)</td>
</tr>
</tbody>
</table>

**Panel C: Cumulative Effect of Day Dummies - Companies in the SBF120 Index**

**Cumulative Effect:**
Day Minus 5 to Day Minus 1

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>(t-statistic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2623</td>
<td>(3.10)</td>
</tr>
<tr>
<td>0.1519</td>
<td>(1.86)</td>
</tr>
<tr>
<td>-0.0183</td>
<td>(-1.13)</td>
</tr>
<tr>
<td>0.0255</td>
<td>(1.54)</td>
</tr>
<tr>
<td>0.4327</td>
<td>(2.25)</td>
</tr>
<tr>
<td>0.2075</td>
<td>(2.08)</td>
</tr>
</tbody>
</table>

**Panel D: Cumulative Effect of Day Dummies - Companies not in the SBF120 Index**

**Cumulative Effect:**
Day Minus 5 to Day Minus 1

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>(t-statistic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2620</td>
<td>(1.83)</td>
</tr>
<tr>
<td>-0.0871</td>
<td>(-0.85)</td>
</tr>
<tr>
<td>0.1007</td>
<td>(2.06)</td>
</tr>
<tr>
<td>-0.0079</td>
<td>(-0.41)</td>
</tr>
<tr>
<td>0.4180</td>
<td>(2.21)</td>
</tr>
<tr>
<td>-0.2115</td>
<td>(-1.72)</td>
</tr>
</tbody>
</table>

**Panel E: Cumulative Effect of Day Dummies - Companies with Options**

**Cumulative Effect:**
Day Minus 5 to Day Minus 1

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>(t-statistic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2430</td>
<td>(3.23)</td>
</tr>
<tr>
<td>-0.0777</td>
<td>(-0.57)</td>
</tr>
<tr>
<td>-0.0135</td>
<td>(-0.54)</td>
</tr>
<tr>
<td>0.0188</td>
<td>(1.33)</td>
</tr>
<tr>
<td>0.2456</td>
<td>(1.54)</td>
</tr>
<tr>
<td>-0.2082</td>
<td>(-0.77)</td>
</tr>
</tbody>
</table>

**Panel F: Cumulative Effect of Day Dummies - Companies without Options**

**Cumulative Effect:**
Day Minus 5 to Day Minus 1

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>(t-statistic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2750</td>
<td>(2.35)</td>
</tr>
<tr>
<td>-0.0895</td>
<td>(-0.80)</td>
</tr>
<tr>
<td>0.0485</td>
<td>(1.42)</td>
</tr>
<tr>
<td>-0.0073</td>
<td>(-0.34)</td>
</tr>
<tr>
<td>0.4511</td>
<td>(2.76)</td>
</tr>
<tr>
<td>-0.2052</td>
<td>(-1.44)</td>
</tr>
</tbody>
</table>
Appendix

The goal of this appendix is to present a simple model that supports the order of expected payoffs we have postulated; i.e. $a < b < c < d < e < f$. To that end, we consider a market for a single asset with a tick size $\theta$.

At time zero, the book is populated as follows:

<table>
<thead>
<tr>
<th>Bid Price</th>
<th>Ask</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta+4\theta$</td>
<td>1</td>
</tr>
<tr>
<td>$\beta+3\theta$</td>
<td>1</td>
</tr>
<tr>
<td>$\beta+2\theta$</td>
<td>1</td>
</tr>
<tr>
<td>$\beta$</td>
<td></td>
</tr>
<tr>
<td>$\beta-\theta$</td>
<td></td>
</tr>
<tr>
<td>$\beta-2\theta$</td>
<td></td>
</tr>
</tbody>
</table>

That is, the bid price at time zero is $\beta$, the spread is $2\theta$, and at each price level the depth is one. In each trading round, a stochastic trader who wants to trade two units shows up. The trader can be either a buyer or a seller with equal probabilities. If the spread is wider than one tick, the trader submits a limit order, otherwise a market orders. Only quotes, but not the depth, are visible.

We assume that two informed traders show up at time zero and see the above quotes. Without loss of generality, we assume the informed traders would like to sell. We further assume that each would like to sell one unit. If both traders use the same strategy, than each received half of combined expected payoff. Thus, when both use market orders, then one unit is sold at $\beta$ and the second unit at $\beta-\theta$, so the expected payoff is $\beta-0.5\theta$. This value corresponds to $(v-\overline{v})^2b$ in our matrix payoff. When an informed trader uses a limit order, we assume the order sits in the book for two rounds. If after two rounds the order is not executed, it is converted to a market order.

Consider now the case that one trader uses a sell market order while the other uses a sell limit order. After the two orders are submitted, the books look like
The payoff of the market order strategy is $\beta$ and this corresponds to $(v-\bar{v})^2d$ in our matrix payoff. To calculate the expected payoff associated with limit orders, we need to consider 4 possible scenarios, according to the arrival of the stochastic trader in the next two periods. If the stochastic trader in the next period is a buyer, then, because the current spread is greater than one tick, he posts a new bid $\beta$. Next, if the next stochastic trader is also a buyer, then he buys at the ask price, $\beta+\theta$. Otherwise he is a seller and he hits the bid at $\beta$. In that case, the informed trader’s limit order was not executed. The limit order is converted to a market order, and executed at $\beta-\theta$. Thus, the expected payoff of using a limit order, conditional on the first stochastic trader being a buyer

$$0.5(\beta+\theta)+0.5(\beta-\theta)=\beta$$

If the stochastic trader is first a seller, then he post an offer $\beta$. Regardless of what the type of the second stochastic trader is, the limit order of the informed trader is not executed, and it is converted to a market order. If the second stochastic trader was a buyer, the informed can sell at $\beta-\theta$, otherwise he sells at $\beta-2\theta$.

Thus, the expected payoff conditional on the first stochastic trader being a seller is $\beta-1.5\theta$. Therefore

$$(v-\bar{v})^2a=0.5(\beta-1.5\theta)+0.5\beta=\beta-0.75\theta$$

In a similar manner, we compute the expected payoffs of other strategies. The results are

$$(v-\bar{v})^2b=\beta-0.50\theta$$

$$(v-\bar{v})^2c=\beta-0.25\theta$$

$$(v-\bar{v})^2d=\beta$$

$$(v-\bar{v})^2e=\beta$$

$$(v-\bar{v})^2f=\beta+0.25\theta$$

The order is as we postulated in the payoff matrix, i.e., $a<b<c<d\leq e<f$. 
Intermediating Adverse Selection

We propose a parsimonious model of over-the-counter trading under asymmetric information to study the presence of intermediation chains that stand between well informed market participants and uninformed ones. Moderately informed intermediaries can fulfill an important economic role of layering adverse selection over several transactions. Informed market participants may prefer to trade through one or more of these intermediaries as they improve trade efficiency but also reduce the surplus accruing to uninformed traders. Our model makes novel predictions about optimal network formation when adverse selection problems impede the efficiency of trade.

Keywords: Intermediation Chains, OTC Trading Networks, Adverse Selection, Asymmetric Information
JEL Codes: G20, D82, D85
1 Introduction

In over-the-counter (OTC) trading networks securities regularly pass through several intermediate traders before reaching their ultimate buyers. Informed market participants interested in acquiring a security might find optimal to use other institutions, potentially with different levels of financial expertise, as intermediaries between them and less informed parties. In this paper, we propose a parsimonious model of OTC trading under asymmetric information to analyze the endogenous involvement of intermediaries who stand between buyers and sellers.

Our model shows how moderately informed traders can fulfill an important economic role in intermediating trade by “smoothing” adverse selection over several layers of transactions. Gains to trade that would otherwise be destroyed due to adverse selection between two asymmetrically informed traders can be preserved if the less informed party trades with a moderately informed intermediary who then trades with the most informed party. In contrast to standard intermediation theories, our intuition can be extended to explain why trading often involves multiple intermediaries rather than one dominant central dealer. We show that intermediation chains can help sustain high liquidity in financial markets when adverse selection problems impede direct trading between the ultimate buyers and sellers of a security. However, long intermediation chains can also make liquidity more fragile if the degree of uncertainty, or adverse selection, may suddenly change, as they destroy trading surplus that would otherwise be preserved with fewer intermediaries.

The following numerical example helps to illustrate the main intuition on how intermediation creates value in our model. We start with two asymmetrically informed agents who wish to trade a security through bilateral bargaining in order to realize exogenous gains to trade (say, with risk sharing benefits). Here, we deliberately omit details about how bargaining takes place between agents or how trade creates a surplus and instead focus on how intermediation affects the adverse selection problem among agents. Ignoring gains to trade, the value of the security is either $100 or $0 with equal probabilities. Agent $I$ is perfectly informed about the value of the security, but agent $U$ is uninformed. The adverse selection problem agent $U$ faces results from the fact that he values the security at $50, its unconditional expected value, but agent $I$ can condition his trading behavior based on his knowledge that the security is either worth $100 or $0. If the gains to trade available are positive but small relative to the spread in possible valuations, efficient trade will be
impossible to achieve because of this adverse selection problem. Our model formalizes the idea that involving a moderately informed agent $M$ as an intermediary could improve trade efficiency. If agent $M$ receives a signal about the value of the security that is accurate with probability 0.75, he will either value the security at $75 or at $25. Agent $U$ will now face a counterparty whose spread in estimated valuation of the security is smaller than agent $I$’s ($50 rather than $100). Agent $U$ will thus be more willing to trade with such counterparty and realize the gains to trade. Trading with the fully informed agent $I$ will also be easier for agent $M$, with his informative signal in hands, than it was for the uninformed agent $U$. Agent $M$ can thus serve as an intermediary between agents $I$ and $U$ and reduce the inefficiencies in trade that the adverse selection problem causes. In some circumstances, trade efficiency is improved by reallocating the adverse selection problem further over multiple transactions involving long chains of intermediaries rather than just one direct transaction between the initial seller and the final buyer. Our model also speaks to how the surplus from trade is split among all involved traders.

Our paper contributes to the literature on trade-facilitating intermediation. We know that intermediaries may enhance trade efficiency by economizing on transaction costs (see Townsend 1978) and on monitoring costs (see Diamond 1984), or by alleviating search frictions (see Rubinstein and Wolinsky 1987, Yavaş 1994, Duffie, Gárleanu, and Pedersen 2005). And since Myerson and Satterthwaite (1983), we also know that the involvement of an uninformed third party who subsidizes transactions can eliminate asymmetric information problems in bilateral trade. Biglaiser (1993) also shows that a middleman who offers a warranty on the quality of the good can improve welfare. The paper features a dynamic model with middlemen acquiring costly expertise that will later allow them to detect the true quality of goods produced by informed sellers and sold to uninformed buyers. Each middleman is involved in multiple trades through time, hence has more incentives to acquire expertise than any other potential buyer. A middleman also can offer uninformed buyers a warranty on the quality of the goods that is credible thanks to his future presence in the market. Reputational concerns, through the informal warranty offered, are essential for the existence of a welfare-enhancing equilibrium. Li (1998) also studies a model in which intermediaries can acquire perfect information about the quality of goods sold by informed sellers. In that paper, middlemen can be trustworthy in equilibrium, not because of reputational concerns as in Biglaiser (1993), but because of the existence of a sufficiently large mass of informed buyers who discipline cheating
middlemen by forcing them to hold on to low-quality goods. Contrary to these models, our model considers the possibility that the intermediary’s information set differs from that of the fully informed trader and from that of the uninformed trader. In our static model without warranties or reputational concerns, the involvement of an intermediary who is either fully informed or totally uninformed would not improve trade efficiency. The idea that moderately informed intermediaries can solve adverse selection problems, without the need for credible warranties or the threat of disciplinary actions, simply by layering it over more than one transaction is unique to our paper.

Babus (2012) allows for incomplete information about traders’ histories and reputational concerns in a trading network model where agents meet sporadically. In equilibrium, a central intermediary becomes involved in all trades. This unique intermediary interacts repeatedly with traders and can heavily penalize anyone who were to default on his obligations. Our model instead predicts the existence of multiple intermediaries who are all needed to reallocate a large adverse selection problem into several layers of transactions where each trader is only slightly better or worse informed than his counterparty. Our paper is also related to Duffie, Malamud, and Manso (2012) who endogenize the acquisition of information by traders, taking as given the search for counterparties, or to Gofman (2011) who studies the inefficiencies in resource allocation that arise when traders face (non-informational) bargaining frictions in a sparse OTC network. Instead, in our paper we endogenize the trading network, taking as given the existence of an adverse selection problem.

Empirically, many papers document that interdealer trading, a key feature of our model, accounts for a substantial fraction of trading in financial markets (see Gould and Kleidon 1994, Reiss and Werner 1995, Lyons 1996, Vogler 1997, Hansch, Naik, and Viswanathan 1998, Reiss and Werner 1998). For example, Lyons (1996) estimates that interdealer trading accounts for up to 85% of transactions in foreign exchange markets. More recently, Weller (2013) shows that a median number of 2 intermediaries are involved between an initial seller and a final buyer of gold, silver, and copper futures contracts. Furthermore, 10% of intraday transactions require the involvement of at least 5 intermediaries.

While some of these empirical papers find evidence that inventory risk sharing helps explain some of the interdealer trading we observe in financial markets, some key features of interdealer trading still remain unexplained. Manaster and Mann (1996) find that the relationship between trader inventories and transaction prices we observe in the data violates standard predictions of
inventory control models such as Ho and Stoll (1983). Manaster and Mann (1996) conclude that a central characteristic of the market maker sector is the fact that intermediaries have “heterogeneous levels of information and/or trading skill,” elements that are usually absent from inventory control theories. Our model features elements of inventory management and asymmetric information simultaneously. Intermediaries are effectively averse to holding inventories (i.e., non-zero positions) since they are not the efficient holders of the security, that is, those who realize the gains to trade. Yet, information asymmetries may prevent them from offloading the security to potential buyers and creating a surplus.

In the next section, we construct a simple adverse selection problem between two asymmetrically informed traders. Then, in Section 3 we show that adding a moderately informed intermediary can reduce the trading inefficiencies that adverse selection triggers between the initial seller and the final buyer. In our model not only does the intermediary make trade more efficient overall, but he also allows the informed trader to extract more value from the uninformed trader who is made worse off by the involvement of a middle man. We extend those arguments in Section 4 by showing that, for more elevated levels of adverse selection or uncertainty, the involvement of multiple intermediaries may help preserve gains to trade that would otherwise be lost with fewer intermediaries. The last section concludes.

2 A Model of Adverse Selection

Consider two agents who would benefit from trading a security: the current owner who values the security at $v$ and a potential buyer who values it at $v + \Delta$. Gains from trade of $\Delta$ are realized when the security ends up in the hands of the buyer; hence, trade is efficient if it takes place with probability 1. The common value $v$ can either be high, $v_h$, or low, $v_l$, with equal probabilities. We use $\sigma$ to denote the spread $(v_h - v_l)$, which will later serve to characterize the adverse selection problem.

Although the role intermediation plays in our model is relatively simple, multi-layered bargaining problems with asymmetric information are usually complex to study given the potential for multiple equilibria arising from the various types of off-equilibrium beliefs. However, there exists a simple structure that will allow us to illustrate our main intuition in a transparent manner. First,
we significantly simplify the equilibrium analysis by having the least informed agent making an
ultimatum offer when trading. This assumption eliminates the need for equilibrium refinement in
solving a signalling game and ensures the uniqueness of our equilibrium. Second, we assume that
the seller has no information about the value of the security when he quotes a price whereas the
buyer receives a perfect signal $s = v$. This information structure is not essential for our results,
but it simplifies the interpretation of the most transparent version of our results.

These two assumptions imply that, without an intermediary, the seller chooses to quote the
buyer one of two prices:

$$p_l \equiv v_l + \Delta,$$

or

$$p_h \equiv v_h + \Delta.$$

The buyer would accept to pay the low price $p_l$ for the security with probability 1. He would,
however, only accept to pay the high price $p_h$ if he were to learn that the security is worth $v_h$.

Hence, if the seller chooses to quote a price $p_l$, the seller’s expected surplus is:

$$p_l - E[v] = \Delta - \sigma/2,$$

while the buyer’s expected surplus is $\sigma/2$. The total surplus created by trade is $\Delta$. If the seller
instead chooses to quote a price $p_h$, the total surplus drops to $\Delta/2$ as trade only takes place with
probability $1/2$. However, the price the seller collects in case of trade is higher, yielding an expected
surplus of:

$$\frac{1}{2}(p_h - v_h) = \Delta/2$$

for the seller and zero for the buyer. Hence, for trade to take place with probability 1, the seller
has to quote the low price $p_l$ rather than the high price $p_h$, which only happens if:

$$\Delta - \sigma/2 \geq \Delta/2$$

$$\Leftrightarrow \sigma \leq \Delta.$$

Efficient trade requires $\sigma$, which quantifies the degree of adverse selection when the seller quotes
the low price, to be small. If adverse selection is too high, trade breaks down with probability 1/2 and half of the surplus from trade is lost.

Derivations in Glode, Green, and Lowery (2012) with an informed proposer show that the boundary for efficient trade would also be \( \sigma \leq \Delta \) if the informed buyer were the proposer and if we imposed a restriction from Grossman and Perry (1986) on off-equilibrium beliefs. For \( \sigma > \Delta \), trade would break down with some probability because the cost of adverse selection, which would then take place through a signalling game the informed proposer is playing, would swamp the gains to trade. The convenience of focusing on the situation in which the informed agent responds to a quoted price rather than making an informed bid for the security originates from the fact that we do not have to solve for such signalling game and impose restrictions on beliefs in order to achieve a unique equilibrium bargaining outcome.

3 Intermediated Trading

In this section we consider the existence of an intermediary, who values the security at \( v \) just like the seller does but who also receives a signal about \( v \) that is accurate with probability \( \mu \in (\frac{1}{2}, 1) \). Since \( \mu > \frac{1}{2} \), the intermediary is better informed than the seller and since \( \mu < 1 \) he is less informed than the buyer. Adding an intermediary in this situation thus does not help realize gains to trade with the seller more quickly, nor does it bring new information to the table. However, as we show below the intermediary allows for higher trade efficiency by splitting the adverse selection problem into two layers of transactions. Ultimately, his involvement affects trade efficiency and each agent’s ability to extract rents. In this scenario, before reaching its ultimate buyer the security has to pass through the intermediary — the uninformed trader first bargains with the intermediary who, if he buys the security, then bargains with the fully informed trader. Each layer of transaction involves a smaller adverse selection problem than with direct trading. Consistent with how we modeled the scenario without an intermediary, we assume that whoever owns the security and is trying to sell it makes a take-it-or-leave-it proposal to his counterparty. Using the derivations below, we will be able to characterize the types of trading network that buyers and sellers would prefer to implement, given the degree of adverse selection between these agents.

Later, we will solve for what happens when the intermediary tries to acquire the security from
the seller, but for now we focus on the situation in which the intermediary already owns the security and bargains with the buyer. Since the intermediary’s information is dominated by the buyer’s, this trade follows the same logic as above in the sense that the intermediary faces an adverse selection problem. As above, the intermediary can ask the buyer either for:

\[ p_l = v_l + \Delta, \]

or

\[ p_h = v_h + \Delta. \]

The buyer would accept to pay the low price \( p_l \) for the security with probability 1. He would, however, only accept to pay the high price \( p_h \) if he were to learn that the security is worth \( v_h \). The expected surplus each agent extracts then depends on the signal the intermediary observes.

Conditional on observing a high signal, the intermediary expects the security to be worth \( \mu v_h + (1 - \mu) v_l \). When quoting the low price \( p_l \), the intermediary’s expected surplus is thus:

\[ p_l - \mu v_h - (1 - \mu) v_l = \Delta - \mu \sigma, \]

and the buyer’s expected surplus is \( \mu \sigma \). Instead, if the seller quotes the high price \( p_h \), trade only takes place once the buyer observes \( v_h \), which occurs with probability \( \mu \). The intermediary then gets an expected surplus of \( \mu \Delta \) and the buyer gets no surplus. Hence, when the intermediary observes a high signal, efficient trade takes place at a price \( p_l \) between the intermediary and the buyer only if:

\[ \Delta - \mu \sigma \geq \mu \Delta \]

\[ \iff \sigma \leq \left( \frac{1 - \mu}{\mu} \right) \Delta. \]

Now, if the intermediary observes a low signal instead, the expected value of the security becomes \( (1 - \mu) v_h + \mu v_l \). By quoting the low price \( p_l \), the intermediary’s expected surplus is:

\[ p_l - (1 - \mu) v_h - \mu v_l = \Delta - (1 - \mu) \sigma, \]
and the buyer’s expected surplus is \((1 - \mu)\sigma\). If the seller quotes the high price \(p_h\), trade only takes place once the buyer observes \(v_h\), which occurs with probability \((1 - \mu)\). The intermediary then gets an expected surplus of \((1 - \mu)\Delta\) and the buyer receives no surplus. Hence, when the intermediary observes a low signal, efficient trade takes place at a quoted price \(p_l\) between the intermediary and the buyer only if:

\[
\Delta - (1 - \mu)\sigma \geq (1 - \mu)\Delta \\
\Leftrightarrow \sigma \leq \left(\frac{\mu}{1 - \mu}\right)\Delta.
\]

Note that it is easier to have efficient trade between an intermediary who observes a low signal and the buyer than between the uninformed and informed traders. It is, however, harder to have efficient trade between an intermediary who observes a high signal and the buyer than between the uninformed and informed traders. This ordering can be restated by the following set of inequalities:

\[
\left(\frac{1 - \mu}{\mu}\right) < 1 < \left(\frac{\mu}{1 - \mu}\right),
\]

when \(\mu \in (\frac{1}{2}, 1)\). It results from the fact that efficient trade requires a discount in the price quoted to the buyer and such discount increases with the expected value of the security, conditional on the information collected by the agent trying to sell the security. A high signal makes it less attractive for the seller to quote a low price \(p_l\), and a low signal makes it more attractive.

Now, we can show what happens when the seller trades with the intermediary, anticipating how trade will subsequently take place between the intermediary and the buyer. We can also compare the efficiency of these outcomes to what would happen without an intermediary. Depending on the degree of information asymmetry about the value of the security three different cases may arise in intermediated trading. In the first two cases, with low and high levels of uncertainty respectively, no intermediary will improve the efficiency of trade over direct trading. In the third case with moderate uncertainty, however, a trading network centered around a specific type of intermediary will allow for more efficient trading.

**Case 1:** \(\sigma \leq \left(\frac{1 - \mu}{\mu}\right)\Delta\)
When $\sigma \leq \left( \frac{1-\mu}{\mu} \right) \Delta$, trade always take place between an intermediary who owns the security and the buyer, which means that regardless of his signal the intermediary quotes $p_l$ to the buyer. Knowing this, the seller then quotes the intermediary $p_l$ and his expected surplus is

$$p_l - E[v] = \Delta - \sigma/2.$$ 

The intermediary gets no surplus from the trade but the buyer extracts $\sigma/2$, just as in the case in which direct trading takes place with probability 1. The total surplus extracted is $\Delta$.

In this case, trading would be efficient if no intermediary was involved as $\sigma \leq \Delta$. The total surplus would then be $\Delta$, with $\Delta - \sigma/2$ going to the uninformed trader and $\sigma/2$ going to the informed trader. The involvement of an intermediary thus leaves the seller’s payoff unchanged but can potentially decrease the efficiency of trade and the buyer’s payoff (if $\mu > \frac{\Delta}{\Delta + \sigma}$). Hence, direct trading is weakly preferred to intermediated trading by both the seller and the buyer.

**Case 2: $\sigma > \left( \frac{\mu}{1-\mu} \right) \Delta$**

The other extreme case occurs when $\sigma > \left( \frac{\mu}{1-\mu} \right) \Delta$. Here, the adverse selection is so severe that even if the intermediary observes a low signal, he quotes a high price $p_h$, and trade breaks down whenever $v = v_l$. The expected value of the security to the intermediary is

$$p_h^\mu \equiv \mu p_h + (1 - \mu) v_l,$$

if he receives a high signal and

$$p_l^\mu \equiv (1 - \mu) p_h + \mu v_l,$$

if he receives a low signal.

If the seller quotes $p_l^\mu$, trade always takes place between the intermediary and the buyer and the seller’s expected surplus from trade is:

$$p_l^\mu - E[v] = (1 - \mu) \Delta - \left( \mu - \frac{1}{2} \right) \sigma.$$
The intermediary then gets an expected surplus of:

\[ \frac{1}{2}p_h^\mu + \frac{1}{2}p_l^\mu - p_l^\mu = \left( \mu - \frac{1}{2} \right) (\Delta + \sigma), \]

and the buyer gets an expected surplus of zero. The total surplus extracted is then \( \Delta/2 \).

The expected surplus for the seller if he quotes \( p_h^\mu \) is:

\[ \frac{1}{2} \left[ p_h^\mu - (\mu v_h + (1 - \mu) v_l) \right] = \frac{\mu}{2} \Delta. \]

The intermediary and the buyer then extract no surplus from the trade. The total surplus extracted is then \( \frac{\mu}{2} \Delta \). The seller consequently asks the intermediary to pay \( p_l^\mu \) rather than \( p_h^\mu \) only if:

\[
(1 - \mu) \Delta - \left( \mu - \frac{1}{2} \right) \sigma \geq \frac{\mu}{2} \Delta \\
\Leftrightarrow \sigma \leq \left( \frac{1 - 3\mu}{\mu - \frac{1}{2}} \right) \Delta.
\]

In such case, we know that trade would be inefficient without an intermediary and the remaining surplus \( \Delta/2 \) would be going to the seller.

**Case 3:** \( \left( \frac{1 - \mu}{\mu} \right) \Delta < \sigma \leq \left( \frac{\mu}{1 - \mu} \right) \Delta \)

In this region an intermediary who owns the security quotes \( p_l \) after observing a low signal and \( p_h \) after observing a high signal. Hence, the security is worth \( p_l \) to the intermediary after receiving a low signal and \( p_h^\mu \) after receiving a high signal.

The expected surplus for the seller when quoting a price \( p_l \) is:

\[ p_l - E[v] = \Delta - \sigma/2. \]

The intermediary’s expected surplus is:

\[ \frac{1}{2}p_l + \frac{1}{2}p_h^\mu - p_l = \frac{\mu}{2} \sigma - \left( \frac{1 - \mu}{2} \right) \Delta, \]
and the buyer's expected surplus is:

$$\frac{1}{2} [\mu v_l + (1 - \mu) v_h + \Delta - p_l] = \left(\frac{1 - \mu}{2}\right) \sigma.$$ 

The total surplus extracted is \(\left(\frac{1+\mu}{2}\right) \Delta\).

The expected surplus for the seller when quoting a price \(p_h^\mu\) is:

$$\frac{1}{2} [p_h^\mu - (\mu v_h + (1 - \mu) v_l)] = \frac{\mu}{2} \Delta,$$

while the intermediary and the buyer extract no surplus from the trade. The total surplus extracted is \(\frac{\mu}{2} \Delta\). The seller will thus prefer to quote \(p_l\) only if:

$$\Delta - \sigma/2 \geq \frac{\mu}{2} \Delta \iff \sigma \leq (2 - \mu) \Delta.$$ 

Overall, when uncertainty does not allow for efficient direct trading, i.e., \(\sigma > \Delta\), intermediated trading may hurt the efficiency of trade, but it may also improve it if the intermediary’s expertise is moderate, that is, if:

$$\frac{\sigma}{\sigma + \Delta} \leq \mu \leq 2 - \sigma/\Delta.$$ 

In such case, the surplus from trade goes from \(\Delta/2\) without an intermediary to \(\left(\frac{1+\mu}{2}\right) \Delta\) with an intermediary. The seller quotes \(p_l\) to the intermediary, which is always accepted, then the intermediary quotes \(p_l\) to the buyer after observing a low signal and \(p_h\) after observing a high signal. The buyer then extracts \(\left(\frac{1-\mu}{2}\right) \sigma\), which is more than the zero surplus he gets without an intermediary. Because trade takes place with probability 1 between the seller and the intermediary, the intermediary is able to extract a surplus \(\frac{\mu}{2} \sigma - \left(\frac{1-\mu}{2}\right) \Delta\), which is strictly positive, hence strictly preferred to not being involved in the trade at all. The seller extracts \(\Delta - \sigma/2\), which makes him worse off than without an intermediary since \(\sigma > \Delta\). In this scenario, the most efficient trading network involves an intermediary with \(\mu = 2 - \sigma/\Delta\) and the resulting probability of realizing the gains to trade reaches \(\frac{3}{2} - \frac{\sigma}{2\Delta}\). However, if instead the intermediary is not moderately informed as defined by the inequalities above, intermediated trade is weakly less efficient than direct trade.
The only agent who might be better off in a trading network that includes an intermediary is the intermediary himself. As in the low uncertainty scenario, direct trading is weakly preferred to intermediated trading by both the seller and the buyer.

To summarize the results so far, neither the initial buyer nor seller will allow an intermediary to be involved when \( \sigma \leq \Delta \). However, when \( \sigma > \Delta \), the informed trader will strictly prefer to set up a trading network with a moderately informed intermediary between him and an uninformed trader, while the uninformed trader will strictly prefer to trade directly with the buyer. Overall, the involvement of a moderately informed intermediary will reduce the trading inefficiencies that adverse selection causes and will increase the total surplus from trade. The proposition below formalizes our main result.

**Proposition 1** When the level of uncertainty in security valuation satisfies \( \sigma \in (\Delta, \sqrt{2} \Delta] \), trading through an intermediary whose expertise level satisfies \( \mu \in \left[ \frac{\sigma}{\sigma + \Delta}, 2 - \sigma / \Delta \right] \) produces a surplus \( \left(1 + \frac{\mu}{2} \right) \Delta \) that is greater than \( \Delta / 2 \), the highest surplus available when trading without an intermediary.

**Proof:** See derivations above.

### 3.1 Parameterized Example

Despite its simplicity, our model of intermediated trade exhibits double-sided adverse selection which gives rise to several cases that depend on parameter values for \( \Delta \), \( \sigma \), and \( \mu \). The small number of key parameters, however, makes our model well-suited for a parameterization.

We first normalize the model by setting the value of gains to trade, \( \Delta \), equal to 1. Trade breaks down with probability \( 1/2 \) whenever \( \sigma > 1 \). The seller then extracts 0.5 while the buyer extracts no surplus. However, a moderately informed intermediary, say with \( \mu = 0.75 \), could help facilitate trade between the uninformed seller (whose \( \mu = 0.5 \)) and the informed buyer (whose \( \mu = 1 \)).

For now, we focus on a parameterization with \( \sigma = 1.2 \). This situation corresponds to Case 3 above, that is, a situation with moderate uncertainty. The seller knows that the intermediary will ask for the low price after receiving a low signal — which the buyer accepts to pay with probability 1 — and will ask for the high price after receiving a high signal — which the buyer accepts to
pay with probability $\mu$. The seller then finds it optimal to quote the intermediary the low price of $p_l$, keeping trade efficient in the first stage and generating a surplus of 0.4 for the seller. The intermediary then extracts 0.325 and the buyer gets 0.15. As explained earlier, the uninformed trader is made strictly worse off by the involvement of the intermediary, even though this trading network improves the surplus generated in equilibrium by 0.375 and makes the intermediary and the informed trader strictly better off.

In Figure 1, we keep the parameterization as in the example above except that we allow for various levels in the uncertainty parameter $\sigma$. We compare how the surplus is allocated across agents in equilibrium for the direct trading network and for the intermediated trading network. The upper left graph shows that direct trading socially dominates intermediated trading for very high or very low levels of $\sigma$, but intermediated trading dominates for intermediate levels of uncertainty, in particular when $1 < \sigma \leq 1.25$. This region is where our numerical example above is located. This is also the only region where one of the end-traders is strictly better off with an intermediary. Outside this region, the intermediary is either trying to extract too much surplus away from end-traders or he is imposing too much adverse selection on them. Trade efficiency is thus more fragile with high levels of uncertainty, or adverse selection, when an intermediary is involved — the trading network that is socially optimal with moderate uncertainty would magnify trading inefficiencies if uncertainty were to suddenly increase.

Figure 2 compares how the surplus is allocated across agents, but this time, we fix $\sigma = 1.2$ and let $\mu$ go from 0.5 to 1. The upper left graph highlights that the involvement of a moderately informed intermediary can improve trade efficiency, relative to inefficient direct trading (when $\sigma > \Delta$). It also shows that a highly informed intermediary can worsen trade efficiency, because his involvement now triggers adverse selection problems in both stages of the intermediated trade. A weakly informed intermediary, on the other hand, is unable to eliminate the adverse selection problem caused by an informed trader, hence the total surplus remains unchanged with or without his involvement. When the intermediary’s $\mu$ is between 0.545 and 0.8, the seller is effectively paying adverse selection premia to both the buyer and the intermediary, whereas he would extract the full surplus available (i.e., 0.5) in a direct trade. The intermediary’s surplus is increasing in his own expertise and reaches a maximum of 0.38 when $\mu = 0.8$, which is also where trade is the most socially efficient possible given that $\sigma > \Delta$. At $\mu = 0.8$, the total surplus is 0.9. The buyer’s
surplus is, however, decreasing in the intermediary’s expertise and reaches a maximum of 0.273 when $\mu = 0.545$.

4 Intermediation Chains

In this section, we show that trading through multiple intermediaries can sometimes preserve gains to trade that would otherwise be lost with fewer intermediaries. A chain of two intermediaries is shown to eliminate part of the breakdown in trade that occurs due to adverse selection between a buyer and a seller in situations where a single intermediary would not help. We also show through a numerical example that the same logic extends such that adding a third intermediary improves
Figure 2. Surplus Extraction and Intermediary’s Expertise. The plots show the impact of an intermediary’s expertise on surplus extraction, for gains to trade of $\Delta = 1$ and uncertainty in security value of $\sigma = 1.2$.

the efficiency of trade when uncertainty is high. However, liquidity is affected more severely by unexpected jumps in uncertainty.

Suppose that $\sigma > \Delta$, meaning that direct trading yields a trading probability $1/2$. We already know that the involvement of one intermediary of type $\mu$ can increase the surplus from trade from $\Delta/2$ to $(1 + \mu)\Delta/2$ if and only if:

$$\frac{\sigma}{\sigma + \Delta} \leq \mu \leq 2 - \frac{\sigma}{\Delta}.$$
However, whenever $\sigma > \sqrt{2\Delta}$, this region for $\mu$ does not exist. Involving a single intermediary would not solve the adverse selection problem; in fact, it might worsen it.

Now, consider instead a situation in which trade between the uninformed seller and the informed buyer first go through an intermediary of type $\mu'$ and then through an intermediary of type $\mu \geq \mu'$. In order to have a unique equilibrium that does not require restrictions on off-equilibrium beliefs or other equilibrium refinements, we assume that the least informed of the two intermediaries, the $\mu'$ trader, receives a signal that is a noisier version of the $\mu$-trader’s signal. Specifically, the signal $s_{\mu'}$ is drawn from the following distribution: it is equal to the signal $s_{\mu}$ with probability $\rho \in (\frac{1}{2}, 1)$ and different with probability $(1 - \rho)$. The precision of such signal thus satisfies:

$$\mu' = \rho \mu + (1 - \rho)(1 - \mu).$$

Since the holder of the security is, as earlier, making an ultimatum price quote to the potential buyer he faces, the informational structure above ensures that the proposer’s information set is always dominated by the responder’s and that bargaining takes place without signalling. This implication allows us to study, in a tractable and intuitive way, a sequential bargaining game with three transactions and with adverse selection among the four agents.

The following proposition summarizes our main result for this section. We discuss its implications below but relegate the full proof of the proposition to the Appendix since the logic is similar to that for the scenario with only one intermediary.

**Proposition 2** When the level of uncertainty in security valuation satisfies $\sigma > \sqrt{2\Delta}$, trading through two intermediaries may produce a surplus $(\frac{\mu + \mu'}{2}) \Delta$ that is greater than $\Delta/2$, the highest surplus available when trading through one or zero intermediary.

**Proof:** See Appendix.

The intuition behind this result is similar to what we have seen in Section 3. Adding layers of transactions among which the difference in information quality between counterparties is small can help reduce trading inefficiencies due to adverse selection. If the fundamental adverse selection problem between the ultimate buyer and seller is large, more layers of transactions and more asymmetrically informed intermediaries are necessary to improve the efficiency of trade. As in
Section 3, the full analysis of this results requires to sequentially solve for optimal trading behavior in each transaction.

First, we need to look at the final transaction between the better-informed intermediary and the fully informed buyer, which is identical to the final transaction in the setup with only one intermediary. From Section 3 we know that in regions where direct trade breaks down with probability \( 1/2 \), trade should also break down between the intermediary and the buyer whenever the intermediary observes a high signal. Inequality (1) also tells us that if \( \mu < \frac{\sigma}{\sigma + \Delta} \), trade also breaks down whenever the intermediary observes a low signal, yielding a maximal probability of realizing the gains to trade of \( 1/2 \). Hence, the only way the involvement of two intermediaries can potentially help preserve more surplus than through direct trading is if the expertise of the better informed intermediary is high enough to satisfy: \( \mu \geq \frac{\sigma}{\sigma + \Delta} \). In such case, the intermediary quotes \( p_l \) to the buyer after observing a low signal and \( p_h \) after observing a high signal.

The trading behavior between the two intermediaries is thus similar to Case 3 in Section 3. The main difference is that the proposer is now receiving a signal that is accurate with probability \( \mu' \) instead of being uninformed. The proposer is, nonetheless, still choosing between quoting a low price \( p_l \) and a higher price \( p_h^{\mu} = \mu p_h + (1 - \mu) v_l \). If the intermediary of type \( \mu' \) always quotes \( p_h^{\mu} \), the surplus from trade is at most \( \Delta/2 \). We show in the proof of the proposition that the most efficient trading price \( p_l \) will be quoted after a low signal whenever \( \sigma \leq \left( \frac{1-(1-\rho)\mu}{1-\mu}\right) \Delta \).

In the case in which the intermediary of type \( \mu' \) quotes \( p_l \) after observing a low signal, but quotes \( p_h^{\mu} \) after observing a high signal, the seller will prefer to quote the low price \( p_l \) whenever \( \sigma \leq (2 - \mu\rho) \Delta \). Since \( \rho < 1 \), this condition on \( \sigma \) is less restrictive than the condition that applied when only one intermediary was involved. It is therefore possible to simultaneously satisfy such condition and the condition that \( \mu \geq \frac{\sigma}{\sigma + \Delta} \) while simultaneously satisfying analogous conditions for the situation with only one intermediary is impossible. When the current conditions are satisfied, trade takes place with probability 1 between the seller and the first intermediary, who then trades the security to a second intermediary with probability 1 after observing a low signal and probability \( \rho \) after observing a high signal. Then, the security is traded to the buyer with probability 1 if the \( \mu \) intermediary holds the security and observes a low signal and with probability \( \mu \) if he holds the security and observes a high signal. As in Section 3 because the uninformed seller is selling the
security at a price \( p_l \) in equilibrium, the surplus he extracts is \( \Delta - \sigma/2 \) and is dominated by the surplus he would extract if there were no intermediary at all.

Overall, when uncertainty is too high for one intermediary to improve trade efficiency, trading through two intermediaries may still produce a surplus from trade of \( \left(\frac{\mu + \rho}{2}\right) \Delta \) that is greater than the surplus created by trading through only one intermediary or through direct trading. However, the improvement in trading efficiency when two intermediaries are involved is only possible for moderate levels of uncertainty, or adverse selection. For example, if \( \sigma > 1.75\Delta \), the adverse selection problem between the ultimate seller and buyer of the security is so severe that no pair of intermediaries can succeed in improving trading efficiency; the necessary condition \( \sigma \leq (2 - \mu \rho) \Delta \) cannot be satisfied. Hence, in that scenario reallocating the adverse selection problem across more layers of transactions might be needed.

### 4.1 Parameterized Example

The following numerical example highlights how adding intermediaries might improve trade efficiency as the adverse selection problem worsens. First, we go back to our earlier parameterization with \( \Delta = 1 \). As we argued earlier, without an intermediary the total surplus from trade would drop to 0.5 whenever \( \sigma > 1 \). Adding a moderately informed intermediary \( \mu = 0.75 \) can mitigate the adverse selection problem and increase the surplus to 0.875 over the region where \( 1 < \sigma \leq 1.25 \).

Now, consistent with the analysis earlier in this section, adding a second intermediary will improve the efficiency of trade in some circumstances. If the adverse selection problem is so severe that \( \sigma > 1.25 \), the surplus with one intermediary of type \( \mu = 0.75 \) is 0.375, which is lower than the surplus from direct trading. Trading through a second intermediary with \( \rho = 0.75 \), implying \( \mu' = 0.625 \), will instead produce a surplus from trade of 0.75 in the region where \( 1.25 < \sigma \leq 1.437 \). Numerically, we can show that a similar improvement in trade efficiency occurs over the region where \( 1.437 < \sigma \leq 1.577 \) if a third intermediary with \( \mu'' = 0.5625 \) is added to the trading network.\(^1\)

Figure 3 replicates Figure 1 by allowing for various levels of \( \sigma \) and comparing surplus allocation, except that it also includes trading networks with two and three intermediaries. The upper left graph shows that adding intermediaries to a trading network allows to sustain higher levels of trade.

\[^1\]The third intermediary’s expertise level, \( \mu'' = 0.5625 \), is computed by assuming that he receives a noisier version of the second intermediary’s signal. Consistent with the case with only two intermediaries, the precision of that weaker signal satisfies: \( \mu'' = \rho' \mu' + (1 - \rho')(1 - \mu') \), which is equal to 0.5625 if we set \( \rho' = \rho = 0.75 \).
efficiency as uncertainty increases. It, however, also shows that direct trading socially dominates trading with one, two, or three intermediaries for very low or very high levels of uncertainty. While trading through different numbers of intermediaries helps preserve more surplus over an intermediate region, it also destroys more surplus than direct trading does if the degree of uncertainty $\sigma$ happens to be outside of that region. In that sense, long intermediation chains that are optimal in moderately uncertain periods would hurt trade efficiency if uncertainty, or adverse selection, were to suddenly jump to a higher level and if adjusting the trading network on the spot was impossible.

**Figure 3. Surplus Extraction and Uncertainty with Multiple Intermediaries.** The plots show the impact of uncertainty on surplus extraction, for gains to trade of $\Delta = 1$ and different trading networks: (i) direct trading, (ii) trading through one intermediary of type $\mu = 0.75$, (iii) trading through two intermediaries of types $\mu' = 0.625$ and $\mu = 0.75$, and (iv) trading through three intermediaries of types $\mu'' = 0.5625$, $\mu' = 0.625$, and $\mu = 0.75$. 


5 Conclusion

This paper shows that a moderately informed intermediary can help solve the adverse selection problem that inhibits efficient trading between an informed agent and an uninformed one. Moreover, as the adverse selection problem becomes more important, a chain of several intermediaries, all with different levels of information quality, may be needed to enhance trade efficiency. Complex trading networks with multiple intermediaries may be socially optimal because reallocating a large adverse selection problem over several layers of transactions reduces the downside each agent faces of being wronged by a better informed agent. However, long intermediation chains that are optimal for a given interval of adverse selection problems make liquidity more fragile when the degree of adverse selection lies outside of that interval.
Appendix

Proof of Proposition 2: As in Section 3 when only intermediary was involved between the buyer and the seller, the analysis of the current scenario with two intermediaries requires to sequentially solve for optimal trading behavior in each stage of transaction. We start with the final trade between the intermediary of type $\mu$ and the fully informed buyer.

$\mu$-Intermediary Trading with Buyer

The final trade between the better-informed intermediary and the fully informed buyer looks exactly like the final trade in the earlier setup with only one intermediary. From Section 3 we know that in regions where direct trade breaks down with probability $1/2$, trade should also break down between the intermediary and the buyer whenever the intermediary observes a high signal. Inequality (1) also tells us that if $\mu < \frac{\sigma}{\sigma + \Delta}$, trade also breaks down whenever the intermediary observes a low signal, yielding a maximal probability of realizing the gains to trade of $1/2$.

Hence, we need $\mu \geq \frac{\sigma}{\sigma + \Delta}$ for two intermediaries to potentially help preserve more surplus than through direct trading. If that condition is satisfied, an intermediary holding the security quotes $p_l$ after observing a low signal and $p_h$ after observing a high signal.

$\mu'$-Intermediary Trading with $\mu$-Intermediary

Here, the proposer is choosing between quoting a low price $p_l$ or a higher price $p_h' = \mu p_h + (1 - \mu) v_l$. Conditional on observing a high signal, his expected surplus is:

$$p_l - E[v|s_{\mu'}] = \Delta - \mu' \sigma,$$

from quoting the low price $p_l$ and

$$Pr(s_{\mu} = v_h|s_{\mu'}) [p_h' - E[v|s_{\mu} = v_h]] = \rho \mu \Delta,$$

from quoting the higher price $p_h'$. Thus, trade is most efficient here when $\sigma \leq \left(\frac{1 - \rho \mu}{\mu'}\right) \Delta$.

Conditional on observing a low signal, the proposer’s expected surplus is:

$$p_l - E[v|s_{\mu'}] = \Delta - (1 - \mu') \sigma,$$
from quoting the low price $p_l$ and

$$Pr(s_\mu = v_h|s_{\mu'}) \left[ p_{h}^\mu - E[v|s_\mu = v_h] \right] = (1 - \rho)\mu\Delta,$$

from quoting the higher price $p_{h}^\mu$. Thus, trade is most efficient here when $\sigma \leq \left( \frac{1 - (1 - \rho)\mu}{(1 - \mu')} \right) \Delta$.

Once again, the most efficient trading price $p_l$ is more likely to be chosen by the proposer after observing a low signal than a high signal; the constraint $\sigma \leq \left( \frac{1 - \rho\mu}{\mu'} \right) \Delta$ being more restrictive than $\sigma \leq \left( \frac{1 - (1 - \rho)\mu}{(1 - \mu')} \right) \Delta$. If this more restrictive condition is satisfied, trade takes place with probability 1 between the two intermediaries, implying that the maximal surplus from trade is $\left( \frac{1 + \mu}{2} \right) \Delta$ overall, just as in the scenario with a single intermediary. Using the earlier assumption that $\mu \geq \frac{\sigma}{\sigma + \Delta}$, this constraint implies that:

$$\frac{\sigma}{\Delta} \leq \frac{1 - \rho\mu}{\rho\mu + (1 - \rho)(1 - \mu)} \leq \frac{1 - \frac{\rho\sigma}{\sigma + \Delta}}{\frac{\rho\sigma}{\sigma + \Delta} + \frac{(1 - \rho)\Delta}{\sigma + \Delta}} = \frac{1 + (1 - \rho)\frac{\sigma}{\Delta}}{(1 - \rho) + \rho\frac{\sigma}{\Delta}}.$$

And this last inequality can be rewritten as:

$$\frac{\sigma}{\Delta} \leq \sqrt{\frac{1}{\rho}} < \sqrt{2},$$

which means that one intermediary would also preserve a surplus of $\frac{1 + \mu}{2} \Delta$. Moreover, the constraint $\sigma \leq \left( \frac{1 - \rho\mu}{\mu'} \right) \Delta$ can be rewritten as:

$$\mu \leq \frac{1 - \mu'\frac{\sigma}{\Delta}}{\rho} < 2 - \sigma/\Delta,$$

where the second inequality follows from $\rho \in (\frac{1}{2}, 1)$ and $\mu' \in (\frac{1}{2}, 1)$. Thus, adding an extra layer of transaction through the involvement of a second moderately informed intermediary does not improve the efficiency of trade over what we would get with a single intermediary, but it can improve it over what we would get through direct trade. Since we are now looking for the possibility that
two intermediaries will improve trade efficiency above and beyond what one intermediary can do, we can rule out the case in which the intermediary of type $\mu'$ quotes $p_l$ regardless of his signal.

**Seller Trading with $\mu'$-Intermediary**

The final step in this analysis is to solve for what happens when the seller trades with the first intermediary, anticipating how trade will subsequently take place among other agents.

We already know that if the intermediary of type $\mu'$ always quotes $p_{lh}$, the surplus from trade is at most $\Delta/2$. We also know that if the intermediary of type $\mu'$ finds it optimal to always quote $p_l$, the most efficient scenario will be such that trade is effectively taking place as if there were only one intermediary. Hence, the only case in which trade efficiency could possibly be improved by the involvement of two intermediaries rather than a single one of them has to satisfy:

\[
\left(1 - \frac{\rho}{\mu'}\right) \Delta < \sigma \leq \left(1 - \frac{(1 - \rho)\mu}{(1 - \mu')}\right) \Delta.
\]

In such case, the intermediary of type $\mu'$ quotes $p_l$ after observing a low signal, but quotes $p_{lh}$ after observing a high signal. The seller is then picking between quoting $p_l$ and collecting an expected surplus of:

\[
p_l - E[v] = \Delta - \sigma/2,
\]

or quoting $p_{lh}^\mu \equiv \rho p_{lh}^\mu + (1 - \rho)((1 - \mu)v_h + \mu v_l)$ and collecting an expected surplus of:

\[
\frac{1}{2}[p_{lh}^\mu - E[v|s_{\mu'} = v_h]] = \frac{1}{2}[\rho p_{lh}^\mu + (1 - \rho)((1 - \mu)v_h + \mu v_l) - \mu' v_h - (1 - \mu')v_l]
\]

\[
= \frac{\rho \Delta}{2}.
\]

The seller will prefer to quote the low price $p_l$ only if:

\[
\Delta - \sigma/2 \geq \frac{\mu \rho}{2} \Delta
\]

\[
\Leftrightarrow \sigma \leq (2 - \mu \rho) \Delta.
\]

Note that since $\rho < 1$ this condition on $\sigma$ is less restrictive than the condition that applies when only one intermediary is involved. It is therefore possible to simultaneously satisfy that condition and the condition that $\mu \geq \frac{\sigma}{\sigma + \Delta}$ even though it was not possible to satisfy the analog of these conditions for the case with only one intermediary. If the current conditions are satisfied,
the security is traded with probability 1 from the seller to the least informed intermediary, who
then trades the security to the most informed intermediary with probability 1 after observing a low
signal and probability $\rho$ after observing a high signal. Then, the security is traded to the buyer with
probability 1 if the $\mu$ intermediary holds the security and observes a low signal and with probability
$\mu$ if he holds the security and observes a high signal. Overall, the surplus from trade is:

$$\frac{1}{2} [\rho + (1 - \rho)\mu] \Delta + \frac{1}{2} \rho \mu \Delta = \left( \frac{\rho + \mu}{2} \right) \Delta,$$

which is greater than the surplus that would be created without a second intermediary. ■
References


Exploratory Trading

Author 90901221*

November 14, 2012

Abstract

Using comprehensive, account-labeled message records from the E-mini S&P 500 futures market, I investigate the mechanisms underlying high-frequency traders' capacity to profitably anticipate price movements. Of the 30 high-frequency traders (HFTs) that I identify in my sample, eight earn positive overall profits on their aggressive orders. I find that all eight of these HFTs consistently lose money on their smallest aggressive orders, and these losses are not explained by inventory management. These losses on small orders, as well as the more-than-offsetting gains on larger orders, could be rationalized if the small orders provided some informational value, and I model how a trader could gather valuable private information by using her own orders in an exploratory manner to learn about market conditions. This “exploratory trading” model predicts that the market response to the trader’s “exploratory” order would help to explain her earnings on her next order, but would not explain any other traders’ subsequent performance. In direct confirmation of the model’s predictions, I find that a simple measure of changes in the orderbook immediately following small aggressive orders placed by the eight HFTs explains a significant additional component of those HFTs’ earnings on subsequent, larger orders, but this information offers little or no additional power to explain other traders’ earnings on subsequent orders. These findings help to clarify nature of the information on which HFTs trade and offer a starting point to address the open questions about social welfare implications of high-frequency trading.

*The views expressed in this paper are my own and do not constitute an official position of the Commodity Futures Trading Commission, its Commissioners, or staff.
1 Introduction

Over the past three decades, information technology has reshaped major financial exchanges worldwide. Physical trading venues have increasingly given way to electronic ones, and trading responsibilities that once fell on human agents have increasingly been delegated to computer algorithms. Automation now pervades financial markets; for example, Hendershott and Riordan (2009) and Hendershott et al. (2011) respectively document the dramatic levels of algorithmic trading on the Deutsche Boerse and the New York Stock Exchange. Much of the algorithmic activity in major markets emanates from so-called “high-frequency traders” (“HFTs”). Although it dominates modern financial exchanges, HFTs’ activity remains largely mysterious and opaque—it is the “dark matter” of the trading universe.

HFTs are distinguished not only by the large number of trades they generate (i.e., their literal high trading frequency), but also by the speed with which they can react to market events. HFTs achieve these remarkable reaction times, typically measured in milliseconds, by using co-location services, individual data feeds, and high-speed computer algorithms. Two further hallmarks of HFTs are their extremely short time-frames for maintaining positions, and their propensity for “ending the trading day in as close to a flat position as possible (that is, not carrying significant, unhedged positions over-night [when markets are closed]).”¹

Empirical study of high-frequency trading has proven challenging, but not impossible. For example, Brogaard et al. (2012) obtain and analyze a NASDAQ dataset that flags messages from an aggregated group of 26 HFT firms, and Hasbrouck and Saar (2011) conduct a complementary analysis by statistically reconstructing “strategic runs” of linked messages in NASDAQ order data. Both of these analyses suggest beneficial effects from HFTs’ activity, but inherent limitations of the underlying data restrict these studies’ scope to explain how and why such effects arise.

Understanding and explaining the impacts of high-frequency trading requires some understanding of what HFTs are actually doing, and of how their strategies work. Even in a market for a single asset, HFTs exhibit considerable heterogeneity, so aggregate HFT activity reveals little about what individual HFTs really do. Data suitable for the study of individual HFTs’ activity are difficult to obtain. Whereas publicly available 13-F forms reveal the behavior of institutional investors at a quarterly frequency, there is no comparable public data that can be used to track and analyze the behavior of individual traders at a second- or millisecond-frequency.

The only fully adequate data currently available for academic research on high-frequency trading

come from regulatory records that the Chicago Mercantile Exchange provides to the U.S. Commodity Futures Trading Commission. Kirilenko et al. (2010) pioneered the use of transaction data from these records to investigate high-frequency trading in their analysis of the so-called “Flash Crash” of May 6, 2010 in the market for E-mini Standard & Poors 500 stock index futures contracts (henceforth, “E-mini”). This work introduced a scheme to classify trading accounts using simple measures of overall trading activity, intraday variation in net inventory position, and inter-day changes in net inventory position. Of the accounts with sufficiently small intra- and inter-day variation in net position, Kirilenko et al. classify those with the highest levels of trading activity as HFTs, and these accounts are archetypes of high-frequency traders.

Kirilenko et al. find that HFTs participate in over one-third of the trading volume in the E-mini market, and subsequent research by Baron et al. (2012) documents the large and stable profits that HFTs in the E-mini market earn. This work provides empirical confirmation of HFTs’ importance, and it offers some crisp descriptions of HFTs’ activity. However, it does not attempt to explain why HFTs act as they do, or how HFTs earn profits. Indeed, no extant empirical research attempts such explanations. In this paper, I address a central aspect of this open problem. HFTs in the E-mini market earn roughly 40% of their profits from the transactions that they initiate—that is, from their so-called “aggressive” orders—and I examine the mechanism underlying HFTs’ capacity to earn these profits.

How do HFTs in the E-mini market make money from their aggressive orders? One possibility is that HFTs merely react to public information faster than everyone else; this premise underlies the models of Biais et al. (2010), Jarrow and Protter (2011), and Cespa and Foucault (2008). A second possibility is that HFTs simply front-run coming demand when they can predict future aggressive orders. However, I find neither of these hypotheses to be consistent with the data.

I identify the HFTs who profit from their aggressive orders, then I investigate how these HFTs manage to do so. I show that the HFTs who profit from their aggressive trading use small aggressive orders to obtain private information that helps to forecast the price-impact of predictable demand innovations. Demand innovations in the E-mini market are easy to predict, but the price-elasticity of supply is not, and price-impact is usually too small for indiscriminate front-running of predictable demand to be profitable. However, the private information about price-impact generated by an

---

2To be more precise, it is extremely easy to predict whether future aggressive orders will be buy orders or sell orders. The dynamic behavior of passive orders resting in the orderbook—analogueous to supply elasticity—is considerably harder to forecast.
HFT’s small aggressive orders enables that HFT to trade ahead of predictable demand at only those times when it is profitable to do so (i.e., when price-impact is large). To elucidate how this works, I develop a theoretical model of what I term “exploratory trading.”

Fundamentally, the model of exploratory trading rests on the notion that an HFT’s aggressive orders generate valuable private information, specifically, information about the price-impact of the aggressive orders that follow. When an HFT places an exploratory order and observes a large price-impact, he learns that supply is temporarily inelastic. If the HFT knows that there is going to be more demand soon thereafter, he can place a larger order (even with a big price-impact) knowing that the price-impact from the coming demand will drive prices up further and ultimately enable him to sell at a premium that exceeds the price-impact of his unwinding order. When an HFT knows that supply is temporarily inelastic, he follows a routine demand-anticipation strategy. The purpose of exploratory trading is not to learn about future demand, but rather to identify the times at which trading in front of future demand will be profitable. Active learning in financial markets is a relatively old idea, dating back at least to Leach and Madhavan’s papers in 1992 and 1993, but exploratory trading is an active-learning mechanism that is new to the academic literature. In section 2, I present a model to formalize the concept of exploratory trading, and I derive the model’s central testable predictions.

Using novel electronic message data at the Commodity Futures Trading Commission, I examine the profitability of individual HFTs’ aggressive orders. I find that eight of the 30 HFTs in my sample profit from their aggressive trading overall and significantly outperform non-HFTs. However, these same eight HFTs all lose money on their smallest aggressive orders. (For brevity, I refer to these eight HFTs as “A-HFTs,” and to the remaining 22 as “B-HFTs.”) Exploratory trading would produce just such a pattern of incurring small losses on exploratory orders then realizing large gains; these descriptive results both motivate further tests and suggest the A-HFTs’ small aggressive orders as natural candidates for potential exploratory orders.

To explicitly test the predictions of the exploratory trading model for the eight A-HFTs, I examine the extent to which information about the changes in the orderbook following small aggressive orders explains the profits that various traders earn on subsequent aggressive orders. The exploratory trading model predicts that information about the changes following an A-HFT’s small aggressive order will explain a significant additional component of the A-HFT’s subsequent performance, but that this information will not explain any additional component of other traders’ subsequent performance. Consistent with these predictions, I find that the orderbook changes immediately following A-HFTs’
small aggressive orders provide significant additional explanatory power for the respective A-HFTs’ performance on their larger aggressive orders, but not for other traders’ performance.

The remainder of this paper is organized as follows: Section 2 presents a simple model of exploratory trading, along with the model’s central predictions, and establishes the empirical agenda. Section 3 describes the data and precisely defines HFTs. Section 4 addresses the overall profitability of HFTs’ aggressive orders and precisely characterizes the A-HFTs, then examines the A-HFTs’ losses on small aggressive orders. Section 5 presents direct empirical tests of the exploratory trading model’s key predictions, section 6 examines the practical significance of exploratory information, and section 7 discusses extensions and implications of the empirical results. Section 8 concludes.

2 Exploratory Trading: Theory

The ultimate objective of this paper is to explain the mechanism underlying HFTs’ capacity to profit from their aggressive orders in the E-mini market, and this section establishes the theoretical framework for my empirical investigation.

As noted in the introduction, demand innovations in the E-mini market are easy to predict from public market data, but the price-elasticity of supply is not. Although there are times when supply is unaccommodating and high future demand forecasts price changes that are large enough to profit from, such times are difficult or impossible to identify by merely observing public market data. In this type of setting, a trader can obtain additional information about supply conditions by placing an “exploratory” aggressive order and observing how prices and supply respond. The additional exploratory information enables the trader to determine whether supply is accommodating (and expected price-impact small) or unaccommodating (and expected price-impact large), and this helps the trader to decide whether he can profit by trading ahead of an imminent demand innovation.

