High frequency trading strategies

Michael Goldstein*, Babson College

Amy Kwan, University of Sydney

Richard Philip, University of Sydney

Current version: December 27, 2018

Abstract

We document an important information channel driving HFT behavior. Examining the order book imbalance immediately before each order submission, cancelation and trade, we show high frequency traders (HFT) use limit order book information to supply liquidity on the thick side of the order book and demand liquidity from the thin side. This strategic behavior is more pronounced during volatile periods and when trading speeds increase. However, by competing with non-HFT limit orders, HFT impose a welfare externality by crowding out slower non-HFT limit orders. Results from a natural experiment that increased market information access speeds further supports our findings.

Keywords: High frequency trading, institutional investors, retail investors

JEL: G12, G28

* Goldstein (corresponding author), email: goldstein@babson.edu; address: 320 Tomasso Hall, 231 Forest Street, Babson Park, MA 02457-0310; tel.: (781) 239 4402; fax: (781) 239 5004; Kwan, email: amy.kwan@sydney.edu.au; address: University of Sydney, Sydney, NSW 2006, Australia; tel.: +61 2 9036 9340; fax: +61 2 9351 6461. Philip, email: richard.philip@sydney.edu.au; address: University of Sydney, Sydney, NSW 2006, Australia; tel.: +61 2 9034 0578; fax: +61 2 9351 6461. We thank Sean Anthonisz, Joseph Barbara, Hedi Benamar, Bidisha Chakrabarty, Kevin Crotty, Doug Foster, Sean Foley, Austin Gerig, Elmer Funke Kupper, Brad Katsuyama, Vincent van Kervel, Adrian Lee, Katya Malinova,Thomas McInish, Maureen O'Hara, Christine Parlour, Talis Putnins, Ryan Riordan, Kevin Roshak, Ioanid Rosu, Andriy Shkilko, Elvira Sojli, Avanidhar Subrahmanyam, Wing Wah Tham, Kumar Venkataraman, Bart Yueshen, Haoxiang Zhu, and Marius Zoican, as well as conference and seminar participants at the Australasian Finance and Banking Conference, Auckland Finance Conference, Centre for International Finance and Regulation (CIFR) Market Microstructure Conference, Curtin University, European Financial Management Association Annual Meeting, Financial Management Association Annual Meeting, Financial Research Network (FIRN) Sydney Market Microstructure Meeting, Queensland University of Technology, and the Northern Finance Association Annual Conference for helpful comments. We would also like to thank CIFR, which funded this research under CIFR grant T013, and the Securities Industry Research Centre of Asia-Pacific (SIRCA) for providing the data used for part of this study.

High frequency trading strategies

Current version: December 27, 2018

Abstract

We document an important information channel driving HFT behavior. Examining the order book imbalance immediately before each order submission, cancelation and trade, we show high frequency traders (HFT) use limit order book information to supply liquidity on the thick side of the order book and demand liquidity from the thin side. This strategic behavior is more pronounced during volatile periods and when trading speeds increase. However, by competing with non-HFT limit orders, HFT impose a welfare externality by crowding out slower non-HFT limit orders. Results from a natural experiment that increased market information access speeds further supports our findings.

Keywords: High frequency trading, institutional investors, retail investors

JEL: G12, G28

1.0 Introduction

Financial markets aggregate trading behaviors of different market agents. Traditionally, retail and institutional investors traded with dealers or market makers making two-sided markets. Recently, high frequency traders (HFT) have become increasingly dominant, influencing financial markets through their trading decisions. Some suggest that HFT reduce the bid-ask spread, increase price efficiency, and increase overall market depth, while others disagree or note these effects may vary with market conditions.[1]

While these are *ex post* effects of HFT trading decisions, much less is known about what affects HFT trading decisions *ex ante* due to the proprietary nature of their business. Some suggest that HFT can anticipate future order flow from other traders and predict future price movements, but how HFTs anticipate the future and what information they use has not been determined.[2] HFT's ability to anticipate the future may also vary with market conditions. Further, little is known about how HFT behavior differs from, let alone affects, the behavior of non-HFT market participants, such as institutions and retail investors, or if HFT behavior has welfare effects on other market participants and under what conditions. While we understand much about how HFT affect financial markets, not much is known about the information channels that drive HFT behavior.

---

[1] Angel, Harris and Spatt (2010), Jones (2013), Harris (2013), Hasbrouck and Saar (2013), Brogaard, Hagstromer, Norden and Riordan (2015), and Malinova, Park and Riordan (2016) suggest HFTs affect spreads. Carrion (2013) and Brogaard, Hendershott and Riordan (2014) suggest HFTs affect price efficiency, and Hasbrouck and Saar (2013) suggest HFTs affect depth. Studies such as Hagstromer and Norden (2013), Hasbrouck and Saar (2013), Menkveld (2013), Malinova, Park and Riordan (2016), and Conrad, Wahal and Xiang (2015) suggest that HFTs provide liquidity during relatively normal periods, while Anand and Venkataraman (2016), Kirilenko, Kyle, Samadi and Tuzun (2016), Brogaard Carrion, Moyaert, Riordan, Shkilko and Sokolov (2017), van Kervel and Menkveld (2017) and Korajczk and Murphy (2016) among others suggest that HFTs withdraw this liquidity during periods of stress.

[2] See, for example, Brogaard, Hendershott and Riordan (2014), Li (2014), Hoffmann (2014), Biais, Foucault and Moinas (2015), Foucault, Hombert and Rosu (2016), Hirschey (2016) and Rosu (2016).

To shed light on these issues, we use a unique broker-level dataset from the Australian Securities Exchange (ASX) and document one such publicly available information channel, namely the order book depth imbalance. Using this dataset, which allows us to classify brokers into three trader types (proprietary HFT firms, institutions and retail), we examine how the order book imbalance affects HFT behavior and that of institutions and retail traders.[3] By reconstructing the full limit order book and measuring the shape of the order book at the time of each order submission, cancelation, amendment, and trade, we find strong evidence that the shape of the limit order book contains information about future price movements, supporting Cao, Hansch and Wang (2009) and Cont, Kukanov and Stoikov (2014). Consistent with the theoretical predictions of Parlour (1998), the order submission strategies for all traders (i.e., HFT, institutions and retail) capture information contained in the limit order book. However, HFT are better at taking advantage of this information, especially at times when the market is volatile and when HFT's speed advantage increases.[4] Our examination of a natural experiment that increased market information access speeds (for a fee) and thus, further segmented the market into fast and slow, provides additional support for these results. Thus, while market participants move the market towards weak form efficiency by trading on order book information, HFTs use their speed advantage strategically in ways when it is most advantageous to do so. In effect, we document the Grossman and Stiglitz (1980) paradox at the microstructure level.

[3] We refer to 'information' as order-related public information, which can be inferred from the limit order book, rather than fundamental, asset-related information; see O'Hara (2015) for a discussion of information in the high frequency world.
[4] Previous studies have suggested HFT increase price efficiency although little is known about the precise channel through which this occurs. Brogaard, Hendershott and Riordan (2014) demonstrate that HFT buy in the direction of permanent price changes through liquidity demanding orders. Carrion (2013) shows that HFT incorporate information from order flow and market-wide returns more efficiently. Several studies, including Chaboud Chiquoine, Hjalmarsson and Vega (2014), Hendershott, Jones and Menkveld (2011), Boehmer, Fong and Wu (2015) find that algorithmic trading, of which HFT is a subset, improves informational efficiency of prices.