The basic model I examine is a simple representation of a market in which demand is easy to predict, but supply elasticity is not. I consider a two-period model with two possible states for supply conditions (accommodating or unaccommodating), and three possible demand innovations in the second period (positive, negative, or zero). The demand innovation is automatically revealed before it arrives in the second period, but the state of supply conditions is only revealed if a trader places an aggressive order in the first period.

In this context, I consider the problem facing a single trader, the “HFT.” In the first period, the
HFT has the opportunity to place an aggressive order and thereby learn about supply conditions. In the second period, regardless of what happened in the first period, the HFT observes a signal about future demand, after which he again has an opportunity to place an aggressive order. The signal of future demand forecasts price innovations much more accurately when combined with information about supply conditions than it does when used on its own. If the HFT places an aggressive order in the first period, he effectively “buys” supply information that he can use in the second period to better decide whether he should place another aggressive order. Consequently, the HFT may find it optimal in the first period to place an order that he expects to be unprofitable, since the information that the order generates will be valuable in the second period.

The rest of this section is devoted to formally developing a model of exploratory trading and deriving the model’s testable predictions. In addition to the basic result about the value of exploratory information sketched above, I address the key issue of why an order generates more information for the trader who submitted it than it does for everyone else. Appendix A contains full mathematical details.

2.1 Baseline Model

In an order-driven market, every regular transaction is initiated by one of the two executing transactors. The transactor who initiates is referred to as the “aggressor,” while the opposite transactor is referred to as the “passor.” The passor’s order was resting in the orderbook, and the aggressor entered a new order that executed against the passor’s preexisting resting order. Assuming that prices are discrete, the lowest price of any resting sell order in the orderbook (“best ask”) always exceeds the highest price of any resting buy order in the book (“best bid”) by at least one increment (the minimal price increments are called “ticks”). A transaction initiated by the seller executes at the best bid, while a transaction initiated by the buyer executes at the best ask; the resulting variation in transaction prices between aggressive buys and aggressive sells is known as “bid-ask bounce.” Hereafter, except where otherwise noted, I will restrict attention to price changes distinct from bid-ask bounce. Empirically, the best ask for the most actively traded E-mini contract almost always exceeds the best bid by exactly one tick during regular trading hours, so movements of the best bid, best ask, and mid-point prices are essentially interchangeable.

If the best bid and best ask were held fixed, a trader who aggressively entered then aggressively exited a position would lose the bid-ask spread on each contract, whereas a trader who passively entered
then passively exited a position would earn the bid-ask spread on each contract. Intuitively, aggressors pay for the privilege of trading precisely when they wish to do so, and passors are compensated for the costs of supplying this “immediacy,” cf. Grossman and Miller (1988). These costs include fixed operational costs and costs arising from adverse selection. Cf. Glosten and Milgrom (1985), Stoll (1989).

An aggressive order will execute against all passive orders at the best available price level before executing against any passive orders at the next price, so an aggressive order will only have a literal price-impact if it eats through all of the resting orders at the best price. In the E-mini market, it is rare for an aggressive order to have a literal price-impact, not only because there are enormous numbers of contracts at the best bid and best ask, but also because aggressive orders overwhelmingly take the form of limit orders priced at the opposite best (which cannot execute at the next price level).

2.1.1 Market Structure

Let time be discrete, consisting of two periods, $t = 1, 2$. This model should be interpreted as a single instance of the hundreds or thousands of similar scenarios that arise throughout the trading day.

Consider an order-driven market with discrete prices, and assume that both the orderbook and order-flow are observable. Conceptually, the flow of aggressive orders is analogous to demand, while the set of passive orders in the orderbook (“resting depth”) is analogous to supply.

2.1.2 Passive Orders

There are two possible states for the behavior of passive orders: accommodating and unaccommodating. Let the variable $\Lambda$ represent this state, which I call the “liquidity state.” The liquidity state is the same in both periods of the model. Denote the accommodating liquidity state by $\Lambda = A$, and the unaccommodating state by $\Lambda = U$. Assume that $\Lambda = U$ with ex-ante probability $u$, and $\Lambda = A$ with complementary ex-ante probability $1 - u$.

The liquidity state characterizes the behavior of resting depth in the orderbook after an aggressive order executes—a generalization of price-impact appropriate for an order-driven market. When an aggressive buy (sell) order executes, it mechanically depletes resting depth on the sell (buy) side of the orderbook. Following this mechanical depletion, traders may enter, modify, and/or cancel passive orders, so resting depth at the best ask (bid) can either replenish, stay the same, or deplete further. The aggressive order’s impact is offset to some extent—or even reversed—if resting depth
replenishes, whereas the aggressive order’s impact is amplified if resting depth depletes further. In the accommodating state ($\Lambda = A$) resting depth weakly replenishes, while in the unaccommodating state ($\Lambda = U$) resting depth further depletes. Intuitively, aggressive orders have a small price-impact in the accommodating state, and a large price-impact in the unaccommodating state.

Although the orderbook is always observable, static features of passive orders in the orderbook do not directly reveal the liquidity state. Because the liquidity state relates to the dynamic behavior of resting depth after an aggressive order executes, this state can only be observed through the changes in the orderbook that follow the execution of an aggressive order.

### 2.1.3 Aggressive Order-Flow

At the end of period 2, traders other than the HFT exogenously place aggressive orders. Let the variable $\varphi \in \{-1, 0, +1\}$ describe this exogenous aggressive order-flow. The variable $\varphi$ is just a coarse summary of the order-flow—It does not represent the actual number of contracts. Intuitively, $\varphi = -1$ represents predictable selling pressure and $\varphi = +1$ represents predictable buying pressure, while $\varphi = 0$ represents an absence of predictable pressure in either direction.

Assume that $\varphi = +1$ and $\varphi = -1$ with equal probabilities $P\{\varphi = +1\} = P\{\varphi = -1\} = v/2$, and $\varphi = 0$ with complementary probability $1 - v$. The value of $\varphi$ does not depend on the liquidity state, $\Lambda$, nor does it depend on the HFT’s actions.

The price-change at the end of period 2, which I denote by $y$, is jointly determined by the exogenous aggressive order-flow and the liquidity state. In the notation of the model,

$$
 y = \begin{cases} 
 \varphi & \text{if } \Lambda = U \\
 0 & \text{if } \Lambda = A 
\end{cases}
$$

(1)

In other words, if the liquidity state is unaccommodating ($\Lambda = U$), aggressive order-flow affects the price, and $y = \varphi$. However, if the liquidity state is accommodating ($\Lambda = A$), aggressive order-flow does not affect the price, and $y = 0$ regardless of the value of $\varphi$.

### 2.1.4 The HFT

The HFT submits only aggressive orders, and these aggressive orders are limited in size to $N$ contracts or fewer. Let $q_t \in \{-N, \ldots, -1, 0, 1, \ldots, N\}$ denote the signed quantity of the aggressive order that
the HFT places in period \( t \), where a negative quantity represents a sale, and a positive quantity represents a purchase.

Assume that the HFT pays constant trading costs of \( \alpha \in (0.5, 1) \) per contract. The lower bound of 0.5 on \( \alpha \) corresponds to half of the minimum possible bid-ask spread, while the upper bound of 1 merely excludes trivial cases of the model in which aggressive orders are always unprofitable. When \( \alpha > u \), the HFT will never place an order in period 2 if he doesn’t know the liquidity state, and I focus on this case to simplify the exposition; results are qualitatively unchanged for \( u \geq \alpha \) (see Appendix A).

I assume that the HFT’s aggressive orders have no literal price-impact. Intuitively, the HFT only trades contracts at the initial best bid/ask. For example, in period 2, if the HFT has learned that the liquidity state is unaccommodating and \( \varphi = +1 \), he will buy all of the contracts available at the best ask. This is one way to interpret the size limitation on the HFT’s orders.

The HFT’s profit from the aggressive order he places in period \( t \) is given by

\[
\pi_t = yq_t - \alpha |q_t|
\]  

(2)

where \( y \) denotes the price-change at the end of period 2. Let

\[
\pi_{total} := \pi_1 + \pi_2
\]

denote the HFT’s total combined profits from periods 1 and 2. Assume that the HFT is risk-neutral and seeks to maximize the expectation of his total profits, \( \pi_{total} \).

2.1.5 Model Timeline

**Period 1** In period 1, the HFT has the opportunity to submit an aggressive order and then observe any subsequent change in resting depth. The HFT cannot observe the liquidity state directly, but he can infer the value of \( \Lambda \) from changes in resting depth if he places an aggressive order; the HFT can conclude that \( \Lambda = u \) if resting depth further depletes following his order, and \( \Lambda = A \) otherwise. If the HFT does not place an aggressive order in period 1, he does not learn \( \Lambda \).

**Period 2** At the start of period 2, the HFT observes the signal of future aggressive order-flow, \( \varphi \). The HFT observes \( \varphi \) regardless of whether he placed an aggressive order in period 1. After the HFT
observes $\varphi$, he once again has an opportunity to place an aggressive order. Finally, after the HFT has the chance to trade, aggressive order-flow characterized by $\varphi$ arrives, and prices change as determined by $\varphi$ and $\Lambda$ in equation (1).

Conceptually, the HFT’s automatic observation of $\varphi$ corresponds to the notion that aggressive order-flow is easy to predict on the basis of public market data. The HFT can always condition his period-2 trading strategy on $\varphi$, but he can condition this strategy on $\Lambda$ only if he placed an aggressive order in period 1.

2.2 Exploratory Information is Valuable

The baseline model of exploratory trading illustrates why exploratory information can be valuable, and it highlights the trade-off between the direct costs of placing an exploratory order and the informational gains from exploration.

2.2.1 Solving the Baseline Model

**Period 2** If the HFT learned the liquidity state during period 1, his optimal aggressive order in period 2 will depend on the values of both $\varphi$ and $\Lambda$. The HFT’s optimal strategy when he knows $\Lambda$ is to set $q_2 = \varphi N$ if $\Lambda = U$, and to set $q_2 = 0$ if $\Lambda = A$. Taking expectations with respect to $\varphi$ and then $\Lambda$, we find

$$E[\pi_2|\Lambda \text{ known}] = Nv(1 - \alpha)^*u^* (1 - u) \quad (3)$$

$$= Nvu(1 - \alpha)$$

If the HFT did not learn the liquidity state during period 1, his (constrained) optimal aggressive order in period 2 will still depend on the value of $\varphi$, but it will only depend on the distribution of $\Lambda$, rather than the actual value of $\Lambda$. The HFT’s optimal strategy when he does not know $\Lambda$ is to set $q_2 = \varphi N$ when $u \geq \alpha$, and to set $q_2 = 0$ when $\alpha > u$. I assumed for simplicity that $\alpha > u$, so

$$E[\pi_2|\Lambda \text{ unknown}] = 0 \quad (4)$$

**Period 1** At the start of period 1, the HFT knows neither $\varphi$ nor $\Lambda$, but he faces the same trading costs ($\alpha$ per contract) as in period 2. Consequently, the HFT’s expected direct trading profits from a
period-1 aggressive order are negative, and given by

\[ E[\pi_1] = -\alpha |q_1| \]

Since there is no noise in this baseline model, and the HFT learns \( \Lambda \) perfectly from any aggressive order that he places in the first period, we can restrict attention to the cases of \( q_1 = 0 \) and \( |q_1| = 1 \).

We obtain the following expression for the difference in the HFT’s total expected profits if he sets \( |q_1| = 1 \) instead of \( q_1 = 0 \):

\[ E[\pi_{total}| q_1 = 1] - E[\pi_{total}| q_1 = 0] = Nvu (1 - \alpha) - \alpha \] (5)

The HFT engages in exploratory trading if he sets \( |q_1| = 1 \), and he does not engage in exploratory trading if he sets \( q_1 = 0 \), so equation (5) represents the expected net gain from exploration. Exploratory trading is optimal for the HFT when this expected net gain is positive.

### 2.2.2 Conditions for Exploratory Trading

The results in section 2.2.1 demonstrate the trade-off between direct trading costs and informational gains at the heart of exploratory trading. By placing a (costly) aggressive order in period 1, the HFT “buys” the perturbation needed to elicit a response in resting depth that reveals the liquidity state. Knowing the liquidity state enables the HFT, in period 2, to better determine whether placing an aggressive order will be profitable. Parameters of the model determine the relative costs and payoffs of exploration.

Recall that when the exogenous aggressive order-flow is described by \( \varphi = 0 \), the HFT does not have any profitable period-2 trading opportunities in either liquidity state. The probability that \( \varphi \neq 0 \), given by the parameter \( v \), represents the extent to which the exogenous aggressive order-flow is predictable. To characterize how various parameters affect the viability of exploratory trading, I consider the minimal value of \( v \) for which the HFT finds it optimal to engage in period-1 (i.e., exploratory) trading. Denoting this minimal value by \( \underline{v} \), we have

\[ \underline{v} = \left( \frac{\alpha}{u} \right) \frac{1}{(1 - \alpha) N} \] (6)

The closer is \( \underline{v} \) to 0, the more conducive are conditions to exploratory trading.
The implications of equation (6) are intuitive. First, higher trading costs (α) tend to discourage exploratory trading. Second, when the HFT can use exploratory information to guide larger orders, the gains from exploration are magnified, so larger values of $N$ tend to promote exploratory trading. Finally exploratory trading becomes less viable when $u$ is smaller. The HFT will take the same action in period 2 when he knows that $\Lambda = U$ as when he doesn’t know $\Lambda$, so when $u$ is small, knowledge of the liquidity state is less valuable because it is less likely to change the HFT’s period-2 actions.\footnote{When $u > \alpha$, the HFT will take the same action in period 2 when he knows that $\Lambda = U$ as when he doesn’t know $\Lambda$, so knowledge of the liquidity state is less likely to change the HFT’s period-2 actions when $u$ is large. In the case of $u > \alpha$, equation (6) becomes $\mathbb{E} = \frac{1}{(1-u)^N}$, and exploratory trading indeed becomes less viable as $u$ approaches 1.}

Given the dearth of exogenous variation in the real-world analogues of $\alpha$, $N$ and $u$, the comparative statics above do not readily translate into empirically testable predictions. However, the model generates a much more fundamental prediction that can be tested empirically: \textit{if an agent is engaging in exploratory trading, then the market response following his exploratory orders should help to explain his performance on subsequent aggressive orders.} The market response after a trader’s exploratory orders should help to forecast price movements, and the trader will tend to follow up by placing further aggressive orders in the appropriate direction when the expected price movement is sufficiently large. Note that because the follow-up orders will tend to be larger than the exploratory orders, the market response after an agent’s exploratory orders should help to explain not only the performance, but also the incidence of his larger aggressive orders.

### 2.3 Private Gains from Exploratory Trading

The baseline model of exploratory trading presented above abstracted away from the details of the HFT’s inference about $\Lambda$. This simplifying assumption does not qualitatively affect the central result about the value of exploratory information, but it obscures why the HFT learns more from placing an aggressive order himself than he does from merely observing an aggressive order placed by someone else.

Factors other than aggressive order arrivals can affect the behavior of resting depth. In particular, a trader may adjust her passive orders in response to new information. Just as a trader might place an aggressive buy order if he believes that prices are too low, so might another trader who shared this belief cancel some of her passive sell orders. As a result, changes in resting depth are typically correlated with aggressive order-flow, even when the aggressive orders do not actually cause those changes. However, changes in resting depth not caused by aggressive orders do not help to forecast
the price impact of future aggressive order-flow. The HFT learns more from aggressive orders that he places himself than he learns from those placed by other traders because he can better infer causal effects from his own orders.

2.3.1 Intuition

An analogy to street traffic illustrates the main intuition for why the HFT obtains additional information from an aggressive order that he himself places. Consider a stoplight that tends to turn green shortly before a car arrives at it. This could arise for two reasons. First, the stoplight could operate on a timer, and cars might tend to approach the stoplight just before it turns green, due (e.g.) to the timing pattern of other traffic signals in the area. Alternatively, the stoplight might operate on a sensor that causes it to typically turn green when a car approaches.

A driver who knows why she arrived at the stoplight at a certain time has a greater capacity to distinguish between the two explanations than does a pedestrian standing at the stoplight. In particular, if a driver knows that the moment of her arrival at the stoplight was not determined by the timing pattern of nearby traffic signals (e.g., if she had been parked, and the stoplight was the first traffic signal that she encountered), she will learn considerably more from her observation of the stoplight than will the pedestrian. Both pedestrian and driver can update their beliefs, but the pedestrian only weights the new observation by the average probability that the driver’s arrival did not depend on the timing pattern of nearby signals.

Much as the driver’s private knowledge about why she approaches the stoplight at a certain moment enables her to learn more than the pedestrian, the HFT’s private knowledge of why he places an aggressive order enables him to learn more from the subsequent market response than he could learn from the response to an aggressive order placed by someone else.

2.3.2 Formalizing the Intuition

To make the preceding intuition more rigorous, consider a variant of the baseline model from section 2.1 in which some trader other than the HFT places an aggressive order at the beginning of period 1. With probability $\rho$, this aggressive order is the result of an unobservable informational shock, and resting depth further depletes following the order, regardless of the liquidity state $\Lambda$. Otherwise (with probability $1 - \rho$) resting depth further depletes after the order if and only if the liquidity state is unaccommodating. Aside from this new aggressive order, all other aspects of the baseline model
remain unchanged.

If the HFT places an aggressive order in period 1, his expected total profits are the same as they were in the baseline model, i.e.,

$$\mathbb{E}[\pi_{\text{total}}|q_1 = 1] = Nvu (1 - \alpha) - \alpha$$

However, the HFT’s expected profits if he does not place an order in period 1 are higher than in the baseline model, because the HFT now learns something from the depth changes following the other trader’s aggressive order. If resting depth weakly replenishes after that order, the HFT learns with certainty that the liquidity state is accommodating (i.e., $\Lambda = A$), so the HFT will not submit an aggressive order in period 2, and his total profits will be zero. Alternatively, if resting depth further depletes following the other trader’s aggressive order, we have

$$\mathbb{P}\{\Lambda = U|\text{resting depth further depletes}\} = \frac{u}{u + \rho(1-u)} + \rho(1 - u)$$

The HFT’s optimal strategy when he does not know $\Lambda$ is to set $q_2 = \varphi N$ when $\frac{u}{u + \rho(1-u)} \geq \alpha$, and to set $q_2 = 0$ otherwise. Taking expectations with respect to $\Lambda$ and $\varphi$, we find that the HFT’s ex-ante expected total profits in this case are given by

$$\mathbb{E}[\pi_{\text{total}}|AO \text{ by someone else}] = \max\left\{Nv\left(\frac{u}{u + \rho(1-u)} - \alpha\right), 0\right\}$$

### 2.3.3 Analysis

The features of the baseline model discussed in section 2.2.2 are qualitatively unchanged in the modified version, but now the “privacy” parameter $\rho$ also exerts an influence. In the limiting case where the depth change following an aggressive order placed by someone else is completely uninformative to the HFT (i.e., $\rho = 1$), equation (7) collapses down to equation (4) from the baseline model. At the opposite extreme, when the HFT learns the liquidity state perfectly from observing another trader’s aggressive order (i.e., $\rho = 0$), the HFT’s expected total profits are unambiguously lower if he places an aggressive order in period 1 himself.

When the HFT can learn more about the liquidity state through mere observation, as he can when $\rho$ is smaller, he has less incentive to incur the direct costs of exploratory trading. Viewed differently, if the HFT does find it optimal to engage in exploratory trading, it must be the case that he obtains
more useful information from the market response to his aggressive orders than he does from the market response to other traders’ aggressive orders. By symmetry, it must also be the case that each other trader obtains no more useful information from the market response to the HFT’s aggressive orders than they do from the market response to another arbitrary trader’s aggressive orders.

### 2.4 Testable Predictions

Before attempting any empirical evaluation of the exploratory trading model’s predictions, two basic issues must be addressed. First, it must be determined which HFTs, if any, actually earn positive and abnormal profits from their aggressive trading. I address this matter in section 4.2, and I identify eight such HFTs, to whom I refer as “A-HFTs.” Next, among the A-HFTs’ aggressive orders, suitable candidates for putative exploratory orders must be identified in some manner. The results from section 2.2.1 suggest that small, unprofitable aggressive orders are prime candidates. In section 4.4, I find that all of the A-HFTs, indeed, tend to lose money on their smallest aggressive orders, consistent with the theory that these orders are placed for exploratory ends.

With these two preliminary matters resolved, I turn to direct empirical tests of the model’s key predictions. As a benchmark, I consider the market response following the last small aggressive order placed by anyone, which is public information. The empirical implications discussed earlier in this section can then be condensed into two central predictions, namely that relative to the public-information benchmark, the market response following an A-HFT’s small aggressive order:

**Predict.1.** Explains a significant additional component of that A-HFT’s earnings on subsequent aggressive orders, but

**Predict.2.** Does not explain any additional component of other traders’ earnings on subsequent aggressive orders

In section 4.3, I make rigorous the notion of “explaining earnings on subsequent aggressive orders,” then in section 5, I introduce an explicit numeric measure of “market response” and formally test the predictions above.

### 3 High-Frequency Trading in the E-mini Market

The E-mini S&P 500 futures contract is a cash-settled instrument with a notional value equal to $50.00 times the S&P 500 index. Prices are quoted in terms of the S&P 500 index, at minimum increments,
“ticks”, of 0.25 index points, equivalent to $12.50 per contract. E-mini contracts are created directly by buyers and sellers, so the quantity of outstanding contracts is potentially unlimited.

All E-mini contracts trade exclusively on the CME Globex electronic trading platform, in an order-driven market. Transaction prices/quantities and changes in aggregate depth at individual price levels in the orderbook are observable through a public market-data feed, but the E-mini market provides full anonymity, so the identities of the traders responsible for these events are not released. Limit orders in the E-mini market are matched according to strict price and time priority; a buy (sell) limit order at a given price executes ahead of all buy (sell) limit orders at lower (higher) prices, and buy (sell) limit orders at the same price execute in the sequence that they arrived. Certain modifications to a limit order, most notably size increases, reset the time-stamp by which time-priority is determined.

E-mini contracts with expiration dates in the five nearest months of the March quarterly cycle (March, June, September, December) are listed for trading, but activity typically concentrates in the contract with the nearest expiration. Aside from brief maintenance periods, the E-mini market is open 24 hours a day, though most activity occurs during “regular trading hours,” namely, weekdays between 8:30 a.m. and 3:15 p.m. CT.

### 3.1 Description of the Data

I examine account-labeled, millisecond-timestamped records at the Commodity Futures Trading Commission of the so-called “business messages” entered into the Globex system between September 17, 2010 and November 1, 2010 for all E-mini S&P 500 futures contracts. This message data captures not only transactions, but also events that do not directly result in a trade, such as the entry, cancellation, or modification of a resting limit order. Essentially, business messages include any action by a market participant that could potentially result in or affect a transaction immediately, or at any point in the future. I restrict attention to the December-expiring E-mini contract (ticker ESZ0). During my sample period, ESZ0 activity accounted for roughly 98% of the message volume across all E-mini contracts, and more than 99.9% of the trading volume.

The price of an ESZ0 contract during this period was around $55,000 to $60,000, and (one-sided) trading volume averaged 1,991,252 contracts or approximately $115 billion per day. Message volume averaged approximately 5 million business messages per day.

---

*Excluded from these data are purely administrative messages, such as log-on and log-out messages. The good-‘til-cancel orders in the orderbook at the start of September 2, and a small number of modification messages (around 2 – 4%) are also missing from these records. Because I restrict attention to aggressive orders, and I only look at changes in resting depth (rather than its actual level), my results are not sensitive to these omitted messages.*
3.2 Defining “High-Frequency Traders”

Kirilenko et al. identify as HFTs those traders who exhibit minimal accumulation of directional positions, high inventory turnover, and high levels of trading activity. I, too, use these three characteristics to define and identify HFTs. To quantify an account’s accumulation of directional positions, I consider the magnitude of changes in end-of-day net position as a percentage of the account’s daily trading volume. Similarly, I use an account’s maximal intraday change in net position, relative to daily volume, to measure inventory turnover. Finally, I use an account’s total trading volume as a measure of trading activity.

I select each account whose end-of-day net position changes by less than 6% of its daily volume, and whose maximal intraday net position changes are less than 20% of its daily volume. I rank the selected accounts by total trading volume, and classify the top 30 accounts as HFTs. The original classifications of Kirilenko et al. and Baron et al. guided the rough threshold choices for inter-day and intraday variation. Thereafter, since confidentiality protocols prohibit disclosing results for groups smaller than eight trading accounts, the precise cutoff values of 6%, 20%, and 30 accounts were chosen to ensure that all groups of interest would have at least eight members. My central results are not sensitive to values of these parameters.

The set of HFTs corresponds closely to the set of accounts with the greatest trading volume in my sample, so the set of HFTs is largely invariant both to the exact characterizations of inter-day and intraday variation in net position relative to volume, and to the exact cutoff values for these quantities. Similarly, changing the 30-account cutoff to (e.g.) 15 accounts or 60 accounts does not substantially alter my results, because activity heavily concentrates among the largest HFTs. For example, the combined total trading volume of the 8 largest HFTs exceeds that of HFTs 9-30 by roughly three-quarters, and the combined aggressive volume of the 8 largest HFTs exceeds that of HFTs 9-30 by a factor of almost 2.5.

3.3 HFTs’ Prominence and Profitability

Although HFTs constitute less than 0.1% of the 41,778 accounts that traded the ESZ0 contract between September 17, 2010 and November 1, 2010, they participate in 46.7% of the total trading volume during this period. In addition to trading volume, HFTs are responsible for a large fraction of message volume. During the sample period, HFTs account for 31.9% of all order entry, order modification and order cancellation messages. The HFTs also appear to earn large and stable profits.
Gross of trading fees, the 30 HFTs earned a combined average of $1.51 million per trading day during the sample period. Individual HFTs’ annualized Sharpe ratios are in the neighborhood of 10 to 11.

The Chicago Mercantile Exchange reduces E-mini trading fees on a tiered basis for traders whose average monthly volume exceeds various thresholds. Trading and clearing fees were either $0.095 per contract or $0.12 per contract for the 20 largest HFTs, and were at most $0.16 per contract for the remaining HFTs. Initial and maintenance margins were both $4,500 for all of the HFTs.

Hereafter, unless otherwise noted, I restrict attention to activity that occurred during regular trading hours. HFTs’ aggressive trading occurs almost exclusively during regular trading hours (approximately 95.6%, by volume), and market conditions during these times differ substantially from those during the complementary off-hours.

4 HFTs’ Profits from Aggressive Orders

Aggressive trading is a tremendously important component of HFTs’ activity. In aggregate, approximately 48.5% of HFTs’ volume is aggressive, and this figure rises to 54.2% among the 12 largest HFTs. Furthermore, many HFTs consistently profit from their aggressive trading. Since the bid-ask spread in the E-mini market rarely exceeds the minimum imposed upon it by the granularity of prices, there is little mystery about how a trader’s passive trades could consistently earn money. By contrast, explaining how a trader who uses only market data could consistently profit on aggressive trades is somewhat difficult.

4.1 Measuring Aggressive Order Profitability

Because all E-mini contracts of a given expiration date are identical, it is neither meaningful nor possible to distinguish among the individual contracts in a trader’s inventory, so there is generally no way to determine the exact prices at which a trader bought and sold a particular contract. As a result, it is typically impossible to measure directly the profits that a trader earns on an individual aggressive order. However, the cumulative price change following an aggressive order, normalized by the order’s direction (+1 for a buy, or −1 for a sell), can be used to construct a meaningful proxy for the order’s profitability. Intuitively, the average expected profit from an aggressive order equals

---

3Explaining the profitability of individual passive trades does not resolve the question of how various HFTs manage to participate in so many passive trades. In equilibrium, we would expect new entrants to reduce the average passive volume of an individual trader until her total profits from passive trades equaled her fixed costs.
the expected favorable price movement, minus trading/clearing fees and half the bid-ask spread. See Appendix B for rigorous justification.

Estimating the cumulative favorable price movement after an aggressive order is straight-forward. Consider a trader who can forecast price movements up to \( j \) time periods in the future, but no further. If the trader places an aggressive order in period \( t \), any price changes that she could have anticipated at the time she placed the order will have occurred by period \( t+j+1 \). Provided that price is a martingale with respect to its natural filtration, the expected change in price from period \( t+j+1 \) onward is zero, both from the period–\( t \) perspective of the trader and from an unconditional perspective. Thus the change in price between period \( t \) and any period after \( t+j \), normalized by the direction of the trader’s order (\(+1\) for a buy, or \(-1\) for a sell), will provide an unbiased estimate of the favorable price movement following the trader’s order.

The remarks above imply that we can derive a proxy for the profitability of an HFT’s aggressive order using the (direction-normalized) accumulated price-changes following that aggressive order out to some time past the HFT’s maximum forecasting horizon. If we choose too short an accumulation window, the resulting estimates of the long-run direction-normalized average price changes following the HFT’s aggressive orders will be biased downward. As a result, we can empirically determine an adequate accumulation period by calculating cumulative direction-normalized price changes over longer and longer windows until their mean ceases to significantly increase. Using too long an accumulation period introduces extra noise, but it will not bias the estimates. I find that an accumulation period, measured in event-time, of 30 aggressive order arrivals is sufficient to obtain unbiased estimates; for all of the empirical work in this paper, I use an accumulation period of 50 aggressive order arrivals to allow a wide margin for error. See Appendix B for further details.

As noted earlier, the bid-ask spread for the E-mini is almost constantly \$12.50 \text{ (one tick)} \) during regular trading hours, and the HFTs in my sample face trading/clearing fees of \$0.095\) to \$0.16 per contract, so the average favorable price movement necessary for an HFT’s aggressive order to be profitable is between \$6.345\) and \$6.41 per contract. Since trading/clearing fees vary across traders, I report aggressive order performance in terms of favorable price movement, that is, earnings gross of fees and the half-spread.
4.2 HFTs’ Overall Profits from Aggressive Orders

To measure the overall mean profitability of a given account’s aggressive trading, I compute the average cumulative price change following each aggressive order placed by that account, weighted by executed quantity and normalized by the direction of the aggressive order. As a group, the 30 HFTs in my sample achieve size-weighted average aggressive order performance of $7.01 per contract. On an individual basis, nine HFT accounts exceed the relevant $6.25 + fees profitability hurdle, and each of these nine accounts exceeds this hurdle by a margin that is statistically significant at the 0.05 level. One of these nine accounts is linked with another HFT account, and their combined average performance also significantly exceeds the profitability hurdle.

Overall, the HFTs vastly outperform non-HFTs, who earn a gross average of $3.19 per aggressively-traded contract. However, these overall averages potentially confound effects of very coarse differences in the times at which traders place aggressive orders with effects of the finer differences more directly related to strategic choices. For example, if all aggressive orders were more profitable between 1 p.m. and 2 p.m. than at other times, and HFTs only placed aggressive orders during this window, the HFTs’ outperformance would not depend on anything characteristically high-frequency.

To control for potential low-frequency confounds, I divide each trading day in my sample into 90-second segments and regress the profitability of non-HFTs’ aggressive orders during each segment on both a constant and the executed quantities of the aggressive orders. Using these local coefficients, I compute the profitability of each aggressive order by an HFT in excess of the expected profitability of a non-HFT aggressive order of the same size during the relevant 90-second segment. With these additional controls, only 27 HFT accounts continue to exhibit significant outperformance of non-HFTs, and only eight of the 27 accounts are among those whose absolute performance exceeded the profitability hurdle.

4.2.1 A-HFTs and B-HFTs

For expositional ease, I will refer to the eight HFT accounts that make money on their aggressive trades and outperform the time-varying non-HFT benchmark as “A-HFTs,” and to the complementary set of HFTs as “B-HFTs.” The eight A-HFTs have a combined average daily trading volume of 982,988 contracts, and on average, 59.2% of this volume is aggressive. The 22 B-HFTs have a combined average daily trading volume of 828,924 contracts, of which 35.9% is aggressive. Gross of fees, the A-HFTs earn a combined average of $793,342 per day, or an individual average of $99,168 per day, while the
B-HFTs earn a combined average of $715,167 per day, or an individual average of $32,508 per day.\(^6\) The highest profitability hurdle among the A-HFTs is $6.37 per aggressively traded contract.

### 4.3 Relative Aggressive Order Profitability: HFT vs. Econometrician

To gain some insight into the factors that affect aggressive trading profits, I examine the extent to which econometric price forecasts explain the realized performance of aggressive orders placed by A-HFTs, B-HFTs, and non-HFTs. The methodology that I develop in this subsection also provides the starting point for my direct tests of the exploratory trading model’s predictions in section 5.

#### 4.3.1 Variables that Forecast Price Movements

Bid-ask bounce notwithstanding, the price at which aggressive orders execute changes rather infrequently in the E-mini market. On average, only about 1 – 3\% of aggressive buy (sell) orders execute at a final price different from the last price at which the previous aggressive buy (sell) order executed, and the price changes that do occur are almost completely unpredictable on the basis of past price changes. However, several other variables forecast price innovations surprisingly well.

In contrast to price innovations, the direction of aggressive order flow in the E-mini market is extremely persistent. On average, the probability that the next aggressive order will be a buy (sell) given that the previous aggressive order was a buy (sell) is around 75\%. In addition to forecasting the direction of future aggressive order flow, the direction of past aggressive order flow also forecasts future price innovations to statistically and economically significant extent, and forecasts based on past aggressive order signs alone are modestly improved by information about the (signed) quantities of past aggressive orders. Price forecasts can be further improved using simple measures of recent changes in the orderbook.

#### 4.3.2 Econometric Benchmark

For each trading day in my sample, I regress the cumulative price-change (in dollars) between the aggressive orders \(k\) and \(k + 50\), denoted \(y_k\), on lagged market variables suggested by the remarks above. Specifically, I regress \(y_k\) on the changes in resting depth between aggressive orders \(k - 1\) and \(k\) at each of the six price levels within two ticks of the best bid or best ask, the signs of aggressive orders \(k - 1\) through \(k - 4\), and the signed executed quantities of aggressive orders \(k - 1\) through \(k - 4\).\(^6\)

\(^6\)All of the preceding descriptive statistics include the small amount of trading activity that occurred outside regular trading hours.
For symmetry, I adopt the convention that sell depth is negative, and buy depth is positive, so that an increase in buy depth has the same sign as a decrease in sell depth. Denoting the row vector of the 14 regressors by $z_{k-1}$, and a column vector of 14 coefficients by $\Gamma$, I estimate the equation

$$y_k = z_{k-1} \Gamma + \epsilon_k$$  (8)

Appendix C presents coefficient estimates and direct discussion of the regression results.

To compute the excess performance of aggressive order $k$, denoted $\xi_k$, I normalize the $k$th regression residual by $\text{sign}_k$, the sign of the $k$th aggressive order:

$$\xi_k = \text{sign}_k \left( y_k - z_{k-1} \hat{\Gamma} \right)$$

As discussed in section 4.1, normalizing the cumulative price-change $y_k$ by the sign of the $k$th aggressive order yields a measure of the $k$th aggressive order’s profitability. Likewise, the quantity $\xi_k$ provides a measure of $k$th aggressive order’s profitability in excess of that expected on the basis of the benchmark econometric specification. I compute the vectors of direction-normalized residuals separately for each of the 32 trading days in my sample, then combine all of them into a single vector for the entire sample period.

4.3.3 Explained Performance

The price movements predicted by (8) explain a substantial component of the performance of aggressive orders placed by A-HFTs, B-HFTs, and non-HFTs alike.$^7$ Looking ahead, this explanatory power validates the use of specification (8) as a basis for the more sophisticated analyses in section 5. Figure 1 and Table 1, below, summarize the overall size-weighted average performance of aggressive orders placed by various trader groups, both in absolute terms, and in excess of the econometric benchmark. Confidence intervals are computed via bootstrap.

$^7$Although the variables in $z_{k-1}$ are all observable before the $k$th aggressive order arrives, the fitted value $z_{k-1} \hat{\Gamma}$ is not literally a forecast of $y_k$ in the strictest sense, as $\hat{\Gamma}$ is estimated from data for the entire day. However, the coefficient estimates are extremely stable throughout the sample period, so thinking of $z_{k-1} \hat{\Gamma}$ as a forecast of $y_k$ is innocuous in the present setting. See section 6.1.
The exploratory trading model developed earlier assumed that A-HFTs ultimately traded ahead of easily predictable demand innovations (when liquidity conditions were suitably unaccommodating), and the explanatory power of equation (8) for the A-HFTs’ performance substantiates this assumption. At the same time, although the econometric controls explain over half of the A-HFTs’ performance on their aggressive orders, the remaining unexplained component of performance is massive. The A-HFTs’ average excess performance is over 50% greater than that of the B-HFTs, and over 10 times greater than that of the non-HFTs. Price forecasts more sophisticated that those from (8) may better explain A-HFTs’ performance; I return to this matter in section 5.

4.4 A-HFTs’ Losses on Small Aggressive Orders

A-HFTs’ aggressive orders tend to become more profitable as order size increases. In fact, despite earning money from their aggressive orders on average, the A-HFTs all tend to lose money on the

---

8This effect appears whether price-changes are measured between the respective last prices at which successive aggressive orders execute (correcting for bid-ask bounce), or between the respective first prices at which they execute, so the positive relationship between executed quantity and subsequent favorable price movements is not simply an artifact of large orders that eat through one or more levels of the orderbook.
smallest aggressive orders that they place. Note also that I refer here to the size of the aggressive orders A-HFTs submit, not the quantity that executes, so the small orders were intentionally chosen to be small, and the large orders intentionally chosen to be large.

The baseline model of exploratory trading in section 2 produces exactly the sort of losses on small aggressive orders and profits on large aggressive orders that the A-HFTs exhibit. The A-HFTs’ differing performance on small and large aggressive orders is consistent with the pattern that we would expect to see if the small orders were generating valuable information that enabled the A-HFTs to earn greater profits from their large orders.

To make precise both the meaning of “small” aggressive orders, and A-HFTs’ losses on them, I specify cutoffs for order size and compute the average performance of A-HFTs’ aggressive orders below and above those size cutoffs. Figure 2 and Table 2, below, display bootstrap confidence intervals for the executed-quantity-weighted average performance of A-HFTs’ aggressive orders weakly below and strictly above various order-size cutoffs.

Figure 2. A-HFT Performance on Small and Larger Aggressive Orders (95% Conf. Intervals)
As shown in Table 2, small aggressive orders represent a substantial fraction of the aggressive orders that A-HFTs place, but these small orders make up very little of the A-HFTs' total aggressive volume. Nevertheless, A-HFTs' losses on these small orders are non-negligible. On average, each A-HFT loses roughly $7,150 per trading day ($1.8 million, annualized) on aggressive orders of size 20 or less; this loss represents approximately 7.2% of an average A-HFT's daily profits.

Although HFTs may tolerate only limited levels of inventory, inventory-management does not adequately explain the A-HFTs' losses on small aggressive orders. We can control for A-HFTs' respective net positions at the times they submit aggressive orders, and restrict attention to only those aggressive orders that move an A-HFT away from a zero net position. Such “non-rebalancing” orders account for over half of the small aggressive orders that A-HFTs place, and they cannot possibly be motivated by inventory management. Nevertheless, A-HFTs still lose money on the smallest of these orders, yet make money on the larger ones. See Table D.1 in Appendix D.

The A-HFTs' qualitative pattern of losses on small aggressive orders and more-than-offsetting gains on larger aggressive orders suggests that the small orders are reasonable candidates for exploratory orders. This finding provides a foundation for direct tests of the exploratory trading model's sharper empirical predictions.

### 5 Explicitly Isolating Exploratory Information

If the exploratory trading model is correct, and if the A-HFTs’ small aggressive orders are indeed exploratory in nature, the two key model predictions presented in section 2.4 must hold. For convenience, I summarize these predictions below.

<table>
<thead>
<tr>
<th>Cutoff</th>
<th>Below Cutoff 95% CI</th>
<th>Above Cutoff 95% CI</th>
<th>AOs Below Cutoff % of All AOs</th>
<th>AOs Below Cutoff % of Aggr. Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(3.78, 3.89)</td>
<td>(7.59, 7.74)</td>
<td>24.31%</td>
<td>0.40%</td>
</tr>
<tr>
<td>5</td>
<td>(4.17, 4.29)</td>
<td>(7.62, 7.78)</td>
<td>43.74%</td>
<td>1.44%</td>
</tr>
<tr>
<td>10</td>
<td>(3.42, 3.55)</td>
<td>(7.71, 7.85)</td>
<td>54.64%</td>
<td>3.09%</td>
</tr>
<tr>
<td>15</td>
<td>(3.79, 3.92)</td>
<td>(7.71, 7.86)</td>
<td>56.75%</td>
<td>3.54%</td>
</tr>
<tr>
<td>20</td>
<td>(4.08, 4.20)</td>
<td>(7.75, 7.90)</td>
<td>60.82%</td>
<td>4.80%</td>
</tr>
</tbody>
</table>
Relative to a benchmark that incorporates the public information about the market response following small aggressive orders placed by anyone, the market response following small aggressive orders placed by an A-HFT:

**Predict.1.** Explains a significant additional component of that A-HFT’s earnings on subsequent aggressive orders, but

**Predict.2.** Does not explain any additional component of other traders’ earnings on subsequent aggressive orders

In this section, I consider a simple numeric characterization of the market response following an aggressive order, and I directly test whether the above predictions of the exploratory trading model hold. I estimate results for the A-HFTs individually, but for compliance with confidentiality protocols, I present cross-sectional averages of these estimates. Empirically, these average results are representative of the results for individual A-HFTs.9

### 5.1 Empirical Strategy: Overview

Though the implementation is slightly involved, my basic empirical strategy is straight-forward. First, I augment the benchmark regression from section 4.3 using

1. Market response information from the last small aggressive order placed by anyone, and

2. Both market response information from the last small aggressive order placed by anyone, AND market response information from the last small aggressive order placed by a specified A-HFT

As I discuss in more detail in the next subsection, the market-response variable that I consider essentially amounts to a measure of the change in orderbook depth that follows an aggressive order.

After estimating both of the specifications above, I find the additional component of performance on larger aggressive orders explained by (2) relative to (1). The market response following an arbitrary small aggressive order is publicly observable. However, because the E-mini market operates anonymously, the distinction between a small aggressive order placed by a particular A-HFT and an arbitrary small aggressive order is private information, available only to the A-HFT who placed the

---

9Throughout the E-mini market, there exist assorted linkages between various trading accounts (as, for example, in the simple case where single firm trades with multiple accounts), so the trading-account divisions do not necessarily deliver appropriate atomic A-HFT units. Though the specifics are confidential, the appropriate partition of the A-HFTs is entirely obvious. For brevity, I use “individual A-HFT” as shorthand to “individual atomic A-HFT unit,” as applicable.
order. Comparing the second specification above to the first isolates the effects attributable to this private information from effects attributable to public information.

Finally, I compare the additional explained performance for the specified A-HFT to the additional explained performance for all other traders. Intuitively, we want to verify that the A-HFT’s exploratory information provides extra explanatory power for the subsequent performance of trader privy to that information (the A-HFT), but not for the performance of traders who aren’t privy to it (everyone else). Note that “everyone else” includes the A-HFTs other than the specified A-HFT.

Some A-HFT accounts and B-HFT/non-HFT accounts belong to the same firms, and various B-HFTs/non-HFTs may be either directly informed or able to make educated inferences about what one or more A-HFTs do. As a result, we should not necessarily expect exploratory information generated by an A-HFT’s small orders to provide no explanatory power whatsoever for all other traders’ performance. However, we should still expect the additional explanatory power for the A-HFT’s performance to significantly exceed that for the other traders’ performance.

5.2 Empirical Implementation

Define an aggressive order to be “small” if that order’s submitted size is less than or equal to a specified size parameter, which I denote by $\bar{q}$.

5.2.1 A Simple Measure of Market Response

I characterize the market response to a small aggressive order using subsequent changes in orderbook depth. I examine the interval starting immediately after the arrival of a given small aggressive order and ending immediately before the arrival of the next aggressive order (which may or may not be small), and I sum the changes in depth at the best bid and best ask that occur during this interval. As in section 4.3, I treat sell depth as negative and buy depth as positive. I also normalize these depth changes by the sign of the preceding small aggressive order to standardize across buy orders and sell orders.

To simplify the analysis and stack the deck against finding significant results, I initially focus only on the sign of the direction-normalized depth changes. Note the direct analogy to the two-liquidity-state setting of the exploratory trading model in section 2.
For a given value of $\bar{q}$, I construct the indicator variable $\Omega$, with $k$th element $\Omega_k$ defined by

$$\Omega_k = \begin{cases} 1 & \text{if } DC(k;\text{any},\bar{q}) > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $DC(k;\text{any},\bar{q})$ denotes the direction-normalized depth change following the last small aggressive order (submitted by anyone) that arrived before the $k$th aggressive order. Similarly, I construct the indicator variable $\Omega^A$, with $k$th element $\Omega^A_k$ defined by

$$\Omega^A_k = \begin{cases} 1 & \text{if } DC(k;\text{AHFT},\bar{q}) > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $DC(k;\text{AHFT},\bar{q})$ denotes the direction-normalized depth change following the last small aggressive order submitted by a specified A-HFT that arrived before the $k$th aggressive order.

### 5.2.2 Estimation Procedure

In the model of exploratory trading presented earlier, exploratory information was valuable only in conjunction with information about future aggressive order flow. Following this result, I incorporate market-response information by using the indicators $\Omega$ and $\Omega^A$ to partition the benchmark regression from section 4.3.

Recall that in section 4.3, I estimated the equation

$$y_k = z_{k-1} \Gamma + \epsilon_k$$

where $y_k$ denoted the cumulative price-change between the aggressive orders $k$ and $k + 50$, and the vector $z_{k-1}$ consisted of changes in resting depth between aggressive orders $k - 1$ and $k$, in addition to the signs and signed executed quantities of aggressive orders $k - 1$ through $k - 4$. Using the indicator $\Omega$, I now partition the equation above into two pieces and estimate the equation

$$y_k = \Omega_k z_{k-1} \Gamma^a + (1 - \Omega_k) z_{k-1} \Gamma^b + \epsilon_k$$

(9)
Next, I use the indicator $\Omega^A$ to further partition (9), and I estimate the equation

$$y_k = \Omega^A_k (\Omega_k z_k - 1 \Gamma^c + (1 - \Omega_k) z_{k-1} \Gamma^d) + (1 - \Omega^A_k) \left( \Omega_k z_k - 1 \Gamma^c + (1 - \Omega_k) z_{k-1} \Gamma^d \right) + \epsilon_k$$

(10)

The variables $y_k$ and $z_{k-1}$ denote the same quantities as before, and the $\Gamma^j$ terms each represent vectors of 14 coefficients.

I estimate (9) and (10) for $\bar{q} = 1, 5, 10, 15, 20$, and for each specification I calculate the relative excess performance of the specified A-HFT, and of all other trading accounts on aggressive orders of size strictly greater than $\bar{q}$. As in section 4.3, I compute the performance of aggressive order $k$ in excess of that explained by each regression by normalizing the $k$th residual from the regression by the sign of the $k$th aggressive order. I now also control for order-size effects directly by regressing the direction-normalized residuals (for the orders of size strictly greater than $\bar{q}$) on the (unsigned) executed quantities and a constant, then subtracting off the executed quantity multiplied by its estimated regression coefficient. Controlling for size effects in this manner makes results more comparable for different choices of $\bar{q}$. Size effects can be addressed by other means with negligible impact on the final results.

For each aggressive order larger than $\bar{q}$ placed by the A-HFT under consideration, I compute the additional component of performance explained by (10) relative to (9) by subtracting the order’s excess performance over (10) from its excess performance over (9); I stack these additional explained components in a vector that I denote by $\Xi_A$. I repeat this procedure to obtain the analogous vector for everyone else, $\Xi_{ee}$.

**5.3 Results**

I evaluate the empirical predictions of the exploratory trading model by comparing the additional explained component of performance for each A-HFT to the additional explained component of performance for all other traders. Specification (10) has more free parameters than (9), but additional explanatory power of (10) due exclusively to the extra degrees of freedom will manifest equally, in expectation, for all traders, so the extra degrees of freedom alone should not cause $\Xi_A$ and $\Xi_{ee}$ to differ significantly.

Figure 3, below, displays the cross-sectional means of $\Xi_A$ and $\Xi_{ee}$ for different values of $\bar{q}$ (see
Appendix Table D.2 for numeric values). Both predictions of the exploratory trading model are borne out in these results. Using information about the market activity immediately following an A-HFT’s smallest aggressive orders (in the form of $\Omega^4$) improves our ability to explain that A-HFT’s performance on larger aggressive orders by a highly significant margin, relative to using only information about the activity following any small aggressive order (in the form of $\Omega$). By contrast, relative to using $\Omega$ alone, incorporating the information in $\Omega^4$ provides little or no significant additional explanatory power for other traders’ performance on larger aggressive orders.

Figure 3. Additional Performance Explained (95% Confidence Intervals)

As a more formal comparison of the gain in explanatory power for the A-HFTs relative to the gain for everyone else, I construct 95% bootstrap confidence intervals for the difference of the pooled means $Mean(\Xi_A) - Mean(\Xi_{ee})$, for $\bar{q} = 1, 5, 10, 15, 20$. Figure 4 summarizes these results, which confirm what the preceding results suggested: the extra component of A-HFTs’ performance on large aggressive orders explained by using $\Omega^4$ in addition to $\Omega$ is significantly greater than the extra component explained for other traders. (See Appendix Table D.3 for numeric values)

Figure 4. [A-HFT Addt'l Explained] - [Everyone Else Addt'l Explained] (95% Conf. Intervals)
Although the extra explanatory power for an average individual A-HFT is significantly greater than that for all other traders, the amount of performance to be explained is also somewhat greater. Comparing extra explanatory power for an individual A-HFT to extra explanatory power for the complementary set of HFTs mitigates this difference. Consistent with the notion that certain B-HFTs may know something about what various A-HFTs are doing, the extra component of performance explained by using $\Omega^A$ in addition to $\Omega$ is larger for the complementary set of HFTs than it is for the broader “everyone except the A-HFT of interest” group. Nevertheless, aside from the case of $\bar{q} = 1$, the average extra explanatory power for an individual A-HFT is still significantly greater than is that for the complementary set of HFTs, as shown in Figure 5, below. (See Appendix Table D.3 for numeric values.)

![Figure 5. [A-HFT Addt'l Explained] - [Other HFTs Addt'l Explained] (95% Conf. Intervals)](image)

### 5.4 Incidence of A-HFTs’ Larger Aggressive Orders

As suggested by the remarks at the end of section 2.2.2, the prediction that exploratory information explains a significant additional component of the A-HFTs’ performance tacitly requires exploratory information to help explain the incidence of the A-HFTs’ larger aggressive orders. In particular, all else being equal, the exploratory trading model predicts that an A-HFT will have a greater tendency to place large aggressive orders when $\Omega^A = 1$ than when $\Omega^A = 0$. A direct test of this prediction about the incidence of the A-HFTs’ larger aggressive orders offers a robustness check on the results in subsection 5.3.

Much as the HFT in the model from section 2 considered the signal of future aggressive order-flow as well as the liquidity state, A-HFTs consider public market data as well as exploratory information to decide when to place large aggressive orders. The size and direction of A-HFTs’ aggressive orders depend on the same variables that forecast price movements, or equivalently on the forecasts of price
movements themselves. On average, the signed quantity of an A-HFT’s aggressive order should be an increasing function of the future price-change expected on the basis of public information. In this context, the exploratory trading model predicts that the expected future price-change will have a larger effect on the signed quantity of an A-HFT’s aggressive orders when $\Omega^A = 1$ than it will when $\Omega^A = 0$.

To test the exploratory trading model’s prediction about the incidence of A-HFTs’ larger aggressive orders, I regress the signed quantities of a given A-HFT’s aggressive orders on the associated fitted values of $y$ from equation (9), partitioned by $\Omega^A$. In other words, for a specified A-HFT and a given value of $\bar{q}$, I estimate the equation

$$q_k = \beta_0 \left(1 - \Omega^A_k\right) \hat{y}_k + \beta_1 \Omega^A_k \hat{y}_k + \epsilon_k$$

where $q_k$ denotes the signed submitted quantity of the A-HFT’s $k$th aggressive order, $\hat{y}_k$ denotes the relevant fitted value of $y_k$ from the public-information regression (9), and $\Omega^A$ is the usual indicator function. I restrict the $\beta$ coefficients to be the same across all A-HFTs.

Table 3, below, displays the coefficient estimates from (11) for various values of $\bar{q}$. A Wald test rejects the null hypothesis $\beta_0 = \beta_1$ at the $10^{-15}$ level for all values of $\bar{q}$. As the exploratory trading model predicts, holding fixed the price-change expected on the basis of public information, the average A-HFT places significantly larger aggressive orders when $\Omega^A = 1$ than when $\Omega^A = 0$.

Table 3. Differential Effects of Predicted Price-Changes on A-HFT Signed Order Size

<table>
<thead>
<tr>
<th>$\bar{q}$</th>
<th>Point Estimates</th>
<th>Standard Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta_0$</td>
<td>$\beta_1$</td>
</tr>
<tr>
<td>1</td>
<td>13.35</td>
<td>15.26</td>
</tr>
<tr>
<td>5</td>
<td>13.41</td>
<td>15.11</td>
</tr>
<tr>
<td>10</td>
<td>13.42</td>
<td>14.97</td>
</tr>
<tr>
<td>15</td>
<td>13.34</td>
<td>15.10</td>
</tr>
<tr>
<td>20</td>
<td>13.23</td>
<td>15.30</td>
</tr>
</tbody>
</table>
6 Practical Significance of Exploratory Information

The losses that A-HFTs incur on their small aggressive orders offer a natural point of comparison for the gains that can be explained from the information generated by those orders. Table 4, below, displays the additional component of A-HFTs’ profits on large aggressive orders directly explained using exploratory information (in the form of $\Omega^4$) as a percentage of A-HFTs’ losses on small aggressive orders.

In one sense, given the extreme simplicity and coarseness of the $\Omega$–operators as representations of exploratory information, the results in Table 4 suggest gains that are surprisingly large in practical terms. At the same time, the gains from exploration should at least weakly exceed the costs, and the additional gains directly explained using $\Omega^4$ fall short of this mark.

Table 4. Extra Explained Gains on Large AOs vs. Losses on Small AOs

| $\bar{q}$ | $(\text{Extra Explained Gains}) / |\text{Losses}|$ |
|-----------|---------------------------------------------|
| 1         | 33.05%                                      |
| 5         | 11.27%                                      |
| 10        | 4.48%                                       |
| 15        | 5.39%                                       |
| 20        | 4.79%                                       |

Representations of exploratory information richer than $\Omega^4$ are extremely easy to construct. For example, an obvious extension would be to consider the not only the sign, but also the magnitude of the direction-normalized depth change following an exploratory order. Regardless of the particular representation of exploratory information used, though, the additional explained component of A-HFTs’ profits on the aggressive orders they place is likely to understate the true gains from exploration. As the simple model in section 2 illustrates, exploratory information is valuable in large part because it enables a trader to avoid placing unprofitable aggressive orders. However, estimates of the additional explained component of profits on A-HFTs’ aggressive orders necessarily omit the effects of such avoided losses. While this bias, if anything, makes the preceding findings of statistical significance all the more compelling, it also complicates the task of properly determining the practical importance of exploratory information.
6.1 Simulated Trading Strategies

To investigate the gains from exploratory information, including the gains from avoiding unprofitable aggressive orders, I examine the effects of incorporating market-response information from small aggressive orders into simulated trading strategies. The key advantage of working with these simulated trading strategies is that avoided unprofitable aggressive orders can be observed directly.

The basic trading strategy that I consider is a simple adaptation of the benchmark regression from section 4.3. I specify a threshold value, and the strategy entails nothing more than placing an aggressive order with the same sign as \( \hat{y}_k \) whenever \( |\hat{y}_k| \) exceeds that threshold. To make this strategy feasible (in the sense of using only information available before time \( t \) to determine the time-\( t \) action) I compute the forecast of future price movement, \( \hat{y}_k \), using the regression coefficients estimated from the previous day’s data. I incorporate market-response information into this strategy by modifying the rule for placing aggressive orders to, “place an aggressive order (with the same sign as \( \hat{y}_k \)) if and only if all three of the following conditions hold:

- \( |\hat{y}_k| \) exceeds its specified threshold,
- The direction-normalized depth-change following the last small aggressive order (placed by anyone) exceeds a specified threshold, and
- The direction-normalized depth-change following the last small aggressive order placed by an A-HFT exceeds a (possibly different) specified threshold.”

Choosing a threshold of \(-\infty\) will effectively remove any of these conditions.

Each strategy yields a set of times to place aggressive orders, and the associated direction for each order. To measure the performance of a given strategy, I compute the average profitability of the indicated orders in the usual manner, with the assumption that these aggressive orders are all of a uniform size.

Relative to A-HFTs’ losses on small aggressive orders, the additional component of A-HFTs’ profits directly explained using \( \Omega^A \) is smallest when \( \bar{q} = 10 \), and I present results for \( \bar{q} = 10 \) to highlight the impact of accounting for avoided losses on estimates of the gains from exploratory information. Results for other values of \( \bar{q} \) are similar.
6.1.1 Three Specific Strategies

All three threshold parameters affect strategy performance, so to emphasize the role of market-response information, I present results with the threshold for $|\hat{y}_k|$ held fixed. Varying the threshold for $|\hat{y}_k|$ does not alter the qualitative results. In particular, it is not possible to achieve the same gains in performance that result from incorporating exploratory information by merely raising the threshold for $|\hat{y}_k|$. The forecast $\hat{y}_k$ uses coefficients estimated from the previous day’s data, and these forecasts exhibit increasing bias as the $z_{k-1}$ observations assume more extreme values.

I consider a range of threshold values for the direction-normalized depth-change following the last small aggressive order placed by anyone, but, for expositional clarity, I present results for three illustrative threshold choices for the direction-normalized depth-change following the last small aggressive order placed by an A-HFT. Specifically, I consider thresholds of $-\infty$ (no A-HFT market-response information), 0 (the same information contained in $\Omega^A$), and 417 (the 99th percentile value). Figure 6 displays the performance of these three strategies over a range of threshold values for the market response following arbitrary small aggressive orders.

![Figure 6. Absolute Gains from Exploratory Information](image)

While the performance gains from incorporating A-HFT exploratory information are obvious, an equally important feature of the results above is more subtle. The A-HFTs’ average gross earnings on aggressive orders over size 10 of $7.78$ per contract are well above the peak performance of the strategy that uses only public information, but substantially below the performance of the strategy that incorporates the A-HFTs’ exploratory information with the higher threshold. This is exactly the pattern that we should expect, given that the former strategy excludes information that is available
to the A-HFTs and the latter strategy includes information that is not available to any individual A-HFT, so these results help to confirm the relevance and validity of this simulation methodology.

6.1.2 Gains from Exploration Relative to Losses on Exploratory Orders

Although the two strategies that incorporate exploratory information from the A-HFTs’ small aggressive orders outperform the strategy that does not, the orders that generated the exploratory information were costly. To compare the gains from this exploratory information to the costs of acquiring it, I first multiply the increases in per-contract earnings for the two exploratory strategies (scaled by the respective number of orders relative to the public-information strategy) by the A-HFTs’ combined aggressive volume on orders over size 10. I then divide these calibrated gains by the A-HFTs’ actual losses on aggressive orders size 10 and under. The resulting ratio is the direct analogue of the percentages in Table 4.