Our results suggest that HFTs use the additional information provided in today's markets adapting the old Wall Street adage of "Don't fight the tape" to "Don't fight the limit order book". HFT executions (either due to market orders or the execution of previously submitted limit orders) occur in the direction of the order book imbalance (i.e., buy when the limit order book is deeper on the bid side than the ask side) but HFTs cancel or amend their limit orders when the order book imbalance moves against them.[5] Using multinomial logistic regression, we investigate a trader's order placement decisions and show that HFTs submit limit orders to the order book primarily when a small favorable depth imbalance exists, i.e., slightly more depth on that side of the book. If the depth imbalance becomes more favorable (i.e., the depth of the order book on the side of the order dominates the other side), the resting limit order is left to execute in the same direction as the imbalance would imply. On the other hand, if the depth imbalance becomes less favorable, HFT are quick to cancel or amend their orders, reducing adverse selection costs. As a result, while it may be true that HFT add liquidity, they do not do so at times when it is most needed: On average, HFTs supply liquidity on the thick side of the order book (where it is not needed) and demand liquidity from the thin side (where it is needed). Notably, HFTs are not simply faster at responding to public news. Splitting our sample into times of high and low information, we find that HFTs are just as successful at responding to information contained in the order book in high and low information environments.

While non-HFTs exhibit similar behavior to HFTs, they are less successful precisely when the market indicates it is most important to do so. Brogaard, Hendershott, and Riordan (2014) find that HFTs trade in the direction of the order book imbalance, but so does everyone

[5] When the bid depth exceeds the ask depth, a trader 'trades in the direction of the order book' if 1) a buy limit order executes or 2) a buy market/marketable limit order executes. Similarly, when the ask depth exceeds the bid depth, a trader 'trades in the direction of the order book' if 1) a sell limit order executes or 2) a sell market/marketable limit order executes.

else – just less successfully.[6] The conditions under which HFTs notably do better than non-HFTs are non-linear: when depth imbalances are very large, suggesting large future price changes, HFTs are more successful at trading in the direction of the imbalance, relative to the other trader types.[7] Importantly, we document that large depth imbalances only persist in the order book for a fraction of a second. Since speed is critical for a trader to capture this fleeting information, HFT are most successful precisely during market conditions when fast trading speeds offer the greatest advantages. While retail and institutional investors are likely to trade for exogenous reasons over longer investment horizons, HFT act opportunistically to capture short term price movements,[8] and trade at the most profitable opportunities where they can use their speed advantage.[9]

The introduction of faster ITCH technology on the ASX in 2012 provides a natural experiment to investigate the effects of a speed change on HFT trading behavior. Using a difference-in-difference framework, we demonstrate that HFT become even more successful at trading in the direction of the order book when they gain a larger speed advantage. However, there is a negative welfare implication for non-HFT institutional and retail traders due to a

[6] While Brogaard, Hendershott and Riordan (2014) relies on trade reports, which only contains the portion of the strategy which executed, we are better able to infer HFT limit order placement strategies (as well as those of retail and institutional market participants) by observing the full limit order book as well as executions.

[7] Empirically, Hirschey (2016) documents that HFTs can anticipate order flow from other investors and Subrahmanyam and Zheng (2016) conclude that HFT manage their limit orders in anticipation of short-term price movements, but do not show how HFT predict future price movements.

[8] As noted in Section II, we also benefit from examining a largely consolidated market, as opposed to previous studies on pieces of fragmented markets. For example, Brogaard et al. (2017) find that HFTs supply liquidity to non-HFTs during extreme price moves in a single security but do not do so when several stocks experience simultaneous extreme price moves on a fragmented market with rebates. Our results on a largely consolidated market with no maker/taker rebates show that HFT are net demanders of liquidity and become even more aggressive at times of high volatility.

[9] Our findings are consistent with several theoretical models, such as Biais, Foucault and Moinas (2015), Li (2014), Hoffmann (2014), Rosu (2016), and Yang and Zhu (2017), who assume that fundamentally uninformed traders (i.e., HFTs) can anticipate the fundamental investor's order flow. Similarly, we show that HFTs, despite being uninformed about the fundamental value, use the information channel of the limit order book better and faster than institutional and retail participants to anticipate fundamental investors' order flow, and thus facilitate price efficiency by incorporating this information and information in the limit order book into the price formation process.