Figure 7 displays the calibrated ratio of additional gains to losses for each exploratory simulated strategy over a range of threshold values for the market response following arbitrary small aggressive orders. Using information from the A-HFTs’ exploratory orders analogous to that in \( \Omega^A \), the additional gains are roughly 15% larger than the losses on exploratory orders. Whereas the extra component of the A-HFTs’ performance directly explained using \( \Omega^A \) represented less than 5% of A-HFTs’ losses on exploratory orders, the analogous estimated performance increases more than offset the costs of exploration once we include the gains from avoiding unprofitable aggressive orders. In the case of the strategy that employs information from the A-HFTs’ exploratory orders with the higher threshold, the estimated gains from exploration exceed the costs by more than one-third.

\(^{10}\)The two strategies that incorporate exploratory information select subsets of the aggressive order placement times generated by the public-information-only strategy. Although the selected orders tend to be more profitable, they are also fewer in number.
7 Discussion

7.1 Broader Opportunities for Exploratory Gains from Aggressive Orders

The empirical results in the preceding sections focused on the information generated by the A-HFTs’ smallest aggressive orders. While their otherwise-perplexing unprofitability made these orders the most obvious starting point for an empirical study of exploratory trading, there is no theoretical reason why these small orders should be the sole source of exploratory information. In the baseline exploratory trading model, it was only to highlight the key aspects of the model that I assumed the HFT’s period-1 order was expected to lose money and served no purpose other than exploration.

In principle, even aggressive orders that an A-HFT expects to be directly profitable could produce valuable, private, exploratory information. To investigate this possibility, I repeat the analysis of section 5.2.2 setting $\bar{q} = 25, 30, 35, 40, 45, 50, 60, 75, 90$. The A-HFTs’ incremental aggressive orders included with each increase of $\bar{q}$ beyond $\bar{q} = 20$ are directly profitable on average, and yet the market response following these orders still provides significantly more additional explanatory power for the A-HFTs’ performance on larger aggressive orders than it provides for that of other traders. Indeed, the additional explained components of the A-HFTs’ performance are markedly larger than those for $\bar{q} = 1, 5, ..., 20$; see Figure 8, below, and see Appendix Table D.4 for numeric results.
7.2 Exploratory Trading and Speed

Further analysis of the exploratory trading model reveals natural connections between exploration and two important concepts of speed.

7.2.1 Low Latency

One common measure of trading speed is latency—the amount of time required for messages to pass back and forth between a trader and the market. While low-latency operation and high-frequency trading are not equivalent, minimal latency is nonetheless a hallmark of high-frequency traders. For a trader who can identify profitable trading opportunities, there is obvious value to possessing latency low enough to take advantage of these opportunities before they disappear. The new insight from the exploratory trading model concerns the more subtle matter of how low latency connects to the identification of such opportunities.

In the model of exploratory trading developed in section 2, the HFT’s inference about $\Lambda$ on the basis of market activity following his aggressive order in period 1 implicitly depends on a notion related to latency. If we suppose that random noise perturbs the orderbook, say according to a Poisson arrival process, then the amount of noise present in the HFT’s observation of the market response in some interval following his aggressive order will depend on the duration of that interval. The duration of this interval will depend in large part upon the rate at which market data is collected and disseminated.
to the HFT, that is, the “temporal resolution” of the HFT’s data. Although this temporal resolution
does not directly depend on the HFT’s latency, the HFT’s latency is implicitly constrained by the
temporal resolution of his market information.

The finer temporal resolution required for low-latency operation enables low-latency traders to
obtain meaningful—and empirically valuable—information about the market activity immediately
following their aggressive orders, and this information degrades at coarser temporal resolutions. The
empirical results from section 5.3 provide a concrete illustration of this effect. The changes in resting
inside depth immediately following an arbitrary aggressive order are less useful for forecasting price
movements than are the analogous changes following an A-HFT’s aggressive order, but the two can
only be distinguished (by the A-HFT) in data with a sufficient level of temporal disaggregation.

7.2.2 High Frequency

Exploratory trading bears a natural relationship to the practice of placing large numbers of aggressive
orders—what might be considered “high-frequency trading” in the most literal sense.

Exploratory information generated by a given aggressive order is only valuable to the extent that
it can be used to improve subsequent trading performance. Because exploratory information remains
relevant for only some finite period, the value of exploratory information diminishes as the average
interval between a trader’s orders lengthens. The exploratory trading model readily captures this
effect if we relax the simplifying assumption that the liquidity state $\Lambda$ remains the same between
periods 1 and 2. Suppose that $\Lambda$ evolves according to a Markov process, such that with probability
$\tau$, a second $\Lambda$ is drawn in period 2 (from the same distribution as in period 1), and with probability
$1 - \tau$, the original value from period 1 persists in period 2. Intuitively, $\tau$ parametrizes the length
of period 1, and this length increases from zero to infinity as $\tau$ increases from zero to unity. As $\tau$
tends towards unity—i.e., as the length of period 1 increases to infinity—the liquidity state in period
1 becomes progressively less informative about the liquidity state in period 2.

As discussed in section 7.1, both theory and empirical evidence suggest that almost any aggressive
order that a trader places generates some amount of exploratory information. Consequently, as a
trader places aggressive orders in greater numbers, he will gain access to greater amounts of exploratory
information. Furthermore, the average time interval between a trader’s aggressive orders necessarily
shrinks as the number of those orders grows, so the exploratory information produced by each order
tends to become more valuable to the trader. These synergistic effects dramatically magnify the
potential gains from exploratory information for traders who place large numbers of aggressive orders.

7.2.3 Latency Détente

There has been much speculation about HFTs engaging in an “arms race” for ever-faster processing and ever-lower latency. If high-frequency trading entailed nothing more than reacting to publicly observable trading opportunities before anyone else, HFTs would indeed face nearly unbounded incentives to be faster than their competitors. While reaction speed is certainly one dimension along which HFTs compete, the empirical evidence of exploratory trading suggests that the A-HFTs, at least, can also compete along another dimension—exploration. Since exploratory trading provides the A-HFTs with private information, a trader who uses only public information will not necessarily be able to dominate the A-HFTs, even if that trader is faster than every A-HFT. Similarly, an A-HFT could potentially compensate for having (slightly) slower reactions than the other A-HFTs by engaging in greater levels of exploration.

7.3 Beyond A-HFTs: Other HFTs and Other Markets

Exploratory trading is not universally relevant to all HFT activity in all markets. Equities markets, for instance, may not exhibit the predictability in demand that makes exploratory trading viable in the E-mini market, so HFTs in these markets might primarily concern themselves with obtaining superior forecasts of demand, or they might employ some completely different technique. However, exploratory trading in the E-mini market depends only on the market’s structure and aggregate dynamics; it does not depend directly on any specific features of the E-mini contract. The prevalence of exploratory trading in other markets is ultimately an empirical matter, but markets similar to the E-mini in size and structure could easily support exploratory trading.

Even in the E-mini market, an important component of HFT activity lies outside the immediate province of the exploratory trading model. Nevertheless, the scope for exploratory trading extends well beyond the aggressive activity of A-HFTs considered thus far. Though I have focused on the A-HFTs up to this point, the B-HFTs could also reap exploratory rewards from their aggressive orders, as could potentially any trader with similar capabilities. The B-HFTs’ overall performance on aggressive orders does not present the same ostensible affront to market efficiency as does that of the A-HFTs, but the B-HFTs’ aggressive orders nonetheless outperform both those of non-HFTs, and the baseline econometric benchmark, by a wide margin. If inventory management or risk-control considerations
force B-HFTs to place unprofitable aggressive orders, exploratory trading could help to explain how
the B-HFTs mitigate the associated losses.

Alternatively, if nothing forces the B-HFTs to place aggressive orders, then the B-HFTs’ consistent
losses from aggressive trading are puzzling in their own right, much as the A-HFTs’ losses on small
aggressive orders were. Although the B-HFTs do not recoup their losses on other aggressive orders as
do the A-HFTs, they make enough from their passive trading to earn positive profits overall. Passive
trading strategies, just like aggressive ones, would benefit from the superior price forecasting that
exploratory information makes possible, so exploratory trading could help to explain the activity of
B-HFTs in this scenario as well.\footnote{Total trading profits from any transaction net to zero, so if a trader earns money on an aggressive order, his passive
counter-party loses money. Since exploratory information is valuable to an aggressor, it follows immediately that it is
also valuable to a passar.}

8 Conclusion

Empirical evidence strongly suggests that the concept of exploratory trading developed in this paper
helps to explain the mechanism underlying certain HFTs’ superior capacity to profitably anticipate
price movements in the E-mini market. The exploratory trading model also illuminates the manner
in which these HFTs benefit from low-latency capabilities and from their submission of large numbers
of aggressive orders.

Exploratory trading is a form of costly information acquisition, albeit an unfamiliar one. HFTs who
engage in exploratory trading are doing something more than merely reacting to public information
sooner other market participants. This raises the possibility that HFTs, through exploratory trading,
uniquely contribute to the process of efficient price discovery. However, exploratory trading differs
from traditional costly information acquisition in several important respects. First, the information
that exploratory trading generates does not relate directly to the traded asset’s fundamental value,
but rather pertains to unobservable aspects of market conditions that could eventually become public,
ex-post, through ordinary market interactions. Also, because exploratory trading operates through the
market mechanism itself, exploration exerts direct effects on the market, distinct from the subsequent
effects of the information that it generates.

Finally, since HFTs appear to trade ahead of predictable demand innovations—albeit in a sophisti-
catedly selective manner—the research of De Long et al. (1990) potentially suggests that HFTs could
have a destabilizing influence on prices if suitable positive-feedback mechanisms exist.
Comprehensive analysis of the theoretical and empirical aspects of these myriad issues lies beyond the scope of this paper, but the theory and evidence presented herein provide a starting point from which to rigorously address the market-quality implications of high-frequency trading going forward.
References


A Exploratory Trading Model Details

A.1 Solving the Baseline Exploratory Trading Model

Let $s_t$ denote the sign of $q_t$.

A.1.1 Solving the Model: Period 2

If $\phi = 0$, the HFT’s optimal choice is to not submit an aggressive order in period 2, or equivalently, to set $|q_2| = 0$. If $\phi \neq 0$, then it is optimal for the HFT to set $s_2 = \phi$ (unless the optimal $|q_2|$ is zero), so we only need to determine the optimal magnitude, $|q_2|$. Because $\pi_2$ is linear in $|q_2|$ when $s_2$ is held fixed, we can restrict attention to corner solutions ($0$ or $N$) for the optimal choice of $|q_2|$ without loss of generality. Note that if $q_2 = 0$, then $\pi_2 = 0$, regardless of the values of $\phi$ and $\Lambda$.

Suppose that the HFT sets $|q_2| = N$. Without loss of generality, assume that $s_2 = \phi \neq 0$. The HFT’s period-2 profits are given by

$$
\tilde{\pi}_2 = \begin{cases} 
N (1 - \alpha) & \text{if } \Lambda = U \\
-N\alpha & \text{if } \Lambda = A
\end{cases}
$$

where the tilde on $\tilde{\pi}_2$ denotes the fact that the HFT’s choice of $q_2$ does not condition on the value of $\Lambda$.

HFT Does Not Know $\Lambda$ If the HFT does not know the value of $\Lambda$, then in the case where $\phi \neq 0$, the HFT’s expected period-2 profit if he sets $|q_2| = N$ is

$$
E [\tilde{\pi}_2 | \phi \neq 0, |q_2| = N] = uN (1 - \alpha) - (1 - u) N\alpha
$$

$$
= (u - \alpha) N
$$

Taking expectations with respect to $\phi$, we find that the ex-ante expectation of $\tilde{\pi}_2$ when the HFT sets $|q_2| = N$ (and $s_2 = \phi$) is given by

$$
E [\tilde{\pi}_2 | q_2 = N] = v (u - \alpha) N
$$
When \( u - \alpha < 0 \), if the HFT did not know \( \Lambda \), he would set \( q_2 = 0 \) rather than \( |q_2| = N \). Hence the ex-ante expectation of \( \tilde{\pi}_2 \) is

\[
E[\tilde{\pi}_2] = \max\{v(u - \alpha) N, 0\}
\]

**HFT Knows \( \Lambda \)** Next, if the HFT *does* know the value of \( \Lambda \), then he will set \(|q_2| = N\) (and \( s_2 = \varphi \)) only when \( \Lambda = U \) and \( \varphi \neq 0 \). Denoting the HFT’s period-2 profits from this strategy by \( \hat{\pi}_2 \), we find

\[
\begin{align*}
E[\hat{\pi}_2 | \varphi \neq 0] &= u(1 - \alpha) N \\
&= (u - \alpha) N + \alpha (1 - u) N \\
E[\hat{\pi}_2] &= v u (1 - \alpha) N \\
&= v(u - \alpha) N + v\alpha (1 - u) N
\end{align*}
\]

Note that

\[
E[\hat{\pi}_2] > \max\{v(u - \alpha) N, 0\}
\]

so the HFT’s expected period-2 profits are strictly greater when he knows \( \Lambda \) than when he doesn’t know \( \Lambda \).

**A.1.2 Solving the Model: Period 1**

At the start of period 1, the HFT knows neither \( \varphi \) nor \( \Lambda \), but he faces the same trading costs, \( \alpha \), as he does in period 2. Consequently, the HFT’s expected direct trading profits from a period-1 aggressive order are negative:

\[
\begin{align*}
E[\pi_1 | q_1] &= E[|q_1| (s_1 y - \alpha) | s_1, q_1] \\
&= |q_1| s_1 E[y] - \alpha |q_1| \\
&= -\alpha |q_1|
\end{align*}
\]

The second equality relies on the assumptions that \( \varphi \) and \( \Lambda \) (and hence \( y \)) are independent of \( s_1 \) and \( q_1 \), while the final equality uses the fact that \( E[y] = 0 \).

Since there is no noise in this baseline model, the HFT learns \( \Lambda \) perfectly from any aggressive order that he places in the first period with \(|q_1| \geq 1\). An aggressive order of size greater than one yields no more information about \( \Lambda \) than a one-contract aggressive order in this setting, but the larger
aggressive order incurs additional expected losses. Thus without loss of generality, we can restrict
attention to the case of \( q_1 = 0 \) and the case of \(|q_1| = 1\).

If the HFT sets \( q_1 = 0 \), he neither learns \( \Lambda \) nor incurs any direct losses in period 1, so his total
expected profits are simply

\[
E[\pi_{total}|q_1 = 0] = E[\tilde{\pi}_2] = \max\{v(u - \alpha)N, 0\}
\]

Alternatively, if the HFT sets \(|q_1| = 1\), his total expected profits are given by

\[
E[\pi_{total}| |q_1| = 1] = -\alpha |q_1| + E[\tilde{\pi}_2] = vu(1 - \alpha)N - \alpha
\]

A.1.3 Remark on the Sequence of Events

The central results of the model would not change if the HFT observed the signal of future aggressive
order-flow before deciding whether to engage in exploratory trading, rather than observing it after
deciding. However, the sequence of events outlined in section 2, in which the HFT must choose whether
or not to explore before he observes \( \varphi \), is more appropriate from an empirical perspective. For the
HFT to learn about the liquidity state after he submits an aggressive order, he must wait for 1) his
order to reach the market and execute, 2) information about that execution to reach other traders, 3)
other traders to decide what to do, 4) other traders’ decisions to reach the market, and 5) information
about the market response to get back to him. Of these five steps, (1), (2), (4) and (5) each take
approximately 3–4 milliseconds, and (3) takes considerably longer, perhaps 3–20 milliseconds, for an
overall total of 15–40 milliseconds. An HFT who has already done his exploration will be able to take
advantage of predictable aggressive order-flow long before an HFT who only engages in exploratory
trading after seeing an order-flow signal.
A.2 Solving the Model of Section 2.3

If the HFT places an order in the first period, it follows immediately from the baseline model results that his expected total profits are given by

\[
E[\pi_{total}|q_1 = 1] = Nvu(1 - \alpha) - \alpha
\]

However, the HFT’s expected profits if he does not place an order in period 1 are higher than in the baseline model, because the HFT now learns something from the depth changes following the other trader’s aggressive order. If resting depth weakly replenishes after that order, the HFT learns with certainty that the liquidity state is accommodating (i.e., \(\Lambda = A\)), so the HFT will not submit an aggressive order in period 2, and his total profits will be zero. Alternatively, if resting depth further depletes following the aggressive order in period 1 (denote this event by \(g_1\)), we have

\[
\begin{align*}
\mathbb{P}\{\Lambda = U | g_1\} &= \frac{\mathbb{P}\{\Lambda = U, \text{ and } g_1\}}{\mathbb{P}\{g_1\}} \\
&= \frac{\mathbb{P}\{g_1 | \Lambda = U\} \mathbb{P}\{\Lambda = U\}}{\mathbb{P}\{g_1 | \Lambda = U\} \mathbb{P}\{\Lambda = U\} + \mathbb{P}\{g_1 | \Lambda = A\} \mathbb{P}\{\Lambda = A\}} \\
&= \frac{1 * \mathbb{P}\{\Lambda = U\}}{1 * \mathbb{P}\{\Lambda = U\} + \rho * \mathbb{P}\{\Lambda = A\}} \\
&= \frac{u}{u + \rho (1 - u)}
\end{align*}
\]

It follows immediately from the analogous result in the baseline model that the HFT’s expected period-2 profits are given by

\[
E[\pi_2| AO \text{ by someone else}] = \max\left\{ Nv \left( \frac{u}{u + \rho (1 - u)} - \alpha \right), 0 \right\}
\]

Since the HFT does not place an order in the first period, his expected total profits equal his expected period-2 profits. Overall, then, the HFT’s expected total profit if he observes an aggressive order placed by someone else in period 1 but does not place an aggressive order himself, is given by

\[
E[\pi_{total}| AO \text{ by someone else}] = \max\left\{ Nv \left( \frac{u}{u + \rho (1 - u)} - \alpha \right), 0 \right\}
\]
B Measuring Aggressive Order Profitability

Calculating round-trip profits using a FIFO or LIFO approach is not a useful way to measure the profitability of individual aggressive orders. Even the most aggressive HFTs engage in some passive trading, so a FIFO/LIFO-round-trip measure would either confound aggressive trades with passive trades, or require some arbitrary assumption to distinguish between inventory acquired passively and inventory acquired aggressively (on top of the already-arbitrary assumption of FIFO or LIFO). A second, more general problem is that a measurement scheme based on inventory round-trips will always combine at least two orders (an entry and an exit), so such measurement schemes do not actually measure the profitability of individual aggressive orders.

In this appendix, I provide rigorous justification for the claim that the average expected profit from an aggressive order in the E-mini market equals the expected favorable price movement, minus trading/clearing fees and half the bid-ask spread. After presenting the formal proof, I discuss details of empirically estimating expected favorable price movement.

B.1 Preliminaries

Trading/clearing fees apply equally to both passively and aggressively traded E-mini contracts, so to simplify the exposition, I will initially ignore these fees. Similarly, I make the simplifying assumption that the bid-ask spread is constant, and identically equal to one tick; for the E-mini market, this assumption entails minimal loss of generality.

In the E-mini market, the profitability of individual aggressive orders can be considered in isolation from passive orders. Because E-mini contracts can be created directly by buyers and sellers, a trader’s net inventory position does not constrain his ability to participate in a given trade\textsuperscript{12}. As long as he can find a buyer, a trader who wishes to sell an E-mini contract can always do so, regardless of whether he has a preexisting long position. More generally, if a trader enters a position aggressively then exits it passively, he could have conducted the passive transaction even if he hadn’t engaged in the preceding aggressive transaction. While a desire to dispose of passively-acquired inventory might motivate a trader to submit an aggressive order, the question of underlying motivation is distinct from the question of whether the aggressive order was directly profitable.

\textsuperscript{12}The one exception would arise in the extremely rare event that a trader who did not qualify for a position-limit exemption held so many contracts (either long or short) that his inventory after the trade would exceed the position limit of 100,000 E-mini contracts. For HFTs, this minor exception can safely be ignored.
B.2 Formal Argument

With these preliminaries established, I turn to the rigorous argument. Consider a trader who executes $J$ aggressive sell orders of size one, and $J$ aggressive buy orders of size one, for some large $J$. Following the remarks above, the trader’s passive transactions can be ignored. Let the average direction-normalized price change after these aggressive orders be $\tilde{\vartheta} \equiv \vartheta \left( \frac{2J}{2J-1} \right)$ ticks for some $\vartheta$ that does not depend on $J$.

First, suppose that the trader always submits an aggressive sell after an aggressive buy, and always submits an aggressive buy after an aggressive sell. Without loss of generality, assume that the trader’s first aggressive order is a buy. The trader’s combined profit from all $2J$ aggressive orders is

\[
\pi_{2J} = -a_1 + b_2 - a_3 + b_4 - \ldots - a_{2J-1} + b_{2J}
\]

\[
= -a_1 + (a_2 - 1) - a_3 + (a_4 - 1) - \ldots - a_{2J-1} + (a_{2J} - 1)
\]

\[
= -a_1 + a_2 - a_3 + a_4 - \ldots - a_{2J-1} + a_{2J} - J
\]

\[
= -a_1 + (a_1 + \zeta_{b,1}) - (a_2 + \zeta_{s,2}) + (a_3 + \zeta_{b,2}) - \ldots
\]

\[
\ldots - (a_{2J-2} + \zeta_{s,J}) + (a_{2J-1} + \zeta_{b,J}) - J
\]

\[
= \sum_{i=1}^{J} (a_{2i-1} + \zeta_{b,i}) - \sum_{j=2}^{J} (a_{2j-2} + \zeta_{s,j}) - a_1 - J
\]

\[
= \sum_{i=1}^{J} a_{2i-1} - \left( a_1 + \sum_{j=1}^{J-1} a_{2j} \right) + \sum_{i=1}^{J} \zeta_{b,i} - \sum_{j=2}^{J} \zeta_{s,j} - J
\]

where $a_k$ and $b_k$ respectively denote the prevailing best ask and best bid at the time the $k$th aggressive order executes, $\zeta_{b,r}$ denotes the change in midpoint price following the $r$th aggressive buy order, and $\zeta_{s,r}$ denotes the change in midpoint price following the $r$th aggressive sell order. Note that

\[
\hat{\vartheta} \equiv \frac{1}{2J} \left( \sum_{r=1}^{J} \zeta_{b,r} + \sum_{r=1}^{J} (-\zeta_{s,r}) \right).
\]

Next, taking expectations, we find
\[ \mathbb{E}[\pi_{2J}] = \sum_{i=1}^{J} \mathbb{E}[a_{2i-1}] - \left( \mathbb{E}[a_1] + \sum_{j=1}^{J-1} \mathbb{E}[a_{2j}] \right) + \sum_{i=1}^{J} \mathbb{E}[\zeta_{b,i}] - \sum_{j=2}^{J} \mathbb{E}[\zeta_{s,j}] - J \]

\[ = J\mathbb{E}[a_1] - \mathbb{E}[a_1] - (J - 1)\mathbb{E}[a_1] + J\mathbb{E}[^\vartheta] - (J - 1) \left( -\mathbb{E}[^\vartheta] \right) - J \]

\[ = (2J - 1) \mathbb{E}[^\vartheta] - J \]

\[ = (2J - 1) \left( \mathbb{E}[^\vartheta] - \frac{2J}{2J - 1} \right) - J \]

\[ = J \left( 2\mathbb{E}[^\vartheta] - 1 \right) \]

where the second equality uses the assumption that midpoint prices follow a martingale with respect to their natural filtration, together with the assumption of a constant bid-ask spread. From the final equality above, it follows immediately that the trader’s average expected profit on an individual aggressive order is given by

\[ \frac{1}{2J} \mathbb{E}[\pi_{2J}] = \mathbb{E}[^\vartheta] - \frac{1}{2} \]

Finally, note that none of the calculations above relied on the assumption that the aggressive orders alternated between buys and sells (this only simplified the notation). It follows immediately from grouping together multiple aggressive orders of the same sign that the result would hold for orders of varying sizes, provided that the overall aggressive buy and aggressive sell volumes were equal.

Under the usual regularity conditions, as \( J \to \infty \), \( ^\vartheta \to A.S. \lim_{J \to \infty} \mathbb{E}[^\vartheta] = \mathbb{E}[^\vartheta]. \)

\[ \Box \]

**B.2.1 Independent Importance of the Result**

Establishing a meaningful technique to estimate the performance of individual aggressive orders was merely a necessary stepping stone for a detailed analysis of HFTs’ performance on aggressive orders, but to the best of my knowledge, I am the first to propose and rigorously justify this technique. Anecdotal evidence from the CFTC suggests that the technique I develop in this paper may have broad applications for academics, regulators and practitioners alike.
B.3 Obtaining Unbiased Estimates

Recall that the discussion in section 4.1 implied that we can estimate the profitability of an HFT’s aggressive order using the (direction-normalized) accumulated price-changes following that aggressive order out to some time past the HFT’s maximum forecasting horizon. If we choose too short an accumulation window, the resulting estimates of the long-run direction-normalized average price changes following an HFT’s aggressive orders will be biased downward. This enables us to empirically determine an adequate accumulation period by calculating cumulative direction-normalized price changes over longer and longer windows until their mean ceases to significantly increase.

Market activity varies considerably in its intensity throughout a trading day, so event-time, which I measure in terms of aggressive order arrivals, provides a more uniform standard for temporal measurements than does clock-time. Empirically, an accumulation period of about 30 aggressive orders suffices to obtain unbiased estimates of the price movement following an HFT’s aggressive order, but I consider results for an accumulation period of 50 aggressive orders to allow a wide margin for error. The mean direction-normalized price changes following individual HFTs’ aggressive orders does not differ significantly for accumulation periods of 50, 200, or 500 aggressive orders, even if we distinguish aggressive orders by size. The same holds true for aggressive orders placed by non-HFTs. Using too long an accumulation period will not bias the estimates, but it will introduce unnecessary noise, so I opt for an accumulation period of 50 aggressive orders.

As I discuss at greater length in section 4.3, future price movements are moderately predictable from past aggressive order flow and orderbook activity, but only at very short horizons. Of the variables that meaningfully forecast future price changes, the direction of aggressive order flow is by far the most persistent, but even its forecasting power diminishes to nonexistence for price movements more than either about 12 aggressive orders or 200 milliseconds in the future. The adequacy of a 30+ aggressive order accumulation period is entirely consistent with these results.

As a simple empirical check on the validity of direction-normalized cumulative price changes as a proxy for the profitability of aggressive orders, I use each HFT’s explicit overall profits and passive trading volume, together with the profits on aggressive orders as measured by the proxy, to back out the HFT’s implicit profit on each passively traded contract. The resulting estimates of HFTs’ respective profits from passive transactions are all plausible from a theoretical perspective, and are comparable to non-HFTs’ implicit performance on passive trades.
C Benchmark Regression Results

In this appendix, I present and discuss results from regression (8) of section 4.3.

Recall that for each trading day in my sample, I regress the cumulative price-change (in dollars) between the aggressive orders \(k\) and \(k + 50\), denoted \(y_k\), on the following variables: changes in resting depth between aggressive orders \(k - 1\) and \(k\) at each of the six price levels within two ticks of the best bid or best ask, the signs of aggressive orders \(k - 1\) through \(k - 4\), and the signed executed quantities of aggressive orders \(k - 1\) through \(k - 4\). For symmetry, I adopt the convention that sell depth is negative, and buy depth is positive, so that an increase in buy depth has the same sign as a decrease in sell depth. I estimate the equation

\[
y_k = z_{k-1}\Gamma + \epsilon_k
\]

\[
:= \gamma_1 d_1^{k-1} + \ldots + \gamma_6 d_6^{k-1} + \gamma_7 \text{sign}_{k-1} + \ldots + \gamma_{10} \text{sign}_{k-4} + \gamma_{11} q_{k-1} + \ldots + \gamma_{14} q_{k-4} + \epsilon_k
\]

where \(d_r^{k-1}\) denotes the change in resting depth at price level \(r\) (\(r = 3\) corresponds to the best bid, \(r = 4\) corresponds to the best ask), \(\text{sign}_l\) denotes the sign of aggressive order \(l\), and \(q_l\) denotes the signed executed quantity of aggressive order \(l\).

Table C.1, below, summarizes the estimates from the regression above, computed over my entire sample. All of the variables are antisymmetrical for buys and sells, and so have means extremely close to zero, but the mean magnitudes in the rightmost column of Table C.1 provide some context for scale.
Comparable results obtain using as few as two lags of aggressive order sign and signed quantity. Linear forecasts of $y_k$ do not benefit appreciably from the inclusion of data on aggressive orders before $k - 4$, or on changes in resting depth prior to aggressive order $k - 1$. Because the price-change $y_k$ is not normalized by the sign of the $k$th aggressive order, it has an expected value of zero, so I do not include a constant term in the regression. Including a constant term in the regression has negligible effect on the results.

Although the last several aggressive order signs do offer rather remarkable explanatory power, the respective distributions of resting depth changes and executed aggressive order quantities have much heavier tails than the distribution of order sign, so price forecasts are substantially improved by the inclusion of these variables.

The positive coefficients on the lagged aggressive order variables and on the depth changes at the best bid and best ask are consistent with the general intuition that buy orders portend price increases,
and sell orders portend price decreases. The negative coefficients on depth changes at the outside price levels require slightly more explanation.

Because the E-mini market operates according to strict price and time priority, a trader who seeks priority execution of his passive order will generally place that order at the best bid (or best ask); however, if the trader believes that an adverse price movement is imminent, he will place his order at the price level that he expects to be the best bid (ask) following the price change. It is relatively uncommon for prices to change immediately after an aggressive order in the E-mini market, but when prices do change, it is extremely rare during regular trading hours for the change to exceed one tick. As a result, the expected best bid (ask) following a price change is typically one tick away from the previous best, so it is not surprising that (e.g.) an increase in resting depth one tick below the best bid tends to precede a downward price change. These features of the E-mini market also shed some light on why changes in depth more than one tick away from the best (i.e., $d_{best \ bid}^{k-1}$ and $d_{best \ ask}^{k+1}$) are not significant predictors of future price movements.\footnote{Similarly, this line of reasoning helps to explain why changes in resting depth prior to aggressive order $k – 1$ do not help to forecast price changes after the $k$th aggressive order.}
D Supplemental Tables of Empirical Results

D.1 Performance of A-HFTs’ Non-Rebalancing Aggressive Orders

Table D.1, below, summarizes the performance of aggressive orders that move the submitting account away from a neutral inventory position. A net position of zero is the most likely “neutral” inventory target, but I allow for the possibility that a given HFT has an arbitrary constant target inventory position, and I restrict attention to aggressive orders that move the HFT away from that target inventory level.

Table D.1. Performance of A-HFTs’ Non-Rebalancing Aggressive Orders

<table>
<thead>
<tr>
<th>Cutoff</th>
<th>Below Cutoff 95% CI</th>
<th>Above Cutoff 95% CI</th>
<th>AOs Below Cutoff % of All AOs</th>
<th>AOs Below Cutoff % of Aggr. Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(3.24, 3.42)</td>
<td>(6.90, 7.12)</td>
<td>15.00%</td>
<td>0.25%</td>
</tr>
<tr>
<td>5</td>
<td>(3.64, 3.79)</td>
<td>(6.95, 7.17)</td>
<td>27.25%</td>
<td>0.89%</td>
</tr>
<tr>
<td>10</td>
<td>(2.48, 2.63)</td>
<td>(7.13, 7.35)</td>
<td>34.06%</td>
<td>1.94%</td>
</tr>
<tr>
<td>15</td>
<td>(2.79, 2.96)</td>
<td>(7.13, 7.35)</td>
<td>35.01%</td>
<td>2.14%</td>
</tr>
<tr>
<td>20</td>
<td>(3.01, 3.19)</td>
<td>(7.21, 7.44)</td>
<td>37.16%</td>
<td>2.81%</td>
</tr>
</tbody>
</table>
D.2 Additional Performance Explained by Exploratory Information

Table D.2, below, presents cross-sectional averages of the estimated additional gross earnings per contract on aggressive orders of submitted size greater than $\bar{q}$ explained by regression (10) in excess of that explained by regression (9). The extra explanatory power of (10) reflects the contribution from the private component of information (available to the A-HFT under consideration) manifested in $\Omega^A$. Numbers reported for the A-HFTs are averages over the estimates for individual A-HFTs. The membership of the “all others” depends upon the particular A-HFT being excluded, and the numbers reported for “All Others” are averages over these slightly different groups. Units are cents per contract, and confidence intervals are constructed by bootstrap.

<table>
<thead>
<tr>
<th>$\bar{q}$</th>
<th>Point Estimate</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$= 1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-HFTs</td>
<td>0.179</td>
<td>(0.069, 0.296)</td>
</tr>
<tr>
<td>All Others</td>
<td>0.034</td>
<td>(0.018, 0.048)</td>
</tr>
<tr>
<td>$= 5$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-HFTs</td>
<td>0.659</td>
<td>(0.480, 0.850)</td>
</tr>
<tr>
<td>All Others</td>
<td>0.082</td>
<td>(0.055, 0.109)</td>
</tr>
<tr>
<td>$= 10$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-HFTs</td>
<td>0.397</td>
<td>(0.221, 0.570)</td>
</tr>
<tr>
<td>All Others</td>
<td>0.074</td>
<td>(0.041, 0.104)</td>
</tr>
<tr>
<td>$= 15$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-HFTs</td>
<td>0.533</td>
<td>(0.355, 0.705)</td>
</tr>
<tr>
<td>All Others</td>
<td>0.101</td>
<td>(0.071, 0.133)</td>
</tr>
<tr>
<td>$= 20$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A-HFTs</td>
<td>0.623</td>
<td>(0.453, 0.799)</td>
</tr>
<tr>
<td>All Others</td>
<td>0.113</td>
<td>(0.077, 0.147)</td>
</tr>
</tbody>
</table>
D.3 Differences in Additional Performance Explained by Exploratory Information

Table D.3, below, presents cross-sectional averages of the difference in mean additional gross earnings per contract on aggressive orders of submitted size greater than $\bar{q}$ explained by regression (10) in excess of that explained by regression (9), between the indicated groups. Units are cents per contract, and confidence intervals are constructed by bootstrap.

Table D.3. Additional Explained Performance of A-HFTs vs. Everyone Else and vs. Other HFTs (Cents per Contract)

<table>
<thead>
<tr>
<th>$\bar{q}$</th>
<th>Point Estimate</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A-HFTs vs. Everyone Else</td>
<td>0.145</td>
</tr>
<tr>
<td></td>
<td>A-HFTs vs. Other HFTs</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>$\bar{q} = 5$</td>
<td>A-HFTs vs. Everyone Else</td>
</tr>
<tr>
<td></td>
<td>A-HFTs vs. Other HFTs</td>
<td>0.458</td>
</tr>
<tr>
<td></td>
<td>$\bar{q} = 10$</td>
<td>A-HFTs vs. Everyone Else</td>
</tr>
<tr>
<td></td>
<td>A-HFTs vs. Other HFTs</td>
<td>0.235</td>
</tr>
<tr>
<td></td>
<td>$\bar{q} = 15$</td>
<td>A-HFTs vs. Everyone Else</td>
</tr>
<tr>
<td></td>
<td>A-HFTs vs. Other HFTs</td>
<td>0.323</td>
</tr>
<tr>
<td></td>
<td>$\bar{q} = 20$</td>
<td>A-HFTs vs. Everyone Else</td>
</tr>
<tr>
<td></td>
<td>A-HFTs vs. Other HFTs</td>
<td>0.426</td>
</tr>
</tbody>
</table>
D.4 Results for Extended Values of $\bar{q}$

Table D.4, below, presents results for extended values of $\bar{q}$. Each large row corresponds to a single value of $\bar{q}$, and within a given row, the topmost two sub-rows present averages of the estimated additional gross earnings per contract on aggressive orders of submitted size greater than $\bar{q}$ explained by regression (10) in excess of that explained by regression (9). Numbers reported for the A-HFTs are averages over the estimates for individual A-HFTs. The membership of the “all others” depends upon the particular A-HFT being excluded, and the numbers reported for “All Others” are averages over these slightly different groups. The bottommost two sub-rows for each value of $\bar{q}$ present cross-sectional averages of the difference in mean additional gross earnings per contract on aggressive orders of submitted size greater than $\bar{q}$ explained by regression (10) in excess of that explained by regression (9), between the indicated groups. Units are cents per contract, and confidence intervals are constructed by bootstrap. The abbreviation “EE” stands for “Everyone Else.”
Table D.4. Additional Explained Performance for Extended Values of $\bar{q}$

<table>
<thead>
<tr>
<th>$\bar{q}$</th>
<th>Point Est.</th>
<th>95% CI</th>
<th>A-HFTs vs.</th>
<th>Point Est.</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>A-HFTs</td>
<td>0.624</td>
<td>(0.437, 0.806)</td>
<td>EE</td>
<td>0.450</td>
</tr>
<tr>
<td></td>
<td>EE</td>
<td>0.175</td>
<td>(0.135, 0.213)</td>
<td>Other HFTs</td>
<td>0.360</td>
</tr>
<tr>
<td>30</td>
<td>A-HFTs</td>
<td>0.697</td>
<td>(0.520, 0.894)</td>
<td>EE</td>
<td>0.551</td>
</tr>
<tr>
<td></td>
<td>EE</td>
<td>0.147</td>
<td>(0.106, 0.182)</td>
<td>Other HFTs</td>
<td>0.454</td>
</tr>
<tr>
<td>35</td>
<td>A-HFTs</td>
<td>0.733</td>
<td>(0.551, 0.920)</td>
<td>EE</td>
<td>0.590</td>
</tr>
<tr>
<td></td>
<td>EE</td>
<td>0.143</td>
<td>(0.106, 0.184)</td>
<td>Other HFTs</td>
<td>0.489</td>
</tr>
<tr>
<td>40</td>
<td>A-HFTs</td>
<td>0.850</td>
<td>(0.646, 1.037)</td>
<td>EE</td>
<td>0.688</td>
</tr>
<tr>
<td></td>
<td>EE</td>
<td>0.162</td>
<td>(0.120, 0.205)</td>
<td>Other HFTs</td>
<td>0.580</td>
</tr>
<tr>
<td>45</td>
<td>A-HFTs</td>
<td>0.850</td>
<td>(0.659, 1.053)</td>
<td>EE</td>
<td>0.688</td>
</tr>
<tr>
<td></td>
<td>EE</td>
<td>0.162</td>
<td>(0.119, 0.205)</td>
<td>Other HFTs</td>
<td>0.560</td>
</tr>
<tr>
<td>50</td>
<td>A-HFTs</td>
<td>1.003</td>
<td>(0.810, 1.219)</td>
<td>EE</td>
<td>0.804</td>
</tr>
<tr>
<td></td>
<td>EE</td>
<td>0.199</td>
<td>(0.145, 0.253)</td>
<td>Other HFTs</td>
<td>0.643</td>
</tr>
<tr>
<td>60</td>
<td>A-HFTs</td>
<td>1.181</td>
<td>(0.985, 1.381)</td>
<td>EE</td>
<td>0.964</td>
</tr>
<tr>
<td></td>
<td>EE</td>
<td>0.218</td>
<td>(0.162, 0.272)</td>
<td>Other HFTs</td>
<td>0.786</td>
</tr>
<tr>
<td>75</td>
<td>A-HFTs</td>
<td>1.073</td>
<td>(0.860, 1.293)</td>
<td>EE</td>
<td>0.902</td>
</tr>
<tr>
<td></td>
<td>EE</td>
<td>0.171</td>
<td>(0.112, 0.230)</td>
<td>Other HFTs</td>
<td>0.757</td>
</tr>
<tr>
<td>90</td>
<td>A-HFTs</td>
<td>1.040</td>
<td>(0.821, 1.242)</td>
<td>EE</td>
<td>0.920</td>
</tr>
<tr>
<td></td>
<td>EE</td>
<td>0.120</td>
<td>(0.053, 0.187)</td>
<td>Other HFTs</td>
<td>0.746</td>
</tr>
</tbody>
</table>
Table D.5. Size-Weighted Average Performance of Aggressive Orders Below and Between Size Thresholds (Dollars per Contract)

<table>
<thead>
<tr>
<th>$q$</th>
<th>A-HFTs</th>
<th>B-HFTs</th>
<th>Non-HFTs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All AOs ≤ $q$</td>
<td>Incremental AOs</td>
<td>All AOs ≤ $q$</td>
</tr>
<tr>
<td>1</td>
<td>3.84</td>
<td>-</td>
<td>4.37</td>
</tr>
<tr>
<td>5</td>
<td>4.23</td>
<td>4.38</td>
<td>4.66</td>
</tr>
<tr>
<td>10</td>
<td>3.49</td>
<td>2.84</td>
<td>4.66</td>
</tr>
<tr>
<td>15</td>
<td>3.85</td>
<td>6.39</td>
<td>4.71</td>
</tr>
<tr>
<td>20</td>
<td>4.14</td>
<td>4.95</td>
<td>4.77</td>
</tr>
<tr>
<td>25</td>
<td>4.41</td>
<td>6.65</td>
<td>4.81</td>
</tr>
<tr>
<td>30</td>
<td>4.79</td>
<td>6.79</td>
<td>4.87</td>
</tr>
<tr>
<td>35</td>
<td>4.99</td>
<td>7.00</td>
<td>4.88</td>
</tr>
<tr>
<td>40</td>
<td>5.28</td>
<td>6.69</td>
<td>4.91</td>
</tr>
<tr>
<td>45</td>
<td>5.42</td>
<td>7.04</td>
<td>4.92</td>
</tr>
<tr>
<td>50</td>
<td>5.61</td>
<td>6.90</td>
<td>4.95</td>
</tr>
<tr>
<td>60</td>
<td>5.87</td>
<td>7.00</td>
<td>4.98</td>
</tr>
<tr>
<td>75</td>
<td>6.12</td>
<td>7.20</td>
<td>5.00</td>
</tr>
<tr>
<td>90</td>
<td>6.38</td>
<td>7.20</td>
<td>5.01</td>
</tr>
</tbody>
</table>
Table D.6. Aggressive Orders Below Size Thresholds: Fractions of All Aggressive Orders and Aggressive Volume

<table>
<thead>
<tr>
<th>$\bar{q}$</th>
<th>A-HFTs</th>
<th></th>
<th>B-HFTs</th>
<th></th>
<th>Non-HFTs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% of All AOs</td>
<td>% of Aggr. Volume</td>
<td>% of All AOs</td>
<td>% of Aggr. Volume</td>
<td>% of All AOs</td>
</tr>
<tr>
<td>1</td>
<td>24.31%</td>
<td>0.40%</td>
<td>39.48%</td>
<td>4.77%</td>
<td>53.79%</td>
</tr>
<tr>
<td>5</td>
<td>43.74%</td>
<td>1.44%</td>
<td>76.09%</td>
<td>16.13%</td>
<td>83.26%</td>
</tr>
<tr>
<td>10</td>
<td>54.64%</td>
<td>3.09%</td>
<td>84.10%</td>
<td>23.95%</td>
<td>89.87%</td>
</tr>
<tr>
<td>15</td>
<td>56.75%</td>
<td>3.54%</td>
<td>88.01%</td>
<td>30.03%</td>
<td>91.57%</td>
</tr>
<tr>
<td>20</td>
<td>60.82%</td>
<td>4.80%</td>
<td>90.70%</td>
<td>35.82%</td>
<td>93.27%</td>
</tr>
<tr>
<td>25</td>
<td>62.38%</td>
<td>5.37%</td>
<td>92.36%</td>
<td>40.29%</td>
<td>94.41%</td>
</tr>
<tr>
<td>30</td>
<td>64.62%</td>
<td>6.39%</td>
<td>94.14%</td>
<td>46.30%</td>
<td>94.98%</td>
</tr>
<tr>
<td>35</td>
<td>65.82%</td>
<td>7.02%</td>
<td>94.97%</td>
<td>49.56%</td>
<td>95.23%</td>
</tr>
<tr>
<td>40</td>
<td>68.27%</td>
<td>8.51%</td>
<td>96.03%</td>
<td>54.31%</td>
<td>95.66%</td>
</tr>
<tr>
<td>45</td>
<td>69.29%</td>
<td>9.20%</td>
<td>96.32%</td>
<td>55.76%</td>
<td>95.80%</td>
</tr>
<tr>
<td>50</td>
<td>71.07%</td>
<td>10.55%</td>
<td>97.66%</td>
<td>63.29%</td>
<td>97.09%</td>
</tr>
<tr>
<td>60</td>
<td>73.81%</td>
<td>12.97%</td>
<td>98.54%</td>
<td>69.12%</td>
<td>97.36%</td>
</tr>
<tr>
<td>75</td>
<td>76.65%</td>
<td>16.01%</td>
<td>99.02%</td>
<td>73.14%</td>
<td>97.64%</td>
</tr>
<tr>
<td>90</td>
<td>80.68%</td>
<td>21.11%</td>
<td>99.20%</td>
<td>74.86%</td>
<td>97.83%</td>
</tr>
</tbody>
</table>
Competition of High-Frequency Market Makers and
Market Quality *

March 25, 2013

Abstract
High-frequency trading has attracted controversial discussions by legislators, regulators and investors alike leading to calls for legislative and regulative intervention. The first entries of big international high-frequency traders to the Swedish equity market in 2009 using NASDAQ OMXS tickdata, offers a unique instance to empirically examine how competition affects market qualities. Competition among high-frequency traders coincides (i) with an increase of intraday volatility of about 25%, but (ii) interestingly with no effect on interday volatility, (iii) with a decrease in order-execution time (difference between an incoming market order or marketable limit order and the standing limit order that the trade is executed against) by about 15%, and (iv) with an increase in market share of high-frequency traders, but (v) with no effect on overall volume. We provide results for both entries and exits, and offer several potential explanations for this first empirical evidence on competition.

Keywords: competition, high-frequency trading, market maker, duopoly, monopoly

JEL Classification: G12, G14, G15, G18, G23, D4, D61

*I am grateful for valuable continuous support by Pete Kyle, Michael Halling and Per Stromberg. I thank seminar participants at the SHOF Finance Workshop. I also thank Anders Anderson, Patrick Augustin, Laurent Bach, Marieke Bos, Hamid Boustanifar, Montasser Ghachem, Jungsuk Han, Jan Schnitzler and Ulf von Lilienfeld-Toal for respected suggestions. I am indebted to NASDAQ OMX for providing the data as well as Petter Dahlstrom and Henrik Talborn for fruitful discussions.
1 INTRODUCTION

High-frequency traders (HFTs)\(^1\) are market participants that are distinguished by the high speed with which they can react on incoming news, the low inventory on their books and the large number of trades they execute.\(^2\) The sheer size of about 50% to 85% of today’s daily volume in equity markets manifests its importance for academic research and public discussion in particular with the rise of calls for legislative and regulatory intervention.\(^3,4\) While empirical research has focused on important concerns such as liquidity, price discovery or volatility effects of HFT, a clear separation of HFTs that allows studying the impact of high-frequency competition has not been possible due to data limitations. This key concern of potential effects of competition between HFTs has neither been approached empirically nor compared to existing empirical merits. Competition, however, potentially causes or influences HFT effects on markets considering for instance that HFTs compete for the same trades. Does competition ultimately improve market quality and dynamics and therefore benefit investors who use and rely upon financial markets? A comprehensive understanding of HFT competition is therefore quite relevant for the efficient functioning of financial markets and proper potential regulatory action.

In this paper, we examine the effect of competition between HFTs to shed some light on the impact on market quality using trade ticker-level NASDAQ OMXS data. The first entries of big international HFTs to the Swedish stock market in 2009 offers a unique instance of changing intertemporal competition as HFT competition varies among stocks and time. In particular, we conduct a difference-in-difference study (see section 2) to exploit the changes between monopolistic and duopolistic HFT within individual stocks of NASDAQ OMXS 30 composed of Sweden’s thirty largest companies. All HFTs are big international well established banks or hedge funds that are

---

\(^1\)Henceforth, the shortcut HFT will be used for ”high-frequency trading” and ”high-frequency trader”, whereas HFTs will stand for ”high-frequency traders”. 

\(^2\)The SEC (2010) report defines HFTs as market participants that end the day with close to zero inventories, frequently submit and cancel limit orders, use colocation facilities and highly efficient algorithms, and have short holding periods.

\(^3\)Through highly competitive and quick market platforms, the advantage of technology such as co-location or the use of ultra-quick algorithms, HFTs changed and influenced financial markets substantially (Jain (2005)). The TABB Group, a leading financial markets research and advisory company finds a HFT share 73%, whereas Brogaard et al. (2012) estimate it to be about 85% (see Table 1).

\(^4\)These controversial views span topics from price manipulation, speed of trading, systemic risk due a high correlation of algorithmic strategies, to price discovery and liquidity. The quality of liquidity HFTs potentially provide is of particular concern as HFTs replaced traditional market makers.
also significant players in the American equity market. We observe entries and exits, measured by actual trades, for each individual stock and trader.\textsuperscript{5} Contrary to previous literature, we can observe trader identities and therefore distinguish between different HFTs.\textsuperscript{6} Our findings suggest unequivocally ambiguous results on market quality. First, intraday hourly volatility increases severely by an average of over 25\% and five-minute volatility by an average of about 12\%, interdaily (both measured from opening to closing and closing to closing prices), however, shows no sign of increase or decrease. These results hold for both entries and exits noting that for the latter one, it shows a decrease in intraday volatility. Second, order-execution time defined as the time difference (in seconds) between an incoming market order or marketable limit order and the standing limit order that the trade is executed against, decreases on average by about 15\%, which is also reflected in a significant reduction of its standard deviation. Surprisingly there is no significant positive effect for exits, only a marginally significant increase in its standard deviation. Finally, even though the HFTs proportion of total volume increases and decreases significantly after entry and exit respectively, there is unexpectedly no effect on total volume.

Granting these findings about competition and market quality, there are several plausible interpretations. First, competition increases intraday volatility since HFTs compete for the same trades. We find that HFTs under competition trade in two-thirds of the cases on the same side of the market (Figure 4). Second, HFTs trade quicker and reduce therefore the time limit orders wait to be executed significantly. Third, there is no effect of competition on overall volume. While HFT volume indeed increases as suggested by theory (Li (2012)), there seems to be a crowding out of other investors such as slow market makers. Fourth, this crowding out evidence can also be seen when HFTs exit a specific stock, because order-execution time does not increase significantly after exit, contrary to a large significant decrease after entry. A potential reason is that these slower traders that were crowded out before do not return immediately or leave the market eventually.\textsuperscript{7}

Furthermore, while there are quite large effects on intraday volatility, there does not seem to be

\textsuperscript{5}Throughout the rest of the paper, when referring to entry or exit, we will use the terminology in the sense outlined above: entry as the change from HFT monopoly to HFT duopoly within a specific stock and exit as the change from HFT duopoly to HFT monopoly within a specific stock.

\textsuperscript{6}See for example Brogaard et al. (2012) or Hasbrouck and Saar (2012), who work with a NASDAQ dataset that flags messages from 26 HFTs and is the most comprehensive HFT database available for researchers in recent years.

\textsuperscript{7}A famous example that exited the market in 2010 as new rules and technology made profitability difficult is LaBranche Specialist, an old specialist on the NYSE.
any on interday volatility. Both for opening to close volatility and close to close volatility there is no effect of competition. This is not entirely surprising, however, considering the zero daily inventories of HFTs and their short investment horizon. Therefore, we can confirm the theory based empirical prediction of increased volatility within a day, but not interdaily.\footnote{See for example Martinez and Rosu (2013) or Li (2012) for theoretical explanations of increased volatility.} Our findings of decreased order-execution time and the increased HFT volume could be related to a crowding out story of slow investors such as traditional market makers. Since HFT market making can respond quicker and follow a potentially more sophisticated strategy, slow traditional traders are not successful in placing their orders. While in the monopoly case, HFTs can be very selective in the trades undertaken and might still give some room for potential slow traders, competition will decrease the chance of being profitable for slow market makers even further. This leads to less trading of these traders in the stock and no immediate return under HFT monopoly, as it is suggested by the non-statistically significant increase in order-execution time after HFTs’ exit.

Interpreting these results as evidence on the causal effect of competition on market dynamics is only valid if competition can be treated as exogenous to the dependent variables examined. Therefore, it is crucial to raise the issue of what drives the cross-stock variation in HFTs market participation. Since HFTs might enter and exit more than once into trading in one specific stock (Table 1) over the sample period, there are several possible reasons. The first could be that HFTs take down their trading algorithms for update purposes, to fix bugs, or to replace unprofitable codes. HFTs generally do their first trade within the first minutes of every trading day, but might also be absent for days or might stop trading entirely. Therefore, the second reason could be that HFTs start or stop participating in a specific stock because of a new trading strategy. Clearly, there are several possible alternatives that competition is not necessarily exogenous. It could be that the variables examined, volatility, market speed or volume, or some omitted variables in a particular stock, drive directly competition. We address this demur in several ways. To rule out time trends and cross-stock differences, we control for day fixed effects and stock fixed effects in all regressions. Controlling for the lags and leads of the variable in question ensures that no increase or decrease prior or past competition is a driving competition. Also, other controls that potentially could trigger competition are subsequently included in the regressions, but have no
statistically significant effect on changes in volatility, market speed or volume triggered through HFT competition. The relative evenly distributed entries and exits over the sample period are also suggesting that there is no particular market reason for HFTs to enter or exit a specific stock. Placebo regression, where entries and exits are set randomly set within the sample period find no effect on any of the variables of interest, thus giving additional comfort about non-spurious effects of competition on market dynamics. Furthermore, the results do not seem to suffer from a selection issue as entries and exits seem to be fairly well distributed among stocks. The two exceptions do not drive or change any of the results. Excluding Scania AB, which accounts for about 10% for all entries and exits, just improves significance. Dropping Nokia Corporation and Lundin Petroleum AB, which serve exclusively as controls, has no statistically significant effect on the results. Another potential concern could be that the dependent variables of interest are significantly different across stocks before the first trade of the day. This, however, is not the case (Table 3). As there are no difference prior the first trade looking at daily measures would only underestimate the effect of competition. A final supporting argument for the validity of our identification is that HFTs cannot observe their opponent’s identity.

There are a number of objections and limitations to our findings, and some curtailing to the conclusions we draw. First, we only consider stock trades that were undertaken at the NASDAQ OMXS, which is the biggest trading platform in Sweden and accounts for about 80% of total trading volume, and might therefore wrongly assume that there are no other active HFTs in the market. Given the advanced access for algorithmic traders this seems to be unlikely, but would, if anything, change the interpretation of our results to increased competition. Second, we do not observe the orderbook and cannot draw any conclusions on other market-microstructure measures, such as latency, cancelation rate, market depth etc. Even though these measures are important, we are interested in the actual realized market effect for which the trade ticker database provides an excellent base. Third, the Swedish stock market is efficient and mature, but by far not as big and liquid as the US market which raises potential concerns about the importance of our findings. However, HFTs in our sample are big international HFTs that have a big market share on the American market. Additionally, stocks listed on the Swedish stock market and OMX30

\[9\] We refer to the appendix section A for detailed information.

\[10\] HFTs list their activities and sometimes market shares on their web pages. Due to a confidentiality agreement
are comparable to liquid US stocks (Comfortably comparable to the lower 50 percentile of the S&P 500.), but not super-liquid stocks where most HFT takes place on the American market. If anything, competition will lead to an underestimation of the effect of competition as HFT activity increases with increased liquidity.

There is a developing theoretical literature that argues ambiguously about good and evil of high-frequency trading. While empirical work commonly provides evidence to support the view of increased market efficiency or show that it is actually not harmful, theoretical work suggest that there might be some other market impacts as well. In today’s markets, HFTs both provide and take liquidity. Theoretical models, however, differ. While Jovanovic and Menkveld (2011) and Gerig and Michayluk (2010) think of HFTs as liquidity providers, Martinez and Rosu (2013) and Foucault et al. (2011) model HFTs as liquidity takers. Cvitanic and Kirilenko (2010) provide loosely related theoretical evidence by showing that markets with HFTs have thinner tails and more mass around the center of the distribution of transaction prices. And Li (2012) shows that HFT competition increases trading aggressiveness, efficiency, market depth and contribute substantially to volume and variance. Empirical evidence is also scarce. Jovanovic and Menkveld (2012) show that HFTs react quicker on new hard information and are therefore less subject to adverse selection. Hendershott et al. (2011) investigate algorithmic trading, a broader classification than HFT using the automation of quotes on the New York Stock Exchange as an exogenous event and find a positive effect on liquidity. Boehmer and Wu (2013) find similar evidence by exploiting co-location services across different countries. Brogaard et al. (2012) and Hendershott and Riordan (2013) find that HFT benefits price discovery and efficiency. There is, however, less consensus about the impact on volatility as Boehmer and Wu (2013) points out. Hasbrouck and Saar (2012) discover an amplified volatility effect due to runs on linked messages in the order-book, whereas Kirilenko et al. (2011) mention that HFTs may have exacerbated the flash crash in May 2010, but did not cause it.\footnote{Kirilenko et al. (2011) is unique in the sense that it makes use of the first adequately identified data made available for researchers by the U.S. Commodity Futures Trading Commission.} Our findings are also closely related to Boehmer and Wu (2013) who document an increased short-term volatility by algorithmic trading. Hirschey (2011) uncovers differences among HFTs studying anticipatory trading (NASDAQ equity data). Baron

\footnote{with NASDAQ OMXS, we unfortunately cannot reveal their names or exact number given that they are less than ten.}
et al. (2012) find that HFTs earn large and stable profits, while Clark-Joseph (2012) examines the profitability of HFTs’ aggressive orders on the same data. Huh (2013) argues that in markets with HFTs being liquidity providers and takers, the ability to use machine-readable public information is crucial for HFTs. An attempt to distinguish between liquidity providers or HFT market makers and liquidity takers or alternatively aggressive HFTs, is made by Hagstrmer and Norden (2012). While Biais and Woolley (2011) provide a comprehensive overview of good and evil of HFT, the crucial question how competition among high-frequency traders market quality has been empirically untouched.

The rest of the paper proceeds as follows. Section 3 describes methodology to exploit cross-sectional variation between stocks. Section 4, describes our NASDAQ OMXS data, before we give a comprehensive overview of our findings in section 5. We conclude in section 6.

2 High-Frequency Traders

In this section we discuss and provide statistics of high-frequency market makers.

2.1 Market Making

Market makers traditionally provide required amounts of liquidity to the security’s market when price pressure or other non-fundamental trading activity moves the market, and bring a short-term buy-and sell-side imbalances back into equilibrium. In return, these market makers are granted various trade execution advantages. The old structure where stock exchanges employ several competing official market makers, which were required to place orders on both side of the market and were obligated to buy and sell at their displayed bids and offers, has changed dramatically in recent years. Through highly competitive and quick market platforms, the advantage of technology such as co-location or the use of ultra-quick algorithms raised new market makers, high-frequency traders making it increasingly difficult for traditional market makers to stay profitable. In 2010 one of the oldest market makers at the NYSE, LaBranche Specialist, exited the market. The new high-frequency traders, however, are not easily categorized or regulated.

This has been in focus of the Securities Exchange Commission (SEC) who views its mandate as acting on behalf of companies trying to raise capital for long-term projects and investors with long horizons. Market makers or day traders have only legitimacy if they contribute to
long-term investors’ interest. In the SEC 2010 Concept Release, high-frequency market makers are categorized into four types with four different strategies: passive, arbitrage, structural and directional market making.

HFTs that entered the Swedish market first are difficult to be strictly categorized. We believe that all have a directional market making component, but differ in aggressiveness.

Figure fig:inventINTRA pictures the average intraday inventory over all stocks and days for five minute bins. Inventory is defined as the cumulative turnover divided by total turnover within each five minute bucket. While the left graph views average inventory over all days and stocks with a monopolistic high-frequency trader, the right hand graph shows inventories under competition for both the incumbent and entrant. Trading takes place from 9am to 5:30pm.

[Insert Figure 3 about here!]

Figure 4 images the average intraday fraction of trades that were executed on the same side of the market (over all stocks and days) for five minute and sixty minute bins under competition. Trades are executed on the same side of the market if within each bin both the entrant and the incumbent buy or sell. The measure is constructed by assigning one if both high-frequency traders trade on the same side of the market and zero if not. The average ratio of trading on the same side of the market as their competitor is $2/3$. The dark shaded bars are hourly averages.

[Insert Figure 4 about here!]

From the previous graphs we can conclude that in the pre-closing period HFTs seem to exclusively trade on the same side of the market. Figure 5 pictures pre-closing trading activities. Pre-closing takes place from 5:20 to 5:30 and determines the closing price by maximizing tradable volume. Timestamps within the closing period reflect the order time and not the actual transaction time. Average turnover per trade, average total stock turnover and average high-frequency trading turnover are printed.

[Insert Figure 5 about here!]
3 METHODOLOGY

We aim to compare measures of market quality such as intraday and interday volatility, volume or liquidity under HFT monopoly and HFT duopoly. The first entries of big international high-frequency traders to the Swedish equity market, offers a unique instance to empirically examine how in fact competition affects market qualities by exploiting cross-sectional differences among stocks. Entries and exits into trading in one specific stock (Table 1) is consistent with the difference-in-difference tests outlined by Bertrand et al. (2004). This approach permits us to interpret our findings as evidence on the causal effect of competition on market dynamics. Having in mind the limitations outlined above, we treat competition as exogenous to the dependent variables examined.

This difference-in-difference test setting allows for multiple time periods and multiple treatment groups, and is summarized in the following equation:

\[ y_{ist} = \beta_1 d_{is} + X_{ist} \Gamma + p_t + m_s + u_{ist}, \]  

with \( i \) indexing entry (the change from HFT monopoly to HFT duopoly or vice versa, or both), \( s \) being the security and \( t \) the time. \( d_{is} \) is an indicator if an HFT entry affected security \( s \) at time \( t \). \( p_t \) are daily time fixed effects and \( m_s \) are security fixed effects. \( X_{ist} \) is the vector of covariates and \( u_{ist} \) is the error term. The dependent variable is \( y_{ist} \).

In all of the above tests, we rely on the use of entries and exits into a specific stock of HFTs as the measure of competition. We refer to entry when there is one additional HFT trading in a specific stock at a specific time (change from monopoly to duopoly) and to exits are when one HFT trading in a specific stock at a specific time stops trading (change from duopoly to monopoly). Note that there can be multiple entries and exits over time of the same HFT within the same stock (HFT fixed effects are included in our analysis.). For this changing intertemporal competition across stocks and time, we provide results for both entries and exits, but also entry and exit together. However, these difference-in-difference estimations with entries and exits summed up to one event (standardized on the entry, i.e. exits were relabeled; one could think of it as reverse entries), does not allow for controls such as lags and leads. If there would be an entry in one stock

\(^{12}\)For a detailed description of our data, see section 4.
and an exit in another around the same date, it would not be clear to which event we should assign the control group to and therefore only create spurious effects.

We have also used an alternative way to measure competition. The Herfindahl index as a measure of competition shows very similar but more significant results (the results will be available in the online appendix).

4 DATA

The ticktrading data comes from NASDAQ OMX Nordic and incorporates all trading information of all trades executed at the Stockholm stock exchange (NASDAQ OMXS). We focus on the OMXS30 index that hosts the thirty biggest public companies in Sweden, because we observe that HFTs trade solely in liquid stocks and restrict their trading activity to Sweden’s major securities. As a second data source for our daily measures such as volatility, we draw from COMPUSTAT GLOBAL. As a final source, we use daily relative time-weighted order-execution spreads provided by NASDAQ.

The key distinction of this database is that it allows identifying proprietary traders that are members at the stock exchange, down to a level showing the channels through which they execute their trades. Big HFTs naturally will execute their trades taking advantage of the cheapest and fastest access; the algorithmic trading accounts. For non-proprietary trading, identities are not precise and might be aggregated. The numbers of identities I observe for these traders should therefore be understood as the minimum number of traders. There are about 500 algorithmic trader identities, but the actual number is assumed to be much larger. While the big HFTs, which we grasp as high-frequency market makers, are few (less than ten)\(^{13}\) with all having about 10% of market share when considered as a monopolist or duopolist, other traders that execute through algorithmic accounts are many and small (the next biggest trader accounts for at most 0.5% of trading volume).

We attempt to provide a comprehensive overview over the sample data by showing summary statistics from three different angles, by stock, by HFTs, and by treatment and control group.\(^{14}\)

\(^{13}\)We cannot release neither names nor numbers due to confidentiality agreements with NASDAQ OMXN. We show, however summary statistics for the two most different HFTs in Table 2.

\(^{14}\)An overview and description of the most important variables is presented in appendix section ??.
Table 1 gives an overview and key statistics for all thirty stocks traded in the OMXS30. We provide the mean and the standard deviation of daily averages for the number of trades, volume, turnover and relative time-weighted order-execution spreads. The number of stock trades per day varies between averages of 1247 to 6103 across stocks. The average relative order-execution spread in our sample is about 0.09% to 0.24%. Column three shows how often a specific stock occurs as a control, column four gives the number of changes from HFT monopoly to HFT duopoly and column five the changes from HFT duopoly to HFT monopoly. Event and controls are fairly well distributed among the securities with two exceptions. Excluding Scania AB, which accounts for about 10% for all entries and exits, just improves significance. Dropping Nokia Corporation and Lundin Petroleum AB, which serve exclusively as controls, has no statistically significant effect on the results. The number of unique trading days that are considered for each stock, before and after entry or exit, is shown in column six.

[Insert Table 1 about here!]

Table 2 shows summary statistics for the two most different high-frequency traders in the market, high-frequency trader A and high-frequency trader B. Statistics are reported for the daily fraction of HFT trades of the entire market, the absolute number of daily HFT trades, the fraction of total daily volume, the fraction of daily HFT trades of algorithmic trades executed, the fraction of aggressive trades (The aggressive side of the trade is an incoming market order or marketable limit order that is executed against a standing limit order.), and aggressiveness imbalance constructed as the difference between aggressive buy transactions minus aggressive sell transaction. Descriptive statistics on the timing and impact of the trades are also listed. Statistics are provided for the fraction of HFT trades that involved a price change, HFT buy and sell transactions that involve a price decrease, HFT by and sell transactions that lead to a price increase, HFT buy and sell trades that are executed before a price increase an HFT buy and sell trades that are executed before a price decrease (last trade on a specific price level in the orderbook is executed by an HFT). Algorithmic trades are trades that were executed through an algorithmic trading account. This is the cheapest and fastest way to trade at NASDAQ OMXS. Volume is the number of securities traded. All statistics are based on daily observations for the
day before and after event (treatment and control group).

The only blatant difference is in aggressiveness. While HFT A executed 91% of its trades aggressively, HFT B shows 35%. We believe, however, that this difference in fact only reflects a specific way of executing trades as we elaborate in a more detailed discussion below. Statistics on actual trades show that there is a minor difference in how often a trade initiates a price change. HFT A, the more aggressive trader, initiates a price change in 10% of its trades, while HFT B initiates a price change in about 20% of the cases; both with a fairly large standard deviation. We do believe that aggressiveness is not as informative as literature seems to treat it. Aggressiveness, often associated with an identifying characteristic of HFT, is a rather misleading characteristic to identify HFTs as it might just reflect different ways of executing trades.\textsuperscript{15} HFTs with a high aggressiveness might follow a "snake strategy", which means that the algorithm places quickly marketable orders when some anomalies such as deviations from trends are observed. Contrary, low aggressiveness might appear when a trader follows a strategy by following the market, placing and cancelling orders, but will appear as less aggressive as the executed trades are limit orders. In fact, Table 2 shows detailed characteristics for the two most different HFTs in aggressiveness in the sample.

Trade timestamps and message timestamps are in milliseconds and then ranked within each millisecond.

[Insert Table 2 about here!]

This table lists descriptive statistics for all stocks and days that serve as the control group and for all stocks an days in the treatment group. Panel A shows statistics for entries (the change from HFT monopoly to HFT duopoly) and Panel B for exits (the change from HFT duopoly to HFT monopoly). Order-execution time is the time difference (in seconds) between an incoming market order or marketable limit order and the standing limit order that the trade is executed against. The hourly and five minute volatility is calculated from hourly and five minutes intraday returns (printed in squared percentages). The Max-Close, Min-Close, Open-Close and Close-Close

\textsuperscript{15}The aggressive side of the trade is an incoming market order or marketable limit order that is executed against a standing limit order.
volatility is calculated as squared percentages, the percentage change squared. Max-Close is the squared change from the maximum price within a day to the closing price. Min-Close is the squared change from the minimum price within a day to the closing price of the day. Open-Close shows the squared change from the opening price to the closing price of the day. Close-Close is the inter-day volatility and calculated from the squared change from the previous day’s closing price to today’s closing price. Further, the table shows volatility of the first five minutes of the day (calculated from minute returns), the number of securities traded (volume), the absolute number of daily HFT trades, the fraction of daily HFT trades of algorithmic trades executed and the daily relative time-weighted spread. Algorithmic trades are trades that were executed through an algorithmic trading account. There is no real difference between the control and treatment group, which should not come as a surprise given stocks serve as observations in both the control and the treatment group. The effects on order-execution time, volatility and volume we isolate in the regressions are quite visible.

[Insert Table 3 about here!]

5 EMPIRICAL RESULTS

Table 4 displays estimated coefficients of our entry difference-in-difference tests of hourly volatility computed from hourly intraday returns (column 1-4), five minute volatility based on 5 minutes intraday returns (column 5), max to close volatility computed as the squared change from the maximum price within a day to the closing price (column 6), min to close volatility computed as the squared change from the minimum price within a day to the closing price (column 7), open to close volatility shows the squared change from the opening price to the closing price of the day (column 8) and close to close volatility calculated from the squared change from the previous day’s closing price to today’s closing price (column 9). Additional controls, besides the level variables (indicator for the treated security and time fixed effects), are stock fixed effect, volume, order-execution time (Order-execution time is the time difference (in seconds) between an incoming market order or marketable limit order and the standing limit order that the trade is executed against.) and lagged variables. Standard errors are clustered on stock level and
reported in parentheses. Our findings suggest unequivocally ambiguous results on market quality. Intraday hourly volatility increases severely by an average of over 20% and five-minute volatility by an average of nearly 20%, interdaily (both measured from opening to closing and closing to closing prices), however, shows no sign of increase or decrease. These results hold for both entries and exits noting that for the latter one, it shows a decrease in intraday volatility (Table 5). We also provide results considering both entries and exits as one event (one can think of exits as reverse entries) in Table 6.

Table 7 displays estimated coefficients of our entry difference-in-difference tests of order-execution time measured by the time difference (in seconds) between an incoming market order or marketable limit order and the standing limit order that the trade is executed against (column 1-4) and the order-execution time daily standard deviation (column 5). Additional controls, besides the level variables (indicator for the treated security and time fixed effects), are stock fixed effect, volume, volatility (computed as intraday volatility of hourly returns) and lagged variables. Standard errors are clustered on stock level and reported in parentheses. The order-execution time decreases on average by about 15%, which is also reflected in a significant reduction of its standard deviation. Surprisingly there is no significant positive effect for exits (Table 8), only a marginally significant increase in its standard deviation. Table 9 combines entries and exits and shows marginally significant estimates.

Table 10 displays estimated coefficients of our entry difference-in-difference tests of volume measured as the number of securities traded (column 1-4) and the fraction of daily HFT volume (column 5). Additional controls, besides the level variables (indicator for the treated security and time fixed effects), are stock fixed effect, order-execution time (order-execution time is the time difference (in seconds) between an incoming market order or marketable limit order and the standing limit order that the trade is executed against.), volatility (computed as intraday...
volatility of hourly returns) and lagged variables. Standard errors are clustered on stock level and reported in parentheses. Even though the HFTs proportion of total volume increases and decreases significantly after entry and exit respectively, there is unexpectedly no effect on total volume. We treat this as an indication that there is a crowding out effect as we outlined earlier in the paper.

[Insert Tables 10, 11, 12 about here!]

6 CONCLUSION

High-frequency traders (HFTs) carry out a crucial part of critical importance for financial markets. We find that competition among high-frequency traders coincides with a stark increase of intraday volatility, but interestingly with no effect on interday volatility. We also find a decrease in bid-ask time (difference between an incoming market order or marketable limit order and the standing limit order that the trade is executed against) and an increase in market share of high-frequency traders, although with no effect on overall volume. We offer an attempt to draw causal conclusions by exploiting cross-sectional variations of stocks and conducting difference-in-difference tests. This paper provides results for both entries and exits (understood as changes from monopoly to duopoly and vice versa (daily)), and offers several explanations in favor of our findings. To briefly sum up the discussion, HFT competition has a stark impact on short-term volatility as HFTs compete for the same prices. Their investment horizon, however, is short and therefore there is no effect on long-term volatility. There is a decrease in bid-ask time, which reflects that HFT market making responds quicker and follows a potentially more sophisticated strategy and thereby increases market quality. The decrease in bid-ask time, increase in HFT market ratio and the seemingly steady volume suggest crowding out of other slower investors; potentially other market maker, which become unsuccessful in placing their orders.

Through highly competitive and quick market platforms, the advantage of technology such as co-location or the use of ultra-quick algorithms, HFTs changed and influenced financial markets substantially taking up to 85% of today’s equity market volume. HFTs tend end their end the day with close to zero inventories, frequently submit and cancel limit orders, use colocation facilities
and highly efficient algorithms, and have short holding periods. These changes provoked an intensive discussion by legislators, regulators and investors leading to controversial views that span topics from price manipulation, speed of trading, systemic risk due a high correlation of algorithmic strategies, to price discovery and liquidity. The quality of liquidity HFTs potentially provide is of particular concern as HFTs replaced traditional market makers. Our findings contribute to this discussion and gives new insights of how HFTs affect markets. Calls for more regulatory action in the HFT industry may deserve a new perspective given these new findings about the effects of competition between HFTs.
References


Hirschey, N. H. (2011). Do high-frequency traders anticipate buying and selling pressure?


Hendershott, T. & Riordan, R., "Algorithmic Trading and Information", *NET Institute*, 2009


A INSTITUTIONAL AND MARKET BACKGROUND

The NASDAQ OMXS (Stockholm) had about an 80% market share in 2009 with the majority of the trading volume in NASDAQ OMXS 30, listing Sweden’s largest public companies. The closest competitor was BATS Chi-X Europe with about 10% to 15% of market share in 2009, followed by Burgundy and Turquoise with less 5%.

The limit order book market is open Monday to Friday from 9am to 5:30pm, CET, except red days. There is one exception though, trading closes at 1pm if the following day is a public holiday. Both opening and closing prices are set by call auctions. Priority rank of an order during the trading day is price, time and visibility.

To access the market, financial intermediaries have four different possibilities. (i) A broker account, which is mostly used by institutional investors or non-automated trading. (ii) An order routing account that allows customers of the exchange member intermediary to rout their orders directly to the market. This is mostly used by direct banks such as internet banks. (iii) A programmed account is typically used to execute orders through an algorithm such as a big sequential sell or buy order. (iv) Finally, there is algorithmic trading account which is the quickest and the cheapest in terms of transaction costs and thus a natural choice for high-frequency traders.

There are about one hundred financial firms (members) registered at NASDAQ OMXS.

An important detail about NASDAQ OMXS is that members cannot place small hidden orders. The rule for being able to hide orders depends on the average daily turnover of a specific stock, but must be at least 50,000EUR. This, however, increases with turnover and reaches for example for one million euro a minimum order size of 250,000EUR. As a result, HFTs have no incentive to hide their orders.
Table 1: Summary Statistics of Sample Stocks

This table presents summary statistics for the NASDAQ OMX30 around the events from the time period between July 1st, 2009 and December 31st, 2009. It lists the ISIN code, the company's name, number of daily trades, daily volume (in 1000 units), daily turnover (in 1000 SEK) and the relative time-weighted spread. Column three shows how often a specific stock occurs as a control, column four gives the number of changes from HFT monopoly to HFT duopoly and column five the changes from HFT duopoly to HFT monopoly. The number of unique trading days for each stock is shown in column six (Note that a stock may serve as a control for more than one event per day.).

<table>
<thead>
<tr>
<th>ISIN Code</th>
<th>Security Name</th>
<th>Control</th>
<th>Entry</th>
<th>Exit</th>
<th>No Days</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH0012221716</td>
<td>ABB Ltd</td>
<td>32</td>
<td>3</td>
<td>2</td>
<td>39</td>
<td>2373</td>
<td>1120</td>
<td>2046</td>
<td>1444</td>
<td>405181</td>
<td>188669</td>
<td>0.177</td>
<td>0.047</td>
</tr>
<tr>
<td>FI0000900681</td>
<td>Nokia Corporation</td>
<td>48</td>
<td>0</td>
<td>0</td>
<td>48</td>
<td>1545</td>
<td>575</td>
<td>12909</td>
<td>509</td>
<td>111319</td>
<td>47498</td>
<td>0.112</td>
<td>0.013</td>
</tr>
<tr>
<td>GB0009895292</td>
<td>AstraZeneca PLC</td>
<td>29</td>
<td>5</td>
<td>4</td>
<td>39</td>
<td>2438</td>
<td>762</td>
<td>1328</td>
<td>416</td>
<td>420658</td>
<td>132163</td>
<td>0.134</td>
<td>0.058</td>
</tr>
<tr>
<td>SE0000101032</td>
<td>Atlas Copco AB A</td>
<td>57</td>
<td>2</td>
<td>1</td>
<td>60</td>
<td>3391</td>
<td>981</td>
<td>5238</td>
<td>1658</td>
<td>492667</td>
<td>158129</td>
<td>0.133</td>
<td>0.045</td>
</tr>
<tr>
<td>SE0000103814</td>
<td>Electrolux AB B</td>
<td>79</td>
<td>1</td>
<td>1</td>
<td>80</td>
<td>3156</td>
<td>1367</td>
<td>2619</td>
<td>1363</td>
<td>430898</td>
<td>226622</td>
<td>0.133</td>
<td>0.048</td>
</tr>
<tr>
<td>SE0000106270</td>
<td>Hennes &amp; Mauritz AB, H &amp; M B</td>
<td>57</td>
<td>4</td>
<td>4</td>
<td>66</td>
<td>4364</td>
<td>1809</td>
<td>2086</td>
<td>804</td>
<td>841452</td>
<td>329115</td>
<td>0.090</td>
<td>0.045</td>
</tr>
<tr>
<td>SE0000107419</td>
<td>Investor AB B</td>
<td>59</td>
<td>0</td>
<td>1</td>
<td>61</td>
<td>1845</td>
<td>514</td>
<td>1952</td>
<td>738</td>
<td>250733</td>
<td>94882</td>
<td>0.175</td>
<td>0.053</td>
</tr>
<tr>
<td>SE0000108227</td>
<td>SKF, AB B</td>
<td>52</td>
<td>2</td>
<td>3</td>
<td>59</td>
<td>2905</td>
<td>1064</td>
<td>3203</td>
<td>1499</td>
<td>363891</td>
<td>175180</td>
<td>0.119</td>
<td>0.028</td>
</tr>
<tr>
<td>SE0000108656</td>
<td>Ericsson, Telefonab. L M B</td>
<td>62</td>
<td>5</td>
<td>5</td>
<td>71</td>
<td>6103</td>
<td>2063</td>
<td>17464</td>
<td>9140</td>
<td>122518</td>
<td>644684</td>
<td>0.107</td>
<td>0.034</td>
</tr>
<tr>
<td>SE0000112724</td>
<td>Svenska Cellulosa AB SCA B</td>
<td>55</td>
<td>1</td>
<td>2</td>
<td>59</td>
<td>2269</td>
<td>840</td>
<td>2154</td>
<td>911</td>
<td>209161</td>
<td>88860</td>
<td>0.120</td>
<td>0.037</td>
</tr>
<tr>
<td>SE0000113250</td>
<td>Skanska AB B</td>
<td>57</td>
<td>2</td>
<td>3</td>
<td>64</td>
<td>2140</td>
<td>890</td>
<td>2010</td>
<td>1008</td>
<td>216641</td>
<td>111395</td>
<td>0.137</td>
<td>0.039</td>
</tr>
<tr>
<td>SE0000115446</td>
<td>Volvo, AB B</td>
<td>21</td>
<td>2</td>
<td>3</td>
<td>28</td>
<td>4099</td>
<td>1038</td>
<td>7135</td>
<td>2418</td>
<td>487457</td>
<td>161347</td>
<td>0.107</td>
<td>0.056</td>
</tr>
<tr>
<td>SE0000122467</td>
<td>Atlas Copco AB B</td>
<td>49</td>
<td>2</td>
<td>5</td>
<td>58</td>
<td>1257</td>
<td>445</td>
<td>1192</td>
<td>515</td>
<td>99839</td>
<td>42432</td>
<td>0.182</td>
<td>0.056</td>
</tr>
<tr>
<td>SE0000148884</td>
<td>Skandinaviska Enskilda Banken A</td>
<td>57</td>
<td>4</td>
<td>4</td>
<td>65</td>
<td>4547</td>
<td>1591</td>
<td>10718</td>
<td>4579</td>
<td>506003</td>
<td>207893</td>
<td>0.163</td>
<td>0.094</td>
</tr>
<tr>
<td>SE0000163594</td>
<td>Securitas AB B</td>
<td>64</td>
<td>4</td>
<td>4</td>
<td>71</td>
<td>1701</td>
<td>798</td>
<td>1989</td>
<td>1088</td>
<td>135058</td>
<td>74850</td>
<td>0.150</td>
<td>0.042</td>
</tr>
<tr>
<td>SE0000171100</td>
<td>SSAB AB A</td>
<td>56</td>
<td>2</td>
<td>3</td>
<td>63</td>
<td>2817</td>
<td>984</td>
<td>2899</td>
<td>1081</td>
<td>315455</td>
<td>114821</td>
<td>0.197</td>
<td>0.067</td>
</tr>
<tr>
<td>SE0000193120</td>
<td>Svenska Handelsbanken A</td>
<td>46</td>
<td>5</td>
<td>5</td>
<td>56</td>
<td>2189</td>
<td>938</td>
<td>1697</td>
<td>662</td>
<td>324175</td>
<td>123245</td>
<td>0.190</td>
<td>0.101</td>
</tr>
<tr>
<td>SE0000202624</td>
<td>Getinge AB B</td>
<td>60</td>
<td>2</td>
<td>3</td>
<td>65</td>
<td>1550</td>
<td>534</td>
<td>887</td>
<td>481</td>
<td>114466</td>
<td>59056</td>
<td>0.164</td>
<td>0.059</td>
</tr>
<tr>
<td>SE0000242455</td>
<td>Swedbank AB A</td>
<td>41</td>
<td>5</td>
<td>5</td>
<td>54</td>
<td>5301</td>
<td>2029</td>
<td>10825</td>
<td>4537</td>
<td>718231</td>
<td>325445</td>
<td>0.228</td>
<td>0.140</td>
</tr>
<tr>
<td>SE0000255648</td>
<td>ASSA ABLOY AB B</td>
<td>53</td>
<td>4</td>
<td>4</td>
<td>64</td>
<td>2311</td>
<td>941</td>
<td>2124</td>
<td>1072</td>
<td>256223</td>
<td>129367</td>
<td>0.126</td>
<td>0.035</td>
</tr>
<tr>
<td>SE0000308280</td>
<td>SCANIA AB B</td>
<td>8</td>
<td>14</td>
<td>13</td>
<td>47</td>
<td>1247</td>
<td>560</td>
<td>851</td>
<td>325</td>
<td>77448</td>
<td>31022</td>
<td>0.245</td>
<td>0.096</td>
</tr>
<tr>
<td>SE0000310536</td>
<td>Swedish Match AB</td>
<td>32</td>
<td>9</td>
<td>8</td>
<td>50</td>
<td>1443</td>
<td>477</td>
<td>1041</td>
<td>408</td>
<td>151999</td>
<td>58474</td>
<td>0.144</td>
<td>0.050</td>
</tr>
<tr>
<td>SE0000314312</td>
<td>Telc2 AB ser. B</td>
<td>54</td>
<td>2</td>
<td>2</td>
<td>58</td>
<td>2277</td>
<td>901</td>
<td>2043</td>
<td>1136</td>
<td>202572</td>
<td>108746</td>
<td>0.190</td>
<td>0.013</td>
</tr>
<tr>
<td>SE0000412371</td>
<td>Modern Times Group MTG AB B</td>
<td>67</td>
<td>5</td>
<td>4</td>
<td>74</td>
<td>1531</td>
<td>543</td>
<td>359</td>
<td>157</td>
<td>112574</td>
<td>48718</td>
<td>0.175</td>
<td>0.046</td>
</tr>
<tr>
<td>SE0000427361</td>
<td>Nordea Bank AB</td>
<td>47</td>
<td>7</td>
<td>6</td>
<td>60</td>
<td>3595</td>
<td>1449</td>
<td>9173</td>
<td>3640</td>
<td>671154</td>
<td>271461</td>
<td>0.145</td>
<td>0.036</td>
</tr>
<tr>
<td>SE0000667891</td>
<td>Sandvik AB</td>
<td>50</td>
<td>6</td>
<td>5</td>
<td>61</td>
<td>3423</td>
<td>929</td>
<td>5485</td>
<td>1856</td>
<td>433416</td>
<td>143590</td>
<td>0.129</td>
<td>0.046</td>
</tr>
<tr>
<td>SE0000667925</td>
<td>TeliaSonera AB</td>
<td>33</td>
<td>5</td>
<td>4</td>
<td>42</td>
<td>2591</td>
<td>1178</td>
<td>8998</td>
<td>5185</td>
<td>428673</td>
<td>264543</td>
<td>0.164</td>
<td>0.077</td>
</tr>
<tr>
<td>SE0000695876</td>
<td>Alfa Laval AB</td>
<td>35</td>
<td>7</td>
<td>7</td>
<td>54</td>
<td>2206</td>
<td>680</td>
<td>2240</td>
<td>1000</td>
<td>195334</td>
<td>84058</td>
<td>0.114</td>
<td>0.033</td>
</tr>
<tr>
<td>SE0000825820</td>
<td>Lundin Petroleum AB</td>
<td>74</td>
<td>0</td>
<td>0</td>
<td>74</td>
<td>1836</td>
<td>506</td>
<td>1452</td>
<td>487</td>
<td>87361</td>
<td>28854</td>
<td>0.169</td>
<td>0.035</td>
</tr>
<tr>
<td>SE0000869646</td>
<td>Bolident AB</td>
<td>60</td>
<td>1</td>
<td>0</td>
<td>61</td>
<td>4306</td>
<td>1521</td>
<td>5251</td>
<td>2089</td>
<td>431122</td>
<td>180909</td>
<td>0.150</td>
<td>0.063</td>
</tr>
<tr>
<td>Total/Mean</td>
<td>1494</td>
<td>128</td>
<td>100</td>
<td>1751</td>
<td>2776</td>
<td>1666</td>
<td>3941</td>
<td>4786</td>
<td>358919</td>
<td>335370</td>
<td>0.150</td>
<td>0.067</td>
<td></td>
</tr>
</tbody>
</table>
Table 2: Summary Statistics of High-Frequency Traders

This table shows summary statistics for the two most different high-frequency traders in the market, high-frequency trader A and high-frequency trader B. Statistics are reported for the daily fraction of HFT trades of the entire market, the absolute number of daily HFT trades, the fraction of total daily volume, the fraction of daily HFT trades of algorithmic trades executed, the fraction of aggressive trades (The aggressive side of the trade is an incoming market order or marketable limit order that is executed against a standing limit order.), and aggressiveness imbalance constructed as the difference between aggressive buy transactions minus aggressive sell transaction. Descriptive statistics on the timing and impact of the trades are also listed. Statistics are provided for the fraction of HFT trades that involved a price change, HFT buy and sell transactions that involve a price decrease, HFT by and sell transactions that lead to a price increase, HFT buy and sell trades that are executed before a price increase an HFT buy and sell trades that are executed before a price decrease (last trade on a specific price level in the orderbook is executed by an HFT). Algorithmic trades are trades that were executed through an algorithmic trading account. This is the cheapest and fastest way to trade at NASDAQ OMX. Volume is the number of securities traded. All statistics are based on daily observations for the day before and after event (treatment and control group).

<table>
<thead>
<tr>
<th></th>
<th>High-Frequency Trader A</th>
<th></th>
<th></th>
<th>High-Frequency Trader B</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
<td>SD</td>
<td>Mean</td>
<td>Median</td>
<td>SD</td>
</tr>
<tr>
<td>HFT Trades / Total Trades</td>
<td>0.1001</td>
<td>0.0832</td>
<td>0.0639</td>
<td>0.0956</td>
<td>0.0757</td>
<td>0.0749</td>
</tr>
<tr>
<td>HFT Trades (per Day and Stock)</td>
<td>279</td>
<td>193</td>
<td>258</td>
<td>266</td>
<td>190</td>
<td>229</td>
</tr>
<tr>
<td>HFT Volume / Total Volume</td>
<td>0.1033</td>
<td>0.0829</td>
<td>0.0729</td>
<td>0.0549</td>
<td>0.0379</td>
<td>0.0494</td>
</tr>
<tr>
<td>Closing Inventory (fraction)</td>
<td>0.0019</td>
<td>0.0000</td>
<td>0.0010</td>
<td>0.0024</td>
<td>0.0000</td>
<td>0.1090</td>
</tr>
<tr>
<td>HFT Trades / Algorithmic Trades</td>
<td>0.3020</td>
<td>0.2799</td>
<td>0.1528</td>
<td>0.2592</td>
<td>0.2349</td>
<td>0.1575</td>
</tr>
<tr>
<td>Aggressive Trades (fraction)</td>
<td>0.9106</td>
<td>0.9836</td>
<td>0.2345</td>
<td>0.3459</td>
<td>0.2672</td>
<td>0.2123</td>
</tr>
<tr>
<td>Aggressiveness Imbalance</td>
<td>-0.0271</td>
<td>-0.0243</td>
<td>0.1538</td>
<td>-0.0036</td>
<td>0.0000</td>
<td>0.1295</td>
</tr>
<tr>
<td>Trades Initiate a Price Changes (fraction)</td>
<td>0.0956</td>
<td>0.0710</td>
<td>0.0815</td>
<td>0.2169</td>
<td>0.2018</td>
<td>0.1687</td>
</tr>
<tr>
<td>Buy Trades Initiate a Price Decrease (fraction)</td>
<td>0.0117</td>
<td>0.0085</td>
<td>0.0139</td>
<td>0.0625</td>
<td>0.0455</td>
<td>0.0846</td>
</tr>
<tr>
<td>Sell Trades Initiate a Price Decrease (fraction)</td>
<td>0.0337</td>
<td>0.0226</td>
<td>0.0350</td>
<td>0.0427</td>
<td>0.0308</td>
<td>0.0379</td>
</tr>
<tr>
<td>Sell Trades Initiate a Price Increase (fraction)</td>
<td>0.0127</td>
<td>0.0096</td>
<td>0.0143</td>
<td>0.0697</td>
<td>0.0485</td>
<td>0.0897</td>
</tr>
<tr>
<td>Buy Trades Initiate a Price Increase (fraction)</td>
<td>0.0305</td>
<td>0.0201</td>
<td>0.0320</td>
<td>0.0380</td>
<td>0.0291</td>
<td>0.0343</td>
</tr>
<tr>
<td>Buy Trades Before a Price Decrease (fraction)</td>
<td>0.0402</td>
<td>0.0361</td>
<td>0.0267</td>
<td>0.0471</td>
<td>0.0394</td>
<td>0.0320</td>
</tr>
<tr>
<td>Sell Trades Before a Price Decrease (fraction)</td>
<td>0.0634</td>
<td>0.0584</td>
<td>0.0365</td>
<td>0.0558</td>
<td>0.0510</td>
<td>0.0336</td>
</tr>
<tr>
<td>Sell Trades Before a Price Increase (fraction)</td>
<td>0.0450</td>
<td>0.0403</td>
<td>0.0300</td>
<td>0.0506</td>
<td>0.0432</td>
<td>0.0351</td>
</tr>
<tr>
<td>Buy Trades Before a Price Increase (fraction)</td>
<td>0.0631</td>
<td>0.0588</td>
<td>0.0353</td>
<td>0.0525</td>
<td>0.0460</td>
<td>0.0330</td>
</tr>
</tbody>
</table>
Table 3: Summary Statistics of the Control and Treatment Group

This table lists descriptive statistics for all stocks and days that serve as the control group and for all stocks an days in the treatment group. Panel A shows statistics for entries (the change from HFT monopoly to HFT duopoly) and Panel B for exits (the change from HFT duopoly to HFT monopoly). Order-execution time is the time difference (in seconds) between an incoming market order or marketable limit order and the standing limit order that the trade is executed against. The hourly and five minute volatility is calculated from hourly and five minutes intraday returns (printed in squared percentages). The Max-Close, Min-Close, Open-Close and Close-Close volatility is calculated as squared percentages, the percentage change squared. Max-Close is the squared change from the maximum price within a day to the closing price. Min-Close is the squared change from the minimum price within a day to the closing price of the day. Open-Close shows the squared change from the opening price to the closing price of the day. Close-Close is the inter-day volatility and calculated from the squared change from the previous day’s closing price to today’s closing price. Further, the table shows volatility of the first five minutes of the day (calculated from minute returns), the number of securities traded (volume), the absolute number of daily HFT trades, the fraction of daily HFT trades of algorithmic trades executed and the daily relative time-weighted spread. Algorithmic trades are trades that were executed through an algorithmic trading account. This is the cheapest and fastest way to trade at NASDAQ OMX. All statistics are based on daily observations.

<table>
<thead>
<tr>
<th>Panel A: Entry</th>
<th>Control Group</th>
<th>Treatment Group Before Entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order-Execution Time (sec)</td>
<td>Obs Mean 66.366 Median 52.000 SD 51.788</td>
<td>Obs Mean 61.438 Median 39.500 SD 59.647</td>
</tr>
<tr>
<td>60min Vola</td>
<td>1137 0.375 Median 0.218 SD 0.506</td>
<td>121 0.301 Median 0.166 SD 0.340</td>
</tr>
<tr>
<td>5min Vola</td>
<td>1146 0.048 Median 0.041 SD 0.031</td>
<td>122 0.041 Median 0.031 SD 0.033</td>
</tr>
<tr>
<td>Max-Min Intraday Change Squared</td>
<td>1100 2.690 Median 0.880 SD 4.851</td>
<td>122 2.411 Median 0.801 SD 4.604</td>
</tr>
<tr>
<td>Open-Close Change Squared</td>
<td>1146 2.730 Median 1.082 SD 4.456</td>
<td>122 3.191 Median 0.857 SD 4.926</td>
</tr>
<tr>
<td>Close-Close Change Squared</td>
<td>1101 3.405 Median 1.291 SD 5.866</td>
<td>122 4.452 Median 1.467 SD 7.465</td>
</tr>
<tr>
<td>Volume (in 1000)</td>
<td>1146 3882 Median 2187 SD 4490</td>
<td>122 4456 Median 1916 SD 7303</td>
</tr>
<tr>
<td>HFT Volume (%)</td>
<td>1146 0.105 Median 0.085 SD 0.073</td>
<td>122 0.086 Median 0.061 SD 0.073</td>
</tr>
<tr>
<td>Trades (#)</td>
<td>1146 2775 Median 2350 SD 1585</td>
<td>122 3132 Median 2525 SD 2183</td>
</tr>
<tr>
<td>Algorithmic Trades (#)</td>
<td>1146 879 Median 743 SD 522</td>
<td>122 1013 Median 850 SD 651</td>
</tr>
<tr>
<td>HFT of Algorithmic (%)</td>
<td>1146 0.331 Median 0.318 SD 0.158</td>
<td>122 0.339 Median 0.320 SD 0.140</td>
</tr>
<tr>
<td>Bid-Ask Spread (SEK)</td>
<td>1146 0.150 Median 0.134 SD 0.063</td>
<td>122 0.125 Median 0.101 SD 0.074</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Exit</th>
<th>Control Group</th>
<th>Treatment Group After Exit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order-Execution Time (sec)</td>
<td>Obs Mean 63.518 Median 48.997 SD 55.147</td>
<td>Obs Mean 57.069 Median 37.001 SD 63.961</td>
</tr>
<tr>
<td>60min Vola</td>
<td>1113 0.358 Median 0.209 SD 0.468</td>
<td>100 0.290 Median 0.184 SD 0.313</td>
</tr>
<tr>
<td>5min Vola</td>
<td>1121 0.047 Median 0.041 SD 0.031</td>
<td>101 0.040 Median 0.028 SD 0.035</td>
</tr>
<tr>
<td>Max-Min Intraday Change Squared</td>
<td>1076 3.046 Median 1.600 SD 5.353</td>
<td>101 1.515 Median 0.545 SD 4.454</td>
</tr>
<tr>
<td>Min-Close Change Squared</td>
<td>1076 2.647 Median 0.995 SD 4.644</td>
<td>101 2.830 Median 0.879 SD 4.543</td>
</tr>
<tr>
<td>Close-Close Change Squared</td>
<td>1076 3.405 Median 1.288 SD 5.800</td>
<td>101 4.413 Median 1.747 SD 7.657</td>
</tr>
<tr>
<td>First 5min Vola</td>
<td>1118 0.038 Median 0.021 SD 0.054</td>
<td>100 0.042 Median 0.024 SD 0.059</td>
</tr>
<tr>
<td>Volume (in 1000)</td>
<td>1121 3905 Median 2186 SD 4730</td>
<td>101 4533 Median 1904 SD 6089</td>
</tr>
<tr>
<td>HFT Volume (%)</td>
<td>1121 0.106 Median 0.086 SD 0.074</td>
<td>101 0.093 Median 0.068 SD 0.072</td>
</tr>
<tr>
<td>Trades (#)</td>
<td>1121 2768 Median 2361 SD 1591</td>
<td>101 3232 Median 2657 SD 2275</td>
</tr>
<tr>
<td>Algorithmic Trades (#)</td>
<td>1121 891 Median 757 SD 511</td>
<td>101 1090 Median 886 SD 799</td>
</tr>
<tr>
<td>HFT of Algorithmic (%)</td>
<td>1121 0.329 Median 0.316 SD 0.151</td>
<td>101 0.351 Median 0.329 SD 0.149</td>
</tr>
<tr>
<td>Bid-Ask Spread (SEK)</td>
<td>1121 0.147 Median 0.133 SD 0.061</td>
<td>101 0.116 Median 0.099 SD 0.065</td>
</tr>
</tbody>
</table>
Table 4: Competition Effects of HFT Entry: Intra- and Inter-Day Volatilities

This table displays estimated coefficients of the following regression: \( y_{ist} = \beta_1 d_{is} + X_{ist} \Gamma + p_t + m_s + u_{ist} \), which allows for multiple time periods and multiple treatment groups (Bertrand, Duflo, and Mullainathan (2004)). With \( i \) indexing entry (the change from HFT monopoly to HFT duopoly), \( s \) being the security and \( t \) the time. \( d_{is} \) is an indicator if an HFT entry affected security \( s \) at time \( t \). \( p_t \) are daily time fixed effects and \( m_s \) are security fixed effects. \( X_{ist} \) is the vector of covariates and \( u_{ist} \) is the error term. The dependent variable, \( y_{ist} \), is hourly log volatility computed from hourly intraday returns (column 1-4), five minute log volatility based on 5 minutes intraday returns (column 5), max to close log volatility computed as the squared change from the maximum price within a day to the closing price (column 6), open to close log volatility shows the squared change from the opening price to the closing price of the day (column 7) and close to close log volatility calculated from the squared change from the previous day’s closing price to today’s closing price (column 8). Additional controls, besides the level variables (indicator for the treated security and time fixed effects), are stock fixed effect, volume, order-execution time (Order-execution time is the time difference (in seconds) between an incoming market order or marketable limit order and the standing limit order that the trade is executed against.) and lagged variables. Standard errors are clustered on stock level and reported in parentheses.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>HFT Entry</td>
<td>0.289***</td>
<td>0.295***</td>
<td>0.274***</td>
<td>0.302***</td>
<td>0.127***</td>
<td>0.392**</td>
<td>-0.096</td>
<td>-0.185</td>
</tr>
<tr>
<td></td>
<td>(0.090)</td>
<td>(0.090)</td>
<td>(0.086)</td>
<td>(0.094)</td>
<td>(0.046)</td>
<td>(0.176)</td>
<td>(0.214)</td>
<td>(0.239)</td>
</tr>
<tr>
<td>Treatment Dummy</td>
<td>-0.091</td>
<td>-0.110</td>
<td>-0.090</td>
<td>-0.106</td>
<td>-0.013</td>
<td>-0.161</td>
<td>0.202</td>
<td>0.347</td>
</tr>
<tr>
<td></td>
<td>(0.078)</td>
<td>(0.076)</td>
<td>(0.074)</td>
<td>(0.080)</td>
<td>(0.032)</td>
<td>(0.149)</td>
<td>(0.167)</td>
<td>(0.220)</td>
</tr>
<tr>
<td>Volume(t)</td>
<td>0.007**</td>
<td>0.004*</td>
<td>0.004*</td>
<td>0.005***</td>
<td>0.005</td>
<td>0.013***</td>
<td>0.017***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.005)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
</tr>
<tr>
<td>Order-Execution Time(t)</td>
<td>-0.267***</td>
<td>-0.269***</td>
<td>0.065*</td>
<td>-0.208**</td>
<td>-0.180</td>
<td>-0.258**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.064)</td>
<td>(0.037)</td>
<td>(0.037)</td>
<td>(0.093)</td>
<td>(0.127)</td>
<td>(0.118)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,882</td>
<td>1,882</td>
<td>1,882</td>
<td>1,882</td>
<td>1,905</td>
<td>1,750</td>
<td>1,834</td>
<td>1,804</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.3971</td>
<td>0.4092</td>
<td>0.4192</td>
<td>0.4194</td>
<td>0.7534</td>
<td>0.4575</td>
<td>0.3248</td>
<td>0.3373</td>
</tr>
<tr>
<td>Stock FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Time FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>HFT FE</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Cluster Stock</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1
Table 5: Competition Effects of HFT Exit: Intra- and Inter-Day Volatilities

This table displays estimated coefficients of the following regression: $y_{ist} = \beta_1 d_{is} + X_{ist}\Gamma + p_t + m_s + u_{ist}$, which allows for multiple time periods and multiple treatment groups (Bertrand, Duflo, and Mullainathan (2004)). With $i$ indexing exit (the change from HFT duopoly to HFT monopoly), $s$ being the security and $t$ the time. $d_{is}$ is an indicator if an HFT exit affected security $s$ at time $t$. $p_t$ are daily time fixed effects and $m_s$ are security fixed effects. $X_{ist}$ is the vector of covariates and $u_{ist}$ is the error term. The dependent variable, $y_{ist}$, is hourly log volatility computed from hourly intraday returns (column 1-4), five minute log volatility based on 5 minutes intraday returns (column 5), max to close log volatility computed as the squared change from the maximum price within a day to the closing price (column 6), open to close log volatility shows the squared change from the opening price to the closing price of the day (column 7) and close to close log volatility calculated from the squared change from the previous day’s closing price to today’s closing price (column 8). Additional controls, besides the level variables (indicator for the treated security and time fixed effects), are stock fixed effect, volume, order-execution time (Order-execution time is the time difference (in seconds) between an incoming market order or marketable limit order and the standing limit order that the trade is executed against.) and lagged variables. Standard errors are clustered on stock level and reported in parentheses.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>HFT Exit</td>
<td>-0.239**</td>
<td>-0.264**</td>
<td>-0.258***</td>
<td>-0.266**</td>
<td>-0.125**</td>
<td>-0.535***</td>
<td>-0.010</td>
<td>0.106</td>
</tr>
<tr>
<td></td>
<td>(0.114)</td>
<td>(0.108)</td>
<td>(0.106)</td>
<td>(0.116)</td>
<td>(0.060)</td>
<td>(0.176)</td>
<td>(0.284)</td>
<td>(0.222)</td>
</tr>
<tr>
<td>Treatment Dummy</td>
<td>0.298***</td>
<td>0.291***</td>
<td>0.255**</td>
<td>0.258**</td>
<td>0.080</td>
<td>0.325</td>
<td>-0.208</td>
<td>0.094</td>
</tr>
<tr>
<td></td>
<td>(0.102)</td>
<td>(0.096)</td>
<td>(0.094)</td>
<td>(0.097)</td>
<td>(0.063)</td>
<td>(0.223)</td>
<td>(0.174)</td>
<td>(0.137)</td>
</tr>
<tr>
<td>Volume(t)</td>
<td>0.006**</td>
<td>0.004*</td>
<td>0.004*</td>
<td>0.004***</td>
<td>0.006</td>
<td>0.015***</td>
<td>0.018***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.006)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td></td>
</tr>
<tr>
<td>Order-Execution</td>
<td>-0.268***</td>
<td>-0.269***</td>
<td>0.042</td>
<td>-0.167*</td>
<td>-0.237*</td>
<td>-0.280***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time(t)</td>
<td>(0.057)</td>
<td>(0.057)</td>
<td>(0.036)</td>
<td>(0.097)</td>
<td>(0.121)</td>
<td>(0.116)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,630</td>
<td>1,630</td>
<td>1,630</td>
<td>1,630</td>
<td>1,650</td>
<td>1,516</td>
<td>1,596</td>
<td>1,562</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.4049</td>
<td>0.4159</td>
<td>0.4261</td>
<td>0.4261</td>
<td>0.7673</td>
<td>0.4561</td>
<td>0.3184</td>
<td>0.3189</td>
</tr>
<tr>
<td>Stock FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Time FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>HFT FE</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Cluster Stock</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1
Table 6: Competition Effects of HFT Exit and Exit: Intra- and Inter-Day Volatilities

This table displays estimated coefficients of the following regression: \( y_{ist} = \beta_1 d_{is} + X_{ist} \Gamma + p_t + m_s + u_{ist} \), which allows for multiple time periods and multiple treatment groups (Bertrand, Duflo, and Mullainathan (2004)). With \( i \) indexing exit and exit (the change from HFT monopoly to HFT duopoly and vice versa), \( s \) being the security and \( t \) the time. \( d_{is} \) is an indicator if an HFT exit affected security \( s \) at time \( t \). \( p_t \) are daily time fixed effects and \( m_s \) are security fixed effects. \( X_{ist} \) is the vector of covariates and \( u_{ist} \) is the error term. The dependent variable, \( y_{ist} \), is hourly log volatility computed from hourly intraday returns (column 1-4), five minute log volatility based on 5 minutes intraday returns (column 5), max to close log volatility computed as the squared change from the maximum price within a day to the closing price (column 6), open to close log volatility shows the squared change from the opening price to the closing price of the day (column 7) and close to close log volatility calculated from the squared change from the previous day’s closing price to today’s closing price (column 8). Additional controls, besides the level variables (indicator for the treated security, indicator for the entry type and time fixed effects), are stock fixed effect and volume, order-execution time (Order-execution time is the time difference (in seconds) between an incoming market order or marketable limit order and the standing limit order that the trade is executed against.). Standard errors are clustered on stock level and reported in parentheses.

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HFT Entry x Date (Dummy)</td>
<td>0.246***</td>
<td>0.260***</td>
<td>0.247***</td>
<td>0.256***</td>
<td>0.119**</td>
<td>0.409**</td>
<td>-0.038</td>
<td>-0.173</td>
</tr>
<tr>
<td></td>
<td>(0.082)</td>
<td>(0.078)</td>
<td>(0.075)</td>
<td>(0.082)</td>
<td>(0.044)</td>
<td>(0.168)</td>
<td>(0.218)</td>
<td>(0.213)</td>
</tr>
<tr>
<td>Treatment (Dummy)</td>
<td>-0.111</td>
<td>-0.146*</td>
<td>-0.136*</td>
<td>-0.141*</td>
<td>-0.001</td>
<td>-0.188</td>
<td>0.155</td>
<td>0.327</td>
</tr>
<tr>
<td></td>
<td>(0.082)</td>
<td>(0.078)</td>
<td>(0.074)</td>
<td>(0.078)</td>
<td>(0.035)</td>
<td>(0.147)</td>
<td>(0.181)</td>
<td>(0.228)</td>
</tr>
<tr>
<td>Entry or Exit</td>
<td>0.148*</td>
<td>0.152*</td>
<td>0.129</td>
<td>0.127</td>
<td>-0.025</td>
<td>-0.005</td>
<td>-0.299***</td>
<td>-0.042</td>
</tr>
<tr>
<td></td>
<td>(0.082)</td>
<td>(0.080)</td>
<td>(0.078)</td>
<td>(0.078)</td>
<td>(0.033)</td>
<td>(0.147)</td>
<td>(0.081)</td>
<td>(0.131)</td>
</tr>
<tr>
<td>Volume(t)</td>
<td>0.007**</td>
<td>0.005*</td>
<td>0.005</td>
<td>0.005***</td>
<td>0.007</td>
<td>0.016***</td>
<td>0.017***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.005)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
</tr>
<tr>
<td>Order-Execution Time(t)</td>
<td>-0.235***</td>
<td>-0.235***</td>
<td>0.075**</td>
<td>-0.147*</td>
<td>-0.119</td>
<td>-0.186</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.057)</td>
<td>(0.036)</td>
<td>(0.085)</td>
<td>(0.104)</td>
<td>(0.111)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>2,119</td>
<td>2,119</td>
<td>2,119</td>
<td>2,119</td>
<td>2,145</td>
<td>1,982</td>
<td>2,064</td>
<td>2,039</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.3988</td>
<td>0.4109</td>
<td>0.4190</td>
<td>0.4190</td>
<td>0.7585</td>
<td>0.4468</td>
<td>0.3153</td>
<td>0.3184</td>
</tr>
<tr>
<td>Stock FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Time FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>HFT FE</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Cluster Stock</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
*** p < 0.01, ** p < 0.05, * p < 0.1
Table 7: Competition Effects of HFT Entry: Order-execution Time

This table displays estimated coefficients of the following regression: $y_{ist} = \beta_1 d_{is} + X_{ist} \Gamma + p_t + m_s + u_{ist}$, which allows for multiple time periods and multiple treatment groups (Bertrand, Duflo, and Mullainathan (2004)). With $i$ indexing entry (the change from HFT monopoly to HFT duopoly), $s$ being the security and $t$ the time. $d_{is}$ is an indicator if an HFT entry affected security $s$ at time $t$. $p_t$ are daily time fixed effects and $m_s$ are security fixed effects. $X_{ist}$ is the vector of covariates and $u_{ist}$ is the error term. The dependent variable, $y_{ist}$, is log order-execution time measured by the median time difference between an incoming market order or marketable limit order and the standing limit order that the trade is executed against (column 1-4) and the log order-execution time daily standard deviation (column 5). Additional controls, besides the level variables (indicator for the treated security and time fixed effects), are stock fixed effect, volume, volatility (computed as intraday volatility of hourly returns) and lagged variables. Standard errors are clustered on stock level and reported in parentheses.

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1) Order-Execution Time</th>
<th>(2) Order-Execution Time</th>
<th>(3) Order-Execution Time</th>
<th>(4) Order-Execution Time</th>
<th>(5) Order-Execution Time (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HFT Entry</td>
<td>-0.128**</td>
<td>-0.136***</td>
<td>-0.102**</td>
<td>-0.153***</td>
<td>-0.082*</td>
</tr>
<tr>
<td></td>
<td>(0.049)</td>
<td>(0.042)</td>
<td>(0.042)</td>
<td>(0.051)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>Treatment Dummy</td>
<td>0.082</td>
<td>0.110*</td>
<td>0.103*</td>
<td>0.081</td>
<td>0.053*</td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
<td>(0.061)</td>
<td>(0.058)</td>
<td>(0.058)</td>
<td>(0.031)</td>
</tr>
<tr>
<td>Volume(t)</td>
<td>-0.009***</td>
<td>-0.009***</td>
<td>-0.009***</td>
<td>-0.009***</td>
<td>-0.003***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Log Volatility(t)</td>
<td>-0.087***</td>
<td>-0.087***</td>
<td>-0.087***</td>
<td>-0.087***</td>
<td>0.032***</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.011)</td>
<td>(0.011)</td>
<td>(0.011)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,905</td>
<td>1,905</td>
<td>1,882</td>
<td>1,882</td>
<td>1,882</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.5938</td>
<td>0.6487</td>
<td>0.6584</td>
<td>0.6588</td>
<td>0.3589</td>
</tr>
<tr>
<td>Stock FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Time FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>HFT FE</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Cluster Stock</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Table 8: Competition Effects of HFT Exit: Order-execution Time

This table displays estimated coefficients of the following regression: $y_{ist} = \beta_1 d_{is} + X_{ist} \Gamma + p_t + m_s + u_{ist}$, which allows for multiple time periods and multiple treatment groups (Bertrand, Duflo, and Mullainathan (2004)). With $i$ indexing exit (the change from HFT duopoly to HFT monopoly), $s$ being the security and $t$ the time. $d_{is}$ is an indicator if an HFT exit affected security $s$ at time $t$. $p_t$ are daily time fixed effects and $m_s$ are security fixed effects. $X_{ist}$ is the vector of covariates and $u_{ist}$ is the error term. The dependent variable, $y_{ist}$, is log order-execution time measured by the median time difference between an incoming market order or marketable limit order and the standing limit order that the trade is executed against (column 1-4) and the log order-execution time daily standard deviation (column 5). Additional controls, besides the level variables (indicator for the treated security and time fixed effects), are stock fixed effect, volume, volatility (computed as intraday volatility of hourly returns) and lagged variables. Standard errors are clustered on stock level and reported in parentheses.

\[
egin{array}{lcccccc}
\text{VARIABLES} & \text{(1)} & \text{(2)} & \text{(3)} & \text{(4)} & \text{(5)} \\
\text{HFT Exit} & 0.022 & 0.057 & 0.026 & 0.154** & 0.063 \\
 & (0.048) & (0.051) & (0.050) & (0.058) & (0.039) \\
\text{Treatment Dummy} & -0.185** & -0.182** & -0.150* & -0.327*** & -0.066** \\
 & (0.084) & (0.085) & (0.084) & (0.084) & (0.031) \\
\text{Volume}(t) & -0.009*** & -0.008*** & -0.008*** & -0.008*** & -0.002*** \\
 & (0.002) & (0.002) & (0.002) & (0.002) & (0.001) \\
\text{Log Volatility}(t) & -0.084*** & -0.083*** & 0.010 & -0.083*** & -0.010 \\
 & (0.011) & (0.011) & (0.008) & (0.011) & (0.008) \\
\text{Observations} & 1,650 & 1,650 & 1,630 & 1,630 & 1,630 \\
\text{R-squared} & 0.5830 & 0.6324 & 0.6148 & 0.6447 & 0.4654 \\
\text{Stock FE} & YES & YES & YES & YES & YES \\
\text{Time FE} & YES & YES & YES & YES & YES \\
\text{HFT FE} & NO & NO & YES & YES & YES \\
\text{Cluster Stock} & YES & YES & YES & YES & YES \\
\end{array}
\]

Robust standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Table 9: Competition Effects of HFT Entry and Exit: Order-execution Time

This table displays estimated coefficients of the following regression: \( y_{ist} = \beta_1 d_{is} + X_{ist} \Gamma + p_t + m_s + u_{ist}, \) which allows for multiple time periods and multiple treatment groups (Bertrand, Duflo, and Mullainathan (2004)). With \( i \) indexing exit and entry (the change from HFT monopoly to HFT duopoly and vice versa), \( s \) being the security and \( t \) the time. \( d_{is} \) is an indicator if an HFT exit affected security \( s \) at time \( t \). \( p_t \) are daily time fixed effects and \( m_s \) are security fixed effects. \( X_{ist} \) is the vector of covariates and \( u_{ist} \) is the error term. The dependent variable, \( y_{ist} \), is log order-execution time measured by the median time difference between an incoming market order or marketable limit order and the standing limit order that the trade is executed against (column 1-4) and the log order-execution time daily standard deviation (column 5). Additional controls, besides the level variables (indicator for the treated security, indicator for the entry type and time fixed effects), are stock fixed effect and volume, volatility (computed as intraday volatility of hourly returns). Standard errors are clustered on stock level and reported in parentheses.

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1) Order-Execution Time</th>
<th>(2) Order-Execution Time</th>
<th>(3) Order-Execution Time</th>
<th>(4) Order-Execution Time</th>
<th>(5) Order-Execution Time (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HFT Entry or Exit x Date (Dummy)</td>
<td>-0.098*</td>
<td>-0.115**</td>
<td>-0.088*</td>
<td>-0.175***</td>
<td>-0.066</td>
</tr>
<tr>
<td>(0.049)</td>
<td>(0.044)</td>
<td>(0.044)</td>
<td>(0.046)</td>
<td>(0.046)</td>
<td></td>
</tr>
<tr>
<td>Treatment (Dummy)</td>
<td>0.041</td>
<td>0.092</td>
<td>0.087</td>
<td>0.056</td>
<td>0.072*</td>
</tr>
<tr>
<td>(0.066)</td>
<td>(0.063)</td>
<td>(0.062)</td>
<td>(0.065)</td>
<td>(0.039)</td>
<td></td>
</tr>
<tr>
<td>Entry or Exit</td>
<td>-0.126*</td>
<td>-0.138**</td>
<td>-0.134**</td>
<td>-0.142**</td>
<td>-0.065**</td>
</tr>
<tr>
<td>(0.065)</td>
<td>(0.062)</td>
<td>(0.059)</td>
<td>(0.058)</td>
<td>(0.024)</td>
<td></td>
</tr>
<tr>
<td>Volume(t)</td>
<td>-0.009***</td>
<td>-0.008***</td>
<td>-0.008***</td>
<td>-0.003***</td>
<td></td>
</tr>
<tr>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Volatility(t)</td>
<td>-0.079***</td>
<td>-0.079***</td>
<td>-0.079***</td>
<td>0.031***</td>
<td></td>
</tr>
<tr>
<td>(0.010)</td>
<td>(0.010)</td>
<td>(0.010)</td>
<td>(0.008)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>2,145</td>
<td>2,145</td>
<td>2,119</td>
<td>2,119</td>
<td>2,119</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.5835</td>
<td>0.6339</td>
<td>0.6422</td>
<td>0.6432</td>
<td>0.3388</td>
</tr>
<tr>
<td>Stock FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Time FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>HFT FE</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Cluster Stock</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1
Table 10: Competition Effects of HFT Entry: Trading Volume

This table displays estimated coefficients of the following regression: $y_{ist} = \beta_1d_{is} + X_{ist}\Gamma + p_t + m_s + u_{ist}$, which allows for multiple time periods and multiple treatment groups (Bertrand, Duflo, and Mullainathan (2004)). With $i$ indexing entry (the change from HFT monopoly to HFT duopoly), $s$ being the security and $t$ the time. $d_{is}$ is an indicator if an HFT entry affected security $s$ at time $t$. $p_t$ are daily time fixed effects and $m_s$ are security fixed effects. $X_{ist}$ is the vector of covariates and $u_{ist}$ is the error term. The dependent variable, $y_{ist}$, is volume measured as daily turnover (in 10 000 SEK) (column 1-4) and the fraction of daily HFT volume (column 5). Additional controls, besides the level variables (indicator for the treated security and time fixed effects), are stock fixed effect, order-execution time (Order-execution time is the time difference (in seconds) between an incoming market order or marketable limit order and the standing limit order that the trade is executed against.), volatility (computed as intraday volatility of hourly returns) and lagged variables. Standard errors are clustered on stock level and reported in parentheses.