"crowding out" effect on non-HFT limit orders executions: The probability of execution for institutional and retail limit orders submitted to the best bid and ask prices decreases when HFT gain a larger speed advantage. Conditioning limit order executions on the shape of the order book, we show that it is the probability of favorable executions (i.e. non-HFT limit order trading in the direction of the order book imbalance), which falls. HFT's speed advantage has a welfare effect on the slower institutional and retail market participants. Over 40% of HFT profits arise from market order executions when the size of the order book imbalance at the time of trade is in the top decile of all order book imbalances. This finding suggests a large portion of HFT profits come from picking off stale limit orders from slower traders. While all traders attempt to trade in the direction of the imbalance, slower institutional and retail traders miss out on the most profitable trades.

Finally, our results have important implications for studies on the market making role of HFT in equity markets. Hendershott, Jones and Menkveld (2011), Hasbrouck and Saar, (2013), and Brogaard, Hagstromer, Norden and Riordan (2015) and others find that HFT market making increases market depth, but employ traditional measures of market depth, aggregated across both bid and ask prices. However, aggregated measures of market depth do not capture the amount of depth available on the side of the limit order book where it is most needed or wanted, making welfare comments more difficult – a trader submitting a buy market order is more concerned about the depth available on the ask side of the limit order book, rather than aggregated depth over both bid and ask prices.[10] By looking at each side of the book separately, we show HFTs on average supply depth on the thick side of the order book but demand depth from the thin side of the order book, i.e., they may add depth, but not on the side that is thin, and cancel limit orders

[10] Whether depth increases are beneficial depends in part on whether depth is added on the side where it is needed: adding depth to the already thick side of a limit order book is likely of little benefit to the market and could disadvantage prior limit order submitters, while adding depth to the thin side of a market may add useful liquidity.

from the thin side of the order book, providing a notably different picture than results from aggregated studies. These asymmetric findings suggest that caution should be taken in interpreting results that aggregate market depth across both the bid and ask sides.

2.0 Data and variable construction

2.1 Data and sample selection

We obtain full order book and trade data for stocks in the S&P/ASX 100 index from the AusEquities database provided by the Securities Industry Research Centre of Asia Pacific.[11] The securities contained in our dataset are the most liquid and actively traded on the Australian Securities Exchange (ASX), the dominant stock exchange for Australian equities, with over 90% market share of on-market traded volume in 2012 (Aitken, Chen and Foley, 2017). The ASX operates as a continuous limit order book between approximately 10:00 am and 4:00 pm, matching orders based on price and time priority, with a randomized open and a randomized close.

We analyze six months of order level data for the period January 3, 2012 to June 30, 2012. To avoid the randomized open and close, we include only trades and orders entered between 10:10:00 and 16:00:00 to ensure that our sample is not contaminated by the opening and closing call auctions. We assume that all outstanding orders remaining in the limit order book at the end of the trading day are cancelled. For each order book event, the data contain the stock symbol, date and time of event to the millisecond level, order size and price, order identification number and an identifier for the submitting broker and event type, which consists of submit, trade, amend, or cancel. Additionally, we use the order identification number to trace subsequent amendments, executions or cancelations back to the original order entry, allowing for a full

[11] The S&P/ASX 100 index contains the 100 largest stocks listed on the ASX by market capitalization. In 2012, approximately 2,050 companies are listed on the ASX with a total market capitalization of approximately AUD 1.5 trillion. The 100 stocks in the index comprise approximately 65% of total market capitalization.

reconstruction of the limit order book and the tracking of the order and its queue position through time.

Our period also covers the introduction on April 2, 2012 of ASX ITCH, which is an ultra-low latency protocol for accessing ASX market information available to all market participants for a monthly fee. ASX ITCH was designed to meet the requirements of speed sensitive traders and increased market information access speeds by up to seven times existing connections (ASX, 2013). Thus, the introduction of ASX ITCH is likely to create larger benefits for HFT, whose strategies rely on fast response times when new information arrives to the market and provides a natural experiment for our results.

Data from the ASX offer several advantages over other exchanges. For example, in our dataset, broker identifiers are assigned into three trader categories: proprietary HFT firms (*HFT*), *Institutions*, and *Retail*. (Appendix 1 provides more details on our three-step classification procedure.) Thus, we do not need to rely on HFT proxies such as message to trade ratios, which are negatively correlated with the true measure of HFT activity (see Yao and Ye, 2017; Ye, 2017). While we are able to identify the larger HFT firms from this data, we acknowledge some smaller proprietary HFT firms could trade through institutional brokers and thus, *Institutions* could also contain some proprietary HFT activity. These smaller HFT traders are unlikely to change overall institutional volume imbalances, although perhaps these smaller HFT firms could influence trading imbalances based on the number of trades. Thus, we report our results both in terms of the volume of shares traded and number of trades. In addition to flow from their institutional clients, some major global firms and large domestic financial institutions may also have order flow coming from institutional clients, their own proprietary HFT desks and retail investors. We classify these firms as *Institutions* as their primary business comes from institutional clients.[12] The misclassification of HFT or retail as institutions tends to bias against detecting differences between our trader categories. Even so, we continue to find economic and statistically significant differences between our trader categories, adding further robustness to our results.

Additionally, because we can replay the full limit order book, we do not have to rely on trade classification algorithms, such as Lee and Ready (1991), to determine whether a trade is buyer or seller initiated. Not relying on algorithms provides notably better inferences: Ellis, Michaely and O'Hara (2000) report that the Lee and Ready (1991) rule misclassifies approximately 20% of all trades. We also can rely on the granularity of the data to compute the depth imbalance proxy for trading strategies.[13] Following Upson, Johnson and McInish (2015), we aggregate all trade reports at the same price, in the same trade direction, from the same broker, and reported in the same millisecond timestamp into one marketable order.

Finally, in comparison to U.S. and European equity markets, the ASX is less fragmented, operating as a virtual monopoly in Australian equities during our period with over 90% of the daily trading volume, so our data comprises almost the entire market. This dominant market share is an advantage over studies on previous exchanges with data on smaller portions of the market. For example, Brogaard, Hendershott and Riordan (2014) examine HFT trading on Nasdaq, which accounts for approximately 20% of U.S. equities market share during their sample period. Similarly, van Kervel and Menkveld (2017) study HFT on Nasdaq OMX, which

[12] Our sample of proprietary HFT contribute approximately 8% of ASX dollar volume for stocks in the S&P/ASX 100 index. ASIC reports that HFT traders, which include HFT from large global firms and domestic financial institutions, account for 27% of total dollar volume in S&P/ASX 200 securities (ASIC, 2013). However, their estimates also include trades executed in dark pools, the largest of which are run by large global firms.

[13] We note, however, that we only observe trading strategies at the firm level due to the proprietary nature of their business. We are unable to identify the extent to which HFT firms follow multiple strategies (see Boehmer, Li and Saar, 2017), so we are observing the net results of their combined strategies, and not any one particular strategy. Nonetheless, the end result of the mix of these strategies is such that it appears that they are conditioning their trades on the shape of the limit order book, which enhances each of these strategies by ensuring their trading decisions encompass all available information.

accounts for approximately two thirds of the market share in Swedish stocks. Brogaard et al. (2017) focuses exclusively on the liquidity demanded and supplied through the Nasdaq exchange, which represents only 30-40% of all trading activity of the sample stocks, so it is possible that HFT are supplying liquidity on Nasdaq while demanding liquidity from other trading venues. Fragmentation could matter: Van Kervel (2015) shows in a fragmented market setting that consolidated measures of liquidity could overestimate the actual amount available, and show that when HFT market makers observe a trade on one venue, they cancel outstanding limit orders on all other venues to reduce their adverse selection costs.[14] As such, we are one of the few studies to analyze HFT in a largely unfragmented market.