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HFT Entry</td>
<td>-0.997</td>
<td>-0.997</td>
<td>-2.245</td>
<td>0.155</td>
<td>0.078***</td>
</tr>
<tr>
<td></td>
<td>(1.528)</td>
<td>(1.528)</td>
<td>(1.385)</td>
<td>(2.679)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Treatment Dummy</td>
<td>2.793</td>
<td>2.793</td>
<td>3.678*</td>
<td>4.393*</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>(1.831)</td>
<td>(1.831)</td>
<td>(1.941)</td>
<td>(2.326)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Order-Execution Time(t)</td>
<td>-13.996***</td>
<td>-13.896***</td>
<td>-0.026***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.371)</td>
<td>(3.246)</td>
<td>(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Volatility(t)</td>
<td>1.795**</td>
<td>1.804**</td>
<td>-0.004***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.717)</td>
<td>(0.716)</td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,905</td>
<td>1,905</td>
<td>1,882</td>
<td>1,882</td>
<td>1,882</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.6796</td>
<td>0.6796</td>
<td>0.7214</td>
<td>0.7217</td>
<td>0.6438</td>
</tr>
<tr>
<td>HFT FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Time FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>HFT FE</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Cluster Stock</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 11: Competition Effects of HFT Exit: Trading Volume

This table displays estimated coefficients of the following regression: $y_{ist} = \beta_1d_{is} + X_{ist}\Gamma + p_t + m_s + u_{ist}$, which allows for multiple time periods and multiple treatment groups (Bertrand, Duflo, and Mullainathan (2004)). With $i$ indexing exit (the change from HFT duopoly to HFT monopoly), $s$ being the security and $t$ the time. $d_{is}$ is an indicator if an HFT exit affected security $s$ at time $t$. $p_t$ are daily time fixed effects and $m_s$ are security fixed effects. $X_{ist}$ is the vector of covariates and $u_{ist}$ is the error term. The dependent variable, $y_{ist}$, is volume measured as daily turnover (in 10 000 SEK) (column 1-4) and the fraction of daily HFT volume (column 5). Additional controls, besides the level variables (indicator for the treated security and time fixed effects), are stock fixed effect, order-execution time (Order-execution time is the time difference (in seconds) between an incoming market order or marketable limit order and the standing limit order that the trade is executed against.), volatility (computed as intraday volatility of hourly returns) and lagged variables. Standard errors are clustered on stock level and reported in parentheses.

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HFT Exit</td>
<td>4.316*</td>
<td>4.316*</td>
<td>4.195*</td>
<td>1.542</td>
<td>-0.078***</td>
</tr>
<tr>
<td></td>
<td>(2.244)</td>
<td>(2.244)</td>
<td>(2.078)</td>
<td>(3.098)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>Treatment Dummy</td>
<td>0.739</td>
<td>0.739</td>
<td>-1.311</td>
<td>2.172</td>
<td>0.081***</td>
</tr>
<tr>
<td></td>
<td>(2.552)</td>
<td>(2.552)</td>
<td>(2.133)</td>
<td>(3.992)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>Order-Execution Time(t)</td>
<td>-13.277***</td>
<td>-13.104***</td>
<td>-0.025***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.686)</td>
<td>(3.493)</td>
<td>(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Volatility(t)</td>
<td>1.776**</td>
<td>1.777**</td>
<td>-0.004**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.551)</td>
<td>(0.550)</td>
<td>(0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,650</td>
<td>1,650</td>
<td>1,630</td>
<td>1,630</td>
<td>1,630</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.6937</td>
<td>0.6937</td>
<td>0.7298</td>
<td>0.7302</td>
<td>0.6567</td>
</tr>
<tr>
<td>HFT FE</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Time FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Cluster Stock</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Table 12: Competition Effects of HFT Entry and Exit: Trading Volume

This table displays estimated coefficients of the following regression: \( y_{ist} = \beta_1 d_{is} + X_{ist} \Gamma + p_t + m_s + u_{ist} \), which allows for multiple time periods and multiple treatment groups (Bertrand, Duflo, and Mullainathan (2004)). With \( i \) indexing exit and exit (the change from HFT monopoly to HFT duopoly and vice versa), \( s \) being the security and \( t \) the time. \( d_{is} \) is an indicator if an HFT exit affected security \( s \) at time \( t \). \( p_t \) are daily time fixed effects and \( m_s \) are security fixed effects. \( X_{ist} \) is the vector of covariates and \( u_{ist} \) is the error term. The dependent variable, \( y_{ist} \), is volume measured as daily turnover (in 10000 SEK) (column 1-4) and the fraction of daily HFT volume (column 5). Additional controls, besides the level variables (indicator for the treated security, indicator for the entry type and time fixed effects), are stock fixed effect and order-execution time (Order-execution time is the time difference (in seconds) between an incoming market order or marketable limit order and the standing limit order that the trade is executed against.), volatility (computed as intraday volatility of hourly returns). Standard errors are clustered on stock level and reported in parentheses.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>HFT Entry or Exit x Date (Dummy)</td>
<td>-2.294</td>
<td>-3.034**</td>
<td>0.419</td>
<td>0.077***</td>
</tr>
<tr>
<td></td>
<td>(1.398)</td>
<td>(1.208)</td>
<td>(2.621)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Treatment (Dummy)</td>
<td>5.448**</td>
<td>5.465***</td>
<td>6.255**</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(2.088)</td>
<td>(1.858)</td>
<td>(2.274)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Entry or Exit</td>
<td>-0.926</td>
<td>-2.116</td>
<td>-1.759</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(1.598)</td>
<td>(1.596)</td>
<td>(1.552)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Order-Execution Time(t)</td>
<td>-12.682***</td>
<td>-12.542***</td>
<td>-0.027***</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>(2.964)</td>
<td>(2.817)</td>
<td>(0.004)</td>
<td></td>
</tr>
<tr>
<td>Log Volatility(t)</td>
<td>1.964***</td>
<td>1.965***</td>
<td>-0.005***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.712)</td>
<td>(0.712)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>2,145</td>
<td>2,119</td>
<td>2,119</td>
<td>2,119</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.6879</td>
<td>0.7241</td>
<td>0.7246</td>
<td>0.6470</td>
</tr>
<tr>
<td>Stock FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Time FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>HFT FE</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Cluster Stock</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td></td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Figure 1: Stylized Motivating Example

This figure presents a motivating example how entry (the change from monopoly to duopoly) typically looks like within a specific stock. It shows high-frequency trader (HFT) A as the incumbent and HFT B as the entrant. Daily ratios of HFT A and B of total trades (participation rate) and total volume are plotted on the y-axis.
Figure 2: Summary Statistics of Key Variables

This figure shows graphically deviations from pre-entry and post-exit averages for key variables. The left column shows effects on the mean for entries (changes from HFT monopoly to duopoly) of HFT competitors and the right column presents the effects for HFT exits restoring monopoly. The top graphs show relative changes in HFT trading participation. The graphs in the middle present relative changes in intraday volatility (60 minutes) and the bottom graphs document relative changes in means of median order-execution times. The hollow circles represent means for the upper and lower 50% percentile of most liquid stocks respectively.
Figure 3: Intraday Average Inventory of High-Frequency Traders

This figure pictures the average intraday inventory over all stocks and days for five minute bins. Inventory is defined as the cumulative turnover divided by total turnover within each five minute bucket. While the left graph views average inventory over all days and stocks with a monopolistic high-frequency trader, the right hand graph shows inventories under competition for both the incumbent and entrant. Trading takes place from 9am to 5:30pm.
Figure 4: Competition over the Same Trades

This figure images the average intraday fraction of trades that were executed on the same side of the market (over all stocks and days) for five minute and sixty minute bins under competition. Trades are executed on the same side of the market if within each bin both the entrant and the incumbent buy or sell. The measure is constructed by assigning one if both high-frequency traders trade on the same side of the market and zero if not. The average ratio of trading on the same side of the market as their competitor is $2/3$. The dark shaded bars are hourly averages. The market is open from 9am to 5:30pm.
This figure pictures pre-closing trading activities. Pre-closing takes place from 5:20 to 5:30 and determines the closing price by maximizing tradable volume. Timestamps within the closing period reflect the order time and not the actual transaction time. Average turnover per trade, average total stock turnover and average high-frequency trading turnover are printed.

![Figure 5: Pre-Closing Trading](image-url)
Figure 6: Dynamic Impact of Entry and Exit on Volatility

This figure shows point estimates for five days before and five days after the event from the difference-in-difference estimation. We consider five days before and five days after the event. The plotted coefficients originate from following regression:

\[ y_{ist} = \beta_1 d_{is}^{−5} + \beta_2 d_{is}^{−4} + \cdots + \beta_{10} d_{is}^{0} + X_{ist} \Gamma + p_t + m_s + u_{ist}, \]

which allows for multiple time periods and multiple treatment groups (Bertrand, Duflo, and Mullainathan (2004)). With \( i \) indexing entry (the change from HFT monopoly to HFT duopoly and vice versa), \( s \) being the security and \( t \) the time. \( d_{is}^j \) is an indicator for the distance, \( j \), if an HFT entry or exit affected security \( s \) at time \( t \). \( p_t \) are daily time fixed effects and \( m_s \) are security fixed effects. \( X_{ist} \) is the vector of covariates and \( u_{ist} \) is the error term. The dependent variable, \( y_{ist} \), is hourly volatility computed from hourly intraday returns. On the left, we show the volatility increase after entry and on the right we show the volatility decrease after exit.
Figure 7: Dynamic Impact of Entry and Exit on Order-execution Time

This figure shows point estimates for five days before and five days after the event from the difference-in-difference estimation. We consider five days before and five days after the event. The plotted coefficients originate from following regression:

\[ y_{ist} = \beta_1 d_{i5}^{t-5} + \beta_2 d_{i4}^{t-4} + \cdots + \beta_{10} d_{i5}^{t} + X_{ist} \Gamma + p_t + m_s + u_{ist}, \]

which allows for multiple time periods and multiple treatment groups (Bertrand, Duflo, and Mullainathan (2004)). With \( i \) indexing entry (the change from HFT monopoly to HFT duopoly and vice versa), \( s \) being the security and \( t \) the time. \( d_{is}^j \) is an indicator for the distance, \( j \), if an HFT entry or exit affected security \( s \) at time \( t \). \( p_t \) are daily time fixed effects and \( m_s \) are security fixed effects. \( X_{ist} \) is the vector of covariates and \( u_{ist} \) is the error term. The dependent variable, \( y_{ist} \), is hourly volatility computed from hourly intraday returns. On the left, we show the volatility increase after entry and on the right we show the volatility decrease after exit.
Asset Pricing Frictions in Fragmented Markets*

January 21, 2013

Abstract

We study the consequences of trading fragmentation and speed on liquidity and asset prices. Exchanges invest in speed-enhancing technologies and price trading services to attract investors. Investors trade due to idiosyncratic preference shocks. We show how the resulting market organization affects asset liquidity and the composition of participating investors. In a consolidated market, speed investments raise liquidity and prices. When markets fragment, liquidity and asset prices can move in opposite directions. We also show how mechanisms that protect execution prices, such as the SEC’s trade-through rule, can decrease price levels and trading volume relative to unregulated markets. Our results suggest that recent regulatory reforms in secondary markets may have unintended negative consequences for public corporations.

JEL Codes: G12, G15, G18, D40, D43, D61.

---

*An earlier version of this paper circulated under the title “Long-Run Liquidity”.
Economists widely accept that market frictions can significantly affect liquidity, and thus have effects on asset prices, aggregate investments and welfare. Technological frictions, for example, can prevent buyers and sellers from being continuously matched over time and trading institutions’ market power can limit investor participation in financial markets. During any given day, traders (e.g., households, dealers, or hedge funds) must incorporate these frictions as a constraint in their investment plans. However, one can expect that, over a longer period, the existence of such frictions can provide incentives to developing institutions that would profit by alleviating them.\(^1\) In fact, consistent with this intuition, the last decade has witnessed profound transformations in secondary markets for securities: The speed at which investors trade has greatly increased (see Figure 2), and the number of new stock and derivatives exchanges in the U.S. and Europe has grown significantly, rendering markets more “fragmented” (see Figure 3). Do these transformations enhance liquidity and trading activity? What are the consequences for asset prices? The answers to these questions are the subject of heated debates in academic and policy circles.\(^2\)

We address these issues by studying a tractable framework that characterizes why and how investors with liquidity needs value trading speed, how investment in speed affects competition between trading institutions and investor market participation, and how the interaction between investor and trading institution choices affect long-run liquidity and asset prices. We show how aggregate outcomes critically depend on financial markets’ competitive structure and the linkages between trading venues. In particular, we show that when markets fragment, liquidity and asset prices can move in opposite directions.

We consider an economy where a continuum of investors (“farmers”) consume a constant stream of dividends (“perishable fruits”) from their random endowment of a real asset (a non-stochastic Lucas (1978) “tree”). Investors can differ in their valuation of the dividend over time due to idiosyncratic preference shocks, which creates potential gains from trade. The size of these shocks can differ from investor to investor in a mean-preserving fashion, as they would if some agents were more leveraged. Investors cannot buy or sell the asset directly; they can only do so by participating in costly trading venues, generically denoted here as exchanges. Participating investors gain repeated access to a given exchange to trade, but with a certain time delay. Trading speed allows investors to readjust holdings more effectively in response to shocks and thus realize higher gains from trade. Exchanges

---

\(^1\) Amihud, Mendelson, and Pedersen (2006), for example, argue that “Alleviating frictions is costly ... and the institutions which alleviate frictions may be able to earn rents. For instance, setting up a market requires computers, trading systems, clearing operations, risk and operational controls, legal documentation, marketing, information and communication systems, and so on.”

\(^2\) See, for example, the discussion in the SEC’s 2010 Concept Release on the U.S. market structure, and the discussions following the recent creation of the Joint Commodity Futures Trading Commission–SEC Advisory Committee on Emerging Regulatory Issues.
therefore have incentives to earn rents by investing in trading technologies that enhance speed.

When all trading is consolidated in a single exchange, speed alleviates market contact frictions, enhances asset liquidity and thus raises the equilibrium asset price. This result is both intuitive and consistent with empirical findings. However, market power gives a monopolist exchange the ability to limit investor participation by raising access fees. Investors with identical mean valuations for the asset but unequal leverage may rationally make different market participation decisions. As access costs increase, only the more highly levered investors (for whom gains from trade are large) join the market. Investor participation frictions thus alter the composition of traders in the market, which distorts the market clearing price. In particular, the equilibrium asset price can increase with participation costs and be higher than its frictionless Walrasian counterpart.

Do frictions generate the same outcomes in economies with fragmented trading? We show in a simple two-exchange economy that speed-driven competition further reduces average trading delays, encourages investor participation, and increases traded volumes. All these factors enhance asset liquidity vis-à-vis the consolidated market. However, the equilibrium effect on asset prices may be surprisingly different from the single-exchange case. Competition between exchanges, while leaving the relation of tradable assets to market participants unaltered, results in lower market access fees and thus increases the fraction of participating investors. As this fraction increases, the market average leverage decreases and the marginal participant (the one that clears the market) sell her endowment to a pool of investors whose average gains from trade become smaller. Provided the asset supply is not “too large,” the marginal investor’s valuation decreases and fragmentation can thus lead to lower asset prices.

These results suggest an interesting empirical relation: When an entrant exchange breaks a monopoly and trading fragments, asset liquidity may increase and the asset price level may decrease. The relative strength of asset liquidity and investor composition effects on the asset price depend on the ability of exchanges to vertically differentiate their liquidity services (Gabszewicz and Thisse (1979)) and thus relax cost-based competition compared to a Bertrand-like outcome.

Our analysis is closely related to recent empirical work that investigates the impact of trading fragmentation. For example, O’Hara and Ye (2011) find evidence that fragmentation has

---

3See, for example, the single-exchange empirical results of Hendershott, Jones, and Menkveld (2011) who study fast (algorithmic) trading.

4This effect of costs on asset prices is of different nature from that stated in Vayanos (1998), where trading costs affect trade sizes and holding periods but not the composition of investors in the market.

5If asset supply is sufficiently large relative to investor demand, the asset price reflects the minimum investor valuation regardless of the market structure.
a beneficial effect on widely used measures of “market quality” related to liquidity, such as bid-ask spreads and execution delays.\textsuperscript{6} These findings seem to support policymakers in the U.S. and other countries who have encouraged fragmentation in recent years.\textsuperscript{7} While our theoretical results are consistent with the findings on liquidity, they also suggest that one should be cautious when interpreting such findings as evidence that fragmentation is enhancing financial markets. Fragmentation may have broader consequences for asset prices, and thus have unintended impacts on related economic variables such as the cost of capital for corporations in primary markets.

When an asset trades in different markets, its price can in principle be different in each market, depending on the degree of integration between venues. Arbitrageurs can, of course, work to move prices closer to each other, but their ability to do so is subject to well-recognized frictions. Policymakers have thus designed mechanisms that address this issue directly, motivated by “protecting investors” from unfavorable prices. In the U.S. equity market such a mechanism is applied via the SEC’s “trade-through” rule (Rule 611 in Regulation National Market System, hereafter Reg NMS), which essentially mandates the integration of price formation across markets and gives investors access to the “national best” price independently of their trading location\textsuperscript{8} (see Appendix A). Although many observers agree that the trade-through rule had a profound impact on U.S. equity markets since its implementation in 2007, its precise effects on financial markets are less clear. We show that a trade-through rule affects competition between exchanges, redistributes investors across them and, importantly, favors the most illiquid market. The resulting distortions in competition affect asset prices. We show that the national best price lies in between the prices that would result in speed-differentiated exchanges were they perfectly segmented. Moreover, when speed investment costs are moderate, the national best price is lower than the volume-weighted average price (VWAP) in segmented markets.

By considering countries with different financial markets, our results provide several international asset pricing predictions. Consider, for example, an asset with an identical payoff structure that trades in different countries. Our results suggest that the trading price should be higher in economies where there are single exchanges (e.g. China, Spain and Brazil) than in economies with fragmented markets (e.g., the U.K, France and the Netherlands), and

\textsuperscript{6}Foucault and Menkveld (2008) and Degryse, Jong, and Kervel (2011) find similar results analyzing European markets. The latter study distinguishes between dark and lit fragmentation, and reports that only lit fragmentation enhances liquidity.

\textsuperscript{7}For example, the SEC motivates Reg NMS by stating that: “Mandating the consolidation of order flow in a single venue would create a monopoly and thereby lose the important benefits of competition among markets. The benefits of such competition include incentives for trading centers to create new products, provide high quality trading services that meet the needs of investors, and keep fees low.”

\textsuperscript{8}Following this concept, the Canadian market regulator, the Investment Industry Regulatory Organization of Canada, adopted a similar rule in 2011.
lowest in countries with fragmented markets and price protection (the U.S. or Canada).

At the core of the model lies the idea that, everything else being equal, all investors are (at least weakly) better off by trading faster, but not all investors value speed equally. Trading venues can thus make efforts to differentiate their trading platforms to cater to different clienteles. For example, a retail investor or pension fund manager may not be very sensitive to speed changes measured in small sub-second time intervals. On the other hand, for sophisticated investors such as derivatives market makers or equity index arbitrageurs, speed may be central to their business model. This fact motivates the modeling of heterogeneous agents. We capture investor heterogeneity in a parsimonious fashion by making private valuation processes heteroskedastic, which we interpret as originating in investor-specific leverage levels.\footnote{Leverage is a universal feature of financial markets, the banking sector, and real markets (e.g. mortgages in real estate equity). Admittedly, our analysis of leverage is stylized since we do not model the origin of leverage differences. Frazzini and Pedersen (2010) study a setting where leverage differences arise due to institutional constraints.}

Using our building idea, we seek to analyze several market designs (summarized in Table I) within an integrated framework. To understand the shared underlying economics, consider the basic three-stage sequence represented in Figure 1. Each stage represents a conceptually different period in the spirit of Alfred Marshall’s (1890) theory of production classification\footnote{According to Marshall, some but not all production inputs can change in the short-run. Investor participation but not trading technologies, both inputs of a hypothetical liquidity production function, can adjust in our model’s short-run. Our trading period corresponds to what Marshall called the market period, a period during which all input supplies are fixed.}:

- **Trading period:** All determinants of liquidity are taken as given by investors. One or more continuous-time markets for the asset open and participants trade.
- **Short-run:** Given trading technologies, one or two exchanges set access fees. Investors make participation and trading location decisions, and pay fees accordingly.
- **Long-run:** Given the number of exchanges, speed investment decisions are made.

For example, the trading period can represent any given trading day, the short-run may represent one or more quarters, and the long-run one or more years. In Section IV.B we add a regulation component to the long-run analysis to study market linkages.

To understand the pricing consequences of market designs and investments, we derive instructive decompositions of equilibrium prices and calibrate the model using U.S. equity data from the 2000s. In our framework, deviations of the market price from a frictionless Walrasian counterpart are captured by two quantities. The first is an illiquidity discount (ILD) that reflects the degree of imperfection in market contact, and is driven by trading...
technologies, as well as its shadow valuation, which depends on the marginal investor in the market. The second, which we label limited participation distortion (LPD), corresponds to the difference between the market and Walrasian prices in the absence of market contact frictions. This term arises solely because of the exchange or exchanges’ market power, and partly reflects how levered market participants are.

To illustrate the calibration results, let us consider the effect of investments in trading technologies, keeping the number of markets fixed. We find that in the long-run, investments increase the asset price nearly 3.2% in a single exchange economy, but only 1.3% on average in a duopolistic economy. These values reflect the fact that speed is more valuable to highly levered investors, who populate the monopolist market. Consider next an economy where an entrant exchange breaks a monopoly. We find that the equilibrium asset price falls by nearly 12% in the short-run, and 14% in the long-run. The effect in the short-run is dominated by the LPD, which sharply decreases in the transition to a duopoly. In the long-run, investments allow exchanges to improve asset liquidity, reducing the ILD. Moreover, vertical differentiation of their liquidity services allows exchanges to relax Bertrand competition and increase access fees, which partially offsets the initial effect on the LPD. Finally, consider a fragmented market where a trade-through rule is implemented. We find that the national best price is in the long-run 128 basis points (bps) lower than the unregulated average price.

Turning to the market environment, we find that in the long-run prices are negatively related to the cost of trading technologies. Interestingly, such costs only affect market participation when markets are fragmented. An increase in investors average leverage raises asset prices in both the short and long-run. The effect is stronger in the long-run due to the fact that exchanges can extract higher rents from investors by investing in faster platforms, rendering the asset more liquid.

Our model also provides novel predictions about trading fragmentation levels. Using a simple Herfindahl-Hirschman Index (HHI) to measure fragmentation, we find that its level (i) decreases as technology becomes cheaper, (ii) increases with positive technology shocks, (iii) decreases with more frequent preference shocks, and (iv) is higher under price protection than in unregulated markets. Turning to trading volume, we find that fragmentation can
Table I: Organization of the Liquidity and Asset Markets

<table>
<thead>
<tr>
<th>Trading Linkages</th>
<th>Consolidated</th>
<th>Fragmented</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liquidity Markets (&quot;Exchanges&quot;)</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Asset Markets</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

have significant equilibrium effects. In the long-run equilibrium, fragmented markets achieve approximately 88% of the Walrasian volume, which is more than twice the volume of a monopolist exchange. The volume effect of investor protection is negative but moderate.

Although we largely concentrate on exchanges' investments, we also analyze the effect of investors acquiring fast trading technologies directly (Section III.B). When this is possible, speed investments naturally depend on investors' characteristics, yielding a market with heterogeneous frequency traders. We show that in this case, asset prices may *increase* with asset supply: Unlike in Walrasian analysis where the market structure is given, an increase in supply can incentivize investors to become more technologically sophisticated, rendering assets more liquid, and increasing prices in equilibrium. The economic intuition is related to Acemoglu (1998) in the context of labor economics and directed technological change.

In summary, this paper makes four key points about secondary markets' frictions and asset pricing. First, investors with identical mean valuations for the asset but heterogeneous leverage can rationally make different market participation decisions. When exchange institutions have market power, the resulting distortion in the composition of active participants can drive the asset price higher than its frictionless level. This effect stresses the fact that equilibrium valuations do not directly reflect the stream of cash flows but, rather, the consumption stream that asset ownership generates. Second, the relationship between liquidity and price evolutions crucially depends on the competitive structure in financial markets. Liquidity can increase over time because of technical progress or increased competition between markets. However, when markets fragment, asset prices can evolve in the opposite direction. This result may shed new light on recent empirical results that link fragmentation with enhanced market quality. Third, regulations that protect investor executions distort competition between trading venues and thus have equilibrium consequences on asset prices. The result here should be of interest for policy makers evaluating optimal regulations in this regard. Finally, our results illustrate that differences in leverage can have important asset-pricing effects and can amplify or diminish illiquidity effects. Understanding their effects on financial markets is crucial, given the increasingly important role of institutions in portfolio and investment decisions.

**Relation to the Literature** Garbade and Silber (1977) provide early work on the issue
of separated markets and the role of speed.\textsuperscript{11} Theoretical analyses of fragmentation include those of Mendelson (1987), Pagano (1989), and Madhavan (1995). These early papers focus on the trade-off between liquidity externalities and market power. Our focus on differentiation is in the spirit of Harris (1993), who argues that markets fragment partly because not all traders solve the same problem. Importantly, these papers analyze liquidity and price informativeness. We contribute to this literature by characterizing the effect of fragmentation on equilibrium asset price levels. To the best of our knowledge, we also provide the first formal analysis of the effects of investor protection on asset prices.

This paper centers around liquidity and the pricing implications of institutions’ responses to alleviate frictions. A different focus is adopted by Pagnotta and Philippon (2012) in a companion paper. These authors study the welfare implications of speed-driven competition, the entry of exchanges, and optimal market design. Since our focus is on positive pricing and quantity implications, our study examines price formation where investors are also able to acquire technologies to reduce their individual “distance” to a market center. Appendix B provides a version of the model with generalized asset holdings and preferences.

Exchanges can differentiate in areas other than speed.\textsuperscript{12} For example, Santos and Scheinkman (2001) study competition in margin requirements, while Foucault and Parlour (2004) analyze competition in listing fees.\textsuperscript{13} These papers do not analyze speed differentiation and thus justifiably consider static frameworks. Our focus on technological speed reflects its prominent role in modern asset markets and its direct relation with secondary market liquidity and explains our effort in developing a suitable dynamic model where speed plays an explicit role.

Our paper complements the vast literature that analyzes the asset-pricing effects of illiquidity. Important theoretical work in this area includes that of Amihud and Mendelson (1986), Constantinides (1986), Vayanos (1998), Lo, Wang, and Mamaysky (2004), Eisfeldt (2004), Acharya and Pedersen (2005), among many others.\textsuperscript{14} We contribute to this literature by developing links between the origin of liquidity in secondary markets (i.e. competitive structure, investor participation and technology) and asset price levels. Our trading model

\textsuperscript{11}Silber and Garbade offer a historical perspective of speed, that surprisingly resembles in spirit many current developments. They analyze the impact of the domestic telegraph system between the New York Stock Exchange (NYSE) and the Philadelphia Stock Exchange during the 1840s and the transatlantic cable between New York and London in 1866.

\textsuperscript{12}Models of vertically differentiated oligopolies have been pioneered by Gabszewicz and Thisse (1979) and Shaked and Sutton (1982). Unlike classical models, we endogenize the value of “quality” (trading delays here) through a microfounded trading game. In particular, we enrich the basic setup by having investors “consuming” a differentiated product first (liquidity) and a homogeneous product (asset cash flows) second.


\textsuperscript{14}Amihud, Mendelson, and Pedersen (2006) provide a thorough survey.
approach is closest to the part of this literature that incorporates search-like frictions, which has been fostered by Duffie, Garleanu, and Pedersen (2005).\textsuperscript{15} In particular, our trading model is in the spirit of Lagos and Rocheteau (2009), whose model we extend by incorporating heterogeneously levered agents.\textsuperscript{16} We complement this literature by endogenizing several important aspects of the market structure. To the best of our knowledge, this paper is the first within this class to analyze liquidity and pricing of a single security across different trading venues.\textsuperscript{17}

Key to our results are participation decisions and the composition of investors in the market.\textsuperscript{18} Consistent with the model’s outcomes, using microdata, Mankiw and Zeldes (1991) and Vissing-Jorgensen (2002) find that the preferences of stock market participants differ greatly from those of non-market participants.

The remainder of the paper is organized as follows. Section I presents the empirical motivation for our model. Section II introduces the theoretical model in the case of a single trading venue and Section III derives the theoretical results for asset prices in this setting. Section IV analyzes competition among a given set of trading venues, the allocation of investors across these venues, and the resulting asset prices. It also analyzes the effects of investor protection on asset prices. Section V presents empirical implications for asset prices, trading volume and fragmentation, as well as a calibration that illustrates the size of the effects. Section VI discusses relationships between our results and previous findings. Section VII concludes the paper. Appendix A describes different investor protection regulations and Appendix B develops an extension of the trading model with unrestricted asset portfolios. All proofs are provided in the Online Appendix.

I Empirical Trends That Motivate the Theory

This section presents a brief account of the empirical facts that motivate our theory.\textsuperscript{19} It also illustrates some of the speed choices that investors face in modern markets. See Appendix

\textsuperscript{15}Trejos and Wright (2012) discuss this growing literature, see the references therein.
\textsuperscript{16}While the search literature in finance mainly concentrates on over-the-counter markets, Weill (2007) and Biais, Hombert, and Weill (2012) also use a related framework to analyze trading in exchanges. Garleanu, Pedersen, and Poteshman (2009) also uses Poisson contact times to analyze the effect of trading frictions on prices.
\textsuperscript{17}Vayanos and Wang (2007) and Weill (2008) use multiple assets but not multiple trading venues.
\textsuperscript{18}Contributions to the analysis of participation costs in financial markets include the works of Brennan (1975), Merton (1987), Allen and Gale (1994), Chatterjee and Corbae (1992), and more recently, Huang and Wang (2009). Within the search tradition, Afonso (2011) studies a version of the model of Vayanos and Wang (2007) with flow investor entry in the presence of thick market and congestion externalities.
\textsuperscript{19}The Securities and Exchange Commission (2010) presents an informative summary of these trends and a discussion of the challenges they pose to market quality.
A for a discussion of some of the important regulations that affect the evolution of speed and fragmentation.

Trading Speed. Major market centers have made costly investments in fast, computerized trading platforms to reduce order execution and communication latencies. This process has gone beyond stock exchanges to include futures, options, bonds, and currencies. Although these investments were chiefly first observed in the U.S., they have more recently been undertaken on a global scale. This trend accelerated during the second half of the 2000s, as Figure 2 illustrates with the average execution speed in the NYSE. The driving forces underlying this speed frenzy are likely to be different from those of other historical periods. In the human-driven trading era, for example, higher execution speeds helped reduce moral hazard problems between, say, floor brokers and investors. However, this aspect has become less relevant today.

In our model, some investors decide to pay a premium for speed. In real markets this idea relates to several dimensions of the investment process. Some important examples are the following.

• Colocation. Colocation is a service offered by trading centers that operate their own data centers and by third parties that host the matching engines of trading centers. The trading center rents rack space to investors, which enables them to place their servers in close physical proximity to a trading center’s matching engine, which helps minimize network and other types of delays.

• Exchanges versus alternative venues. According to the SEC classification, U.S. investors can opt to direct their trading orders to registered exchanges (such as the NASDAQ), Electronic Communication Networks, dark pools, or Broker-Dealer Internalizers. Although alternative and over-the-counter venues have also made significant technological progress, organized exchanges typically offer investors the fastest communication and trading responses.

• Markets with speed restrictions. In foreign exchange markets, for example, many sophisticated traders concentrate on Electronic Broking Services and Reuters inter-dealer brokerage platforms, both of which have minimum quote life or minimum fill ratios. One exchange that does not have such a minimum is Currenex, which is therefore particularly attractive to high-frequency traders.

• Inter-market connectivity. The provision of connections between financial centers plays a key role in modern markets. Take Chicago-New York as an example, an essential
route for derivative traders. In 2011 a firm named Spread Networks invested approximately $300 million in a new fiber optic cable that links these cities through the straightest possible route, saving about 100 miles over existing routes. This allows the company to shave 6 milliseconds off their delay, for a total delay of 15 milliseconds. Clients of these superior connectivity services are willing to pay premium fees for their use.

**Fragmentation.** Securities trading, especially in North America and Europe, has become significantly more fragmented over the last decade. Figure 3 illustrates this trend. Traditional markets in the U.S. and Europe have lost market share to (usually faster) entrants such as Direct Edge, BATS and Chi-X. For instance, the fraction of NYSE-listed stocks traded at the NYSE decreased from 80% in 2004 to just over 20% in 2009. Similar patterns have been observed in other asset classes. For example, there are more than 10 options exchanges in the U.S. alone. Overall, fragmentation has increased so dramatically that market participants now keep track of fragmentation indexes across asset classes and countries. Figure 3 also illustrates that the levels of fragmentation in other regions of the world are still well behind those of North America and Europe. Given that several developing countries are currently trying to foster competition between exchanges, the results in this paper may help them anticipate effects on prices, liquidity and trading volumes.

---

**Figure 2: NYSE Average Order Execution Speed (Seconds)**
Source: NYSE SEC Rule 605 reports (small orders executed at market; 5% tails excluded from the average).

---

20 The success of this business model motivated McKay Brothers, a leading provider of low-latency wireless transport equipment, to contract the creation of a $300 million microwave-based network of towers connecting the same cities to Aviat Networks.

21 See, for example, the Fidessa fragmentation indexes
Figure 3: Changes in Fragmentation Levels in Global Equity Markets

The Herfindahl-Hirschman Index of market concentration (the sum of squares of market shares; e.g., for the NYSE in 2012, the sum of squares of market shares in NYSE-listed stocks trading for the NYSE, Nasdaq, BATS, DirectEdge and others) for a cross-section of equity exchanges. Market share is defined as the proportion of exchange-listed shares traded through the exchange itself and excludes dark pools.

Sources: Exchanges websites, RBC Global Asset Management, Bloomberg, and HSBC Research.

II Trading Model

This section provides explicit micro foundations for how investors demand speed in financial markets. We first analyze the problem of an investor trading in a single cost-free market and characterize the ex ante value of speed. We then characterize an equilibrium where investors decide whether to participate in such market by paying an access fee and a single exchange sets the fee to maximize profits.

A Preferences and Technology

We start by describing the main building blocks of our model: investor preferences and trading technology. The preferences need to incorporate heterogeneity to create gains from trade as well as interesting participation decisions among exchanges. The trading technology must capture the role of speed in financial markets.

Time is continuous and runs forever. The model has a continuum of heterogeneous investors, two goods, and one asset. The measure of investors is normalized to one and their preferences are quasilinear. The numéraire good (cash) has a constant marginal utility normalized to one and can be freely invested at a constant rate of return $r$. We restrict asset holdings to
The asset is in fixed supply, \( \bar{a} \), which is normalized to lie in \([0, 1]\) so that it also represents the expected endowment of each investor. That is, before trading, each investor is endowed with one unit of the asset with probability \( \bar{a} \).

One unit of asset pays a constant dividend equal to one of a perishable non-tradable good. The flow utility that an investor derives from holding \( a_t \) units of the asset at time \( t \) is

\[
    u_{\sigma,\epsilon_t}(a_t) = (\mu + \sigma \epsilon_t) a_t,
\]

where \((\sigma, \epsilon_t)\) denotes the type of investor. This type is defined by a fixed component \( \sigma \) and a time-varying (random) component \( \epsilon_t \). The fixed component \( \sigma \in [0, \bar{\sigma}] \) is known at time zero and distributed according to the twice-differentiable cumulative distribution \( G \), with a log-concave density function \( g \) that is positive everywhere. The type \( \epsilon_t \in \{-1, +1\} \) changes randomly over time. The times when a change can occur are distributed exponentially with parameter \( \gamma \). Conditional on a change, the temporary shocks are independent and identically distributed, and the value \( \epsilon = 1 \) occurs with probability \( \phi(1) = \phi \in (0, 1) \).

Our paper focuses on the trading technology for the asset. For clarity, we describe here the case where all investors trade at the same speed (later we endogenize speed choices and consider markets with different speeds). The market where investors trade the asset is characterized by a constant contact rate \( \rho \). Conditional on being in contact, the market is Walrasian and clears at a price \( p \). Investors that are not in contact simply keep their holdings constant.

Our assumptions about technology and preferences imply that the value function of a class-\( \sigma \) investor, with current valuation \( \epsilon(t) \), and current asset holdings \( a \) at time \( t \) is

\[
    V_{\sigma,\epsilon_t}(a, t) = \mathbb{E}_t \left[ \int_t^T e^{-r(s-t)} u_{\sigma,\epsilon_s}(a) ds + e^{-r(T-t)} (V_{\sigma,\epsilon_T}(a_T, T) - p_T (a_T - a)) \right],
\]

where the realization of the random type at time \( s > t \) is \( \epsilon(s) \), \( T \) denotes the next time the investor makes contact with the market, and \( a_T \) corresponds to optimal time \( T \) holdings. Expectations are defined over the random variables \( T \) and \( \epsilon(s) \) and are conditional on the current type \( \epsilon(t) \) and the speed of the market \( \rho \).

\[22\]See Appendix B for a generalization of the model where the equilibrium price is found relaxing this assumption.

\[23\]All random variables are defined on a probability space \((X, \mathcal{F}, \mathbb{P})\), and all random variables at time \( t \) are measurable with respect to the filtration \( \{\mathcal{F}_t : t \geq 0\} \) representing the information commonly available to investors.
B Trading Equilibrium and the Value of Speed

We characterize here the trading equilibrium with free market participation where all investors join the market. We show that the asset price remains constant during the trading game, and the value functions are thus time independent and satisfy

\[
r V_{\sigma,\epsilon}(a) = u_{\sigma,\epsilon}(a) + \gamma \phi (\epsilon' \neq \epsilon) [V_{\sigma,\epsilon'}(a) - V_{\sigma,\epsilon}(a)] + \rho [V_{\sigma,\epsilon} (a_{\sigma,\epsilon}^*(p)) - V_{\sigma,\epsilon}(a) - p (a_{\sigma,\epsilon}^*(p) - a)].
\]  

(3)

To analyze Equation 3, we then need to characterize the demand functions \( a_{\sigma,\epsilon}^* (p) \). Note that, on average, a proportion \( \phi \) of investors are of trading type \( \epsilon = +1 \) representing natural buyers, who are on the long side of the market when \( \phi \geq \bar{\phi} \). Following Duffie, Garleanu, and Pedersen (2005), we concentrate on the case where supply is short hereafter and treat the complementary case in Appendix B.\(^{24}\) To simplify the exposition we consider a symmetric shock structure by setting \( \phi = 1/2 \). Since supply is short, low-\( \sigma \) types always sell their entire holdings when they contact the market. Moreover, there is a marginal type \( \hat{\sigma} \) that is indifferent between buying and not buying when \( \epsilon = 1 \). Thus, we prove that when \( \phi > \bar{\phi} \) the demand system is as follows.

**Lemma 1.** The demand functions are \( a_{\sigma,\epsilon}^* = 0 \) when \( \epsilon = -1 \) or when \( \sigma < \hat{\sigma} \), and \( a_{\sigma,\epsilon}^* = 1 \) when \( \epsilon = +1 \) and \( \sigma \geq \hat{\sigma} \), where

\[
\hat{\sigma} (p, \rho) \equiv \left( 1 + \frac{\gamma}{r + \rho} \right) (rp - \mu).
\]  

(4)

Clearly, there cannot be an equilibrium where \( a_{\sigma,-} = 1 \) since in that case total demand would exceed \( \bar{\sigma} \) preventing market clearing. Note also that the asset holdings of types \( \sigma < \hat{\sigma} \) are non-stationary since they never purchase the asset. A type \( \sigma < \hat{\sigma} \) sells its holdings on first contact with the market and never holds the asset again. The value of \( \hat{\sigma} \) then naturally classifies investors into two categories, which we label as temporary investors \( (\sigma < \hat{\sigma}) \) and active \( (\sigma \geq \hat{\sigma}) \) investors.\(^{25}\) Over time the assets move from the low-\( \sigma \) types to the high-\( \sigma \) types and then keep circulating among high-\( \sigma \) types in response to \( \epsilon \) shocks and trading opportunities.

It is easy to see that the price remains constant along the transition path. Given the ex ante supply of the asset \( \bar{a} \) and the market contact rate \( \rho \), the gross supply of assets is always

\(^{24}\)Although binary asset holdings simplify the analysis greatly, this assumption is not without loss of generality. Like in Duffie, Garleanu, and Pedersen (2005), asset indivisibility generates an equilibrium price that decreases in the contact rate \( \rho \) when asset supply is large, making such case somewhat less compelling. We know from Lagos and Rocheteau (2009) that this extensive margin considerations are trivial when asset holdings are unrestricted. We provide an extension of our framework with \( a \geq 0 \) in Appendix B.

\(^{25}\)An alternative classification involves small- and high-volume investors, which would be natural under a generalization with unrestricted asset holdings.
\( \rho \bar{a} \). All negative trading types \( \epsilon = -1 \) want to hold \( a = 0 \) and they represent half of the traders. The trading types \( \epsilon = +1 \) want to hold one unit if \( \sigma > \hat{\sigma} \) and nothing if \( \sigma < \hat{\sigma} \). The gross demand from high types is then always equal to \( \rho (1 - G(\hat{\sigma})) / 2 \). The market clearing condition, which holds at all times, is therefore

\[
\frac{1}{2} [1 - G(\hat{\sigma})] = \bar{a}. \tag{5}
\]

Let \( \alpha_{\sigma,\epsilon}(a) \) be the share of class-\( \sigma \) investors with trading type \( \epsilon \) currently holding \( a \) units of asset. Using the demand system in Lemma 1, and the market clearing condition 5, we can now characterize the steady-state distribution among types \( \sigma > \hat{\sigma} \) and the single-market equilibrium price as follows.

**Proposition 1.** The single market, full-participation trading equilibrium is as follows.

- **When** \( \bar{a} < \frac{1}{2} \) **there is a unique equilibrium price given by**

\[
p = \frac{\mu}{r} + \frac{G^{-1}(1 - 2\bar{a})}{r} \left( \frac{r + \rho}{r + \gamma + \rho} \right) \tag{6}
\]

- **Transition dynamics:** The price remains constant while asset holdings shift from low \( \sigma \)-types to high \( \sigma \)-types. Low types \( \sigma < \hat{\sigma} \) sell their initial holdings and never purchase the asset again. High types \( \sigma \geq \hat{\sigma} \) buy when \( \epsilon = 1 \) and sell when \( \epsilon = -1 \).

- **The distribution of holdings among high \( \sigma \)-types converges to the steady-state distribution of well-allocated assets** \( \alpha_{\sigma,+}(1) = \alpha_{\sigma,-}(0) = \frac{1}{2} \frac{2(\epsilon+\gamma)}{\gamma+\rho} \) and misallocated assets \( \alpha_{\sigma,+}(0) = \alpha_{\sigma,-}(1) = \frac{1}{4} \frac{\gamma}{\gamma+\rho} \).

- **When** \( \rho \to \infty \), allocations and the equilibrium price converge to the Walrasian outcome \( \alpha_{\sigma,+}(0) = \alpha_{\sigma,-}(1) = 0 \) and \( p_W = \frac{1}{r} \left[ \mu + G^{-1}(1 - 2\bar{a}) \right] \).

With full participation the value of \( \hat{\sigma} = G^{-1}(1 - 2\bar{a}) \) can be derived directly from the market clearing condition, and thus only depends on the asset supply and the distribution of investor types. Consequently, the frictionless Walrasian price depends only on these variables.

Our next task is then to compute the value of speed for investors. We proceed in two steps: we first compute the steady-state value functions for active investors, and later compute the ex ante values, taking into account the transition dynamics. Consider the steady-state value functions for any type \( \sigma > \hat{\sigma} \). Given the Bellman equation of Equation 3 and Lemma 1, for
the types holding assets, we have
\[ rV_{\sigma^+} (1) = \mu + \sigma + \frac{\gamma}{2} [V_{\sigma^-} (1) - V_{\sigma^+} (1)] \]
\[ rV_{\sigma^-} (1) = \mu - \sigma + \frac{\gamma}{2} [V_{\sigma^+} (1) - V_{\sigma^-} (1)] + \rho (p + V_{\sigma^-} (0) - V_{\sigma^-} (1)). \] (7)

For the types not holding the assets, we have
\[ rV_{\sigma^+} (0) = \frac{\gamma}{2} [V_{\sigma^+} (0) - V_{\sigma^-} (0)] \] (9)
\[ rV_{\sigma^-} (0) = \frac{\gamma}{2} [V_{\sigma^-} (0) - V_{\sigma^+} (0)] + \rho (V_{\sigma^+} (1) - V_{\sigma^+} (0) - p). \] (10)

It is convenient at this point to consider the change of variables given by
\[ s (\rho) \equiv \frac{\rho}{r + \gamma + \rho}. \] (11)

The variable \( s \in [0, 1] \), which we denote effective speed, is economically more informative than \( \rho \) alone, since it reflects the market’s ability to reallocate the asset between investors for a given preference shock rate \( \gamma \).

We can find the ex ante participation value, denoted \( W \). For a given investor of type \( \sigma \) joining a market with effective speed \( s \) and a marginal investor given by \( \hat{\sigma} \), this function is
\[ W (\sigma, \hat{\sigma}, s) \equiv \frac{\bar{a}}{2} \sum \epsilon V_{\sigma, \epsilon} (1) + \frac{1 - \bar{a}}{2} \sum V_{\sigma, \epsilon} (0). \] (12)

The interpretation of the right-hand side of Equation 12 is straightforward: The investor is endowed with one unit of the asset with probability \( \bar{a} \), and the probability of a temporary type \( \epsilon \) is one-half. The no-trade outside option of any investor is
\[ W_{\text{out}} = \frac{\mu \bar{a}}{r}. \] (13)

Naturally, given the symmetric shock structure, \( W_{\text{out}} \) does not depend on \( \sigma \).

Given the system of Equations 7-10, we have the following.

**Lemma 2.** The function \( W \) for a class-\( \sigma \) investor is given by
\[ W_{\text{Ex ante value}} (\sigma, \hat{\sigma}, s) = \frac{\mu \bar{a}}{r} + \frac{s \bar{a} \hat{\sigma}}{r} + \frac{s}{2r} \max (0, \sigma - \hat{\sigma}) \] (14)

where \( s \) is defined by Equation 11, and the marginal type \( \hat{\sigma} (p, \rho) \) is as in Equation 4.
The intuition is that the market participation gains, $W$ net of the outside option $W_{out}$, is composed of two parts. The (transient) ownership value $u\hat{a}+s\hat{a}\hat{\delta}$, independent of $\sigma$ is the value that can be achieved by all types $\sigma < \hat{\sigma}$ with the “sell and leave” strategy. The second part, $\frac{s}{2\hat{\sigma}} \max (0; \sigma - \hat{\sigma})$, represents the surplus of trading forever, which depends on the type $\sigma$. Importantly, the latter part is super-modular in $(s, \sigma)$. Hence, the value assigned to any effective speed $s$ increases with the investor type.

C Equilibrium with Costly Participation

We can now analyze investors’ participation decisions and formally define an equilibrium with costly participation. Let $q_i$ be the cost of accessing venue $i$, and let $P(q, s|\sigma)$ be the participation mapping onto $\{0, 1, 2, ..., I\}$, where $P = 0$ means staying out, $P = 1$ means joining venue 1, and so forth. Staying out costs nothing, so we take $q_0 = 0$. Recall that $G$ was the ex ante distribution of permanent types. Let $\tilde{G}_i(\sigma)$ be the measure of trader types lower than $\sigma$ in market $i$. If all potential investors join market $i$, as before, we simply have $\tilde{G}_i = G$. In the generic case, however, we have $\tilde{G}_i \leq G$ since some investors may not participate. Indeed, we shall see that in the multiple-venue model the distribution $\tilde{G}$ is typically discontinuous. We then have the following.

Definition 1. An equilibrium is a set of market access fees $q^* = (q_1^*, ..., q_I^*)$, effective speeds $s = (s_1, ..., s_I)$, participation decisions $\{P(q, s|\sigma), \sigma \in [0, \bar{\sigma}]\}$ by investors, asset prices $p_i(q, s, \tilde{G}_i)$ and marginal types $\hat{\sigma}_i(q, s, \tilde{G}_i)$ such that for all $i \leq I$

- $q_i^* \in \arg \max_{q_i} \int_{0}^{\bar{\sigma}} P(q_i, q_{i+1}, s|\sigma) \ dG(\sigma)$
- $P(q, s|\sigma) \in \arg \max_{x \in \{0, ..., I\}} [W(\sigma, \hat{\sigma}_i, s_i) - q_i]$
- $p_i(q, s, \tilde{G}_i)$ and $\hat{\sigma}_i(q, s, \tilde{G}_i)$ solve $\hat{\sigma}_i = (rp_i - \mu) \left(1 + \frac{s}{r + p_i}\right)$ and $\frac{1}{2} \left(\tilde{G}_i(\bar{\sigma}) - \tilde{G}_i(\hat{\sigma}_i)\right) = a\tilde{G}_i(\bar{\sigma})$, where $\tilde{G}_i(\sigma) \equiv \int_{0}^{\sigma} 1 \{P(q, s|x) = i\} \ dG(x)$

This equilibrium concept naturally extends to the case in which the speed $s$ is also endogenous. With only one venue, the investor with marginal trading type $\hat{\sigma}$ must simply be indifferent between joining and not joining the market. So we must have $W(\hat{\sigma}, \hat{\sigma}, s) - W_{out} = q$ and therefore $q = \frac{s\hat{\sigma}\hat{\delta}}{\bar{\sigma}}$. Consequently, all types below $\hat{\sigma}$ are indifferent between joining and staying out since they obtain zero net participation gains. Let $\delta$ be the mass of investors that join, sell, and leave. Market clearing requires $\delta$ to equal $(1/2\pi - 1)(1 - G(\hat{\sigma}))$. This condition holds at an interior solution as long as $\delta < G(\hat{\sigma})$, or in other words, as long as $\frac{G(\hat{\sigma})}{1 - G(\hat{\sigma})} > \frac{1}{2\pi} - 1$. In the remainder of the paper we assume that either $\bar{a}$ is close enough to
1/2 or that there is a sufficient mass of low-type investors to ensure the existence of interior solutions.

To derive analytical results, when convenient we assume the following distribution.

**Assumption 1 (A.1).**

\[
G(\sigma) = 1 - e^{-\frac{\sigma}{\nu}}, \quad \nu > 0
\]  

(15)

Note that under A.1 the Walrasian price is simply given by \( \frac{\mu}{r} [\mu - \nu \log (2\sigma)] \).

Using Equation 11, we can express the equilibrium price 6 as

\[
p = \frac{\mu}{r} + \frac{\hat{\sigma}}{r} \left( \frac{r + \gamma s}{r + \gamma} \right).
\]  

(16)

With costly participation we generally have \( \hat{\sigma} > G^{-1}(1 - 2\sigma) \). It is worth noting that the equilibrium price 16 differs from the benchmark liquidity-adjusted price in the literature (e.g., Duffie, Garleanu, and Pedersen (2005)) in two key aspects:

1. Investors participation decisions that affect \( \hat{\sigma} \) are determined endogenously. Naturally, these decisions will be interrelated with the exchange(s) optimization problem in equilibrium.

2. Market contact frictions captured by \( s \) will also be endogenously determined.

Consequently, we can explicitly analyze how \( (\hat{\sigma}, s) \) are jointly determined as a function of: investor characteristics, the state of technology, the market competitive structure, and regulation. This is the main objective of the following two sections.

**D Discussion of Assumptions**

The \( \epsilon \)-shocks can capture time-varying liquidity demands, financing costs, hedging demands, or specific investment opportunities (see Duffie, Garleanu, and Pedersen (2007) for a discussion). The important point is that these shocks affect the private value of the asset, not its common value. For instance, a corporate investor may need to sell financial assets to finance a real investment. A household may do the same for the purchase of a durable good or a house. The parameter \( \sigma \) then simply measures the size of these shocks and thus the volatility of the private value process. One can interpret \( \sigma \) as capturing leverage levels for a given investor. For example, as a group, retail investors generally have lower leverage levels than institutional investors. Thus, retail investors would correspond to lower values of \( \sigma \) in the model. Moreover, note that individual preference types do not drive the endowment
process in any way. Random ownership of a tree share may represent, for example, the outcome of an unmodeled labor market where there is some sort of “stock compensation.”

The parameter $\gamma$ measures the mean reversion of the utility flow process and is assumed for simplicity to be the same for all investors. In the context of delegated management, the shock frequency can represent the sum of the shocks affecting all the investors in a given fund or brokerage house. We introduce heterogeneity in $\sigma$ and not in $\gamma$ because the key point in our analysis is the link between gains from trade and speed. It is important to understand that a higher value of $\gamma$ implies lower gains from trade. Investors with a high value of $\gamma$ are not eager to trade since they can simply wait for their type to mean-revert. In particular, a high value of $\gamma$ would not capture the idea of fleeting trading opportunities. This idea is better captured by a high value of $\sigma$.

The non-Walrasian feature of the market is captured in a parsimonious fashion by a single market contact rate $\rho$. Real markets are of course more sophisticated, but for tractability reasons we abstract from modeling the explicit connections between the exchange and, say, a population of potential market makers. One could add bargaining with market makers and bid–ask spreads, but this would not likely bring new insights beyond those of Duffie, Garleanu, and Pedersen (2005) and Lagos and Rocheteau (2009). A market mechanism similar to ours is considered in the monetary economy of Rocheteau and Wright (2005), which they label competitive equilibrium. We also abstract from liquidity (i.e., thick market) externalities. While this abstraction is not without loss of generality and liquidity externalities may still be relevant for some exchange-traded assets, technology may arguably help realize them, even when several trading venues coexist.\footnote{The fact that large cap stocks in the U.S. currently trade in nearly 50 different trading venues suggests that liquidity externalities are not as important or prevalent as they were in the pre-electronic “human trade” era.}

Because the focus of exchange differentiation is on trading speed, we interpret $q$ as ex ante costs related to a speed decision that affect multiple trading rounds rather than transaction cost related to a given trade volume. Such costs can include, for example, co-location fees at a particular trading center (see Section I for a discussion). The online appendix of Pagnotta and Philippon (2012) analyzes a model with transaction costs.

**III Asset Prices in a Consolidated Market**

In this section we derive equilibrium asset prices in an environment with a single profit-maximizing exchange. The single exchange seeks to maximize profit by selecting a given access fee and a given trading speed. Creating or adopting a new trading platform takes
After the exchange chooses a trading speed and access fee, investors decide whether to participate. Once these decisions are made, trading starts at time zero.

We specialize our definition of short and long-run periods in the introduction according to the timing illustrated in Figure 4. After the exchange chooses a trading speed and access fee, investors decide whether to participate. Once these decisions are made, trading starts at time zero. After analyzing the benchmark case, in the remainder of this section we study the case where investors, as opposed to the exchange, have the ability to reduce trading delays.

### A Speed Investments and Asset Price

**Short Run Exchange Program and Asset Price** Consider the case of a market that in the short-run operates under an exogenously given “default” speed $\rho > 0$ and seeks to maximize profits by selecting an access fee. In this case the single exchange behaves like a classic monopolist and total profits are given by $\pi = q \left( 1 - G(\hat{\sigma}) + \delta \right)$, which we can write using the market clearing condition 5 as $\pi = \frac{q}{\bar{a}} \left( 1 - G(\hat{\sigma}) \right)$. Note that if $\bar{a} = 1/2$ we obtain $\delta = 0$, the simplest case to analyze. When $\bar{a}$ is less than $1/2$, we simply need to remember that a mass $\delta$ of investors joins, sells, and becomes inactive. The program of the monopolist then has a first-order condition given by

$$\hat{\sigma}_{\text{con}} = \left[ 1 - G(\hat{\sigma}_{\text{con}}) \right],$$

where the subscript $\text{con}$ denotes a market with consolidated trading. Importantly, the implied marginal type depends only on the distribution function $G$. Denoting the price with consolidated trading as $p_{\text{con}}$ and using a superscript $S$ to denote the short-run, we obtain

$$p_{\text{con}}^S = \frac{\mu}{r} + \hat{\sigma}_{\text{con}}^S \times \left( \frac{r + \gamma s_0}{r + \gamma} \right),$$

where as before $s_0 = \frac{\rho}{r + \gamma + \rho}$ and $\hat{\sigma}_{\text{con}}^S$ is given by Equation 17. Note that under A.1 we have $\hat{\sigma}_{\text{con}}^S = \nu$. 

---

*Figure 4: Timing Consolidated Market*

After the exchange chooses a trading speed and access fee, investors decide whether to participate. Once these decisions are made, trading starts at time zero.
**Long Run Exchange Program and Asset Price** In the long-run the exchange can adapt its trading technology \( s \) at a cost \( C(s) \), affecting the quality of its liquidity service and potentially extracting higher rents from investors. Profit maximization of the exchange requires solving

\[
\max_{(q,s) \geq 0} q \left( 1 - G(\hat{\sigma}(q,s)) \right) - C(s).
\]  

(19)

Where convenient, we assume the following cost function.\(^\text{27}\)

**Assumption 2 (A.2).** The cost of the contact rate \( \rho \) is given by \( c \times \max \{ \rho - \bar{\rho}; 0 \} \), where \( c > 0 \). Equivalently, the cost of the effective speed \( s \) is

\[
C(s) = c \times \max \left\{ \left( r + \gamma \right) \frac{s}{1-s} - \bar{\rho}; 0 \right\}.
\]

(20)

In words, under A.2 investment costs are linear in speed improvements \( \rho - \bar{\rho} \) and thus, by Equation 11, a convex function of \( s \).

The analysis of the solution to the program 19 yields the characterization of the equilibrium price \( p_{con} \) in the long-run. To gain intuition on the differences between the equilibrium and the benchmark frictionless price \( p_W \), we define the following quantities.

**Definition 2.** The limited participation distortion (LPD) of the market price in exchange \( i \) is given by

\[
LPD \equiv \lim_{c \to 0} [p_i - p_W].
\]

Let \( \lambda_i \equiv \frac{r + \gamma_i}{r + \gamma} \). The illiquidity discount (ILD) of price \( p_i \) is given by

\[
ILD \equiv \hat{\sigma}_i (1 - \lambda_i).
\]

The LPD represents the difference between the equilibrium and Walrasian prices in the absence of market contact frictions. That is, it captures the distortion in the value of the marginal investor due to the exchange market power. Naturally, this quantity vanishes when the market access cost approaches zero. The ILD captures the value of the losses that the market marginal investor faces due to her temporary inability to rebalance her portfolio after

\(^{27}\)Since we are focusing on long-term investment in trading infrastructures, we concentrate on fixed costs of speed enhancements which represent items such as hardware, the development of matching algorithms, integration with clients systems, among others, as opposed to variable costs such as energy or maintenance. The cost of a given technology also does not depend on the number of traders that participate in the exchange. This is, of course, a simplification but, arguably, is does not involve much loss of generality in electronic markets, as opposed to the era of trading floors.
a preference shock. Note that $1 - \lambda_i$ can be seen as a measure of market illiquidity, and has a maximum value of zero when the effective speed $s_i$ equals one.\textsuperscript{28} Thus, in the remainder of this section we assume A.1 and A.2, as given by Equations 15 and 20, to obtain explicit expressions.

**Proposition 2.** The long-run equilibrium price in a consolidated market is given by

$$
p_{con}^L = \frac{1}{r} \left[ \mu - \nu \log (2\pi) \right] \left( \text{Walrasian Price} \ (p_W) \right) + \frac{\nu}{r} \left[ 1 + \log (2\pi) \right] \left( \text{Limited Participation Distortion} \right) - \frac{\nu}{r} \left( \gamma \sqrt{\frac{2rc}{(r + \gamma) \nu}} \right) \left( \text{Illiquidity Discount} \right)
$$

(21)

When market contact frictions are small, the market consolidated price $p_{con}$ is higher than the frictionless price $p_W$.