[Insert Table 1]

Table 1, Panel A reports the summary statistics for the 94 stocks, which appear in the S&P/ASX 100 index over the full sample period. *Market capitalization* is measured on January 3, 2012, the first trading day in the sample, and is expressed in billions of AUD. All other variables are measured on a daily basis and averaged across the sample period. The average stock has a market capitalization of $13.52 AUD billion and volume weighted trade price of $11.43. The average daily dollar volume is $27.8 AUD million and the average number of trades is 2,264. Given that the minimum pricing increment on the ASX is $0.01 for stocks priced above $2.00, an average daily time-weighted quoted spread (*Spread*) of 1.02 cents indicates that many stocks in the sample are likely to be spread constrained. While twelve stocks in our sample have an average stock price under $2, our results are robust to removing these stocks from the sample.

Table 1, Panel B reports the summary statistics for all trader types, *HFT*, *Institutions* and *Retail*. Consistent with the prior literature, we find that *HFT* monitor the limit order book more

[14] Chen et al. (2017) finds similar results empirically with the introduction of speed differentials in Canada.

actively. Relative to *Institutions* and *Retail*, *HFT* have a higher percentage of order cancelations, and their median submission to cancel time is significantly lower. The average number of active and passive executions is approximately equal for *HFT*, whereas *Institutions* and *Retail* are predominantly limit order traders.[15]

In the following sections, we investigate how *HFT* incorporate information contained in the order book into their trading strategies.

2.2 Depth imbalance

Previous work has established that imbalances are related to future price directions, and that imbalances can affect order submission strategies. Chordia, Roll and Subrahmanyam (2002) and Chordia and Subrahmanyam (2004) document a strong relationship between trade imbalances and future returns. Using more granular limit order book data, Cao, Hansch and Wang (2009) and Cont, Kukanov and Stoikov (2014) find strong evidence that order imbalances between the buy and sell schedules of the limit order book are significantly related to future stock returns. Cont, Kukanov and Stoikov (2014) show that price changes over short time intervals are mainly driven by imbalances between the depth available at the best bid and ask prices. Specifically, large buying (selling) pressure on the bid (ask) price predicts future price rises (falls). Further, Ranaldo (2004) examines how the state of the limit order book can affect a trader's order submission strategy. Based on these studies, we use the information contained in the state of the limit order book to proxy for strategic trading.

To measure the shape of the limit order book prior to an order book event, we calculate depth imbalance (*DI*) as the difference between the depth available at the best bid and ask prices,

[15] The total number of passive executions exceeds the total number of aggressive executions as we aggregate all trade reports in the same trade direction, and reported in the same millisecond timestamp into one aggressive execution (see Upson, Johnson, and McInish (2015)).

as a proportion of the total depth available at the best bid and ask prices. Specifically, for each order book event (i.e., submission, trade, amendment or cancelation) we determine:

$$DI_t = \frac{\sum_{i=1}^{n} DepthBid_{i,t} - \sum_{i=1}^{n} DepthAsk_{i,t}}{\sum_{i=1}^{n} DepthBid_{i,t} + \sum_{i=1}^{n} DepthAsk_{i,t}}$$

where $\sum_{i=1}^{n} DepthBid_{i,t}$ ($\sum_{i=1}^{n} DepthAsk_{i,t}$) is the total depth available at the top *n* bid (ask) price levels immediately before the order book event, *t*. (We compute *DI* immediately before the time of the event to avoid capturing the volume of the order book event itself; see Appendix 2 for a detailed example of the construction of *DI*.) For our main results, we calculate *DI* based on the depth available at the top five bid and ask prices ($n = 5$); for robustness, we also test our results using one (n=1) and three (n=3) price levels and all results continue to hold.[16] Our measure of *DI* is bounded between -1 and 1, where a value close to -1 (1) indicates that the depth available at the ask (bid) price levels is much larger than the bid (ask) depth available, so this measure could be interpreted as a relative buying interest index.

For some tests, we multiply *DI* by an indicator for whether the order is a buy or sell to remove the effects of trade direction. We refer to this directionally adjusted *DI* measure as *Adjusted DI*. When *Adjusted DI* is positive, the trade or order event occurs in the direction of the depth imbalance (e.g., an active buy execution from a buy market order or a passive buy execution from a buy limit order executes when the bid depth exceeds the ask depth); when it is negative, the trade or order event occurs in the opposite direction of the depth imbalance (e.g., a buy market or limit order executes when the ask depth exceeds the bid depth).

[16] For n=5, we define the best five continuous bid (ask) prices as the best bid (ask) and then the next four possible continuous prices below (above) that price, whether or not there are any orders at that price: if the best bid is 20.17, the next four prices are 20.16, 20.15, 20.14, and 20.13, even if the depth at 20.16 and 20.13 is zero. For n=3, we use a similar definition; for n=1 we use just the best bid (ask).

Since the basic unit of analysis is an order, for trades there would be two entries – an active execution of the order that was liquidity demanding (a market order or marketable limit order that caused the execution to occur) and a passive execution (of the resting limit order that executed against the active order). Of course, the trader type need not be constant: For example, an institutional market order to sell could execute against a resting HFT limit order to buy. If the depth on the bid side is greater than the depth on the ask side, this example would result in an active execution with a negative *Adjusted DI* for the institution and a passive execution with a positive *Adjusted DI* for the HFT.

Table 1, Panel C summarizes the average *Adjusted DI* immediately before active executions, passive executions, order submissions, amendments and cancelations for each trader type. For active executions, we find that all trader types (*HFT, Institutions,* and *Retail)* on average trade in the direction of the imbalance, indicating that traders are more likely to submit a market buy (sell) order when the bid (ask) depth is much larger than the ask (bid) depth. Comparing the magnitude of *Adjusted DI*, *HFT* submit market orders when *Adjusted DI* is much larger, 0.156, compared to 0.027 and 0.025 for *Institutions* and *Retail*, respectively.

While an active execution is under the discretion of the trader, a passive execution relies on an unpredictable incoming market order. As such, we expect a larger *Adjusted DI* for active executions, relative to passive executions, for traders acting on order book information. Consistent with this intuition, *Adjusted DI* at the time of passive executions is 0.083 for *HFT* (compared to 0.156 for active market order executions), and is negative for *Institutions*, and *Retail* (-0.029, and -0.009, respectively). A negative *Adjusted DI* indicates that *Institutions* and *Retail* limit orders are picked off the thin side of the limit order book. In contrast, HFT typically supply liquidity to the thick side of the order book, which requires strategic submission and cancelations of limit orders.

We find that *HFT* submit limit orders when there is a moderate *Adjusted DI* (*Adjusted DI* = 0.063) but strategically cancel their limit orders when the order book moves against their resting limit orders, indicated by a lower *Adjusted DI* (0.017), which reduces the adverse selection costs of *HFT*. In contrast, *Retail* traders both submit and cancel their orders, in general, in the same direction of the book at a much more moderate *Adjusted DI* (0.043 for submissions; 0.028 for cancels), while on average, *Institutions* submit very slightly on the opposite side (-0.004) and cancel very slightly on the same side (0.002). We investigate HFT trading behaviors and their potential impact on non-HFT trading in the next section.

3.0 Empirical Results

3.1 Depth imbalance, future stock prices and aggregate trading volumes

Before examining HFT trading behavior, we first establish the information content of order book depth imbalances. To investigate whether depth imbalances contain information about the future stock price, we start by ranking trades into deciles based on the depth imbalance immediately before the trade for each stock-day. For each transaction, we also calculate future returns by comparing the midpoint of the best bid and ask prices at the time of the trade with the bid-ask midpoint 10 trades in the future. Figure 1, Panel A presents the average future return for trades from each depth imbalance decile. We observe a strong positive relation between the size and direction of the depth imbalance and future stock returns indicating that depth imbalances in the order book can predict future stock returns.[17] Specifically, as the number of buyers relative to

[17] Cont, Kukanov and Stoikov (2014) find that depth imbalances predict future short term price changes, but only use depth imbalances at the best bid and ask prices for their main study and do not examine differences between HFT and non-HFT. Using a sample period before the growth of HFT, Cao, Hansch and Wang (2009) find that order imbalances behind the best bid and offer contribute to approximately 22% of price discovery. In Appendix 3, we

sellers in the limit order book increase, the relative level of stock prices in the future also increases.