The equilibrium long-run price has three components. The first is the Walrasian price $p_W$, which corresponds to the limit price where market contact frictions vanish and where market participation is costless for investors. The second and third terms correspond to the limited participation and illiquidity adjustments. With a single exchange, the LPD depends on the distribution of investors, as implied by Equation 17, both in the short and long-run, so we have $\hat{\sigma}_{con}^L = \hat{\sigma}_{con}^S (= \nu$ under A.1). The relative value of the distortion also depends on the asset supply and is largest when $\bar{a} = 1/2$. This is because the single exchange chooses $\hat{\sigma}$ irrespective of $\pi$, while $\hat{\sigma}_W$ decreases with asset supply.\textsuperscript{29} The size of the ILD depends on the extent of illiquidity, captured by $1 - \lambda_{con}$, and its shadow price $\hat{\sigma}_{con} = \nu$ corresponding to the marginal investor type. Note that the ILD approaches zero as market contact frictions become small. This occurs, for example, when $c$ approaches zero. In such cases we have $p_{con} > p_W$, which is a reflection of the levered composition of market participants in the market equilibrium.

**B Heterogeneous Frequency Trading and Asset Demand Curves**

We characterized in Section III.A the consolidated market equilibrium price in the long-run, where the key investment role is played by the exchange. In modern financial markets, institutional investors also undertake costly investments to ensure fast responses to trading signals and fast communication. For example, a given investment desk may acquire better

\textsuperscript{28}Since $s \in (0, 1)$, whenever $\gamma \gg r$, $\lambda$ also lies in the interval $(0, 1)$

\textsuperscript{29}This holds provided $\pi$ is sufficiently large, that is $\hat{\sigma}_{con} > G^{-1}(1 - 2\pi)$. Otherwise, we must have $\hat{\sigma} = G^{-1}(1 - 2\pi)$
Figure 5: Heterogenous Investor Speeds

Panel A shows that investors that participate in the market are ex ante identical in their "distance" from the market, $\frac{1}{\rho - \theta}$. Panel B shows that endogenizing technology choice leads to $\sigma$-dependent heterogeneity in contact rates.

Consider again a market with contact speed $\rho$, which investors take as given. We now assume that all investors experience an additional time delay, possibly due to equal geographical distance from the market center. In particular, the effective market contact rate for any investor is given by $\rho - \bar{\theta}$, with $\bar{\theta} < \rho$. Before joining the market, investors can increase their effective contact rate by investing in a latency reduction technology $\theta \in [0, \bar{\theta}]$, at a type-independent cost $c_I \theta$ with $c_I > 0$. Whenever investors select a technology $\theta$, their agent-specific effective speed then becomes

$$s(\rho,\theta) = \frac{\rho - (\bar{\theta} - \theta)}{r + \gamma + \rho - (\bar{\theta} - \theta)}.$$  \hfill (22)

In this environment investors’ pre-trade decisions include both a participation and a technology acquisition component:

$${\mathcal{P}}: [0, \bar{\sigma}] \rightarrow \{ 0, 1 \} \times [0, \bar{\theta}].$$

Figure 5 illustrates that, before trading starts, the effective distance from the market is investor specific. Anticipating investor behavior, the exchange then maximizes profits by
selecting an optimal market access fee $q$, as depicted in Figure 6.\textsuperscript{30}

In a symmetric equilibrium, all type-$\sigma$ investors choose the same technology $\theta (\sigma)$, so we can write $s (\sigma) \equiv s (\rho, \theta (\sigma))$. Note that once a particular effective rate $s (\sigma)$ has been determined, trading commences and the equilibrium analysis in Section II remains unchanged. Thus, we can express the expected participation value for a type-$\sigma$ investor selecting a speed $s (\sigma)$ as

$$W (\sigma, \hat{\sigma}, s (\sigma)) - W_{\text{out}} = \frac{\bar{a}}{r} s (\sigma) \hat{\sigma} + \frac{s (\sigma)}{2r} \max (0; \sigma - \hat{\sigma}),$$

(23)

which is super-modular in $(\sigma, \theta)$. Because investors will optimally select trading technologies, we can thus express the ex ante expected participation value as

$$\tilde{W} (\sigma, \hat{\sigma}, s (\sigma)) = \max_{\theta \in [0, \bar{\theta}]} \{W (\sigma, \hat{\sigma}, s (\theta (\sigma))) - C_I (\theta)\}.$$

(24)

Consequently, the marginal investor $\hat{\sigma}$ satisfies $\tilde{W} (\hat{\sigma}, \hat{\sigma}, s (\hat{\sigma})) - q = W_{\text{out}}$. Finally, note that the market clearing condition is not affected due to the independence of preference shocks across investor types.

Let $p_{hft}$ denote the equilibrium price in the market with “heterogeneous frequency traders.” By jointly solving the investor speed selection problem and the exchange optimal fee problem, we can show the following.

Proposition 3. When investors select trading technologies, the equilibrium price is

$$p_{hft} = \frac{1}{r} \left[ \mu - \nu \log (2\bar{\pi}) \right] + \frac{\hat{\sigma}_{hft}}{r} \left[ 1 + \nu \log (2\bar{\pi}) \right] - \frac{\hat{\sigma}_{hft} (1 - \lambda_{hft})}{r},$$

(25)

where $\hat{\sigma}_{hft}$ and $\lambda_{hft}$ are expressions given in the Online Appendix. When speed frictions are

\textsuperscript{30}Although one could analyze joint speed choices, to keep the analysis from being overly complicated, we take the market speed as fixed and concentrate on investor choice in this section.
small, investments in speed increase with the asset supply and asset demand curves can slope upward.

According to Proposition 3, investors’ choices affect both the LPD and ILD relative to \( p_{con} \) in Equation 21. In this case, the marginal investor type does not depend solely on the investor type distribution but also on parameters such as the cost of latency reductions. Consequently, we generally have \( \hat{\sigma}_{hft} \neq \hat{\sigma}_{con} = \nu \). The ILD is naturally affected by the change in the value of the marginal investor type, and by the marginal investor’s distance to the market (embedded in \( \lambda_{hft} \)). When market contact frictions are relatively small, the value of the ILD will be generally lower than in the previous section due to the fact that, for our calibrated parameters, we have \( \hat{\sigma}_{hft} < \hat{\sigma}_{con} \).

The equilibrium with heterogeneous frequency trading also has interesting implications for the shape of demand curves. Consider an increase in the asset supply parameter \( \bar{a} \). This change has a direct negative impact on the marginal investor type which tends to lower the asset price 25. However, such an increase has a positive effect on the optimal speed choice \( s_{hft} \), which in turn raises \( \hat{\sigma}_{hft} \) indirectly. The intuition behind this effect is clear when comparing the investor ex ante values given by Equations 14 and 23. In the latter, the value of ownership increases with the investor type \( \sigma \) through the interaction between \( \bar{a} \) and \( s(\sigma) \). In an economy where the asset supply incentive effect on \( s_{hft} \) is strong, one can observe, at least at low frequencies, that demand curves slope upward. An implication of this latter effect is that, in an environment with multiple assets, we would expect to see traders that heavily invest in speed concentrate in large assets (e.g., large market capitalization stocks), potentially widening the relative illiquidity of small assets.

**IV Asset Prices in Fragmented Markets**

In this section we analyze competition between a given set of trading venues, the allocation of investors across these venues, and the resulting asset prices.

Consider two venues, 1 and 2, with speeds \( \rho_1 \) and \( \rho_2 \) (i.e effective speeds \( s_i = \frac{\rho_i}{r+\gamma+\rho_i} \)) and participation fees \( q_1 \) and \( q_2 \), respectively. If both speeds were equal, the exchanges would have to compete à la Bertrand in fees, leaving each venue with zero profits. Exchanges therefore have an incentive to differentiate their intermediation services by offering different speeds and thus relaxing competition in fees. We concentrate on this case and, without loss of generality, take venue 2 as the fast market, so that \( \rho_1 < \rho_2 \). Since fees can be adjusted more easily than trading platforms, it is natural to model exchange competition as
For a given trading regulation, the fast exchange chooses a trading speed and then exchanges set access fees. Investors decide whether to participate and, if so, which market to join. Once these decisions are made, trading starts at time zero.

a sequential game where markets first select a trading technology and subsequently compete on fees. In the remainder of this paper we focus solely on exchanges’ speed choices.

With more than one trading venue, investors can find different prices to buy or sell an asset in each of them. Policymakers are often concerned about “protecting” investors from bad execution and have designed a number of rules that regulate how market prices relate to each other (see the discussion in Appendix A). We consider two types of trading regimes, which capture stylized extreme cases.

Definition 3. Trading regimes $\tau \in \{\text{seg}, \text{prot}\}$ are as follows.

- **Market segmentation** ($\text{seg}$): When a venue refuses to execute trades coming from investors of another venue. There are then two asset markets and two liquidity markets.

- **Price protection** ($\text{prot}$): The market authority enforces a single asset price. There are then one asset market and two liquidity markets.

Under segmentation, an investor joins a market and cannot trade with an investor in another market. Once investors have been allocated, the markets are effectively segmented and equilibrium asset prices can be different. Under price protection, asset prices must be the same in both venues.\(^{31}\) Note that a single price would also prevail if there were limitless arbitrage opportunities for a mass of potential arbitrageurs. In this regard, price protection is equivalent to perfect arbitrage. A more general framework with costly arbitrage (e.g., Shleifer and Vishny (1997)) would deliver equilibrium asset prices that lie somewhere between the low and high segmented prices we derive.

For a given trading regime, a market structure is characterized by a set of speed and fees $(s, q)$. Investors observe these outcomes and decide whether to participate in financial markets, and which trading venue is more attractive. The pre-trade decisions of type-\(\sigma\) investors

---

\(^{31}\)This is our simple way to capture access and trade-through rules in the SEC’s Reg NMS. The distinction between the U.S.’s top-of-the-book and Canada’s full-depth protection becomes trivial here since we only consider unitary orders. See Appendix A for a discussion of investor protection and more details.
are formally described as
\[ P : (s, q) \times [0, \hat{\sigma}] \rightarrow \{0, 1, 2\} . \]

The timing of decisions is illustrated in Figure 7. We will see that investors choosing to trade in the fast market do so by paying a higher access fee. In light of the discussion in Section I, investors decision to pay a speed premium can be interpreted as a decision to colocate with an exchange’s servers.

## A Segmented Markets

Let us start studying investors’ pre-trade decisions given \((s, q)\). When markets are effectively segmented, we need to consider three indifference conditions. First, there is a type \(\hat{\sigma}_1\) who is indifferent between joining market 1 and staying out. This type must satisfy

\[ W(\hat{\sigma}_1, \hat{\sigma}_1, s_1) - W_{\text{out}} = q_1, \]

which, using Equation 14, implies

\[ q_1 = \frac{\bar{a}s_1\hat{\sigma}_1}{r}. \tag{26} \]

The second indifference condition defines the marginal type \(\hat{\sigma}_{12}\), who is indifferent between joining market 1 and market 2. By definition, this type must be such that

\[ W(\hat{\sigma}_{12}, \hat{\sigma}_2, s_2) - q_2 = W(\hat{\sigma}_{12}, \hat{\sigma}_1, s_1) - q_1. \tag{27} \]

The third indifference condition is that the temporary investors are indifferent between joining markets 1 and 2. Therefore we must have

\[ W(\hat{\sigma}_2, \hat{\sigma}_2, s_2) - W_{\text{out}} - q_2 = W(\hat{\sigma}_1, \hat{\sigma}_1, s_1) - W_{\text{out}} - q_1 \]

Given the indifference condition for \(\hat{\sigma}_1\) this implies \(W(\hat{\sigma}_2, \hat{\sigma}_2, s_2) - W_{\text{out}} - q_2 = 0\). Combining these conditions, we obtain \(\frac{s_1}{2r} (\hat{\sigma}_{12} - \hat{\sigma}_1) = \frac{s_2}{2r} (\hat{\sigma}_{12} - \hat{\sigma}_2)\) and \(q_2 = \frac{\bar{a}s_2\hat{\sigma}_2}{r}\), and therefore

\[ \hat{\sigma}_{12} = \frac{r}{\bar{a}} \frac{q_2 - q_1}{s_2 - s_1}. \tag{28} \]

Note that \(\hat{\sigma}_1 < \hat{\sigma}_2 < \hat{\sigma}_{12}\). The set of types that join market 2 cannot be an interval. It is composed of all the types above \(\hat{\sigma}_{12}\) and some types below \(\hat{\sigma}_1\).

Let us consider exchange optimization. The total revenue for slow and fast exchanges is given by \(q_1 (G(\hat{\sigma}_{12}) - G(\hat{\sigma}_1) + \delta_1)\) and \(q_2 (1 - G(\hat{\sigma}_{12}) + \delta_2)\), where, as in Section II, \(\delta_i\) represents the mass of temporary investors joining market \(i\). The mass of temporary traders in each
market is determined by market clearing. Due to market segmentation, we need to consider two different market clearing conditions:

\[ (1 - G(\hat{\sigma}_{12}) + \delta_2) \bar{a} = \frac{1 - G(\hat{\sigma}_{12})}{2} \]

for market 2 and

\[ (G(\hat{\sigma}_{12}) - G(\hat{\sigma}_1) + \delta_1) \bar{a} = \frac{G(\hat{\sigma}_{12}) - G(\hat{\sigma}_1)}{2} \]

for market 1. Using these conditions, we can express the slow and fast exchange gross profit functions as \( \pi^{seg}_2 = q_2 \frac{1 - G(\hat{\sigma}_{12})}{2a} \) and \( \pi^{seg}_1 = q_1 \frac{G(\hat{\sigma}_{12}) - G(\hat{\sigma}_1)}{2a} \).

To simplify the analysis we take market 1’s speed as exogenously given by the economy’s default speed, that is \( s_1 = \bar{s} \), while market 2 chooses an effective speed \( s_2 \) at a cost \( C(s_2) \).

In the speed choice stage, market 2 thus solves

\[
\max_{s_2} \frac{q_2(s_2)}{2\pi} [1 - G(\hat{\sigma}_{12}(s_2))] - C(s_2) \tag{29}
\]

We can naturally extend the equilibrium in Definition 1 to this environment by adding an investment stage. To ensure sub-game perfection we must first find the access fee \( q^*_i \) that maximizes \( \pi^{seg}_i \) given \( q_{j\neq i} \), for \( i, j = 1, 2 \); and where the marginal types are given by Equations 26 and 28. Second, given \( q^*_1 \) and \( q^*_2 \), we must find the speed level \( s^*_2 \) that solves program 29. We shall see in the next section that the equilibrium solution to this problem implies a level of participation and speed in the fast market that exceeds the one in the consolidated market.

Once these objects are determined, we can characterize the equilibrium prices.

**Proposition 4.** In the segmented markets equilibrium, asset prices are given by

\[
p_1 = \frac{1}{\bar{r}} \left[ \mu - G^{-1} (1 - 2\pi) \right] + \frac{1}{\bar{r}} \left[ \hat{\sigma}_1 + G^{-1} (1 - 2\pi) \right] - \frac{1}{\bar{r}} \hat{\sigma}_1 (1 - \lambda_1) \tag{30}
\]

and

\[
p_2 = \frac{1}{\bar{r}} \left[ \mu - G^{-1} (1 - 2\pi) \right] + \frac{1}{\bar{r}} \left[ \hat{\sigma}_2 + G^{-1} (1 - 2\pi) \right] - \frac{1}{\bar{r}} \hat{\sigma}_2 (1 - \lambda_2(s_2)) \tag{31}
\]

where marginal types \( \hat{\sigma}_1 \) and \( \hat{\sigma}_2 \) satisfy Equations 59-61 given in the Appendix. The limited participation distortion (LPD) is higher in the fast market. The ILD in the fast market can be lower or higher than in the slow market.
As in a consolidated market, we can see the prices in the markets of Proposition 4 as the sum of three terms: the Walrasian price, an LPD, and an ILD. The LPD here reflects the change in the value of the marginal investor relative to the Bertrand outcome. Intuitively, when speed differentiation increases, markets are able to increasingly relax competition in access fees. Given that investors with high types choose to trade in the fast market, this distortion is naturally greater in market 2. However, it is interesting to note that when markets are segmented the ILD is not necessarily lower in the faster market. Although the amount of illiquidity is lower in market 2 (i.e. $1 - \lambda_2 < 1 - \lambda_1$), its marginal valuation in that market is higher. In fact, when market contact frictions are small, the ILD is likely to be higher in the fast market.

B Protected Markets

In this section we analyze how price protection in the trading period affects competition between exchanges and equilibrium asset prices.

With price protection asset markets operate under a single market clearing condition. Investors then self-select into trading venues that can be seen as different entry points into a single integrated asset market. Following widely used terminology in U.S. stock markets, we refer to the single clearing price as the national best price, denoted $p_{nb}$. Active participants in markets 1 and 2 are still characterized by the indifference condition 26 for the marginal type $\hat{\sigma}_1$, and the marginal type $\hat{\sigma}_{12}$ is still characterized by Equation 27. However, with a single asset price temporary traders are not indifferent between joining market 1 and joining market 2. Intuitively, temporary traders now have a stronger incentive to join market 1, where they pay a lower access fee, since they can sell their endowment at the same price in either market. We then have $\delta_2 = 0$ and the market clearing condition becomes

\[
\frac{(1 - G(\hat{\sigma}_1) + \delta_1) \tilde{a}}{\text{Total Asset Supply}} = \frac{(1 - G(\hat{\sigma}_1))}{2 \text{ Stationary Asset Demand}} \tag{32}
\]

The redistribution of investors across venues impacts market revenue. In particular, the Online Appendix shows that the gross revenue functions $\pi_1^{prot}$ and $\pi_2^{prot}$ are now $\frac{a}{2\tilde{a}}[1 - G(\hat{\sigma}_1) + 2\pi(G(\hat{\sigma}_{12}) - 1)]$ and $q_2 (1 - G(\hat{\sigma}_{12}))$, respectively. By solving the new two-stage competition game, and imposing market clearing condition 32, we can characterize the national best price.
**Proposition 5.** The single equilibrium price under price protection is

\[
p_{nb} = \frac{1}{r} \left[ \mu - G^{-1} (1 - 2\pi) \right] + \frac{1}{r} \left[ \sigma_{seg}^+ + G^{-1} (1 - 2\pi) \right] - \frac{1}{r} \sigma_{seg}^+ (1 - \lambda_1) + \frac{1}{r} (\sigma_{prot}^+ - \sigma_{seg}^+) \lambda_1.
\]

Under A.1 (Equation 15) the price protection distortion (PPD) is positive.

The first three terms on the right-hand side of Equation 33 in Proposition 5 coincide with the regulation-free price in the slow market, as given by Equation 30. There is an additional term in this case that captures the price distortion of the trading regulation in market 1:

\[
PPD = p_{nb} - p_{seg}^1.
\]

Under general conditions, such a distortion is positive, and the national best price \(p_{nb}\) is greater than \(p_{seg}^1\). The intuition is as follows. Price protection acts as a subsidy to the slow market because its investors are (effectively) allowed to sell their assets to investors in the fast market. This creates, everything else being constant, a larger demand for the slow market, which encourages the slow market to increase its access fee \(q_1\), ultimately raising the value of the marginal type \(\hat{\sigma}_1\). This is why we have \(\hat{\sigma}_1^{prot} > \hat{\sigma}_1^{seg}\). Protection also softens the price elasticity of the marginal type \(\hat{\sigma}_{12}\), which again is good for the slow venue. Thus the slow venue’s profits increase under protection for two reasons: more demand and less price elasticity.

Note that Proposition 5 does not characterize the impact of price protection on \(p_{seg}^2\). We analyze this issue further in the following section.

**V Empirical Implications**

In this section we study the relation between the different asset prices obtained in Sections III and IV, as well as implications for trading volume and fragmentation levels.

Our first task is to compare price outcomes in a consolidated market and in unregulated fragmented markets. To do so, we would like to have a notion of the average price in the segmented markets. A natural choice is to compute the volume-weighted average price (VWAP). Let \(\tau_i\) represent the steady-state expected number of trades in market \(i\) per unit of time, a natural notion of volume. We then define the VWAP as follows.
Definition 4. The VWAP \( p_{vw} \) is given by

\[
p_{vw} = \left( \frac{\tau_1}{\tau_1 + \tau_2} \right) p_1 + \left( \frac{\tau_2}{\tau_1 + \tau_2} \right) p_2. \tag{34}
\]

Note that we can easily derive the expressions for \( \tau_1 \) and \( \tau_2 \). First, the masses of active traders in markets 1 and 2 are given by \( G(\hat{\sigma}_{12}) - G(\hat{\sigma}_1) \) and \( 1 - G(\hat{\sigma}_{12}) \), respectively. From Proposition 1, we know that in the steady state the proportion of asset owners wanting to sell their asset and the proportion of agents wanting to buy it is given by \( \frac{1 - \gamma}{\gamma + \rho} \). With the contact rates for markets 1 and 2 equal to \( \rho_1 \) and \( \rho_2 \), respectively, the average number of trades in each market at any moment is then

\[
\tau_1 = \rho_1 \frac{\gamma}{4(\gamma + \rho_1)} [G(\hat{\sigma}_{12}) - G(\hat{\sigma}_1)] \bar{a} \tag{35}
\]

\[
\tau_2 = \rho_2 \frac{\gamma}{4(\gamma + \rho_2)} [1 - G(\hat{\sigma}_{12})] \bar{a}, \tag{36}
\]

A The Pricing Effect of Trading Fragmentation

Does exchange competition affect asset prices in the long-run? The relationship between \( p_{con} \) and \( p_{vw} \) is not obvious because there are two opposite effects in force. On the one hand, the monopoly restricts entry and induces a high marginal type \( \hat{\sigma}_{con} \), raising the LPD and thereby increasing the equilibrium price. On the other hand, speed increases prices and competition can increase the market average speed. The following result characterizes the relationship.

Proposition 6. (Competition and Asset Prices)

SR. The limited participation distortion (LPD) is always higher in the consolidated market. Thus, for any given market speed, the consolidated price is higher than the VWAP in the short-run.

LR. In the long-run, the VWAP can be higher or lower than the consolidated price. Provided that market contact frictions are moderate, the consolidated price is higher. Under A.1 (Eq. 15), the illiquidity discount (ILD) of the fast market price is always lower than in the consolidated market.

If both the monopolist exchange and the fast market have access to the same speed in the short-run, according to Equations 21, 30, and 31, price differences between \( p_{con} \) and \( p_{vw} \) will be due to the LPD only. Since the monopolist exchange restricts participation to a greater extent, the price in the consolidated market will be higher, regardless of the volume.
distribution across venues 1 and 2. Figure 8 illustrates this fact by comparing the marginal investor $\hat{\sigma}_{con}$ and a volume-weighted marginal type $\hat{\sigma}_{vw}$. In the long-run, markets can invest in speed technologies and the sign of $p_{con} - p_{vw}$ thus depends on the relative strengths of the effects on the LPD and ILD. We have seen in Proposition 2 that the LPD remains unchanged in a consolidated market but the ILD decreases. The LPD in fragmented markets will increase in the long-run because of speed differentiation allowing markets to relax fee competition, but it remains lower than in the consolidated market. The extent of frictions in the fragmented market relative to the consolidated case is likely to decrease and thus reduce the ILD. In particular, under A.1, we have $s_{2}^{seg} > s_{con}$ in the long-run and therefore the ILD is lower in the fast market. 32 Consequently, the sign of the total effect of competition on asset prices depends on parameter values. However, when market contact frictions are small (e.g., stocks, exchange-traded derivatives), the LPD effect dominates, implying $p_{con} > p_{vw}$.

B  The Pricing Effect of Price Protection

Consider an economy where asset markets are fragmented but unregulated. How would adoption of a trade-through-like rule affect prices? A related question is what would the effect on prices be like when the transition is from a consolidated to a regulated fragmented market? Section IV shows that the protected price is higher than the unregulated price in the slow market. We now want to compare the protected market with the consolidated market and VWAP prices. The following result characterizes these relations.

Proposition 7. (Investor Protection and Asset Prices)

The consolidated market price is always higher than the national best price. Under A.1 (Eq. 15), the protected price $p_{nb}$ lies between the prices of the segmented venues, and is lower than the VWAP, provided the speed cost parameters $c$ and $s$ are sufficiently low.

A single exchange will distort participation to a greater extent than in a duopoly, increasing the LPD with respect to the protected market regardless of the time period. This fact is illustrated by the left-pointing arrows in Figure 8: The value of the marginal investor in the consolidated market lies to right of the ones in fragmented markets. Moreover, the speed investments in the long-run will reduce the amount of illiquidity in the consolidated market, increasing $p_{con}$ relative to $p_{nb}$. Speed investments indirectly affect $p_{nb}$ in the long-run, by increasing the value of the marginal investors as the exchanges differentiate themselves from each other and relax fee competition. Note that the marginal investor in the duopolistic economy resides in the slow market and thus its valuation of the asset depends on the default

32 It is easy to show that $\hat{\sigma}_{con} > \hat{\sigma}_{2}^{seg}$ does not depend on A.1.
technology. Consequently, there is no direct effect of investments on the ILD in the protected market, and thus the ILD is lower in the consolidated market, regardless of the time period. Whether the VWAP is higher than the protected price depends on the relative strength of two effects. On the one hand, we know from Proposition 5 that price protection introduces a positive distortion in the asset price of market 1, implying \( p_{1}^{seg} < p_{nb} \). On the other hand, both the LPD and the level of asset liquidity are higher in the fast market than in the protected market, implying \( p_{2}^{seg} > p_{nb} \). The volume distribution then becomes important in comparing \( p_{nb} \) and \( p_{vw} \). When the ability of the fast market to differentiate its liquidity service is substantial, this market can attract greater trading flows and thus \( p_{vw} > p_{nb} \). This occurs, for example, when technology costs and the default speed level are relatively low. Figure 8 shows that this is the case in our baseline calibration.

C  Calibration and Analysis of Price Distortions

In this section we consider a calibration of our model to illustrate the impact of long-run investments and market organization on prices. Assuming A.1–A.2 (Equations 15 and 20), our model involves the nine exogenous parameters listed in Table II. Some of these parameters, we argue, can be calibrated using secondary markets data. In cases where parameters are more difficult to calibrate, we explain our choices and their impact on the
calibration results.

The illiquid asset in the model can represent one of the several asset classes that trade in exchanges, such as stocks, futures, or equity options. In order to be specific, and motivated by the empirical observations in Section I, we concentrate on U.S. equity markets during 2001-2007. This is an important time interval since it corresponds to the period in between the SEC’s decimalization mandate and the final implementation of Reg NMS. Further, 2007 is the last year for which accurate data from the SEC Rule 605 is available, since in later years speeds are rounded down to zero. Stock prices obviously reflect market risk exposure, which is not the focus of this paper. Hence, when interpreting the results it is important to take into account that our main goal is instead quantifying the relative contribution of each of the frictions we identify to distortions to the Walrasian price. For simplicity, we also calibrate our parameters taking the NYSE as being the single trading venue. Although other markets, chiefly the NASDAQ, may have competed with the NYSE during the period, this approach allows us to calibrate parameters applying the simple formulas of Section III to market data without much loss of generality.\(^{33}\)

The asset characteristics are as follows. The interest rate is set equal to 2.5% annually, just below the one-month T-bill rate average value for the period 2001–2007. The annual holding cash flow \(\mu\) is set to 2.4 units of the consumption good, which implies a dividend yield close to that of S&P500 stocks for this period (relative to the baseline Walrasian price). The asset supply \(a\) is normalized to 0.47 so that supply is short. This is almost without loss of generality since the market price does not depend on \(a\) when this parameter is sufficiently close to one half.\(^{34}\)

To calibrate the speed contact parameters, we consider SEC Rule 605 data for the NYSE for the years 2001 and 2007. We interpret the 2001 value as corresponding to our “slow” (default) speed \(\rho_\text{d}\), which is relatively intensive in human trading, and the 2007 value as our “fast” (long-run) speed. The default speed then corresponds to a daily Poisson rate of 1,170, which translates to an average execution speed of approximately 20 seconds. To match an average execution speed of one second in 2007 for the NYSE, we consider a daily contact rate equal to 23,400, the number of seconds in a trading day. In order to annualize these values, as shown in Table II, we multiply daily rates by 252 trading days.

To calibrate the preference switching rate \(\gamma\) we proceed as follows. We normalize the steady state fraction of agents with misallocated assets, equal to \(\frac{1}{4} \frac{\gamma}{\gamma + \rho}\) as given in Proposition 1, to be 5%. Using the value of \(\rho_\text{d}\) we compute the implicit value for \(\gamma\), which yields a daily rate

\(^{33}\) According to the Securities and Exchange Commission (2010), the NYSE executed as much as four-fifths of the volume of NYSE-listed stocks in 2005.

\(^{34}\) From Section II we know that the price is insensitive to asset supply when the marginal investor type in market \(i\), \(\hat{\sigma}_i\), is higher than \(G^{-1} (1 - 2\pi)\).
equal to 292.5. We interpret this value as representing the trading needs of a given broker, as opposed to a single individual investor, that represents a large number of customers with the same \( \sigma \)-type value. Such a broker, depending on its customers order flow, may find it optimal to switch holdings between 0 and 1 positions multiple times during a trading day.\(^{35}\)

Since we assume an exponential distribution of investor permanent types (assumption A.1), we only need to calibrate the value of the average investor type \( \nu \). This parameter, however, is not easy to compute based on market data. Since our calibration results are sensitive to the choice of \( \nu \), we consider a range of different values and analyze the corresponding pricing responses. In particular we take \( \nu \) to lie in the set \( \{ \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1 \} \), and consider \( \nu = \frac{1}{2} \) as the baseline value. To illustrate the economic interpretation of these values, consider the median investor type, \( m(\sigma) \), when \( \nu = \frac{1}{4} \). The annual holding flow utility under a temporary shock \( \varepsilon \) (see equation 1) is \( u_{m(\sigma),\varepsilon}(1) = 2.44 + \frac{\varepsilon}{4} \ln(2) \). Consequently, the private component of the annual utility flow for this investor is approximately 17.5 bps (relative to \( p_W \)). This implies that, when facing a negative or positive temporary shock, the annual flow utility equals 2.27 and 2.61 units of consumption, respectively. Similarly, for the baseline value \( \nu = \frac{1}{2} \), the annual utility flow lies in \( \{ 2.14; 2.78 \} \). Given the lack of direct evidence on the cumulative distribution function \( G \) it is difficult to ascertain whether \( \nu = \frac{1}{2} \) is the most plausible value, so the success of our calibration analysis should be qualified.

The marginal cost of speed investment \( c \) is chosen so that in the long-run, and given the other parameter values, the optimal single-market contact speed implies an average execution speed of approximately one second.\(^{36}\) This value is close to the average trading speed for the NYSE in 2007 according to the SEC Rule 605 data.

\[^{35}\]To illustrate this interpretation, consider a discretization of the steady-state trade volume formula 35: \( \tau = \frac{\nu}{4} \sigma \rho \times \rho \times \text{Number of brokers} \). According to data available at www.nyse.nyx.com, the average number of daily trades for all 3,025 NYSE-listed stocks in 2001 was approximately equal to 452. Taking \( \tau = 452 \), and considering our calibrated parameters, this formula implies that there were, on average, 8 active trading brokers for the average NYSE-listed stock on any given day.

\[^{36}\]Inverting equation 47 in the Appendix, one finds that

\[ c = \frac{\nu (r + \gamma)}{2er (r + \gamma + \rho_{\text{con}})^2}. \]

The value of \( c \) is then recovered considering \( \rho_{\text{con}} = 23,400 \times 252 \).
### Table II: Parameter values in Calibration

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Notation</th>
<th>Baseline Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest rate</td>
<td>( r )</td>
<td>2.5%</td>
</tr>
<tr>
<td>Holding cash flow</td>
<td>( \mu )</td>
<td>2.44</td>
</tr>
<tr>
<td>Default contact rate</td>
<td>( \rho )</td>
<td>2.95 \times 10^5</td>
</tr>
<tr>
<td>Short-run contact rate market 2</td>
<td>( \rho_s )</td>
<td>1.18 \times 10^6</td>
</tr>
<tr>
<td>Long-run contact rate consolidated market</td>
<td>( \rho_{con} )</td>
<td>5.90 \times 10^6</td>
</tr>
<tr>
<td>Switching intensity temporary types</td>
<td>( \gamma )</td>
<td>73,710</td>
</tr>
<tr>
<td>Marginal cost of speed investments</td>
<td>( c )</td>
<td>7.6 \times 10^{-9}</td>
</tr>
<tr>
<td>Asset supply</td>
<td>( \bar{\alpha} )</td>
<td>0.47</td>
</tr>
<tr>
<td>Average investor type (baseline value)</td>
<td>( \nu )</td>
<td>0.5</td>
</tr>
</tbody>
</table>

*The values of parameters \( \{r, \mu, \rho, \rho_s, \rho_{con}, \gamma\} \) correspond to annual rates.

We start considering a short-run equilibrium where fees and market participation are endogenously determined, but trading platforms’ speed are a given. To analyze the case of fragmented markets, we need to include in the parameter space a trading speed for the fast market, denoted \( \rho_2 \). We consider an exchange that has an average execution speed of five seconds, that is, \( \rho_2 = 4 \rho \). This value is close to that reported by Angel, Harris, and Spatt (2011) for the NASDAQ in late 2001, and can be seen as technologically feasible for other markets during the early 2000s. It results, of course, in a slower average execution speed than the one reported for the NYSE around 2007 (around one second).

Table III reports the model-implied value of the price decompositions derived in Sections III and IV. A number of interesting observations are in order. First, we observe that consistent with Propositions 6 and 7, for an asset class such as stocks where market contact frictions are small, we have \( p_{con} > p_{vw} \) and \( p_{vw} > p_{nb} \). Note that in both inequalities, the LPD plays a key role. In fact, without this quantity being endogenized, all the computed prices should be below the frictionless Walrasian value (e.g., Duffie, Garleanu, and Pedersen (2005)), while only \( p_1 \) and \( p_{nb} \) are below \( p_W \). For example, considering the baseline \( \nu = 0.5 \), the absolute value of the ILD in the slow market (33 bps) is higher than the LPD (5 bps). It is also interesting to note that an observer comparing the national best and Walrasian price could conclude that the market is essentially frictionless, while in fact the small aggregate distortion is due to the sum of the values of the three individual distortions we identified. Although the absolute value of the PPD is fairly small in our parametrization, note that it represents nearly half the size of the ILD, and about 42% of the LPD.

Let us now consider the long-run equilibrium of asset prices under the baseline parametrization, depicted in Table IV. Consistent with Proposition 2, the value of the LPD in a consolidated market (driven by the ex ante distribution of types \( G \)) is unchanged. However,
<table>
<thead>
<tr>
<th>Limited Participation Distortion</th>
<th>Illiquidity Discount</th>
<th>Price Protection Distortion</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consolidated ν = 0.25</td>
<td>9.55</td>
<td>-2.04</td>
<td>107.52</td>
</tr>
<tr>
<td>Slow Venue</td>
<td>0.19</td>
<td>-0.16</td>
<td>100.03</td>
</tr>
<tr>
<td>Fast Venue</td>
<td>0.90</td>
<td>-0.09</td>
<td>100.81</td>
</tr>
<tr>
<td>VWAP</td>
<td>0.66</td>
<td>-0.14</td>
<td>100.55</td>
</tr>
<tr>
<td>National Best</td>
<td>0.19</td>
<td>-0.16</td>
<td>0.08  100.11</td>
</tr>
<tr>
<td>Consolidated ν = 0.5</td>
<td>18.98</td>
<td>-4.05</td>
<td>114.94</td>
</tr>
<tr>
<td>Slow Venue</td>
<td>0.38</td>
<td>-0.33</td>
<td>100.06</td>
</tr>
<tr>
<td>Fast Venue</td>
<td>1.78</td>
<td>-0.18</td>
<td>101.60</td>
</tr>
<tr>
<td>VWAP</td>
<td>1.32</td>
<td>-0.28</td>
<td>101.09</td>
</tr>
<tr>
<td>National Best</td>
<td>0.38</td>
<td>-0.33</td>
<td>0.16  100.21</td>
</tr>
<tr>
<td>Consolidated ν = 0.75</td>
<td>28.30</td>
<td>-6.03</td>
<td>122.26</td>
</tr>
<tr>
<td>Slow Venue</td>
<td>0.57</td>
<td>-0.49</td>
<td>100.08</td>
</tr>
<tr>
<td>Fast Venue</td>
<td>2.66</td>
<td>-0.27</td>
<td>102.39</td>
</tr>
<tr>
<td>VWAP</td>
<td>1.96</td>
<td>-0.41</td>
<td>101.62</td>
</tr>
<tr>
<td>National Best</td>
<td>0.57</td>
<td>-0.49</td>
<td>0.24  100.32</td>
</tr>
<tr>
<td>Consolidated ν = 1</td>
<td>37.50</td>
<td>-7.99</td>
<td>129.50</td>
</tr>
<tr>
<td>Slow Venue</td>
<td>0.76</td>
<td>-0.65</td>
<td>100.11</td>
</tr>
<tr>
<td>Fast Venue</td>
<td>3.52</td>
<td>-0.35</td>
<td>103.17</td>
</tr>
<tr>
<td>VWAP</td>
<td>2.60</td>
<td>-0.55</td>
<td>102.15</td>
</tr>
<tr>
<td>National Best</td>
<td>0.76</td>
<td>-0.65</td>
<td>0.31  100.42</td>
</tr>
</tbody>
</table>

This table presents the values of the price distortions and of the equilibrium asset price relative to the frictionless Walrasian price, which is normalized to 100, in the short-run, for four values of the average investor type $\nu$. Parameter values (except $\nu$) correspond to the baseline calibration (Table II). In the short-run, investor participation is variable, but trading technologies are fixed. The Walrasian price is the price in the absence of market contact and investor participation frictions. The VWAP corresponds to that between the slow and fast venues. The national best price correspond to the unique price in fragmented markets under price protection. The LPD and ILD are as in Definition 2.
the ILD decreases significantly: It decreases to 25 bps from 4.05%. This change reflects the effect of speed investments on allocative efficiency. In the case of fragmented trading, note that the LPD increases in all cases, which reflects the enhanced ability of exchanges to increase their access fees due to the relaxation of Bertrand competition through speed differentiation (the average execution speed is now 20 times faster in the fast exchange). This increase reflects the interaction between the time-varying competitive environment due to technological progress and trading regulations.

Finally, note that both in the short and long-run the LPD terms increase monotonically in the value of the average investor type $\nu$. While the effect is large for the consolidated market, since the marginal investor type equals $\nu$, in the other market structures analyzed the effect of changes in $\nu$ are quantitatively small.

## D Comparative Statics of Asset Prices

**Consolidated Market** How is the equilibrium price in a consolidated market affected by changes in the environment? Are the relations affected in the long-run? By differentiating Equations 18 and 21, we can characterize the behavior of the market clearing asset price as follows.

**Proposition 8.** In a consolidated market,

(i) The equilibrium asset price increases in the volatility of investors’ private utility process. The effect is stronger in the long-run.

(ii) The equilibrium asset price decreases in the frequency of preference shocks. The effect is weaker in the long-run, provided the cost of speed is not “too high.”

(iii) The equilibrium asset price decreases in the cost of speed.

Figure 9 graphically displays these relationships. An increase in the average investor type $(\nu)$ increases the value of the marginal investor type. Everything else being constant, this increases the LPD in the short-run. In the long-run, an increase in $\nu$ also makes speed investments more profitable for the exchange, resulting in a lower-ILD equilibrium and further raising the asset price.

In the short-run, the equilibrium asset price decreases with the frequency of the temporary shocks $\gamma$. This effect is intuitive: For a given installed speed capacity, an increase in $\gamma$ increases the proportion of agents with misallocated assets at any given time, rendering the asset less valuable. In the long-run, speed investments facilitate more efficient asset
Table IV: Long Run Price Decomposition (Walrasian Price=100)

<table>
<thead>
<tr>
<th></th>
<th>Limited Participation Distortion</th>
<th>Illiquidity Discount</th>
<th>Price Protection Distortion</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>$\nu = 0.25$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consolidated</td>
<td>9.55</td>
<td>-0.18</td>
<td></td>
<td>109.37</td>
</tr>
<tr>
<td>Slow Venue</td>
<td>0.39</td>
<td>-0.20</td>
<td></td>
<td>100.19</td>
</tr>
<tr>
<td>Fast Venue</td>
<td>1.28</td>
<td>-0.03</td>
<td></td>
<td>101.26</td>
</tr>
<tr>
<td>VWAP</td>
<td>0.99</td>
<td>-0.15</td>
<td></td>
<td>100.91</td>
</tr>
<tr>
<td>National Best</td>
<td>0.39</td>
<td>-0.20</td>
<td>0.10</td>
<td>100.28</td>
</tr>
<tr>
<td><strong>$\nu = 0.5$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consolidated</td>
<td>18.98</td>
<td>-0.25</td>
<td></td>
<td>118.73</td>
</tr>
<tr>
<td>Slow Venue</td>
<td>0.82</td>
<td>-0.41</td>
<td></td>
<td>100.40</td>
</tr>
<tr>
<td>Fast Venue</td>
<td>2.62</td>
<td>-0.04</td>
<td></td>
<td>102.58</td>
</tr>
<tr>
<td>VWAP</td>
<td>2.04</td>
<td>-0.29</td>
<td></td>
<td>101.88</td>
</tr>
<tr>
<td>National Best</td>
<td>0.82</td>
<td>-0.41</td>
<td>0.19</td>
<td>100.60</td>
</tr>
<tr>
<td><strong>$\nu = 0.75$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consolidated</td>
<td>28.30</td>
<td>-0.30</td>
<td></td>
<td>127.99</td>
</tr>
<tr>
<td>Slow Venue</td>
<td>1.24</td>
<td>-0.62</td>
<td></td>
<td>100.62</td>
</tr>
<tr>
<td>Fast Venue</td>
<td>3.96</td>
<td>-0.05</td>
<td></td>
<td>103.91</td>
</tr>
<tr>
<td>VWAP</td>
<td>3.08</td>
<td>-0.44</td>
<td></td>
<td>102.84</td>
</tr>
<tr>
<td>National Best</td>
<td>1.24</td>
<td>-0.62</td>
<td>0.29</td>
<td>100.91</td>
</tr>
<tr>
<td><strong>$\nu = 1$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consolidated</td>
<td>37.50</td>
<td>-0.35</td>
<td></td>
<td>137.15</td>
</tr>
<tr>
<td>Slow Venue</td>
<td>1.66</td>
<td>-0.83</td>
<td></td>
<td>100.84</td>
</tr>
<tr>
<td>Fast Venue</td>
<td>5.28</td>
<td>-0.06</td>
<td></td>
<td>105.22</td>
</tr>
<tr>
<td>VWAP</td>
<td>4.11</td>
<td>-0.58</td>
<td></td>
<td>103.80</td>
</tr>
<tr>
<td>National Best</td>
<td>1.66</td>
<td>-0.83</td>
<td>0.39</td>
<td>101.22</td>
</tr>
</tbody>
</table>

This table presents the values of the price distortions and of the equilibrium asset price relative to the frictionless Walrasian price, which is normalized to 100, in the long-run, for four values of the average investor type $\nu$. Parameter values (except $\nu$) correspond to the baseline calibration (Table II). In the long-run both investor participation and trading technologies are variable. The Walrasian price is the price in the absence of market contact and investor participation frictions. The VWAP corresponds to that between the slow and fast venues. The national best price correspond to the unique price in fragmented markets under price protection. The limited LPD and hte ILD are as in Definition 2.
This figure plots the effect of parameter changes on consolidated market price in the short-run (SR), long-run (LR), and in the frictionless Walrasian case. The parameter changes are relative to the baseline calibration (Table II). Panel A shows the marginal cost ($c$); Panel B shows the default speed ($\rho$); Panel C shows the average investor type ($\nu$), and Panel D the preference switching frequency ($\gamma$).

An increase in the cost of speed decreases the asset price since, all else being equal, the exchange invests less in speed rendering the asset more illiquid. Interestingly, $p_{con}$ is not affected in the long-run by the default speed level. This fact reflects the lack of competitive forces in a single exchange economy, and contrasts with the results for fragmented markets in the following section.

**Fragmented Markets** Figure 10 displays the effect of changes in parameter values on the VWAP and the national best price. As in the consolidated case, an increase in the marginal cost of speed decreases asset prices, both with and without price protection. We can observe in Panel A that the VWAP is more sensitive to this change. This is due to the fact that an increase in $c$ changes the ILD in market 2 as well as the marginal investor types $\hat{\sigma}_1$ and $\hat{\sigma}_2$ due to easier speed differentiation. Only the latter effect impacts the equilibrium price in the protected market.

Interestingly, an increase in the level of default speed $\rho$ decreases the VWAP. This is due
Table V: Trading Volume (Volume in Walrasian market=100)

<table>
<thead>
<tr>
<th></th>
<th>Short Run</th>
<th>Short Run (Protected)</th>
<th>Long Run</th>
<th>Long Run (Protected)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consolidated</td>
<td>31.31</td>
<td></td>
<td>38.65</td>
<td></td>
</tr>
<tr>
<td>Slow Venue</td>
<td>29.01</td>
<td>28.66</td>
<td>28.64</td>
<td>28.37</td>
</tr>
<tr>
<td>Fast Venue</td>
<td>58.23</td>
<td>57.74</td>
<td>59.61</td>
<td>58.82</td>
</tr>
<tr>
<td>Slow + Fast</td>
<td>87.23</td>
<td>86.40</td>
<td>88.26</td>
<td>87.19</td>
</tr>
</tbody>
</table>

This table presents the trading volumes in the steady-state equilibrium relative to the frictionless Walrasian case, which is normalized to 100. The parameter values correspond to the baseline calibration (Table II). In the short-run investor participation is variable, but trading technologies are fixed. The Walrasian case correspond to a market without contact or investor participation frictions. The volume is given by the instantaneous expected transaction rate as in Equations 35 and 36.

An increase in the average investor type increases the equilibrium price in all cases, by raising the value of the marginal investor type, but the effect is strongest in the absence of price protection. This is because an increase in \( \nu \) provides the fast market with stronger incentives to invest in speed, reducing the ILD in market 2, with a direct impact on \( p_{vw} \), but only affecting \( p_{nb} \) indirectly by further raising the value of the marginal investor in market 1. Finally, an increase in the preference switching rate \( \gamma \) has a twofold effect on the asset price. On the one hand, it increases the demand for trading and speed. On the other hand, it negatively impacts the market allocative efficiency and investors’ value functions.

The negative effect is relatively stronger in market 1, where speed remains fixed. Thus, the national best price increases at a slower rate than the VWAP following an increase in this rate.

### E Trading Volume

Table V presents model-implied trading volumes relative to a Walrasian market. The effect of changes in the average investor type \( \nu \) are moderate and thus, for brevity, we report values for the baseline calibration only. Our measure of volume is given by the steady state value.
Figure 10: Comparative Statics of Prices in Fragmented Markets

This figure plots the effect of parameter changes on the VWAP in segmented markets, the single national best price with price protection, and the frictionless Walrasian price. The parameter changes are relative to the baseline calibration (Table II). Panel A shows the marginal cost \((c)\); Panel B shows the default speed \((\rho)\); Panel C shows the average investor type \((\nu)\), and Panel D the preference switching frequency \((\gamma)\).
of transaction rates per unit of time, as in Equations 35 and 36. We can observe that the lack of competition between the exchanges has a first order negative effect on volume: The total volume in a consolidated market is nearly half that in the fragmented market cases in both the short and long-run. However, note that the ability of the single exchange to invest in trading technologies has a large impact on traded values. Everything else being constant, when the average execution speed decreases from 20 seconds to 1 second, the relative transaction rate increases from 31.31 to 38.65.

In fragmented markets, the fast venue displays significantly higher volume than the slow venue. This volume difference increases in the long-run as the degree of speed differentiation increases. Furthermore, consistent with our results in Proposition 5, price protection decreases total investor participation in fragmented markets, generating a fall in traded volumes of approximately 1% of the Walrasian values in both the short and long-run.

\section*{F Trading Fragmentation}

How do parameter changes affect the volume distribution across trading venues in fragmented markets? To answer this question we need a formal notion of trading fragmentation. A simple metric is given by $1 - HHI$, where $HHI$ is the standard Herfindahl–Hirschman Index. Using the stationary equilibrium of the model, we can compute the $HHI$ simply as the sum of the squared terms $(\frac{\tau_i}{\tau_i + \tau_j})^2$, where the trading rates $\tau_1$ and $\tau_2$ are given by Equations 35 and 36, respectively.

Figure 11 displays how fragmentation is affected by the trading environment in both segmented and protected markets. Note that a general pattern arises across panels: Everything else being equal, the level of fragmentation is higher under price protection. This makes sense in light of our discussion in Section IV. Price protection increases the demand for the slow market, reducing the amount of ex post competition between venues.

Let's analyze the effect of the cost parameters. We observe in Panels A and B of Figure 11 that, regardless of the trading regime, both an increase in the marginal speed cost and an increase in the default speed level induce higher trading fragmentation. The intuition behind this effect is that these cost parameter changes reduce markets’ ability to differentiate, and thus exchanges are forced to compete more intensely in fees. As trading fees decrease, the relative volume of the fast market increases, reducing fragmentation.

Panels C and D show that an increase in preference parameters $\nu$ and $\gamma$ induce the opposite effect on trading fragmentation. Everything else equal, an increase in the investor average type results in more speed-sensitive investors, which helps the fast venue attract a higher mass of the population and raises equilibrium trading volumes. Similarly, an increase in the
frequency of preference shocks increases the relative demand of the fast market, which is able to realize a greater fraction of the total gains from trade.

![Figure 11: Comparative Statics of Trading Fragmentation](image1)

This figure plots the effect of parameter changes on fragmentation levels in the segmented and protected markets. Trading fragmentation is measured as 1-HHI. The parameter changes are relative to the baseline calibration (Table II). Panel A shows the marginal cost ($c$); Panel B shows the default speed ($\mu$); Panel C shows the average investor type ($\nu$), and Panel D the preference switching frequency ($\gamma$).

**VI Discussion**

**Liquidity and Participation Frictions.** The quantitative analysis suggests that distortions in participation incentives are important to understand the effects of market structures on liquidity and asset prices. Our analytical and quantitative analysis of participation distortions complements the work of Huang and Wang (2009). Note, however, an important difference: While participation costs that lead to lower market liquidity always decrease prices in their framework, they can increase prices in ours by changing both the mass and composition of investors.

Although speed plays a key role as a differentiating factor for exchanges, the frequency of market contact per se has a smaller effect on prices vis-a-vis participation distortions. For our baseline calibration, the long-run value of the LPD is two times larger than the ILD in the protected market, and seven times larger in the segmented market. The relatively small
value of the ILD is natural here given that our calibration is based on public equity data. It is also consistent with the calibration results of Gârleanu (2009). In contrast, Longstaff (2009) finds large pricing effects of illiquidity by calibrating an asset pricing model using U.S. private equity data.

**Does the Empirical Evidence Support the Model’s Results?** The fact that fast trading enhances liquidity is documented by Hendershott, Jones, and Menkveld (2011) and Hasbrouck and Saar (2010), among others. Boehmer (2005) documents the trade-off between execution speed and costs in U.S. markets and finds that, analogously to market 2 in our model, the NASDAQ is more expensive than the NYSE, but also faster.

Recent studies by O’Hara and Ye (2011) on U.S. markets and by Foucault and Menkveld (2008) and Degryse, Jong, and Kervel (2011) on European markets support the model’s prediction that liquidity increases with exchange competition. Importantly, our results on overall asset price levels suggest caution in interpreting such findings as definitive proof of the “benign” effects of fragmentation on market quality.

To the best of our knowledge, Amihud, Lauterbach, and Mendelson (2003) provide the only direct evidence of the effect of trading consolidation on asset prices. The authors study consolidation of two almost identical equity claims: A stock and a warrant on the stock that is deep in the money at expiration. They find that stock prices appreciate 1.27%, on average, on the warrant expiration. This empirical finding is consistent with Proposition 6. Arguably, their empirical analysis does not provide a direct test of our model’s prediction, which is based on exchange competition. However, since the Tel Aviv Stock Exchange market power is not altered in their study, one could interpret such price appreciation as representing a reduction of the ILD, with the LPD part of the price remaining constant.

**An International Asset Pricing Perspective.** In Section V we compare the pricing outcomes of three stylized economies: (1) a consolidated market, (2) an unregulated fragmented market, and (3) a price-protected fragmented market. We argue that when market contact frictions are small (e.g., for stocks), the model offers a pricing ordering $p_{con} \geq p_{w} \geq p_{nb}$. In our long-run calibration, these prices are equal to 118.73, 101.88, and 100.6, relative to the Walrasian price. How could an econometrician test this hypothesis? The first step would consist in identifying economic areas with market organizations that can be represented by our stylized cases. Incumbent exchanges in Hong Kong, China, Brazil and Spain, for example, experience little or no competition, and can thus be represented by (1) (see Figure 3). The U.S. and Canada are chief examples of financial markets with trade-through rules. Most of continental Europe and the U.K. can be seen as having unprotected
fragmented markets (see Appendix A) and thus are closest to (2). The second step would be to match assets across such areas with almost identical fundamentals (e.g., loadings on risk factors). An econometrician could then test whether within-groups prices are, say, highest in Hong Kong, intermediate in Europe, and lowest in the U.S.

**Is Price Protection Important?** Segmented equity markets are obviously an analytical abstraction here. In real markets arbitrageurs and smart routing technologies work to (at least partially) undo price differentials between markets. Does this fact make a trade-through rule redundant? Interestingly, empirical evidence by Foucault and Menkveld (2008) suggest that the answer is no. These authors study the competition between a London Stock Exchange order book (EuroSETS) and Euronext Amsterdam for Dutch firms and find that, even when there is formal entry barrier to arbitrageurs, the trade-through rate in their sample equals 73%.

Our calibration results suggest that price protection distorts the asset price in the slow venue by approximately 20 bps in the baseline calibration. Although this value may seem low, note that we define the PPD conceptually as \( p_{nb} - p_{seg}^{1} \). This difference may not represent the full effect of the introduction of such regulation. If a given market, say, the U.K., adopts a trade-through rule, the total pricing effect would instead be given by \( p_{nb} - p_{vw} \), which equals 128 bps in our calibration. As mentioned, our VWAP may not accurately represent prices in partially integrated markets, but as an approximation it suggests that the total effect of regulation may in fact be larger than our PPD. Further research on this topic is needed to assess this issue quantitatively.

Beyond its effect on asset prices, price protection affects the nature of competition between exchanges in our model. This is consistent with an earlier discussion by Stoll (2006), who argues the following:

“The casual observer of the heated debate that has surrounded the order protection rule may well wonder what the fuss is all about. After all, we are just talking about pennies. But for the exchanges, it may be a matter of business survival. Pennies matter, but more important, the rule requires the linkage of markets, which threatens established markets and benefits new markets. The battle appears to be over pennies, but in fact, it is over the ability of markets to separate themselves from the pack.”

---

37If one reinterprets the model market choice from an international perspective, say, a choice between two European countries, stock market segmentation is far from negligible (Bekaert, Harvey, Lundblad, and Siegel (2011)).
Furthermore, by endogenizing the entry decisions of exchanges, Pagnotta and Philippon (2012) show that price protection can expand the ex ante number of trading venues.

**Fast and Slow Traders.** The result that equilibrium prices can increase in the asset supply through investors’ speed investments is, to the best of our knowledge, new to the literature. Our result suggests that traders that invest heavily in speed concentrate on big cap assets, such as large S&P500 stocks or the E-mini futures. Industry reports and the evidence of Brogaard (2011) and Kirilenko, Kyle, Mehrmad, and Tuzun (2010) for high-frequency traders support this intuition.

**VII Concluding Remarks**

We study a tractable model that links the organization of financial markets with asset liquidity and prices. The model highlights the importance of competition between exchanges, their incentives to innovate to differentiate their services, and the linkages with investors’ choices. We show that trading fragmentation can improve market quality, as measured by traditional measures of liquidity, while having negative effects on asset prices in the long-run. The model also provides, to the best of our knowledge, the first analysis of the impact of order protection regulations on asset prices.

Our model suffers from several limitations and suggests interesting avenues for further research. First, it would be natural to introduce heterogeneity in investors’ information on the asset common value component. This would provide additional incentives for investors to demand speed. Promisingly, recent theory developments (e.g., Guerrieri, Shimer, and Wright (2010)) have made progress integrating search-like frictions and asymmetric information. Another interesting extension is to consider more sophisticated execution mechanisms. For example, one could exploit the modeling technology of Biais, Hombert, and Weill (2012) to include limit orders.

Empirically, it would be interesting to test the effects of fragmentation on prices more directly. In this regard, recent national regulation reforms that allow the introduction of new trading venues (e.g., South Korea, Australia, Brazil) can provide useful natural experiments for event studies. Such empirical work is important to guide policymakers and to further our understanding of the connections between secondary and primary markets and their effects on the cost of capital and other important variables for corporations.

The economics of the model are also relevant for large classes of derivatives that will start trading in exchanges in the near future. These market will see a growth in the number of
electronic platforms and liquidity will spread between Operating Trading Facilities, Swap Execution Facilities, and other institutions as mandated by regulations such as Dodd-Frank and EMIR. As derivatives and equity market structures continue to converge, understanding liquidity and pricing in multiple derivative markets will require some of the same techniques developed here.

References


Degryse, Hans, Frank De Jong, and Vincent Van Kervel, 2011, The impact of dark trading and visible fragmentation on market quality, . 4, 45


Frazzini, Andrea, and Lasse H. Pedersen, 2010, Betting Against Beta, . 5


Hasbrouck, Joel, and Gideon Saar, 2010, Low-Latency Trading, . 45


Rocheteau, Guillaume, and Randall Wright, 2005, Money in search equilibrium, in competitive equilibrium, and in competitive search equilibrium, *Econometrica* 73, 175–202. 19


Trejos, Alberto, and Randall Wright, 2012, Money and Finance: An Integrated Approach, . 9


50
Appendices

A Regulatory Frameworks

In this section we briefly discuss some regulations that are closely related to the model in the main body of the paper.

Entry of New Exchanges. The SEC introduced the Regulation of Exchanges and Alternative Trading Systems (Reg ATS) in 1998. It was designed with the aim of protecting investors and resolving concerns arising from alternative trading systems. A second objective of this regulation was to foster innovation in the space of trading systems and matching technologies by facilitating the entry of new participants. This intention is reflected in the “lax” control of trading venues that represent less than 5% of the trading volume for any given security.

In Europe, the Markets in Financial Instruments Directive (MiFID) implemented during the last half of the 2000s played a similar role in fostering competition between trading venues.

Investor Protection. There are essentially two approaches to investor protection: the trade-through model and the principles-based model.

Trade-Through Model. Under this approach price is the primary criterion for best execution. Market centers must be connected to one another and prevent trading through better prices available elsewhere, which requires complex connections as well as strong monitoring activity by regulators. In the U.S. the modern form of investor protection was introduced by Reg NMS. In particular, Rule 611 (trade-through) states that prices are quoted gross of trading fees (the SEC places a cap on fees) and only the top of the book is protected: When a big trading order arrives in a given marketplace, only the amount of shares represented by the depth of the book at the national best bid and offer is protected. As an example, suppose that NASDAQ and NYSE are the only market centers and that an investor submits a market order to buy 100,000 shares of a given stock to NASDAQ. Currently the ask price at the NASDAQ is higher than the ask price at the NYSE (where the ask depth is 10,000 shares).
Then NASDAQ can either match the price at the NYSE or the first execution occurs at the NYSE for 10,000 shares. The remaining 90,000 shares “walk up” the book at the NASDAQ.

In Canada, the Order Protection Rule (OPR) implemented by the Investment Industry Regulatory Organization of Canada shares the same spirit, but aims to protect orders throughout the entire order book (as opposed to just the top level).

**Principles-Based Model.** Criteria other than price, such as the type of investor behind the trade, are included in the best execution policy here. Thus, this approach allows for more discretion and less transparency in the assessment of execution results. This approach best represents the spirit of Europe’s MiFID and Japan’s Financial Instruments and Exchange Act. For example, Article 40-2(1) of the Financial Instruments and Exchange Act defines best execution policy as a “method for executing orders from customers ... under the best terms and conditions.” Some of the criteria to be considered are the listing trading venue, price, liquidity, execution probability, and execution speed. In Japan this system does not apply to professional investors. In both Europe and Japan, sell-side best execution policies are not obliged to consider every venue. The monitoring of execution quality is generally left to clients, which can be a problem in countries where investors have inadequate knowledge of financial markets. The claimed advantages of the principles-based approach lie in a much simpler set of linkages between markets and promoting innovation by not forcing uniformity.

**B Generalized Asset Holdings**

In this section we generalize investors’ asset demand holdings allowing for asymmetric shock frequency and general asset holdings $a \geq 0$.

Let the preference shocks be represented by $\epsilon_i \in \{\epsilon_t, \epsilon_h\}$ and let $\phi$ be the probability of the high shock. An investor of type $\sigma$ under state $\epsilon_i$ and with asset holdings $a \geq 0$ enjoys a
utility flow

\[ u_{i\sigma}(a) = \theta_{i\sigma} u(a) \]

where \( \theta_{i\sigma} = \mu + \epsilon_i \sigma \). The adjusted utility in Equation 39 becomes

\[
\overline{u}_{i\sigma}(a) = \frac{(r + \rho) u_{i\sigma}(a) + \gamma E_i [u'_{i\sigma}(a) \gamma + \rho] u(a)}{r + \gamma + \rho} = \bar{u}_{i\sigma}
\]

Optimal portfolio holdings in an interior solution satisfy \( \bar{u}_{i\sigma} u'(a) = r p \). Thus, whenever \( u \) is invertible we have

\[ a_{i\sigma}^* = (u')^{-1} \left( \frac{r p}{\bar{u}_{i\sigma}} \right). \]

The market clearing condition is

\[
\int_{\sigma} \sum_i \phi_i a_{i\sigma}^* (p) \, dG(\sigma) = \bar{u}.
\]

Consequently, the equilibrium price \( p \) solves

\[
\int_{\sigma} \sum_i \phi_i (u')^{-1} \left( \frac{r p (r + \gamma + \rho)}{(r + \gamma + \rho) \mu + (r + \rho) \epsilon_i \sigma + \gamma (2\phi - 1) \sigma} \right) dG(\sigma) = \bar{u}.
\]

**Example: CRRA utility** Let \( u(a) \) be given by \( a^{1-\xi} \). The asset demand function then satisfies \( \bar{u}_{i\sigma} a^{-\xi} = r p \), and thus

\[ a_{i\sigma}^* (p) = (r p)^{-\frac{1}{\xi}} (\bar{u}_{i\sigma})^{\frac{1}{\xi}}. \]

The market clearing condition is

\[
\int_{\sigma} \left\{ \phi (r p)^{-\frac{1}{\xi}} (\bar{u}_{i\sigma})^{\frac{1}{\xi}} + (1 - \phi) (r p)^{-\frac{1}{\xi}} (\bar{u}_{i\sigma})^{\frac{1}{\xi}} \right\} \, dG(\sigma) = \bar{u}
\]

\[
\Rightarrow \int_{\sigma} \left\{ \phi (\bar{u}_{i\sigma})^{\frac{1}{\xi}} + (1 - \phi) (\bar{u}_{i\sigma})^{\frac{1}{\xi}} \right\} \, dG(\sigma) = \bar{u} \left( r p \right)^{\frac{1}{\xi}}
\]

53
To simplify things further, consider the normalization $\mu = 0$, $\phi = 1/2$, and assume A.1. Then, the market clearing condition becomes

$$\frac{1}{2} \left( \frac{r + \rho}{r + \gamma + \rho} \right)^{\frac{1}{2}} \left( \epsilon_I^{\frac{1}{2}} + \epsilon_H^{\frac{1}{2}} \right) \int_\sigma \sigma^{\frac{1}{2}} e^{-\frac{\sigma}{\nu}} d\sigma = \bar{\alpha} (r \rho)^{\frac{1}{2}}.$$ 

Thus, the equilibrium price is given by

$$p = \frac{\nu}{r} \left[ \frac{1}{2\sigma} \left( \frac{r + \rho}{r + \gamma + \rho} \right) \left( \epsilon_I^{\frac{1}{2}} + \epsilon_H^{\frac{1}{2}} \right) \Gamma \left( 1 + \frac{1}{\xi} \right) \right]^{\xi}, \quad (37)$$

where $\Gamma$ denotes the Gamma function.

Note the following properties of equilibrium price 37. First, differently from the consolidated price in Section III, it decreases smoothly in the asset supply $\bar{\sigma}$. Second, whether higher market speed raises the asset price depends on the elasticity of asset demand, driven by parameter $\xi$. Consistent with Proposition 5 in Lagos and Rocheteau (2009), $\xi \in (0, 1)$ is a sufficient condition for $\frac{\partial p}{\partial \rho} > 0$. 

54
Asset Pricing Frictions in Fragmented Markets

-On Online Appendix-

Emiliano Pagnotta

New York University Stern School of Business

This Appendix comprises proofs of propositions and lemmas in the main paper.

Proof of Lemma 1

We follow Lagos and Rocheteau (2009) in expressing the optimal holdings problem recursively as follows\(^{38}\)

\[
a^* (p; \sigma, \epsilon) = \arg \max_{a \in \{0, 1\}} \{ \bar{u}(a; \sigma, \epsilon) - rpa \}
\]

where \(\bar{u}\), the adjusted holding utility, is given by

\[
\bar{u}(a; \sigma, \epsilon) = \frac{(r + \rho) \ u_{\sigma, \epsilon}(a) + \gamma \mathbb{E}[u_{\sigma, \epsilon'}(a) \mid \epsilon]}{r + \rho + \gamma}.
\]

The RHS of Equation 39, denoted as the adjusted holding utility, represents the expected average utility when holding the asset, for a given \(\epsilon\). Note that since \(\epsilon\) is i.i.d. with mean zero, we have \(\mathbb{E}[u_{\sigma, \epsilon'}(a) \mid \epsilon] = \mu a\) for any \(a\) and any \(\epsilon\). This expected utility over \(\epsilon'\) does not depend on \(\sigma\) or \(\epsilon\). This implies that

\[
\bar{u}(a; \sigma, \epsilon) = \left( \mu + \sigma \epsilon \frac{r + \rho}{r + \rho + \gamma} \right) a.
\]

From Equation 38 it is clear that \(\hat{\sigma}\) satisfies

\[\bar{u}(a; \hat{\sigma}, 1) = rpa.\]

\(^{38}\)See Lemma 1 there. The Lemma only needs to be adapted to take into account heterogeneity in \(\sigma\).
Combining Equations 40 and 41 when \( a = 1 \), yields Equation 4. Thus, the displayed demand functions represent optimal holdings.

**Proof of Proposition 1**

Case \( \bar{a} < 1/2 \). Combining Equations 4 and 5 yields \( \hat{\sigma} = G^{-1} (1 - 2\bar{a}) \). Replacing this value back in Equation 4, and re-arranging, delivers the equilibrium price in Equation 6. Also note that

\[
\lim_{\rho \to \infty} p = p_W = \frac{1}{r} \left[ \mu + G^{-1} (1 - 2\bar{a}) \right].
\]

We analyze next equilibrium allocations. Consider first a type \( (\epsilon = +1, a = 1) \). This type is satisfied with its current holding and does not trade even if it contacts the market. Outflows result only from changes of \( \epsilon \) from +1 to -1, which happens with intensity \( \gamma/2 \). There are two sources of inflow: types \( (\epsilon = -1, a = 1) \) that switch to \( \epsilon = +1 \) and types \( (\epsilon = +1, a = 0) \) that purchase one unit when they contact the market.

In steady state, outflows must equal inflows:

\[
2 \rho \alpha \sigma_+ (0) = \gamma \rho \alpha \sigma_- (1) + \frac{\gamma}{2} \alpha \sigma_+ (0).
\] (42)

Dynamics for types \( (\epsilon = -1, a = 0) \) are similar:

\[
\frac{\gamma}{2} \alpha \sigma_- (0) = \rho \alpha \sigma_- (1) + \frac{\gamma}{2} \alpha \sigma_+ (0).
\] (43)

For types \( (\epsilon = +1, a = 0) \) and \( (\epsilon = -1, a = 1) \) trade creates outflows so we have

\[
\left( \frac{\gamma}{2} + \rho \right) \alpha \sigma_+ (0) = \frac{\gamma}{2} \alpha \sigma_- (0)
\] (44)

\[
\left( \frac{\gamma}{2} + \rho \right) \alpha \sigma_- (1) = \frac{\gamma}{2} \alpha \sigma_+ (1)
\] (45)

Finally, the shares must add up to one, therefore

\[
\sum_{\epsilon = \pm, a = 0, 1} \alpha \sigma_{\epsilon} (a) = 1
\] (46)

To see the steady state allocations, add (42) and (45) to get \( \alpha \sigma_- (1) = \alpha \sigma_+ (0) \). This immediately implies \( \alpha \sigma_- (0) = \alpha \sigma_+ (1) \). Using (42), we obtain \( \alpha \sigma_+ (1) = \left( 1 + 2\frac{\rho}{\gamma} \right) \alpha \sigma_- (1) \). We can then solve for the shares of each type \( \alpha \sigma_+ (1) = \frac{\gamma + 2\rho}{4 \gamma + \rho} \) and \( \alpha \sigma_+ (0) = \frac{1}{4 \gamma + \rho} \). Notice also that the market clearing condition among high types is
simply \( \alpha_{\sigma,+} (1) + \alpha_{\sigma,-} (1) = 1/2 \). It is immediate that when \( \rho \to \infty \), the Walrasian limits of these allocations satisfy \( \alpha_{\sigma,-} (1) = \alpha_{\sigma,+} (0) = 0 \).

Case \( \bar{\sigma} \geq 1/2 \). When \( a > 1/2 \), sellers become the short side of the market and the equilibrium price equals
\[
p = \frac{\mu}{r} - \frac{\hat{\sigma}}{r} \left( \frac{r + \rho}{r + \gamma + \rho} \right).
\]
In the knife-edge case \( \bar{\sigma} = 1/2 \), the equilibrium price belongs to the interval
\[
\left[ \frac{\mu - \hat{\sigma}}{r} \left( \frac{r + \rho}{r + \gamma + \rho} \right), \frac{\mu + \hat{\sigma}}{r} \left( \frac{r + \rho}{r + \gamma + \rho} \right) \right].
\]

\[ \square \]

**Proof of Lemma 2**

This lemma is a particular case of Proposition 1 in Pagnotta and Philippon (2012), which applies to the costless participation environment here. We include the derivation of the function \( W \) for completeness. Define \( I_{\sigma,\epsilon} \equiv V_{\sigma,\epsilon} (1) - V_{\sigma,\epsilon} (0) \) as the value of owning the asset for type \((\sigma,\epsilon)\). Then, taking differences of Equations 7-10 we get
\[
r I_{\sigma,-} = \mu - \sigma + \frac{\gamma}{2} (I_{\sigma,+} - I_{\sigma,-}) + \rho (p - I_{\sigma,-})
\]
\[
r I_{\sigma,+} = \mu + \sigma - \frac{\gamma}{2} (I_{\sigma,+} - I_{\sigma,-}) - \rho (I_{\sigma,+} - p)
\]

We can then solve \( r (I_{\sigma,+} - I_{\sigma,-}) = 2\sigma - (\gamma + \rho) (I_{\sigma,+} - I_{\sigma,-}) \) and obtain the gains from trade for type \( \sigma \) in market \( \rho \): \( I_{\sigma,+} - I_{\sigma,-} = \frac{2\sigma}{r + \gamma + \rho} \). We then have \( I_{\sigma,\epsilon} = \frac{\mu + \rho p}{r + \rho} + \epsilon \frac{\sigma}{r + \gamma + \rho} \) and the average values
\[
\bar{V}_{\sigma} (0) = \frac{\rho}{2r} (I_{\sigma,+} - p)
\]
\[
\bar{V}_{\sigma} (1) = \frac{\mu}{r} + \frac{\rho}{2r} (p - I_{\sigma,-})
\]

where \( \bar{V}_{\sigma} (0) \equiv \frac{V_{\sigma,+(0)} + V_{\sigma,-(0)}}{2} \) and \( \bar{V}_{\sigma} (1) \equiv \frac{V_{\sigma,+(1)} + V_{\sigma,-(1)}}{2} \).

Let us now compute the ex ante value functions. Let us first consider types \( \sigma < \hat{\sigma} \). They join the market to sell at price \( p \), and then do not trade again. Averaging over types \( \epsilon = \pm 1 \), we get the ex ante value function \( \hat{W} \) that solves the Bellman equation
\[
r \hat{W} = \mu \bar{a} + \rho \left( p \bar{a} - \hat{W} \right) \implies \hat{W} = \frac{\mu + \rho p}{r + \rho} \bar{a}
\]

Since \( \mu + \rho p = \frac{\mu}{r} (r + \rho) + \rho (p - \frac{\mu}{r}) \) we can rewrite
\[
\hat{W} = \frac{\mu \bar{a}}{r} + \frac{\rho}{r + \rho} (rp - \mu) \frac{\bar{a}}{r}
\]
From the definition of \( \hat{\sigma} \) in Equation 4 and \( s(\rho) \equiv \frac{\rho}{r + \gamma + \rho} \) we then have \( \hat{W} = \frac{\mu \bar{a}}{r} + s(\hat{\sigma}) \hat{\sigma} \). Note that \( \hat{W} \) does not depend on the type \( \sigma \), but only on the price and speed of the market. Of course we also have \( \hat{W} = \bar{a} \tilde{V} \hat{\sigma} \quad (1) \).

Let us now consider the steady state types, \( \sigma > \hat{\sigma} \). Since the probability of owning one unit of the asset is \( \bar{a} \), we have

\[
W(\sigma) = \bar{a} \tilde{V}_\sigma (1) + (1 - \bar{a}) \tilde{V}_\sigma (0).
\]

Using the expression above, we get

\[
W_\sigma = \bar{a} \mu + \bar{a} \rho \frac{\rho}{2r} (\rho - I_{\sigma,-}) + (1 - \bar{a}) \frac{\rho}{2r} (I_{\sigma,+} - \rho) \\
= \frac{\mu \bar{a}}{r} + \bar{a} \rho \frac{\rho}{r + \rho} + \frac{1}{2r} \left( \frac{\rho}{r + \rho} (\mu - \rho) + \frac{\rho}{r + \gamma + \rho} \right) \\
= \frac{\mu \bar{a}}{r} + \bar{a} s(\rho) \hat{\sigma} + \frac{1}{2r} s(\rho) (\sigma - \hat{\sigma}).
\]

Therefore, we have, when \( \sigma > \hat{\sigma} \), we have

\[
W(\sigma, \rho) = \hat{W} + \frac{1}{2} \frac{s(\sigma - \hat{\sigma})}{r}.
\]

\[\square\]

**Proof of Proposition 2**

We proceed in two steps, first analyzing the FOC of Equation 19 with respect to \( q \) and \( s \), and then studying the price decomposition in Equation 21.

**Step 1.** Note that the FOC with respect to \( q \) is the same in the short and in the long-run equilibrium, as given by Equation 17. Under A.1, \( \hat{\sigma}_{con} = \hat{\sigma}_{con}^{S} = \nu \) and thus \( q_{con} = \frac{\bar{a} \nu}{r} \). We can then write the single exchange program as

\[
\max \quad 1 \geq s \geq 0 \quad \frac{s \nu}{2r} \tilde{C}(s) = -C(s).
\]

The interior solution FOC with respect to \( s \) is \((1 - s)^2 = \left( \frac{\nu}{\tilde{c}} \right) 2r c (\gamma + r) \). The optimal effective speed is then

\[
s_{con} = 1 - \left( 2r c e (\gamma + r) \right)^{1/2} \nu^{-1/2}.
\] (47)
Step 2. Using Equation 16, and $\lambda \equiv \frac{r+\gamma s}{r+\gamma}$ we have

$$p_{con}^L = \frac{\mu}{r} + \frac{\hat{\sigma}_{con}}{r}\lambda_{con}$$  \hfill (48)

Adding and subtracting $\frac{1}{r}G^{-1}(1-2\bar{a})$ and $\frac{\hat{\sigma}_{con}}{r}$ to the RHS of Equation 48, and rearranging, we have

$$p_{con}^L = \frac{1}{r} \left[ \mu + G^{-1}(1-2\bar{a}) \right] + \frac{1}{r} \left[ \hat{\sigma}_{con} - G^{-1}(1-2\bar{a}) \right] - \frac{\hat{\sigma}_{con}}{r} (1 - \lambda_{con}) \quad (49)$$

Note that the first term of the RHS of 49 corresponds to the expression for $p_W$ in Proposition 1. Further, note from Equation 47 that $\lim_{c \to 0}s_{con} = 1$, which implies $\lim_{c \to 0}\lambda_{con} = 1$. Using this fact and Definition 2, we have that the limited participation distortion equals

$$\lim_{c \to 0} [p - p_W] = \frac{1}{r} \left[ \hat{\sigma}_{con} - G^{-1}(1-2\bar{a}) \right] = \frac{\nu}{r} \left[ 1 + \ln (2\bar{a}) \right],$$

where the second equality uses A.1. Finally, the illiquidity discount is given by

$$\frac{\hat{\sigma}_{con}}{r} (1 - \lambda_{con}) = \frac{\nu}{r} \left( \gamma \sqrt{\frac{2rc}{r + \gamma \nu}} \right),$$

where the second equality uses A.1 and Equation 47.

Proof of Proposition 3

The decomposition of the equilibrium price is an in Proposition 2. To derive the components of the equilibrium price, we proceed in two steps. First, we derive the first order condition of the investor speed choice problem, which yields $\sigma \mapsto \theta (\sigma)$; and we compute the marginal type $\hat{\sigma}$ for a given participation fee $q$. Second, we solve the exchange optimization program. Finally, we analyze the effect of changes in $\bar{a}$ on $s$.

Step 1. Investors optimization.

Before trading, a type $\sigma$ investor maximizes the RHS of Equation 24, which can be written as
\[
\max_{0 \leq \theta \leq \hat{\sigma}} \left\{ s(\theta) \left[ \frac{\bar{\sigma}}{r} + \frac{1}{2r} \max(0; \sigma - \hat{\sigma}) \right] - c_I \theta \right\}.
\]

The interior first order condition at \( \sigma = \hat{\sigma} \) yields

\[
\theta(\hat{\sigma}) = \max \left\{ 0, \min \left\{ \bar{\sigma} \left( \frac{\alpha (\gamma + r)}{c_I} \right)^{\frac{1}{2}} - (r + \gamma + \rho - \bar{\theta}) \right\} \right\}, \tag{50}
\]

which by Equation 22 implies that, at an interior solution,

\[
s(\hat{\sigma}) = 1 - \left[ \frac{r}{\alpha} (\gamma + r) c_I \right]^{\frac{1}{2}} \hat{\sigma}^{-\frac{1}{2}} = 1 - (y\hat{\sigma})^{\frac{1}{2}}, \tag{51}
\]

where \( y \equiv \frac{r}{\alpha} (\gamma + r) c_I \). The marginal investor \( \hat{\sigma} \) satisfies

\[
q = \bar{W}(\hat{\sigma}, \hat{\sigma}, s(\hat{\sigma})) - W_{out} = s(\hat{\sigma}) \frac{\bar{\sigma}}{r} - c_I \theta(\hat{\sigma}). \tag{52}
\]

Then we have

\[
\bar{W} = \frac{\bar{\sigma}}{r} \left[ \hat{\sigma} - 2(y\hat{\sigma})^{\frac{1}{2}} \right] + c_I (r + \gamma + \rho - \bar{\theta}).
\]

**Step 2. Exchange optimization.**

Using Equation 52 we can write the exchange program as

\[
\max_{\hat{\sigma} \in [0, \bar{\sigma}]} \left\{ \left( \frac{\bar{\sigma}}{r} s(\hat{\sigma}) \right) \hat{\sigma} - c_I \theta(\hat{\sigma}) \right\} [1 - G(\hat{\sigma})].
\]

Using \( \frac{1 - G(\sigma)}{g(\sigma)} = \nu \) by A.1, we can write the first order condition as follows

\[
\nu \left( \frac{\bar{\sigma}}{r} (s(\hat{\sigma}) + \hat{\sigma} s'(\hat{\sigma})) - c_I \theta'(\hat{\sigma}) \right) = \frac{\bar{\sigma}}{r} s(\hat{\sigma}) - c_I \theta(\hat{\sigma}). \tag{53}
\]

Using Equations 50 and 51, we can express Equation 53 as the following polynomial equation

\[
F(\hat{\sigma}) \equiv \hat{\sigma}^{\frac{3}{2}} - 2y^{\frac{1}{2}}\hat{\sigma} + \left( \frac{r}{\alpha} c_I (r + \gamma + \rho - \bar{\theta}) - \nu \right) \hat{\sigma}^{\frac{1}{2}} + \nu y^{\frac{1}{2}}. \tag{54}
\]

The polynomial 54 has three roots. Descartes’ rule of signs suggest that \( F(\hat{\sigma}) \) has two or zero positive roots. Note that this holds regardless of the sign of the \( \hat{\sigma}^{\frac{1}{2}} \) term coefficient. A numerical analysis based on our baseline calibration indicates that \( F \) has two positive roots, only one of which satisfies the condition \( s(\hat{\sigma}) > \frac{\bar{\sigma}}{2} \). Substituting
this solution, $\sigma_{h, f t}$, in Equation 51, and using $\lambda \equiv \frac{r + \gamma}{r + \gamma}$, yields $\lambda_{h, f t}$.

Computing $\frac{ds}{da}$.

Using 51, we can re-express 54 as

$$H(s, \bar{a}) = \frac{r}{\bar{a}}(\gamma + r) c_l (2s - 1) \left(1 - s\right)^2 - \nu s + \frac{r}{\bar{a}} c_l \left(r + \gamma + \rho - \bar{\theta}\right). \quad (55)$$

Given 55, we have $\frac{ds}{da} = -\frac{H_{\bar{a}}}{H_s}$. Computing the corresponding derivatives, we obtain

$$\frac{ds}{da} = \frac{(2s - 1)(1 - s) + (1 - s)^3 \left(1 + \frac{r - \bar{\theta}}{r + \gamma}\right)}{2\bar{a}(3s - 1)}. \quad (57)$$

Note that if $s \approx 1$ then we have $\frac{ds}{da} > 0$. Consequently, when market frictions are relatively small, investors’ investment in low latency technology increase with $\bar{\pi}$.

Proofs of Proposition 4

In the case of segmented markets prices are formed independently. We can then apply the single equilibrium formula in Equation 16 to market $i \in \{1, 2\}$

$$p_i = \frac{\mu}{r} + \frac{\hat{\sigma}_i}{r} \lambda_i. \quad (56)$$

Adding and subtracting $\frac{1}{r} G^{-1} (1 - 2\bar{\pi})$ and $\frac{\hat{\sigma}_i}{r}$ to the RHS of Equation 56, and re-arranging, we have

$$p_i = p_W + \frac{1}{r} \left[\hat{\sigma}_i - G^{-1} (1 - 2\bar{\pi})\right] - \frac{\hat{\sigma}_i}{r} (1 - \lambda_i), \quad (57)$$

which yields Equations 30 and 31. For temporary investors to be indifferent between joining and staying out, we must have $W(\hat{\sigma}_i, \hat{\sigma}_i, s_i) - W_{\text{out}} - q_i = 0$ in each market $i$. Otherwise all the low types would strictly prefer one market to another. For market 1 this condition is satisfied by Equation 26. For market 2 then we must have

$$q_2 = \frac{\bar{a}s_2 \hat{\sigma}_2}{r}. \quad (58)$$

Recall that the type that is indifferent between joining market 1 and 2 is given by Equation 28. When both markets are active we have $\hat{\sigma}_1 < \hat{\sigma}_2 < \hat{\sigma}_{12}$, which implies
that the limited participation distortion is higher in market 2. Note that whether the illiquidity premium is higher in market one depends on whether \( \hat{\sigma}_1 (1 - \lambda_1) \leq \hat{\sigma}_2 (1 - \lambda_2) \). Since \( \hat{\sigma}_1 < \hat{\sigma}_2 \) and \( 1 - \lambda_1 > 1 - \lambda_2 \), the relative size of the illiquidity premiums is not determined a priori.

We now characterize the conditions that determine the equilibrium values of \((\hat{\sigma}_{seg}^1, \hat{\sigma}_{seg}^{12})\).

In the fee competition stage, exchange \( i \in \{1, 2\} \) seek to maximize the profit function \( \pi_i^{seg} \) of Section IV with respect to \( q_i \), taking as given the conditions describing the affiliation of investors to markets 1 and 2, that is Equations 26 and 28. The system of first-order conditions is then given by

\[
1 - G(\hat{\sigma}_{12}) = g(\hat{\sigma}_{12}) (\hat{\sigma}_{12} + \hat{\sigma}_1 k), \tag{59}
\]

\[
G(\hat{\sigma}_{12}) - G(\hat{\sigma}_1) = (g(\hat{\sigma}_1) + k g(\hat{\sigma}_{12})) \hat{\sigma}_1, \tag{60}
\]

where \( k \equiv \frac{\hat{\sigma}_1}{s_2 - s_1} \). Given the equilibrium values of \((\hat{\sigma}_1, \hat{\sigma}_{12})\), from Equation 28 we have

\[
\hat{\sigma}_2 = \frac{1}{s_2} (\hat{\sigma}_{12} (s_2 - s_1) + s_1 \hat{\sigma}_1). \tag{61}
\]

The types \( \hat{\sigma}_1, \hat{\sigma}_2 \) and \( \hat{\sigma}_{12} \) are then defined by Equations 59-61.

Proof of Proposition 5

In the protected market case there is a single price \( p_{nb} \) satisfying

\[
p_{nb} = \frac{\mu}{r} + \frac{\hat{\sigma}_1^{prot}}{r} \lambda_1. \tag{62}
\]

Adding and subtracting \( \frac{1}{r} G^{-1} \left(1 - 2\alpha_1 \right), \frac{\hat{\sigma}_{seg}^{1}}{r}, \) and \( \frac{\hat{\sigma}_{seg}^{12}}{r} \lambda_1 \) to the RHS of Equation 56, and re-arranging, yields Equation 33.

We now show that under A.1 the price protection distortion is positive, which is equivalent to \( \sigma_1^{prot} - \sigma_1^{seg} > 0 \). From Equation (4), a single asset price in both markets implies

\[
\left(1 + \frac{\gamma}{r + \rho_1}\right) \hat{\sigma}_2 = \left(1 + \frac{\gamma}{r + \rho_2}\right) \hat{\sigma}_1. \tag{63}
\]
This means that in the protected case, \( \hat{\sigma}_2 < \hat{\sigma}_1 \). For \( \hat{\sigma}_{12} \) we have

\[
\frac{s_2 \hat{\sigma}_2}{r} + \frac{s_2}{2r} (\hat{\sigma}_{12} - \hat{\sigma}_2) - q_2 = \frac{s_1 \hat{\sigma}_1}{r} + \frac{s_1}{2r} (\hat{\sigma}_{12} - \hat{\sigma}_1) - q_1.
\]

We can write it as

\[
\frac{s_2 - s_1}{2r} \hat{\sigma}_{12} = q_2 - q_1 + \left( \frac{s_1 \hat{\sigma}_1}{r} - \frac{s_2 \hat{\sigma}_2}{r} \right) \left( \hat{a} - \frac{1}{2} \right)
\]

then we can use (63) to write \( \hat{\sigma}_2 = \frac{1 + \frac{\gamma}{r + \rho_2}}{1 + \frac{\gamma}{r + \rho_1}} \hat{\sigma}_1 \). Then

\[
\frac{s_2 - s_1}{2r} \hat{\sigma}_{12} = q_2 - q_1 + \frac{s_1 \hat{\sigma}_1}{r} \left( 1 - \frac{s_2}{s_1} \frac{1 + \frac{\gamma}{r + \rho_2}}{1 + \frac{\gamma}{r + \rho_1}} \right) \left( \hat{a} - \frac{1}{2} \right)
\]

and since \( \hat{\sigma}_1 = \frac{\gamma}{\alpha \rho_1} \) we get

\[
\frac{s_2 - s_1}{2r} \hat{\sigma}_{12} = q_2 - q_1 \left( \frac{1}{2} \hat{a} - \frac{\frac{\rho_1}{r + \rho_1}}{\frac{\rho_2}{r + \rho_2}} \left( \frac{1}{2} \hat{a} - 1 \right) \right).
\]

Re-arranging the above expression yields

\[
\hat{\sigma}_{12}^{\text{prot}} = \frac{2r}{s_2 - s_1} \left( q_2 - \frac{z}{2\hat{a}} q_1 \right),
\]

where \( z \equiv 1 - \frac{1 + \frac{\gamma}{r + \rho_2}}{1 + \frac{\gamma}{r + \rho_1}} (1 - 2\hat{a}) \).

The profits of market 1 are \( q_1 (G(\hat{\sigma}_{12}) - G(\hat{\sigma}_1) + \hat{\sigma}_1) \). To simplify the notation, let \( \alpha \equiv 2\hat{\alpha} \) and \( k \equiv \frac{s_1}{s_2 - s_1} \). Using Equation 32, we can write exchanges’ second stage programs as

\[
\max_{q_1} \pi_1^{\text{prot}} = \frac{q_1}{\alpha} (1 - \alpha + \alpha G(\hat{\sigma}_{12}) - G(\hat{\sigma}_1))
\]

\[
\max_{q_2} \pi_2^{\text{prot}} = q_2 (1 - G(\hat{\sigma}_{12}))
\]

The conditions \( \frac{\partial \pi_1^{\text{prot}}}{\partial q_1} = 0 \) and \( \frac{\partial \pi_2^{\text{prot}}}{\partial q_2} = 0 \) lead to

\[
1 - G(\hat{\sigma}_{12}) = g(\hat{\sigma}_{12}) (\hat{\sigma}_{12} + z k \hat{\sigma}_1)
\]

\[
1 - \alpha + \alpha G(\hat{\sigma}_{12}) - G(\hat{\sigma}_1) = (g(\hat{\sigma}_1) + \alpha z k g(\hat{\sigma}_{12})) \hat{\sigma}_1.
\]
Using A.1, we can express the system above as

\[
\frac{\ddot{\sigma}_{12}}{\nu} = 1 - z_k \frac{\ddot{\sigma}_1}{\nu} \\
\frac{\ddot{\sigma}_1}{\nu} \left( e^{\frac{\ddot{\sigma}_{12}-\ddot{\sigma}_1}{\nu}} + \alpha z_k \right) = e^{\frac{\ddot{\sigma}_{12}-\ddot{\sigma}_1}{\nu}} - \alpha
\]

It is convenient to defined $\Delta \equiv (\ddot{\sigma}_{12} - \ddot{\sigma}_1)/\nu$ and $x \equiv \frac{\ddot{\sigma}_1}{\nu}$, so that we can write the system in $(x, \Delta)$:

\[
(1 + z_k) x = 1 - \Delta \\
e^\Delta - \alpha = (e^\Delta + \alpha z_k) x
\]

The second equation of the system is $1 - x = \frac{\alpha (1 + zk)}{e^\Delta + \alpha zk}$. This leads to a schedule $x$ increasing in $\Delta$. The issue is how it changes with $\alpha$. We study the function on the RHS, namely: $\log \left( \frac{\alpha (1 + zk)}{e^\Delta + \alpha zk} \right) = \log (\alpha) + \log (1 + zk) - \log (e^\Delta + \alpha zk)$. Taking the derivative w.r.t. $\alpha$

\[
\frac{1}{\alpha} + \frac{kz'}{1 + zk} - \frac{\alpha k z' + kz}{e^\Delta + \alpha k z} = \frac{1}{\alpha} - \frac{1}{\alpha + \frac{e^\Delta}{k z}} + kz' \left( \frac{1}{1 + k z} - \frac{1}{\alpha + k z} \right)
\]

since $\frac{e^\Delta}{\alpha} > 1$ we have $\frac{1}{1 + k z} - \frac{1}{\alpha + \frac{e^\Delta}{k z}} > 0$. Similarly $\frac{1}{\alpha} - \frac{1}{\alpha + \frac{e^\Delta}{k z}} > 0$. So $\frac{\alpha (1 + zk)}{e^\Delta + \alpha zk}$ is increasing in $\alpha$. Therefore the equilibrium condition $e^\Delta - \alpha = (e^\Delta + \alpha zk) x$ implies a schedule $x$ increasing in $\Delta$ and decreasing in $\alpha$. The first equilibrium condition $(1 + zk) x = 1 - \Delta$ gives a schedule $x$ decreasing in $\Delta$ and decreasing in $\alpha$. Straightforward analysis then shows that $x$ must be decreasing in $\alpha$. The free price structure corresponds to $\alpha = 1$, while the protected price structure corresponds to $\alpha = 2a < 1$. Therefore, since $\ddot{\sigma}_1 = \nu x$, $\ddot{\sigma}_1$ must be higher under price protection.

\[\square\]

**Proof of Proposition 6**

To prove that the limited participation distortion is always larger is a consolidated market it suffices to show that $\ddot{\sigma}_{con} > \ddot{\sigma}_2$. Notice that given $\ddot{\sigma}_1 < \ddot{\sigma}_2 < \ddot{\sigma}_{12}$, a sufficient condition is $\ddot{\sigma}_{con} \geq \ddot{\sigma}_{12}$. For the single exchange, $\ddot{\sigma}_{con}$ satisfies Equation 17

\[
\ddot{\sigma}_{con} = \frac{1 - G(\ddot{\sigma}_{con})}{g(\ddot{\sigma}_{con})}.
\] (69)

From Equation 59 we have

\[
\ddot{\sigma}_{12} = \frac{1 - G(\ddot{\sigma}_{12})}{g(\ddot{\sigma}_{12})} - \ddot{\sigma}_1 \frac{s_1}{s_2 - s_1}
\]
Since $\hat{s}_1 \frac{\hat{s}_1 - s_1}{s_2 - s_1} \geq 0$, the log-concavity of $g$ implies $\hat{s}_{\text{con}} \geq \hat{s}_{12}$ and thus

$$\frac{1}{r} \left[ \hat{s}_{\text{con}} - G^{-1} (1 - 2\bar{a}) \right] > \frac{1}{r} \left[ \hat{s}_2 - G^{-1} (1 - 2\bar{a}) \right].$$

Note that given the pricing expressions in Section III-IV, $p_{\text{con}} \geq p_2$ depends on whether

$$\frac{\hat{s}_{\text{con}}}{\hat{s}_2} \geq \frac{r + \gamma s_2}{r + \gamma s_{\text{con}}}. \quad (70)$$

As market contact frictions become small ($s$ approaches one), the RHS of inequality 70 approaches one, and the ability of exchanges to differentiate their services decreases, reducing $\hat{s}_2$. On the other hand, Equation 69 shows that $\hat{s}_{\text{con}}$ is unaffected. This implies that inequality 70 holds strictly for a large enough value of $s$.

Finally, the inequality $s_{\text{seg}}^2 > s_{\text{con}}$ follows from Proposition 5 in Pagnotta and Philippon (2012), which we provide below as a Lemma for completeness. This fact implies that in the long-run $1 - \lambda_{\text{seg}}^2 < 1 - \lambda_{\text{con}}$ and thus the illiquidity discounts satisfy: $\hat{s}_{\text{seg}}^2 (1 - \lambda_{\text{seg}}^2) < \hat{s}_{\text{con}} (1 - \lambda_{\text{con}})$.

**Lemma 3.** Under A.1 $s_{\text{seg}}^2 > s_{\text{con}}$.

**Proof.** Under A.1 and with $\alpha = 1$, we have $\hat{s}_{12} = \nu - \frac{s_1}{s_2 - s_1} \hat{s}_1$ and $q_2 = \frac{\nu}{2r} (s_2 - s_1)$. The profits of the fast venue are $\pi_2 = q_2 (1 - G(\hat{s}_{12}))$, and therefore

$$\pi_2 = \frac{\nu}{2r} (s_2 - s_1) (1 - G(\hat{s}_{12}))$$

Note that this system is equivalent to the monopoly case when $s_1 = 0$. The FOC for speed is

$$2rC'(s_2) = \nu (1 - G(\hat{s}_{12})) - \nu (s_2 - s_1) g(\hat{s}_{12}) \frac{\partial \hat{s}_{12}}{\partial s_2}. \quad (71)$$

The consolidated solution is $2rC'(\bar{s}_2) = \nu e^{-1}$. With two active venues we have $\frac{\partial \hat{s}_{12}}{\partial s_2} = \frac{k}{s_2 - s_1} \hat{s}_1 - k \frac{\partial \hat{s}_1}{\partial s_2}$. Then,

$$2rC'(s_2) = \nu (1 - G(\hat{s}_{12})) - \nu g(\hat{s}_{12}) \left[ k \hat{s}_1 - s_1 \frac{\partial \hat{s}_1}{\partial s_2} \right]$$

$$= e^{-\frac{\pi_2}{2r}} \left( \nu - \left[ k \hat{s}_1 - s_1 \frac{\partial \hat{s}_1}{\partial s_2} \right] \right).$$

Using $x \equiv \frac{\hat{s}_1}{\nu}, \Delta \equiv \frac{\hat{s}_{12} - \hat{s}_1}{\nu}$

$$2rC'(s_2) = \nu e^{kx - 1} \left( 1 - kx + s_1 \frac{\partial x}{\partial s_2} \right). \quad (72)$$

Since $C'$ is an increasing function, market 2 chooses a higher speed whenever the RHS of 72 is greater
than $\nu e^{-1}$. That is,

$$e^{kx} \left( 1 - kx + s_1 \frac{\partial x}{\partial s_2} \right) - 1 > 0. \quad (73)$$

Now we derive $\frac{\partial x}{\partial s_2}$. Differentiating the system 67-68 we have

$$(1 + k) dx + d\Delta - \frac{k}{(s_2 - s_1)} ds_2 = 0$$

$$\left( e^\Delta + k \right) dx + e^\Delta (x - 1) d\Delta - \frac{k}{(s_2 - s_1)} ds_2 = 0.$$

After appropriate substitutions we get

$$\frac{\partial x}{\partial s_2} = \frac{k^2 x (1 + e^\Delta (1 - x))}{e^\Delta (1 + \Delta) + k (1 + e^\Delta)}. \quad (74)$$

In order to verify 73, we plug 74 in 73, and define

$$S(k) \equiv e^{kx} \left( 1 - kx + \frac{k^2 x (1 + e^\Delta (1 - x))}{e^\Delta (1 + \Delta) + k (1 + e^\Delta)} \right) - 1. \quad (75)$$

Re-arranging we have

$$S(k) = e^{kx} \left( \frac{e^\Delta (1 + \Delta) + k (1 + e^\Delta) - kxe^\Delta (1 + \Delta - kx)}{e^\Delta (1 + \Delta) + k (1 + e^\Delta)} \right) - 1. \quad (76)$$

To satisfy the inequality we need $S(k) > 0$ for all $k > 0$ and $S(0) = 0$ (corresponding to the monopolist case where $s_1 = 0$). Let $x(k)$ and $\Delta(k)$ denote the solutions to the system 67-68 for a given $k \geq 0$. Since $x(k)$ and $\Delta(k)$ are continuous functions, $S(k)$ is continuous. Using 67-68 one can see that $\lim_{k \to \infty} x(k) = 0$ and $\lim_{k \to \infty} \Delta(k) = \bar{\Delta}$, where $\bar{\Delta}$ is defined by $e^\Delta + \Delta = 2$. Notice that $\lim_{k \to \infty} x(k) = 1 - \bar{\Delta}$. Similarly, $\lim_{k \to 0} x(k) = 1 - \bar{\Delta}$ and $\lim_{k \to 0} \Delta(k) = \bar{\Delta}$, where $\bar{\Delta}$ is defined by $e^\Delta \bar{\Delta} = 1$. Taking limits of 75 we find $\lim_{k \to 0} S(k) = e^0 - 1 = 0$ and $\lim_{k \to \infty} S(k) = e^{1 - \bar{\Delta}} - 1 > 0$.

A sufficient condition for $S(k) > 0$ for all $k > 0$ is to show that the term between brackets in 76 is greater than one. This is the case whenever

$$e^\Delta (1 + \Delta + k) + k + e^\Delta k \left[ (1 - x) + (xk)^2 - x\Delta \right] > e^\Delta (1 + \Delta + k) + k \quad (77)$$

Note from 67 that $1 - x = kx + \Delta$. Then,

$$(1 - x) + (xk)^2 - x\Delta = kx + \Delta (1 - x) + (xk)^2 > 0.$$

We conclude that $S(k) > 0$ for all $k > 0$. 

$\square$
**Proof of Proposition 7**

To prove the first part of the proposition we need to show that \( \hat{\sigma}_{\text{con}} \lambda_{\text{con}} > \hat{\sigma}_1^{\text{prot}} \lambda_1 \). Using the system of Equations 65-66, and following the steps of Proposition 6 it is easy to show that \( \hat{\sigma}_{\text{con}} > \hat{\sigma}_1^{\text{prot}} \). Given the value of the default technology \( s \) we have \( \lambda_{\text{con}} \geq \lambda_1 \), with strict inequality any time that the single exchange invests in speed. We conclude that \( p_{\text{con}} > p_{\text{nb}} \).

Following Proposition 5, under A.1 we have \( \hat{\sigma}_{1}^{\text{seg}} < \hat{\sigma}_1^{\text{prot}} \) and thus \( p_{\text{nb}} > p_1 \). To prove that \( p_{\text{nb}} < p_2 \) it is sufficient to show that \( \sigma_{2}^{\text{seg}} > \sigma_1^{\text{prot}} \). From Equation 28 we have that

\[
\frac{\sigma_{2}^{\text{seg}}}{s_2 - s_1} = \frac{\sigma_{12}^{\text{seg}}}{s_2} + \frac{s_1}{s_2 - s_1}.
\]

The RHS of this expression is according to Equation 59 equal to \( \nu \), and thus we have \( \frac{\sigma_{2}^{\text{seg}}}{\nu} = \frac{s_2 - s_1}{s_2} = \frac{1}{1 + k} \). According to Equation 65 we have \( \frac{\sigma_{12}^{\text{seg}}}{\nu} = \left( \frac{1 - \Delta}{1 + \alpha k} \right) \) and thus, provided \( \alpha \) is large enough, we have \( \sigma_{2}^{\text{seg}} > \sigma_1^{\text{prot}} \).

Finally note that when frictions are relatively large (\( s \) low) and/or costs are low, we have \( s_2 \approx 1 \) and \( s_1 \ll s_2 \). By Equation 28 this increases \( \hat{\sigma}_{12}^{\text{seg}} \). The combined effect of these factors is to lower traded volume in market 1 (Equation 35) and to increase volume in market 2 (Equation 36), which makes \( p_{vw} \) closer to \( p_2^{\text{seg}} \) and thus higher than \( p_{\text{nb}} \).

\( \square \)

**Proof of Proposition 8**

I calculate the comparative statics for the short and long-term consolidated prices.

The effect \( \frac{\partial p_{\text{con}}^S}{\partial \nu} > 0 \) is immediate. The expression \( \frac{\partial p_{\text{con}}^L}{\partial \nu} \) is given by

\[
\frac{\partial p_{\text{con}}^L}{\partial \nu} = \frac{\partial p_{\text{con}}^S}{\partial \nu} + \frac{\partial}{\partial \nu} \left( \frac{r + \gamma s_{\text{con}}}{r + \gamma} \right).
\]

By Equation 47 we have \( \frac{\partial s_{\text{con}}}{\partial \nu} > 0 \) and thus \( \frac{\partial p_{\text{con}}^L}{\partial \nu} > \frac{\partial p_{\text{con}}^S}{\partial \nu} \).

Note that in a consolidated market the effect of a change in \( \gamma \) only affects the illiquidity discount term. In the short-run price, \( \frac{\partial p_{\text{con}}^S}{\partial \gamma} \) is given by \( \frac{\nu(s-1)}{(r+\gamma)^2} \), which is negative since \( s \leq 1 \). In the long-run, the effect is given by

\[
\frac{\partial p_{\text{con}}^L}{\partial \gamma} = - \left( \frac{ce\gamma}{2r} \right)^{\frac{1}{2}} \left( \frac{2r + \gamma}{(r + \gamma)^2} \right)^{\frac{3}{2}},
\]

13
which is clearly negative. It is immediate that the ratio \( \left| \frac{\partial p_{\text{con}}^S}{\partial \gamma} \right| / \left| \frac{\partial p_{\text{con}}^L}{\partial \gamma} \right| \) is greater than one provided
\[
1 - s_{\text{con}} < \frac{r}{r + \gamma} (1 - s),
\]
or equivalently using Equation 47
\[
c < \left( \frac{1 - s}{2r + \gamma} \right)^2 \frac{2rv}{(r + \gamma)} e.\]

Finally, the cost parameter \( c \) effect on the price is given by
\[
\frac{\partial p_{\text{con}}^L}{\partial c} = -\gamma \left( \frac{ev}{2cr (r + \gamma)} \right)^\frac{1}{2} < 0.\]
\[
\square
\]
<table>
<thead>
<tr>
<th><strong>Title:</strong> Fleeting Orders and the Competitive Equilibrium</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author 1:</strong></td>
</tr>
<tr>
<td><strong>First Name:</strong> Shmuel</td>
</tr>
<tr>
<td><strong>Last Name:</strong> Baruch</td>
</tr>
<tr>
<td><strong>Organization:</strong> University of Utah</td>
</tr>
<tr>
<td><strong>Country:</strong> United States</td>
</tr>
<tr>
<td><strong>Email:</strong> <a href="mailto:shmuel.baruch@business.utah.edu">shmuel.baruch@business.utah.edu</a></td>
</tr>
<tr>
<td><strong>Author 2:</strong></td>
</tr>
<tr>
<td><strong>First Name:</strong> Larry</td>
</tr>
<tr>
<td><strong>Last Name:</strong> Glosten</td>
</tr>
<tr>
<td><strong>Organization:</strong> Columbia University</td>
</tr>
<tr>
<td><strong>Country:</strong> United States</td>
</tr>
<tr>
<td><strong>Email:</strong> <a href="mailto:lrg2@columbia.edu">lrg2@columbia.edu</a></td>
</tr>
</tbody>
</table>
Fleeting Orders and the Competitive Equilibrium

Abstract

We consider a dynamic model of strategic liquidity suppliers. Their limit orders are hit by either a "news" (i.e. informed) trader or a noise trader. We show that within this environment we can always find, regardless of the distributions of the news and noise trade quantity, a mixed strategy equilibrium and hence fleeting orders. Furthermore, the resulting random supply schedule converges, as the number of liquidity suppliers increases to infinity, to the competitive supply function.

An oft noted feature of today’s equity markets is that quoting and quickly canceling are common and frequent events—orders are fleeting and quotes flicker. Nanex reports that by the end of 2010, the ratio of messages to trades is on the order of 25.¹ Nanex also reports that by the end of 2011, there were, on average and during the middle of the day approximately 120 quotes per $10,000 transacted.² We have evidence of this general property, though in a less extreme form, going back to 1999, well before Reg NMS and machine trading. Hasbrouck and Saar (2007) document that in 1999 data, on the Island ECN, more than a quarter of quotes were canceled within two seconds.³ Joel Hasbrouck and Gideon Saar speculate on the reasons for

¹http://www.nanex.net/FlashCrash/QuoteRates.html
² http://www.nanex.net/Research/QuoteStuffingBanned/QuoteStuffingBanned.html
³In 2000 we thought two seconds was a short period of time
this rapid cancellation of orders. One possibility is "spoofing" which is a manipulative act designed to get the market thinking there is more buying (or selling) interest than there actually is.\textsuperscript{4} They dismiss this as unlikely, given the ease with which it can be detected.\textsuperscript{5} Another possibility is that the orders are designed to look for "hidden liquidity" by, for example, placing a buy order just below the offer. This could more easily be accomplished by using a "fill or kill" order and furthermore it does not characterize all of the quote submissions and cancellations. In 2000 it was possible to go to the Island website and watch the limit order book and observe these "flickering" quotes, even away from the market. The flickering was apparent even in very liquid securities such as the QQQQ. Hasbrouck (2012) also has evidence that the variance of high frequency quoting does not show a significant time trend from 2001 to 2011. That paper also reports that the volatility of quote changes at one millisecond intervals is nearly five times what would be predicted by thirty-four minute quote change volatility.

This quote behavior does not appear to coincide with the equilibrium in Glosten (1994) or in Biais et al. (2000), both of which predict a well-behaved supply curve which responds to transactions and other new information. It seems rather implausible to think new information was, in 1999, coming in on a second by second basis, or in 2012, on a millisecond by millisecond basis. What this paper explores is the possibility that liquidity suppliers are

\textsuperscript{4}Nanex suspects that today “quote stuffing” may be part of a manipulation. The site shows instances of more that 8,000 quote changes per second, for several seconds, with no trade whatsoever. http://www.nanex.net/aqck/2816.HTML

\textsuperscript{5}Though successful prosecution of manipulation is extremely difficult
playing the liquidity provision game differently. They play the game using mixed strategies and the flickering quotes represent the liquidity providers rapidly submitting quotes, canceling them and resubmitting another round of randomly generated quotes.

This is our first contribution. We show that fleeting orders arising from players using mixed strategies is precisely what we should expect to see. While pure strategy supply schedules may sometimes form an equilibrium, most likely, they will not. Existence of the pure strategy equilibrium often requires rather extreme adverse selection in that the probability of trading with a news trader must be quite high.

It is easy to see why we should see flickering quotes. Once a player observes the quotes of others, she will want to cancel the order and pick a price just below the price right above hers (or someplace else). But everyone knows that everyone else will want to change and they are back to picking another price at random. Thus, fleeting orders are not necessarily a part of a nefarious plan to overload the market (i.e. quote stuffing) nor are they part of a "manipulation." They are, rather, just a part of the way the liquidity provision game should be played.

It is standard in market microstructure models of price determination with private information to assume that the equilibrium liquidity supplied in an electronic limit order book ("LOB") is characterized by a zero-profit condition. This condition, that prices and quantities in the LOB are characterized by an "upper-tail expectation" is described in Glosten (1994).\(^6\) The

\(^6\) and presaged by Rock (1990).
argument presented there is that this is the limit, as the number of players
gets large, of the equilibria of games between liquidity suppliers. This is
formalized for some environments by Biais et al. (2000); Back and Baruch
(2013); and Biais et al. (2013). In particular the latter two make clear that
the standard pure strategy equilibrium, in which liquidity suppliers provide
”supply schedules” may not exist. More importantly, for a common mi-
crostructure model in which there are ”noise traders” and informed traders
arriving randomly to the market, Dennert (1993) shows that one equilibrium
in mixed strategies does not converge to the ”upper tail expectation” char-
acterization. Quite the contrary. In his setting, as the number of liquidity
suppliers gets large, all the submitted offer quotes pile up at the upper end of
the allowable set of prices. Furthermore, and we will show this below, there
are environments in the informed/noise setting (a setting not considered by
Biais et al. (2000) for which pure strategy supply schedule equilibria, if they
exist, are not characterized by a first order condition. Our second contribu-
tion is the result that, quite generally, there does exist a sequence of mixed
strategy equilibria that converges to the competitive LOB equilibrium in a
setting with noise and informed traders.

What we show is that, for the class of market microstructure models with
informed and noise trade, it is easy to find a mixed strategy equilibrium.
If there are \(n\) competing liquidity suppliers an equilibrium involves each of
them picking a price at random and quoting \(\frac{1}{n-1}\) of the maximum noise trade
quantity (for ease of exposition, normalize this largest trade to one). There
is an important feature of this equilibrium. Notice that the number of shares
offered at a price \(p\) or lower is a binomial random variable (the number of
liquidity suppliers that happened to choose a price lower than $p$) divided by $n - 1$. This is approximately the sample mean of $n$ Bernouli random variables. As $n$ gets large, this converges to a constant function of $p$.\footnote{The proof is a little more difficult, since as $n$ changes, the mixing distribution changes so it is not a straight forward application of the law of large numbers} This limit is the number of shares offered at $p$ or lower in the competitive equilibrium.

As noted above, the microstructure model we analyze features noise traders—traders who choose their quantities without reference to the terms of trade for those quantities. While such an assumption seems unrealistic for very large trades at potentially extreme prices, it seems reasonable for the situation we are looking at. Quantities will be bounded, and all traders may use marketable limit orders. In such a case, noise trade seems reasonable. After all, the flickering nature of quotes on a millisecond basis means the investors who submit orders do not know what the terms of trade will be by the time the order gets to be executed. They may base their choice on the expected terms of trade—the terms of trade that typically hold—but that just determines the distribution of the trade size. We take that distribution as given.

We think of the actual operation of the market as a repeated game in which liquidity suppliers provide quotes that last momentarily and may or may not get hit by an arriving marketable order. If they are not hit, the game repeats with the environment unchanged but with a new set of random quotes. If they are hit by a marketable order, it is most convenient to assume that either a bit of news comes out and the quotes are hit by informed news traders with a marketable limit order reacting before anyone else, or there
is no bit of news revealing the order to be uninformed. In this way, we can look at a single round of the game, the stage game, to characterize the equilibrium.

The paper is set out as follows. Section 1 lays out the dynamic game and argues that it is reasonable to analyze the game as a sequence of one period stage games. An example is also presented illustrating the robustness of the equilibrium notion. We also obtain a rather novel result from the example that as the probability of informed trade goes to zero, the stage game expected profit goes to zero but the expected profit in the dynamic game is positive. Section 2 analyzes the game presented in Dennert (1993) and shows, in contrast, that there is a sequence of equilibria converging to the competitive limit order book. Section 3 shows the difficulty of obtaining a pure strategy equilibrium. Section 4 proves, for the general news/noise trader model the existence of a mixed strategy equilibrium that converges to the competitive limit order book. An example illustrates the theory. The equilibrium described in Section 4 is a ”no rents” equilibrium. Section 5 illustrates positive rents equilibria for a special case of the model in Section 4. All of the analysis above assumes that prices are continuous, so Section 6 provides hints as to what will happen when there are discrete prices.

1. Continuous Time Market and Fleeting Orders

We consider a dynamic limit-order market for a single asset with the usual price-time priority, but zero minimum price variation. There are $n$ risk neutral and uninformed liquidity providers, hereafter limit-order traders. Orders arrive at the market instantaneously and trade is reported instantaneously.
However, quotes are disseminated only when trade takes place or after $\Delta t$ units of time, whichever comes first.

Market orders, originating from noise traders, arrive at the market sporadically. More specifically, we assume the cumulative order flow of the noise traders is a compound Poisson processes, $z_t$, with symmetric jumps

$$z_t = \sum_{m=1}^{\tilde{N}_t} \tilde{\epsilon}_m \tilde{q}_m$$

where $\tilde{N}_t$ is a standard Poisson process with intensity $\beta$, $\tilde{\epsilon}_m \in \{-1, 1\}$ indicates whether the $m$-th order is a buy or sell order (we assume equal probabilities), and $\tilde{q}_m$ is the order size of the $m$-th order, drawn from a common distribution $F_Q$. We use the notation $\tilde{q}$ for a generic order size. Accordingly, the symbol $q$ denotes a positive number.

Information relevant to the value of the asset is released to the public at a random time $\tilde{\tau} \sim \exp(\theta)$. Following the announcement, the value of the asset, $\tilde{v}$, is realized. We denote by $F_V$ and $\bar{v}$ the distribution function and the least upper bound of $\tilde{v}$, respectively. The random variable $\tilde{v}$ may be either a continuous or a discrete random variable. If $\tilde{v}$ is unbounded, then $\bar{v} = \infty$. Before the limit-order traders can refresh their quotes, fast news traders pick off all stale bids and offers. I.e.; news traders submit buy or sell orders with limit prices $\tilde{v}$. To wrap up the trading game, we assume the market closes after news traders trade, and the asset liquidates.

Having specified the exogenous strategies of the noise and news traders, we turn our attention to the limit-order traders. The $i$-th trader has, at time $t$, a collection of outstanding orders that contains all buy and sell limit orders that were submitted in the past and were not executed or canceled prior to
time $t$. We denote this collection by $b^i_t$.

At any time a trader can send messages to the exchange. A message contains instructions to add new limit orders and/or cancel existing ones. After the exchange executes the instructions, the updated collection of limit orders is $b^{i+}_t$.

We sort the orders in $b^i_t$ according to their price, and summarize the result in a non-decreasing price schedule $P^i_t : R \setminus \{0\} \to R^+$ with the interpretation that $P^i_t(q)$ and $P^i_t(-q)$ are the marginal prices of the $q$-th unit that the $i$-th trader offers and bids, respectively.

Analogously, we can express the sorted collection of orders in a non-decreasing function $S^i_t : R^+ \to R \setminus \{0\}$. In its positive range, $S^i_t(p)$ is the number of shares the $i$-th trader offers at prices smaller or equal to $p$, and in its negative range, $S^i_t(p)$ is the number of shares the $i$-th trader bids at prices greater or equal than $p$. We let $S_{-i}^t = \sum_{j \neq i} S^t_j$. Informationally, $P^i_t$ and $S^i_t$ are equivalent, and we use them interchangeably.\(^8\)

We look for a stationary equilibrium with fleeting orders in which time precedence plays no role. In this equilibrium, traders cancel their limit orders immediately after quotes are disseminated by the exchange and replace them with new ones. The fresh orders are viewed as random by other market participants and to emphasize this uncertainty, we write $\tilde{S}^t_{-i}$. The equilibrium

---

\(^8\)Formally, we take supply function $S$ to be right continuous left limit (resp. left continuous right limit) in its positive (resp. negative) range. We take price schedule $P$ to be left continuous right limit (resp. right continuous left limit) in its positive (resp. negative) domain. In its positive domain $P(q) = \inf\{p : S(p) \geq q\}$, and in its negative domain $P(-q) = \sup\{p : S(p) \leq -q\}$. We can reconstruct $S$ from $P$ in a similar manner.
is stationary in the sense that the distributions of $\tilde{S}_{t-i}$, is independent of time.

To keep our exposition succinct, in the following we compute the $i$-th trader’s best response using an artificial payoff function that is not smaller than the one prescribed by the game. More specifically, we assume that ties are always broken in favor of the $i$-th trader. Since ties occur in our model only when noise traders trade, this is an advantage. Our approach is valid if ties occur with zero probability. In that case, the optimal strategy for the artificial payoff function is also the best response. We emphasize that we only alter the payoff function, but we do not restrict the $i$-th trader’s strategy.

Suppose a market buy order of size $\tilde{q}$ arrives at time $t < \tilde{\tau}$. The order walks up the book picking off limit orders until the order is filled up. If the $i$-th trader offers his/her $q$-th unit at $p$, then the trader sells this unit (i.e. sells at least $q$) if and only if $S_{t-i}^b(p-)$, the number of shares other traders offer at prices strictly smaller than $p$, plus $q$ is still smaller than the size of the incoming order $\tilde{q}$; i.e. $S_{t-i}^b(p-) + q \leq \tilde{q}$. Analogously, suppose the incoming order is a sell order, and the $i$-th trader bids his/her $q$-th unit at $p$. The trader buys this unit (i.e. buys at least $q$) if and only if $-S_{t-i}^s(p+)$, the number of shares other traders bid at prices strictly better than $p$, plus $q$ is still smaller than the size of the incoming sell order; i.e. $-S_{t-i}^s(p-)+q \leq \tilde{q}$.

Thus, the payoff of the $i$-th trader at time $t < \tilde{\tau}$ is

$$\pi_0(P_t^i, S_{t-i}^d(p), \tilde{q}, \tilde{v}) =$$

$$\int_0^\infty I_{\{q+S_{t-i}^s(p(q)) \leq \tilde{q}\}}(P_t^i(q) - \tilde{v})dq + \int_0^\infty I_{\{q-S_{t-i}^b(p(q)) \leq \tilde{q}\}}(\tilde{v} - P_t^i(-q))dq$$

Market orders in our model are only submitted by noise traders. Therefore, the $i$-th trader’s payoff at time $t < \tilde{\tau}$ is $\pi_0(P_t^i, S_{t-i}^d(p), d\tilde{z}_t, \tilde{v})$, which is typically
zero except at those times when a noise order arrives; i.e. when \(d\tilde{z}_t \neq 0\). We integrate out the random variables \(\tilde{v}\), the size of the jumps, and \(\tilde{S}_{t-i}\), and get the expected payoff when trading with noise traders:

\[
\bar{\pi}_0(P_t^i) \equiv E \left[ \pi_0(P_t^i, \tilde{S}_{t-i}(p), d\tilde{z}_t, \tilde{v}) \middle| \Delta N_t > 0 \right]
\]

Note that because \(\tilde{S}_{t-i}\) has a stationary distribution, the functional \(\bar{\pi}_0\) is independent of time.

News traders pick off all stale limit orders; i.e. news traders submit limit orders with unbounded size. The payoff at time \(\tilde{\tau}\) is:

\[
\pi_1(P_{\tilde{\tau}}^i, \tilde{v}) = \int_0^\infty (P_{\tilde{\tau}}^i(q) - \tilde{v})I_{\{\tilde{v} \geq P_{\tilde{\tau}}^i(q)\}} dq + \int_0^\infty (\tilde{v} - P_{\tilde{\tau}}^i(-q))I_{\{\tilde{v} \leq P_{\tilde{\tau}}^i(-q)\}} dq
\]

Note that again the functional \(\pi_1\) is time independent. We integrate \(\tilde{v}\) out to get the expected payoff at time \(\tau\):

\[
\bar{\pi}_1(P_{\tilde{\tau}}^i) = E[\pi_1(P_{\tilde{\tau}}^i, \tilde{v})|\tau]
\]

The expected payoff of the \(i\)-th trader at time \(s\) is

\[
\Pi_i(s) = E \int_s^\tau \pi_0(P_t^i, \tilde{S}_{t-i}(p), d\tilde{z}_t, \tilde{v})dN_t + \pi_1(P_t^i, \tilde{v})
\]

\[
= E \int_s^\tau \bar{\pi}_0(P_t^i)\beta dt + \bar{\pi}_1(P_t^i)
\]

\[
= \int_s^\infty e^{-\theta(t-s)} \left[ \bar{\pi}_0(P_t^i)\beta + \theta \bar{\pi}_1(P_t^i) \right] dt
\]

\[
= \int_s^\infty (\beta + \theta)e^{-\theta(t-s)} \left[ (1 - \mu)\bar{\pi}_0(P_t^i) + \mu \bar{\pi}_1(P_t^i) \right] dt
\]

where for the second equality, we integrate out all random variables except \(\tilde{\tau}\), which we integrate out in the third equality. The forth equality is a change of variable, where \(\mu \equiv \theta/\theta + \beta\).
The profit flow is
\[ \tilde{\pi}(P_i) = [(1 - \mu)\tilde{\pi}_0(P_i) + \mu\tilde{\pi}_1(P_i)] \] (2)
and it is time independent.

To sum up, if \( \tilde{S}_{-i} \) is stationary, and the allocation rules favor the \( i \)-th trader, then the expected payoff flow is history independent. To find a subgame perfect equilibrium we analyze the following stage game.

**The Stage Game:** In the stage game the book is initially empty. Next, the limit-order traders simultaneously submit price schedules. With probability \( \mu \) a noise trader trades, the size of the order is \( \tilde{q} \), and with probability \( (1 - \mu) \) news traders trade. Allocations are determined by the price priority rule, and ties are broken using an unspecified random mechanism.

Given that the minimum price variation is zero, a simple infinitesimal undercutting argument implies that every trader achieves the maximum of (2). Indeed, if a trader knows that there is a positive probability that a tie will occur at \( p' \), then the trader offers/bids the same number of shares, \( \Delta S_i(p') \), at a marginally better price. Because the new price is only marginally different than \( p' \), the profit remains the same. I.e., each trader can attain the maximum of (2). The infinitesimal undercutting reasoning also implies that in any equilibrium of the stage game ties occur with zero probability. Thus, the equilibrium strategy maximizes (2).

Let \( \tilde{P}_i \) be a symmetric mixed-strategy equilibrium of the stage game. Consider the dynamic game. Assume that each of the \( j \neq i \) limit-order
traders uses a stationary fleeting order strategy; i.e. each time the exchange disseminates quotes, the traders replace their quotes with new ones drawn from the same mixing distribution of the stage game equilibrium. A standard pointwise maximization (for each $t$ maximize the integrand) implies that each $P_i$ in the support of $\tilde{P}_i$ belongs to the argmax set of (1). Moreover, it is also optimal for the $i$-th trader to cancel outstanding quotes each time the exchange disseminates quotes, and replace them with new random quotes drawn from the mixing distribution of the stage game equilibrium. Because, in the stage game, ties occur with probability zero, the argmax of (1) is also the set of best responses. We conclude that the mixed strategy equilibrium of the stage game is a Nash equilibrium with fleeting orders of the dynamic game. The equilibrium is sub game perfect because history does not play any role in this equilibrium. Finally, if we denote by $\pi^*$ the value of (2) when using the equilibrium strategies, then the equilibrium expected expected payoff in the dynamic game is

$$\Pi^* = \int_s^\infty (\beta + \theta)e^{-\theta(t-s)}\pi^* \, dt = \frac{\pi^*}{\mu}$$

**Example:** Let $n = 2$, $\tilde{v}$ be either 0 or 1 with equal probabilities, $\tilde{q}$ be either 1 lot, with probability $3/4$, or two lots, with probability $1/4$.

We start with the offer side of the book. We postulate that an equilibrium in mixed strategies exists in which each of the limit-order traders offers one lot at a random price, with support $(ask, 1)$, and a second lot either at the same price (with probability $l$) or at one (with probability $1 - l$). Thus, the unknowns are the constants $ask$, $l$, $\pi^*$, and the mixing distribution function
$M_a$. For $p \in (ask, 1)$, we have

$$\frac{\pi^*}{2} = \mu \cdot (p - 1) + \frac{1 - \mu}{2} \cdot (p - 0.5)(1 - M_a(p) + M_a(p)0.25(1 - l)) \quad (3)$$

$$0 \leq \frac{\mu}{2} \cdot (p - 1) + \frac{1 - \mu}{2} \cdot (p - 0.5)(1 - M_a(p))0.25, \quad \text{with equality if } l > 0 \quad (4)$$

Examining the right side of (3): with probability $\mu$, a news event occurs and with probability 0.5 the news trader buys in which case the limit order at $p$ will lose $(1 - p)$. With probability $\frac{(1 - \mu)}{2}$ a noise trader buyer arrives and the profit will be $(p - 0.5)$. A limit order at $p$ will transact with a noise trader if either the other limit order is placed higher than $p$, which occurs with probability $(1 - M_a(p))$; or the other limit order is at $p$ or lower, that limit order only offers one lot, and the noise trader buys two lots all of which happens with probability $M_a(p)(1 - l)0.25$.

The first equation states that the expected payoff associated with offering the first lot is $\pi^*/2$, independent of the price. Thus, the limit-order trader is willing to randomize the limit order price. The right hand side of (4) is the expected profit when a second lot is offered (i.e. if $l > 0$). The expected profit must be zero if the second unit is offered at a price smaller than one, since the right hand side of (4) is zero at $p = 1$.

To compute the equilibrium, we start with the guess $l = 0$. Thus, equation (3) is reduced to

$$\pi^* = \mu \cdot (p - 1) + (1 - \mu)2(p - 0.5)(1 - M_a(p) + M_a(p)0.25)$$

At $p = 1$, $M_a(p) = 1$, and we get $\pi^* = \frac{1 - \mu}{2}0.5 \cdot 0.25$. We plug $\pi^*$ back into (3) and solve for $M_a(p)$. We then find $ask$ by solving $M_a(ask) = 0$. We verify
that as long as $\mu \geq 1/9$, (4) holds. However, when $\mu < 1/9$, it is “profitable” to offer a second lot at prices smaller than one; i.e. the right hand side of (4) is strictly positive.

To find the equilibrium when $\mu < 1/9$, we guess $l > 0$, and use (4) to solve for $M_a(p)$. We then find the lower support of $M_a$ by solving $M_a(\text{ask}) = 0$. At $p = \text{ask}$, (3) gives us

$$\pi^* = \mu(\text{ask} - 1) + (1 - \mu)(\text{ask} - 0.5)$$

We then solve (3) for $l$ and verify that $l$ is a constant. We note that our initial guess $l > 0$ holds only if $\mu < 1/9$.

When pasting the two solutions, we get $M_a$, $\text{ask}$, $l$ and $\pi^*$ that are continuous in $\mu$. They are given by

$$M_a(p) = \frac{p - \text{ask}}{(1 - \text{ask})(2p - 1)}; \text{ ask} \leq p < 1$$

where the lower support is

$$\text{ask} = \begin{cases} \frac{1 + 7\mu}{2 + 6\mu} & \mu \leq 1/9 \\ \frac{5 + 3\mu}{8} & \mu > 1/9 \end{cases}$$

The probability that a second lot is offered at the same price as the first one is

$$l = \begin{cases} \frac{1 - 9\mu}{1 + 3\mu} & \mu \leq 1/9 \\ 0 & \mu > 1/9 \end{cases}$$

and the expected profit is

$$\pi^* = \begin{cases} \frac{3\mu(1 - \mu)}{6\mu + 2} & \mu \leq 1/9 \\ \frac{1 - \mu}{8} & \mu > 1/9 \end{cases}$$
Symmetrically, each limit-order trader bids a single lot at a random price $p$ distributed $M_b(p) = M_o(1 - p)$ and bids the second lot at the same price with probability $l$ and at zero with probability $1 - l$.

**Lemma 1.** *An equilibrium with the above mentioned properties exists.*

The proof of the lemma is in the appendix. Even though both traders offer shares at 1 (with positive probability, and when $\mu \geq 1/9$ with probability one), the tie breaking rule is not required because the noise trader’s order is always executed at prices smaller than one. In addition, the mixing distribution is continuous, and therefore ties occur with probability zero.

We note that as $\mu$ goes to zero, the expected profit in the stage game goes to zero. This means that the profit flow in the dynamic game goes to zero. However, as $\mu$ goes to zero, the number of transactions before a news event gets arbitrarily large, and the aggregate profit in the dynamic game is strictly positive:

$$\lim_{\mu \to 0} \Pi^* = \lim_{\mu \to 0} \frac{\pi^*}{\mu} = \lim_{\mu \to 0} \frac{3\mu(1 - \mu)}{\mu(6\mu + 2)} = \frac{3}{2}$$

This concludes the example.

What if the equilibrium of the stage game is in pure strategies? If $S_{-i}$ is continuous so ties never occur, we can implement the equilibrium in the dynamic game in exactly the same way. However, limit orders are now forecastable. So there is no need to cancel orders just to replace them with identical orders. Thus, if there is an equilibrium in pure strategies, then traders send messages to the exchange only after trade takes place. That is, traders only replenish executed orders. In Section 5 we provide an example in which we can find an equilibrium in pure strategies.
2. Convergence

We saw that if the stage game has a mixed strategy equilibrium, then there is an equilibrium with fleeting orders in which quotes are random and short lived. This calls into question the assertion that the competitive equilibrium is a viable description of quotes with a large number of limit-order traders. We will show that as their number increases, the total equilibrium random supply function converges, in mean square, to the competitive supply function.

Because the stage game equilibrium is played repeatedly, we don’t lose generality if we focus on the stage game. In addition, thanks to the symmetry of the game, we can examine the offer side of the book separately from the bid side. Therefore, for the remaining of the paper, we analyze only the offer side.

We define the functions $v(p) = E[\tilde{v}|\tilde{v} > p]$, and

$$G(p) = \frac{\mu(v(p) - p)(1 - F_V(p))}{0.5(1 - \mu)(p - E\tilde{v})}$$

(5)

We need the following technical result.

**Lemma 2.** The equation $G(p) = 1$ has a solution $p_c > E\tilde{v}$. In the interval $(p_c, \bar{v})$, the function $G(p)$ is continuous, strictly decreasing, and

$$G(p_c) = 1$$

$$\lim_{p \uparrow \bar{v}} G(p) = 0$$

Consequently, in the interval $(0, 1)$, the inverse function $G^{-1}$ is also strictly decreasing.
Theorem 1. Assume $\tilde{q} \equiv 1$. The stage game has a symmetric equilibrium in mixed strategies in which each limit-order trader offers $1/(n-1)$ at a random price with distribution function

$$M_n(p) = (1 - G(p))^{1/(n-1)}, \ p \in (p_c, \tilde{v}) \quad (6)$$

In this equilibrium, the limit-order traders earn zero profit.

Proof. From Lemma 2 it follows that the mixing distribution is well defined. Suppose other traders follow the strategy stated in the theorem, and consider the problem of the $i$-th trader. Let $\tilde{p}_j$ be $j$-th trader’s random offering price. Clearly to offer shares at prices strictly smaller than $p_c$ is suboptimal because the trader can offer the shares at $p_c$ and still be ahead of all other offers in the book.

The expected profit associated with offering the $q$-th unit, for $q \in (0, \leq 1/(n-1)]$, at $p \geq p_c$ is

$$\mu E[(p - \tilde{v})I_{\{p<\tilde{v}\}}] + (1 - \mu)E[I_{\{q + \tilde{S}_{-i}(p) < 1\}}](p - E\tilde{v})$$

$$= \mu(p - v(p))(1 - F_V(p)) + \frac{(1 - \mu)}{2}(1 - \text{Prob}(p > \max_{j \neq i} \tilde{p}_j))(p - E\tilde{v})$$

$$= \mu(p - v(p))(1 - F_V(p)) + \frac{(1 - \mu)}{2}(1 - M_n(p)^{n-1})(p - E\tilde{v})$$

$$= 0 \quad (7)$$

where for the last equality we use the definition of the mixing distribution. If the trader were to offer strictly more than $1/(n-1)$ units, then to trade with a noise trader, the offering price for the “higher units” has to be better than at least two other random offering prices. Because the probability of undercutting two random prices is strictly smaller than the probability of undercutting one, it follows that the payoff of higher units has to be negative,
We conclude that the \( i \)-th trader is indifferent at which price, in the support of \( M_n(p) \), to offer the first \( 1/(n-1) \) units. In particular, it is optimal to offer them as a block at a single random price.

The equilibrium in Theorem 1, however, is not unique. Dennert (1993) looks at a special case of Theorem 1 in which the liquidation value is either -1 or 1, and reports that in equilibrium each limit-order trader offers one share at a random price.\(^9\) In this equilibrium, to gain from trade (i.e. trade with a noise trader) a limit-order trader has to post the best offer in the book. As a result, the chances of trading with a noise trader decrease with \( n \), and the mixing distribution shifts to the right as \( n \) increases. In particular, the sequence of equilibria does not converge to the competitive equilibrium.

In contrast with the result in Dennert (1993), in the equilibrium in Theorem 1, to gain from trade, a limit-order trader has to undercut only one of the other limit-order traders. To see that the equilibrium in Theorem 1 converges to the competitive equilibrium, let \( \tilde{S}_n(p) \) denote the total number of shares offered in equilibrium at prices smaller or equal to \( p \), and let \( S_c(p) \) denote the supply function in the competitive equilibrium.

**Theorem 2.** As \( n \) goes to infinity, the equilibrium in Theorem 1 converges,

---

\(^9\)Dennert (1993) models a dealer market, where the active trader shops for the best available price. In limit order markets, offers are already ranked from best to worse by the exchange. Thus, the equilibrium in Dennert can be implemented in a limit order market. More generally, any equilibrium in dealers market in which dealers don’t offer quantity discounts can be implemented in a limit order market.
in mean square, to the competitive equilibrium; i.e.

\[ E(\tilde{S}_n(p) - S_c(p))^2 \rightarrow 0 \]

**Proof.** When \( \bar{q} \equiv 1 \), the competitive supply function is

\[ S_c(p) = I_{\{p \geq p_c\}} \]

In the mixed strategy equilibrium, the total supply of shares is

\[ \tilde{S}_n(p) = \frac{1}{n-1} \sum_{i=1}^{n} I_{\{\bar{p}_i \leq p\}} \]  \hspace{1cm} (8)

Because \( (n-1)\tilde{S}_n(p) \sim B(n, M_n(p)) \), we have

\[ E\tilde{S}_n(p) = \frac{n}{n-1} M_n(p) \xrightarrow{n \rightarrow \infty} I_{\{p \geq p_c\}} = S_c(p) \]

and

\[ Var(\tilde{S}_n(p)) = \frac{n}{(n-1)^2} M_n(p)(1 - M_n(p)) < \frac{n}{4(n-1)^2} \xrightarrow{n \rightarrow \infty} 0 \]

Thus,

\[ E(\tilde{S}_n(p) - S_c(p))^2 = Var(\tilde{S}_n(p)) + \left( E\tilde{S}_n(p) - S_c(p) \right)^2 \xrightarrow{n \rightarrow \infty} 0 \]  \hspace{1cm} \Box
Figure 1: Convergence to the competitive outcome. In both figures, the \( q \) axis is the competitive equilibrium (i.e. \( S_c(0.8) = 1 \)). In the left figure, the wider band is the 95% confidence band when \( n = 10 \); E.g. with probability 0.95 the asking price for the first, second and third \( 1/(n-1) \) units are virtually the competitive price 0.8. On the other hand, the last fraction of a noise order of size one is executed, with 0.95 probability, anywhere between 0.801 and 0.942. The narrow band, in the same figure, is the 95% confidence band when \( n = 1,000 \). The right figure shows the mixing distributions when \( n = 5 \) (the upper curve), \( n = 10 \) (the middle curve), and \( n = 1,000 \) (the lowest curve). In both figures, the liquidation value is either zero or one with equal probabilities, the noise order size is deterministic and equals to one, and \( \mu = 0.6 \).

The convergence of the equilibrium in Theorem 1 cannot be uniform because the competitive supply function is discontinuous at the ask price. The convergence is illustrated in Figure 1. Even with a huge number of limit-order traders \( (n = 1,000) \), the depth at the ask price suffices for only about 80% of the noise order. The remaining 20% of the order executes at a dramatically higher price than the competitive. In the following, we consider continuous economies in which the competitive supply function is smooth and the convergence is uniform (Corollary 1 in Section 4 and in Figure 2.)
3. The Continuous Economy

It is common in the literature to assume that the random variables can take on any real value. This abstraction sometimes make the analysis tractable. We therefore further assume that the distributions of \( \tilde{v} \) and \( \tilde{q} \) are continuous, and the support of \( \tilde{q} \) is \((0, 1)\).\(^{10}\)

The competitive equilibrium is given implicitly by

\[
F_Q(S_c(p)) = 1 - G(p), \quad p \in (p_c, \tilde{v}) \tag{9}
\]

where \( G(p) \) is defined in (5), \( p_c = G^{-1}(1) \) is the competitive ask price, and \( \tilde{v} = G^{-1}(0) \).

Back and Baruch (2013) consider a more general setting than ours. That paper reports that, for some models, when the adverse selection is sufficiently high, there is a pure strategy equilibrium with continuous supply functions that converges to the competitive equilibrium. That result is not, however, general as this example shows. In this section, we follow the same approach, and conjecture that there is an equilibrium in which \( S_{-i} \) is continuous.

We define the profitability function

\[
u(p, q) = \mu(p - v(p))(1 - F_V(p)) + \frac{1 - \mu}{2} (p - E\tilde{v})(1 - F_Q(q)) \tag{10}\]

If \( S_{-i} \) is continuous, then the objective of the \( i \)th trader is to choose a non-decreasing price schedule \( P \) that maximizes

\[
\int_0^\infty u(P(q), q + S_{-i}(P(q))) \, dq \tag{11}
\]

\(^{10}\)This assumption is a normalization that allows us to seamlessly move from the equilibrium mixing distribution to the deterministic competitive supply function.
The objective (11) can be maximized pointwise; i.e., for each \( q \geq 0 \), maximize the function \( p \to u(p, q + S_{-i}(p)) \). The f.o.c. is

\[
\frac{\partial}{\partial p} u(p, q + S_{-i}(p)) \bigg|_{p=p^* (q)} = 0
\]  

(12)

We can now use the symmetric equilibrium condition, namely, \( S_{-i} = (n - 1)S^* \) to derive an o.d.e. that the total supply function function, \( S_n \), satisfies at prices greater than the ask price:

\[
u_p(p, S_n(p)) + \frac{(n - 1)}{n} S'_n(p)u_q(p, S_n(p)) = 0
\]  

(13)

The solution of the o.d.e. is strictly increasing (because \( u_p > 0 \) and \( u_q < 0 \)), and hence the individual supply function \( S^*(p) = S_n(p)/n \) is feasible. Moreover, the sequence of solutions converges to the competitive equilibrium supply function as \( n \) goes to infinity.\(^{11}\)

The pointwise optimization we carried above is valid if \( p \to u(p, q + S_{-i}(p)) \) is quasi-concave. However, if we proceed with the assumption that \( \tilde{q} \) is a standard uniform random variable, we can readily see that the pointwise objective is quasi-convex! Indeed,

\[
\frac{\partial^2}{\partial p \partial q} u(p, q + S_{-i}(p)) = \frac{\mu - 1}{2} < 0
\]

which implies that for every \( p \), the function \( q \to \frac{\partial}{\partial p} u(p, q + S_{-i}(p)) \) is strictly decreasing. Thus, for every \( p \) there is a value, call it \( S^*(p) \), such that for all

\(^{11}\)The competitive supply function satisfies \( u(p, S_c(p)) = 0 \), and therefore, expressed in terms of a differential equation, \( S_c \) is the solution of the o.d.e. \( u_p(p, S_c(p)) + S'_c(p)u_q(p, S_c(p)) = 0 \).
\( q > 0, \)
\[
\frac{\partial}{\partial p} u(p, q + S_\sim(p)) \begin{cases} 
< 0, & \text{if } q > S^*(p) \\
> 0, & \text{if } q < S^*(p) 
\end{cases}
\]

A priori, \( S^*(p) \) may be zero or infinity, however from (12), it follows that \( S^*(p) \) must be the inverse of \( P^* \). Thus,
\[
\frac{\partial}{\partial p} u(p, q + S_\sim(p)) \begin{cases} 
< 0, & \text{if } P^*(q) > p \\
> 0, & \text{if } P^*(q) < p 
\end{cases}
\]

We conclude that the objective function of the pointwise maximization, \( p \to u(p, q + S_\sim(p)) \), is first decreasing and then increasing and hence it is quasi-convex. Thus, when \( \tilde{q} \) is uniformly distributed, \( S^* \) is the not an an equilibrium individual supply function. We will see in Section 5 a different distributional assumption for which \( S^* \) is an equilibrium.

4. Convergence and the Continuous Economy

In this section we show the existence of a sequence of Nash equilibria with a random aggregate supply function that converges to the the competitive supply function in the continuous economy given by (9). To construct the equilibria, we discretize the order size that noise traders use. That is, in the \( n \)th economy there are \( n \) limit-order traders, and the order size of the noise trader is a lattice random variable, \( \tilde{q}_n \) with support
\[
\{1/(n - 1), 2/(n - 1), \ldots, (n - 1)/(n - 1)\}
\]
The demand \( \tilde{q}_n \) is related to the demand in the continuous economy via
\[
\tilde{q}_n \equiv \left[ \frac{\tilde{q}(n - 1)}{n - 1} \right],
\]

23
where \([x]\) is the smallest integer larger than \(x\). In particular,

\[
\text{Prob}(\tilde{q}_n \leq j/(n-1)) = F_Q(j/(n-1)) 
\]

(14)

Note that even though we use a lattice model, the feasible strategies are general and we do not restrict the limit-order traders to discrete orders. However, in the following, we prove the existence of a symmetric equilibrium in which each limit-order trader offers a block of \(1/(n-1)\) units at a single random price, \(\tilde{p}_i\). That is,

\[
\tilde{S}_n(p) = \sum_{i=1}^{n} \frac{1}{n-1} I_{\{\tilde{p}_i \leq p\}}
\]

and thanks to the symmetry, \((n-1)\tilde{S}_n(p) \sim \text{Bin}(n, M_n(p))\), where \(M_n(p)\) is the common mixing distribution. We have the following

**Lemma 3.** Assume \((n-1)\tilde{S}_{-i}(p) \sim \text{Bin}(n, M_n(p))\), and let

\[
K(p) = \text{Prob}(\tilde{q}_n > \tilde{S}_{-i}(p))
\]

(15)

Then

\[
K(p) = 1 - EF_Q(\tilde{S}_{-i}(p))
\]

**Proof.** The definition of the lattice variable \(\tilde{q}_n\) implies that for any point in its support, say \(j/(n-1)\) for some integer \(j \leq (n-1)\), we have \(\tilde{q}_n \leq j/(n-1)\) if and only if \(\tilde{q} \leq j/(n-1)\). Because \((n-1)\tilde{S}_{-i}(p)\) is a binomial random variable, and hence an integer random variable, \(\tilde{S}_{-i}(p)\) takes values only in the support of the \(\tilde{q}_n\). Hence,

\[
\text{Prob}(\tilde{q}_n \leq \tilde{S}_{-i}(p)) = E[\text{E}[I_{\{\tilde{q}_n \leq \tilde{S}_{-i}(p)\}} | \tilde{S}_{-i}]] = E[\text{E}[I_{\{\tilde{q} \leq \tilde{S}_{-i}(p)\}} | \tilde{S}_{-i}]] = E[\text{F}_Q(\tilde{S}_{-i}(p))]
\]

\[\square\]
**Theorem 3.** In the lattice model there exists a symmetric Nash equilibrium in which each limit-order trader offers $1/(n - 1)$ at a random price. The limit-order traders break even, and the distribution function of the random price is given implicitly by

$$M_n(p) = h(G(p)), \ p \in (p_c, \bar{v})$$

where $G$ is defined in (5), and $h(\cdot)$ is the inverse of the function

$$k(h) = 1 - E[F_Q(\tilde{j}/(n - 1))], \ \tilde{j} \sim Bin(n - 1, h)$$

In particular, for every $p \in (p_c, \bar{v})$, we have $K(p) = G(p)$

The proof is in the Appendix.

**Lemma 4.** In the lattice equilibrium, we have

$$E[F_Q(\tilde{S}_{-i})] = F_Q(S_c(p))$$

where $S_c(p)$ is the competitive supply function in the continuous economy.

**Proof.** Outside the support of $M_n(p)$ the identity is obvious. For $p \in (p_c, \bar{v})$, we have

$$1 - F_Q(S_c(p)) = G(p) = K(p) = 1 - E[F_Q(\tilde{S}_{-i})]$$

where the first equality is (9), the second equality is from Theorem 3, and the last equality is (15).

**Corollary 1.** As we increase $n$, the equilibrium mixing distribution, $M_n$, converges uniformly to the competitive supply function, $S_c(p)$. 

25
The proof of the corollary is involved because the transformations between the mixing distribution, the expected random supply, and the competitive supply function are all implicit. The proof is deferred to the appendix. In the uniform example, however, the corollary is immediate. From Lemma 4, we have

\[ S_c(p) = E[ \tilde{S}_n] = M_n(p) \]

In this example the mixing distribution is exactly the competitive supply function, and in particular the mixing distribution is independent of \( n \). The strategy itself depends on \( n \), because the number of units offered is \( 1/(n-1) \).

Finally, we note that in the uniform example, \( E\tilde{S}_n(p) \geq S_c(p) \).\footnote{In fact, Lemma 4 implies that whenever \( F_Q \) is concave (i.e. its density is decreasing), \( E\tilde{S}_n(p) \geq S_c(p) \).}

Endowed with Corollary 1, the convergence result is immediate.

**Theorem 4.** As \( n \) goes to infinity, the equilibrium in lattice economy converges, in mean square, to the competitive equilibrium.

**Proof.** Because \( (n-1)\tilde{S}_n(p) \sim Bin(n, M_n(p)) \), we have

\[ \lim_{n \to \infty} E[\tilde{S}_n(p)] = \lim_{n \to \infty} \frac{n}{n-1} M_n(p) = S_c(p) \]

where the last equality is the Corollary. Additionally,

\[ Var(\tilde{S}_n(p)) = \frac{n}{(n-1)^2} M_n(p)(1 - M_n(p)) < \frac{n}{4(n-1)^2} \xrightarrow{n \to \infty} 0 \]

and therefore

\[ E[(\tilde{S}_n(p) - S_c(p))^2] = Var(\tilde{S}_n(p)) + \left( E[\tilde{S}_n(p)] - S_c(p) \right)^2 \xrightarrow{n \to \infty} 0 \]

\[ \square \]
4.1. Mixed Strategy Equilibrium Example

To illustrate our results, we calculate the mixing distribution when the distribution of the noise trade is the triangle distribution, \(1 - F_Q(q) = (1-q)^2\), and the distribution of \(\tilde{v}\) is arbitrary. We define \(K(p)\) to be the probability that an uninformed trader buys more than the random supply at \(p, \tilde{S}_{-1}(p)\).

The mixing condition, that expected profits are zero for all \(p\), requires that \(K(p) = G(p)\). The random uninformed quantity purchased is a random variable \(\tilde{q}_n\) which is of the form \(\frac{\tilde{K}}{n-1}\) where \(\tilde{K}\) is an integer random variable.

The probability that \(\tilde{K}\) is strictly greater than \(n\) is given by \((1 - \frac{n}{n-1})^2\).

On the other hand, \((n-1)\tilde{S}_{-1}(p)\) is a binomial random variable with mean \((n-1)M(p)\) and variance \((n-1)M(p)(1-M(p))\). Call this random variable \(\tilde{J}\). Then, we have for \(n > 2:\)

\[
K(p) = P\{\tilde{q}_n > \tilde{S}_{-1}(p)\} = E[P\{\tilde{q}_n > \tilde{S}_{-1}(p)|\tilde{S}_{-1}(p)\}] = E[P\{\tilde{K} > \tilde{J}|\tilde{J}\}]
\]

\[
= \frac{E[(n-1 - \tilde{J})^2]}{(n-1)^2}
= \frac{(n-1)^2 - 2(n-1)E[\tilde{J}] + E[\tilde{J}^2]}{(n-1)^2}
= \frac{(n-1)^2 - 2M(p)(n-1)^2 + (n-1)M_n(p)(1 - M_n(p)) + (n-1)^2M_n(p)^2}{(n-1)^2}
= (1 - M_n(p))^2 + \frac{M_n(p)(1 - M_n(p))}{(n - 1)}
\]

Setting \(K(p)\) equal to \(G(p)\) and solving the quadratic equation for \(M_n(p)\) using the smaller root, since the larger root exceeds one, yields the following, for \(n > 2:\)

\[
M_n(p) = 1 + \frac{1}{2(n-2)} - \sqrt{\frac{1}{4(n-2)^2} + \frac{(n-1)G(p)}{(n-2)}}
\]
and for \( n = 2 \), we get \( M_2(p) = 1 - G(p) \).

The total random supply function, \( \tilde{S}_n(p) \) is a binomial random variable with mean \( nM_n(p) \) divided by \( (n-1) \), and hence the expected supply at \( p \) is \( \frac{nM_n(p)}{(n-1)} \).

The appendix provides realizations of the supply function for \( \mu = .05 \) and \( n = 10, 20, 30 \) and \( \tilde{v} \) that is either zero or one. It is notable that the supply function is quite steep at lower prices. This is due to the fact that with the triangle distribution, small trades are very likely to be uninformed. The supply function flattens out at upper prices since large uninformed trades are quite unlikely.

5. Economic Rents

The lattice equilibrium in Theorem 3 is a workhorse model: without making any distributional assumptions about \( \tilde{v} \) and \( \tilde{q} \), the equilibrium converges to the competitive. However, in this equilibrium the strategic limit-order traders break even. This type of equilibrium is easy to work with because once we have verified that to offer \( 1/(n-1) \) units has zero expected profits, it follows immediately that one cannot gain by offering even more units. If we were to look for equilibrium with positive expected profit, then we have to carefully check whether it is optimal or not to offer additional units.

In this section, we present an example of equilibrium with positive expected profit, and the equilibrium converges to the competitive. Interestingly, the example we consider can also be dealt with using the technology

\[ \text{[\text{Footnote text]}\]}

\[ \text{[\text{Footnote text]}\]}

28
developed in Back and Baruch (2013).

5.1. Pure Strategy Positive Rents

For the analysis of economic rents, we simplify the distribution of $\tilde{v}$ and assume that it is equally likely to be zero or one. Following the development above, the supply function solution to the first order condition, $S_n(p)$ must satisfy:

$$0.5\mu + 0.5(1 - \mu)(1 - S_n(p))^2 - (1 - \mu) \left( \frac{n - 1}{n} \right) S'_n(p)(p - 0.5)(1 - S_n(p)) = 0$$

It can be easily verified that:

$$S_n(p) = 1 - \sqrt{\frac{\mu}{1 - \mu} \sqrt{(2p - 1)\frac{n}{n-1} - 1}}$$

is the solution.

This will hold for $p > A_n$, where $S_n(A_n) = 0$. It is easy to see that $(2A_n - 1)\frac{n}{n-1} = \frac{1}{\mu}$. This may or may not be an equilibrium. As with the example in section 3, it is important to check the second order conditions that for all $q > 0$ and at the solution $S_n(p)$:

$$\frac{\partial}{\partial p} u(p, q + S_{-i}(p)) \begin{cases} < 0, & \text{if } q < S_n(p)/n \\ > 0, & \text{if } q > S_n(p)/n \end{cases}$$

In the case at hand $\frac{\partial}{\partial p} u(p, q + S_{-i}(p))$ is given by:

$$0.5\mu + 0.5(1 - \mu)(1 - q - \frac{n - 1}{n} S_n(p))^2 - \frac{n - 1}{n} S'_n(p)(1 - \mu)(1 - q - \frac{n - 1}{n} S_n(p))(p - 0.5)$$

Note that this expression is quadratic and convex in $q$ and that $S_n(p)/n$ is one of its zeros. We need to check that $S_n(p)/n$ derived above is its larger
root. To do so we check that the derivative of the above with respect to \( q \) evaluated at \( S_n(p)/n \) is positive:

\[
-(1 - \mu)(1 - q - \frac{n - 1}{n} S_n(p)) + \frac{n - 1}{2n} S'_n(p)(1 - \mu)(2p - 1) > 0
\]

After substitution for \( S_n(p) \) and \( S'_n(p) \) we note that the derivative above is given by:

\[
\frac{\sqrt{\mu(1 - \mu)}}{2\sqrt{(2p - 1)^{-n/(n-1)}}} [2 - (2p - 1)^{-n/(n-1)}]
\]

This derivative should be positive for all \( p \) and that is determined by the expression in square brackets. This expression is increasing in \( p \) and positive at \( p = 1 \). It will be positive for all \( p \) if it is positive at \( A_n \). Noting the expression for \( A_n \) above, the derivative will be positive if and only if \( 2 - \frac{1}{\mu} > 0 \) or \( \mu > .5 \). We also need to check that the smaller root of the quadratic equation in \( q \) is less than zero. Brute force shows that this is true if \( \mu > .5 \).

Thus, the pure strategy equilibrium is as described above as long as there is sufficient adverse selection, namely \( \mu > .5 \). If on the other hand, \( \mu < .5 \) then, if there is a pure strategy equilibrium, it is not as characterized above. But this is very extreme adverse selection—on average every other trade is the result of news and those who learn of it first. This implausibility leads to an analysis of a mixed strategy equilibrium with positive profits.
Figure 2: Convergence of total supply of shares. The left figure corresponds to the equilibrium in mixed strategies with rents. The outer band is the 95% confidence band when \( n = 10 \); e.g., with probability 0.95 the asking price for the first 0.1 units is anywhere between 0.8 and 0.87 while the asking price for the next 0.1 unit between 0.805 and 0.9. The solid inner band is the 95% confidence band when \( n = 1,000 \). Finally, the dashed curve corresponds to the competitive equilibrium. The right figure corresponds to the pure strategy equilibrium. The solid line is the total supply curve when \( n = 5 \), and the dashed curve is the competitive supply function. The case \( n = 1,000 \) is not shown because it is indistinguishable from the competitive. In both figures, the liquidation value is either zero or one with equal probabilities, the noise order size has the the triangle distribution, and \( \mu = 0.6 \).

5.2. Mixed Strategy, Positive Rents

Given \( \tilde{q} \), we assume the order size is a lattice random variable, \( \tilde{q}_n \), with support

\[
\{ 1/n, 2/n, \ldots, n/n \}
\]

and distribution \( \text{Prob}(\tilde{q}_n \leq j/n) = F_Q(j/n) \). We search for an equilibrium in which each of the limit-order traders offers \( 1/n \) at a random price. In order for this to be an equilibrium, it must be that profits are the same no
matter what price is picked and that the liquidity suppliers have no incentive to supply a greater quantity. In the end, we shall show the analysis for the environment considered above; i.e. $\tilde{v}$ is equally likely to be zero or one and $F_Q$ is the triangle distribution. We shall, however, start with the general analysis as this indicates when this equilibrium will and will not exist. The first condition requires that expected profit be the same for all prices between $A_n$ and 1:

$$\mu(p-v(p))(1-F_v(p))\frac{1}{n}+.5(1-\mu)(p-E[\tilde{v}])P\{\tilde{q}_n > \tilde{S}_{-i}(p)\} = \frac{\pi^*}{2}; A_n \leq p \leq 1$$

The second condition states that it is unprofitable to add additional quantity at a price. Given that $1/n$ has been added at a price $p$ or less, player $i$ knows that the random supply at $p$ is $S_{-i}(p) + 1/n$. It will be unprofitable to add quantity at $p$ if:

$$\mu(p-v(p))(1-F_v(p)) + .5(1-\mu)(p-E[\tilde{v}])P\{\tilde{q}_n \leq \tilde{S}_{-i}(p) + \frac{1}{n}\} \leq 0$$

We fix $\pi^*$ by evaluating the profit expression at $p = 1$:

$$\pi^* = (1-\mu)(1-E[\tilde{v}])P\{\tilde{q}_n > \frac{n-1}{n}\} = (1-\mu)(1-E[\tilde{v}])P\{\tilde{q}_n = 1\}$$

By substituting for the "loss to informed" term in the above we get:

$$\frac{1-E[\tilde{v}]}{p-E[\tilde{v}]}P\{\tilde{q}_n = 1\} \leq P\{\tilde{q}_n = \tilde{S}_{-i}(p) + \frac{1}{n}\}$$

We first note that such an inequality cannot be satisfied for any non-decreasing density of $\tilde{q}$. For example, if $\tilde{q}$ were uniformly distributed, then the probabilities on either side of the equal sign would be equal and we would require $p \geq 1$, which is false. Thus, we will try the triangle distribution.
For the triangle distribution, \( P\{\tilde{q}_n = 1\} = \frac{1}{n^2} \) and hence the profit equation must solve:

\[
\mu(p-v(p))(1-\bar{F}_v(p)) + .5(1-\mu)(p-E[\tilde{v}])P\{\tilde{q}_n > \tilde{S}_{-i}(p)\} = \frac{(1-\mu)(1-E[\tilde{v}])}{2n^2} = \pi^*/2
\]

After noting that \( \tilde{S}_{-i}(p) \) is \( 1/n \) times a binomially distributed random variable with parameters \( n-1 \) and the mixing distribution at \( p, M(p) \), and that \( P[\tilde{q}_n > \tilde{S}_{-i}(p)] = E[(1-\tilde{S}_{-i})^2] \) we get a quadratic equation in \( M(p) \) similar to the one for the zero profit example worked out above. The solution is given by:

\[
M(p) = \frac{(n-1)(2n-1)}{n^2} - \frac{\sqrt{(n-1)^2(2n-1)^2 - 4[1-G(p)] - (1-E[\tilde{v}])}}{2(n-1)(n-2)/n^4}
\]

Armed with the mixing distribution, we can calculate

\[
P\{\tilde{q}_n = \tilde{S}_{-i}(p) + \frac{1}{n}\} = E[(1-\tilde{S}_{-i}(p))^2-(1-\tilde{S}_{-i}(p)-\frac{1}{n})^2] = \frac{2n-1 - 2(n-1)M(p)}{n^2}
\]

The inequality to be satisfied is:

\[
\frac{1-E[\tilde{v}]}{p-E[\tilde{v}]} \leq (2n-1) - 2(n-1)M(p)
\]

After many calculations, this inequality is equivalent to the following:

\[
8n^2(n-1)\frac{\mu}{1-\mu}(v(p) - p)(1-F_v(p)) - (1-p)((n + \frac{(n-2)(1-E[\tilde{v}])}{p-E[\tilde{v}]}) \geq 0
\]

After substituting \( v(p) - p = 1-p \) and \( 1 - F_V(p) = .5 \) and factoring out \( 1-p \) we get:

\[
2n^2(n-1)\frac{\mu}{1-\mu}(2p - 1) - (pn - 1) \geq 0
\]

In order for this to be satisfied for all \( p \) between \( A_n \) and 1, we require that \( \mu > \frac{1}{(2n-1)^2} \) and

\[
A_n \geq \frac{\mu(2n^2(n-1) + 1) - 1}{n(\mu(2n-1)^2 - 1)}
\]
Recall that $A_n$ is defined by $M(A_n) = 0$. Making reference to the expression for $M(p)$, this requires that $1 - G(A_n) - \frac{1}{n^2(2A_n-1)} = 0$ with $G(p)$ given by $\frac{\mu}{1-\mu} \frac{1-p}{p-\frac{1}{2}}$. Solving, we get:

$$A_n = .5(1 + \mu) + \frac{(1 - \mu)}{2n^2}$$

The inequality above then specifies that a convex quadratic function of $\mu$ must be greater than or equal to zero. The quadratic equation has one positive root, $\frac{n-1}{(n+1)(2n-1)}$ and one negative root. The conclusion of the analysis is that for this positive profit equilibrium to exist there must be sufficient adverse selection, specifically, we require

$$\mu \geq \frac{n-1}{(n+1)(2n-1)} \geq \frac{1}{(2n-1)^2}$$

Obviously, as $n$ gets large, the constraint becomes non-binding. That is, the existence of the positive profit equilibrium is more likely, the greater the number of limit-order providers. That in turn means that the profit, while positive, is likely to be small.

One could speculate, based on the example in Section 1 that for smaller $\mu$ a doubly mixed strategy might be an equilibrium. Recall that in that example with two players, a price was chosen at random and then with probability $l$ an extra unit was supplied at that price and with probability $1-l$ an extra unit was supplied at $p = 1$. We have been unable to verify whether or not there is a positive profit equilibrium in which limit-order traders either supply $1/n$ or $2/n$ at random prices.

It is interesting that if $\tilde{v}$ has a uniform distribution, there is not a positive rent equilibrium of the sort we examine here. To see this note that when $(1-p)/2$ was substituted the corresponding term for the uniform distribution
would have been \((1 - p)^2/2\). The first part of the expression would continue to have a factor \(1 - p\) and hence the inequality would not have been satisfied for \(p\) near one.

It is easy to see that \(M(p)\) converges to the competitive mixing distribution, \(1 - \sqrt{G(p)}\). As noted above, the supply at a price \(p\) or below is the sample mean of independent Bernoulli trials with success probability \(M(p)\) and hence, for the reasons given in section 4, the random supply function converges to the competitive supply function.

6. Discrete prices

In all of our analysis, the distribution of the random price is continuous—a player may quote any real number. This is, of course, at odds with Rule 612 of Regulation NMS which requires quoting on pennies for stocks that trade over one dollar and on hundredths of a penny for stocks that trade under a dollar.\(^{14}\) We think of this model with continuous prices as a convenient abstraction, but that means our analysis is most appropriate for high priced stocks and perhaps for stocks that trade less than, but very close to a dollar. What we require is that the tick size be small relative to the possible realizations of news.

The convenience is clear. We need not consider secondary precedence

\(^{14}\)The rule applies to quotes. Transactions on subpennies are allowed on exchanges due to "midquote" orders which are undisplayed orders at the midpoint of the bid ask spread. Furthermore the new retail liquidity program allows for undisplayed orders, available only to brokers representing retail trade, at fractional pennies at least $.0005 better than the market quote.
rules because almost surely no two players will quote the same price. It would be straightforward to model the stage game with discrete prices and a random allocation rule. However, extending the stage game to the dynamic analysis would not be straightforward at all. The key problem is, of course, that with time precedence, a player may choose not to cancel and quote if standing pat would put him in a position of quoting at a reasonable price with high precedence. We illustrate this with two examples presented without detail.

Consider two players A and B quoting on two possible prices; i.e. the tick size is very large relative to the variation in news. Furthermore, suppose the noise trader trades one share, as in section 2, while $\tilde{v}$ is equally likely to be zero or one. If both players pick the same price, then one is picked to trade with a coin toss.\footnote{We do not supply the details of the analysis at this point. It may be in later versions, or we may make it the focus of a new paper.} For reasonable parameter values, there will be a mixed strategy equilibrium of the stage game in which each player quotes one share and randomly picks between the higher price and the lower price. Furthermore, parameters can be chosen so that there are positive expected profits.

Suppose that the market reveals that A has chosen the high price and B the low price and no noise trader arrives so there is not a transaction. If B (at the low price) stands pat, then A’s lot cannot be improved since B enjoys time precedence. Similarly, B will not move since A enjoys time precedence at the higher price. In this case, there will be no cancelation and requote—it is an equilibrium for both to stand pat. This also turns out to be the case
if they both happen to pick the same price, even if it is the high one. If the parameters are such that a mixed strategy equilibrium exists, then both prefer random precedence at a high price to execution at the low price. In this case of extreme price discreteness, quotes will not be fleeting. Initially, limit-order submitters will randomize and then it is an equilibrium to stand pat until there is a transaction.

Suppose, now that there are three prices between one half and one. For reasonable parameter values there is a positive profit mixed strategy equilibrium in which A and B randomize across the three prices. Once the quotes are revealed, some constellations of quotes will lead to cancellation and re-quote; others will not. For example, if A randomly chooses the highest price and B the lowest, then B will want to cancel and move higher, but then A will cancel and undercut. That is, one would expect these quotes to be fleeting. On the other hand, if A and B both quote the middle price, there will be no movement.

The purpose of this discussion is to suggest that how ephemeral limit orders are will depend upon how large the tick size is relative to the size of news innovations. The smaller the tick size, the more cancellation and requoting one might expect. One might then expect greater cancellation rates for stocks priced just under a dollar (when the tick is $.0001) and fewer cancellations for stocks priced just over a dollar (when the tick is $.01). Bartlett III and McCrary (2013) find this to be the case. For stocks priced just under a dollar the rate of best bid and offer changes per second is twice what it is for stocks priced just over a dollar. This result is consistent with our model and discussion, but a more rigorous test would look at not just
cancellations at the market quotes but away from the market as well. That poses a data challenge.

7. Conclusion

We have analyzed a dynamic model of liquidity provision using limit orders. We argue that this dynamic game is reasonably thought of as repeated stage games in which limit-order providers enter quotes randomly. After seeing the results, they then cancel and requote randomly in the next stage game. In a general noise trader model with informed “news” traders the stage game is shown to always have a mixed strategy equilibrium in which the limit-order traders earn zero profits. Furthermore, this mixed strategy equilibrium converges, as the number of limit-order traders gets large to the competitive limit-order book.

We also look for positive profit equilibria and conclude that while a pure strategy equilibrium of the stage game may exist, existence probably requires extreme adverse selection. For more reasonable levels of adverse selection, a mixed strategy positive profit equilibrium exists for some environments. This equilibrium also converges to the competitive as the number of limit-order traders gets large. Our analysis is less complete here. Further searches for positive profit mixed strategy equilibria might be profitable.

Our results lean heavily on the assumption of continuous prices. We speculate on the effect of a positive minimum tick, but a more detailed analysis lies in the future. Preliminary results suggest that a larger tick will lead to fewer cancellations and re quoting consistent with recent evidence.
Appendix A. Proofs

Proof of Lemma 1. To verify that we have found an equilibrium, we consider the offer side of the stage game, the bid side is symmetric. We assume the second trader uses the equilibrium strategy, and consider the problem of the first trader. Offering more than 2 lots cannot be optimal because noise traders buy at most two. We therefore focus on price schedules $P_1 : [0, 2] \to R^+$. We note that shares offered at prices greater than one are never executed because trader 2 offers two lots at prices smaller or equal to one.

The expected profit, conditional on $\tilde{q} > 0$, is

$$\frac{\mu}{2} \int_0^2 I_{\{P_1(q) \leq 1\}} (P_1(q) - 1) \, dq + \frac{1 - \mu}{2} \int_0^2 \text{Prob}(q + \tilde{S}_2(P_1(q) - ) \leq \tilde{q})(P_1(q) - 0.5) \, dq$$

$$= \int_0^1 I_{\{P_1(q) \leq 1\}} \left[ \frac{\mu}{2} (P_1(q) - 1) + \frac{1 - \mu}{2} \left( 1 - M_n(P_1(q)) + \frac{M_n(P_1(q))(1-l)}{4} \right) (P_1(q) - 0.5) \right] \, dq \quad (I)$$

$$+ \int_1^2 I_{\{P_1(q) \leq 1\}} \left[ \frac{\mu}{2} (P_1(q) - 1) + \frac{1 - \mu}{2} \cdot \frac{1 - M_n(P_1(q))}{4} (P_1(q) - 0.5) \right] \, dq \quad (II)$$

The integrand of (I) is $\pi^*/2$ as long as $P_1(q) \in (\text{ask}, 1]$, and strictly less otherwise. The integrand of (II) is zero if (i) $P_1(q) = 1$ or (ii) $\mu \leq 1/9$ and $P_1(q) \in (\text{ask}, 1]$. The integrand of (II) is strictly negative otherwise.

We conclude that it is also optimal for the first trader to offer a single lot at any price in $(\text{ask}, 1]$. In particular it is optimal to randomize. Also, there is no harm in offering the second lot at one, and when $\mu \leq 1/9$ it is also optimal to offer the second lot at the same price at which the first lot is offered. In particular, it is optimal to randomize and with probability $l$ to offer both lots at the same price. Thus, we have verified the equilibrium. □
Proof of Lemma 2: In the interval $(E\tilde{v}, \bar{v}]$, the function $G(p)$ is continuous. To see this, note that the denominator in (5) is continuous. The numerator in (5) is
\[ \mu \int_{p}^{\bar{v}} (v - p)I_{\{p<v\}}dF_{V}(v) \]
which is continuous in $p$ whether $\tilde{v}$ is a discrete or continuous random variable. Thus, $G(p)$ is continuous in $(E\tilde{v}, \bar{v}]$.

In the interval $(E\tilde{v}, \bar{v}]$, the function $G(p)$ is strictly decreasing. Indeed, the derivative of the numerator is $\mu(F_{V}(p) - 1) < 0$. The denominator is clearly increasing in $p$. Hence we conclude that $G(p)$ is strictly decreasing.

Because $\lim_{p\downarrow E\tilde{v}} G(p) = 0$, and $\lim_{p\downarrow E\tilde{v}} G(p) = \infty$, it follows that a solution to the equation $G(p) = 1$ exists. Because $G$ is strictly decreasing, its inverse is also strictly decreasing. \qed

Proof of Theorem 3. We first show that the mixing distribution in Theorem 3 is well defined. Because $F_{Q}$ is a distribution, clearly, $k(0) = 1$ and $k(1) = 0$. Also,
\[
k(h) = 1 - EF_{Q}(\tilde{j}/(n - 1)) = 1 - EE[I_{\{\tilde{q}<\tilde{j}/(n-1)\}}]\tilde{j}] = 1 - EE[I_{\{(n-1)\tilde{q}<\tilde{j}\}}]\tilde{q}] = EB((n - 1)\tilde{q}; n - 1, h)
\]
where $B(x; n, h)$ is the distribution function of a binomial random variable with $n$ Bernoulli trials, each with a probability of successes $h$. Since $B(x; n, h)$ is strictly decreasing with $h$, it follows that also the expectation, $k(h)$, is strictly decreasing. \footnote{Formally, let $j = \lfloor x \rfloor$ be the floor of $x$, then
\[
B(j, n, h) = (n - j) \binom{n}{j} \int_{0}^{1-h} t^{n-j-1}(1-t)^{j} dt
\]}
From Lemma 2, we know that $G$ is strictly decreasing, and hence $M_n(p) = h(G(p))$ is strictly increasing. To conclude that $M_n$ is a distribution with support $[G^{-1}(1), G^{-1}(0)]$, we verify:

\[ M_n(G^{-1}(1)) = h(1) = 0 \]

and

\[ M_n(G^{-1}(0)) = h(0) = 1 \]

Therefore, $M_n$ is a distribution function.

Next, we apply $k(h)$ to both sides of the definition of $M_n(p)$ to get that in $(G^{-1}(p), G^{-1}(0))$, we have $K(p) = G(p)$. More generally,

\[
K(p) = \begin{cases} 
1 & p < G^{-1}(1) \\
G(p) & G^{-1}(p) \leq p < G^{-1}(0) \\
0 & G^{-1}(p) \leq p
\end{cases}
\]

Consider now the problem of the $i$th trader, assuming all other limit-order traders follow the strategy stated in the theorem. For $q \in (0, 1/(n-1)]$, we have

\[
Prob(\tilde{q}_n \geq q + \tilde{S}_{-1}(P_i(q))) = Prob(\tilde{q}_n \geq \tilde{S}_{-1}(P_i(q))) = K(p)
\]

where the second equality is the definition of $K(p)$. Hence, the expected profits associated with the “first” $1/(n-1)$ units, each unit may be offered at

and hence the probability is strictly decreasing. Informally, when we increase the probability of success in each trial, then the probability of having a total of $j$ or less successes strictly decreases - while this is not true for the probability of having exactly $j$ successes, it is true for the cumulative probability.
different prices, is
\[
\int_0^{1/(n-1)} \mu(P_i(q) - v(P_i(q)))(1 - F_V(P_i(q))) + \frac{1-\mu}{2} (P_i(q) - \bar{\mu}) K(p) \, dq
\]
\[
(A.1)
\]
We consider the value of the integrand for different \( p \)'s. For \( p < p_c \), the integrand is negative because \( p_c \) is the ask price of the competitive equilibrium. For \( p > \bar{v} \), the integrand is zero.

For \( p \) in the support of the mixing distribution, we use the definitions of \( G(p) \) (equation (5)) and \( M_n(p) \) to conclude that the integrand of the objective (A.1) is zero. Therefore, the expected gain on the first \( 1/(n-1) \) unit is at most zero, and exactly zero if the units are offered in the interval of prices \( (p_c, \bar{v}) \).

We need to show that it is suboptimal to offer more than \( 1/(n-1) \) units. But this is obvious because the chances that additional units will be picked by the noise traders are strictly smaller than the probability that the first \( 1/(n-1) \) units are. Since the profitability of the latter is zero, it follows that it is suboptimal to offer more than \( 1/(n-1) \) units.

We conclude that it is optimal to offer \( 1/(n-1) \) in the support of \( M_n \) and in particular it is optimal to offer the entire block at the same random price with distribution \( M_n \). Finally, as we have seen, the expected profit is zero.

\( \Box \)

\textit{Proof of Corollary 1.} The proof is in steps.

Step 1: Given \( \delta > 0 \) and \( \epsilon > 0 \), there exists an \( N \), independent of \( p \), such that for all \( n > N \), we have \( \text{Prob}(|\tilde{S}_{-i}(p) - M_n(p)| > \delta) \leq \epsilon/4 \). In particular, for every \( p \), \( \tilde{S}_{-i}(p) - M_n(p) \) converges to zero in probability.
Indeed, we have \((n - 1)\tilde{S}_{-i}(p) \sim B(n - 1, M_n(p))\). Take \(N > 1/(\epsilon\delta^2)\), then from Chebyshev’s Inequality

\[
\text{Prob}(|\tilde{S}_{-i}(p) - M_n(p)| > \delta) \leq \frac{\text{Var}(\tilde{S}_{-i}(p))}{\delta^2} = \frac{M_n(p)(1 - M_n(p))}{(n - 1)\delta^2} \leq \frac{1}{4n\delta^2} < \frac{\epsilon}{4}
\]

Step 2: \(E F_Q(\tilde{S}_{-i}(p)) - F_Q(M_n(p))\) uniformly converges to zero. Let \(\epsilon > 0\) be given. We need to show that there is an \(N\) such that for all \(n > N\) we have \(|EF_Q(\tilde{S}_{-i}(p)) - F_Q(M_n(p))| < \epsilon\).

The distribution function \(F_Q\) is continuous in the closed interval \([0, 1]\) and hence it is uniformly continuous. Thus, there exists a \(\delta > 0\), associated only with \(\epsilon\), such that if \(|q_1 - q_2| < \delta\), then \(|F_Q(q_1) - F_Q(q_2)| < \epsilon/2\).

We take now \(N > 1/(\epsilon\delta^2)\) (as in Step 1). Now,

\[
|EF_Q(\tilde{S}_{-i}(p)) - F_Q(M_n(p))| \\
\leq E|F_Q(\tilde{S}_{-i}(p)) - F_Q(M_n(p))| \\
= E|F_Q(\tilde{S}_{-i}(p)) - F_Q(M_n(p))|I\{|\tilde{S}_{-i}(p) - M_n(p)| \leq \delta\} \\
+ E|F_Q(\tilde{S}_{-i}(p)) - F_Q(M_n(p))|I\{|\tilde{S}_{-i}(p) - M_n(p)| > \delta\} \\
\leq \frac{\epsilon}{2} + 2\text{Prob} \left( \left| \tilde{S}_{-i}(p) - M_n(p) \right| > \delta \right) < \epsilon
\]

Step 3: We use Lemma 4 to replace, in Step 2, \(EF_Q(\tilde{S}_{-i}(p))\) with \(F_Q(S_c(p))\) and conclude that \(F_Q(S_c(p)) - F_Q(M_n(p)) \rightarrow 0\) uniformly.

Step 4: We are now ready to show that \(S_c(p) - M_n(p) \rightarrow 0\) uniformly. In other words, we need to show that given \(\epsilon\), there is an \(N\), independent of \(p\), such that

\[
|S_c(p) - M_n(p)| < \epsilon
\]
The inverse distribution function, $F_Q^{-1}(x)$ is continuous in $[0, 1]$ and hence uniformly continuous. Thus, there is a $\delta > 0$ such that $|x_1 - x_2| < \delta$ implies $|F_Q^{-1}(x_1) - F_Q^{-1}(x_2)| < \epsilon$. From Step 3, we know that there is an $N$ that depends only $\delta$ such that for $n > N$, we have

$$|F_Q(S_c(p)) - F_Q(M_n(p))| < \delta$$

Thus, for $n > N$, we also have

$$|S_c(p) - M_n(p)| = |F_Q^{-1}(F_Q(S_c(p))) - F_Q^{-1}(F_Q(M_n(p)))| < \epsilon$$

where the inequality follows from the uniform continuity of $F_Q^{-1}$.

\[\square\]

References