[Insert Figure 1]

Next, we examine how all market participants respond to depth imbalances. For each depth imbalance decile, we calculate the percentage of total volume that is buyer or seller initiated. Given that depth imbalances predict future returns, we expect strategic traders to trade in the direction of the order book imbalance. Specifically, we expect more aggressive buying (i.e., more buyer initiated trades) when a large positive depth imbalance exists and more aggressive selling when there are large negative imbalances. Consistent with strategic trading, our results in Figure 1, Panel B confirm a strong positive (negative) relationship between the size of the depth imbalance and the percentage of buyer (seller) initiated trade volume.

Table 2, Panel A, presents the associated depth imbalance levels and corresponding average volume of shares traded for each depth imbalance decile. We find that the market participants are more active when a large depth imbalance exists in the order book. The average number of shares traded is over 340,000 shares for each of the most extreme depth imbalance deciles (i.e., deciles 0 and 9) but approximately 270,000 shares when the order book is relatively balanced (i.e, deciles 4 and 5). This finding supports the theoretical prediction of Cespa and Vives (2017), who show that traders demand more liquidity when the market becomes less liquid. Interestingly, we find that *HFT* are more active in the extreme deciles while *Institutions* and *Retail* are more active when only a moderate depth imbalance exists.

---

show that depth on levels 2 to 5 contain additional information on future price changes in a high frequency world. Comparing the adjusted R-square between a restricted model, which regresses returns against the depth imbalance at the best bid and ask prices, and an unrestricted model, which regresses returns against the depth at the best bid and ask prices and the depth imbalance for levels 2 to 5 of the order book, we find that the unrestricted model outperforms the restricted model for over 85% of our stock-day regressions. In Appendix 3 we also show that depth imbalance causes a permanent, rather than transitory, price change. Specifically, we find that the difference in average returns between the trades with the most positive depth imbalances and trades with the most negative depth imbalances increases monotonically over a 200 trade horizon.

[Insert Table 2]

3.2 Depth imbalance and volume imbalances by trader type

Our previous analysis shows that in aggregate, traders buy aggressively when there is a large positive depth imbalance and sell aggressively when a large negative depth imbalance exists. To investigate whether the relation between depth imbalance and trading volumes differs by trader, for each trader type, we calculate the amount of buyer and seller initiated volumes, in each decile, as a percentage of total market volume. Panel A of Table 2 and Figure 2, Panels A to C present the results separately for *HFT*, *Institutions*, and *Retail*, respectively. Consistent with the full sample results from Figure 1, Panel B, we observe a general positive (negative) relationship between depth imbalance and aggressive buying (selling) for all trader types, indicating that all traders trade in the direction of the depth imbalance.

[Insert Figure 2]

Comparing between the panels, *HFT* are more successful than *Institutions* and *Retail* when depth imbalances are very positive or very negative. Figure 2, Panel A shows that *HFT* buy (sell) most aggressively when depth imbalance is the most positive (negative). For *Institutions* (Panel B) and *Retail* (Panel C), the percentage of buyer (seller) initiated trades increases with the size of the positive (negative) depth imbalance for moderate levels of imbalances. However, in the extremes (i.e., when depth imbalance is very positive or very negative), both *Institutions* and *Retail* are less successful at trading in the direction of the imbalance. *Retail*, in particular, is less successful in trading in the direction of the depth imbalance when the imbalance is very positive (for buys) or very negative (for sells).[18] This result is consistent with the notion that *HFTs* are more likely to trade opportunistically on order book information compared to *Institutions* and

[18] In results shown in Appendix 4, we find that *HFT* active and passive executions have larger price impacts and realized spreads, relative institutional and retail executions.

Retail, who tend to trade for reasons exogenous to the order book. Thus, we expect *HFTs* to trade most aggressively when large, profitable opportunities exist, which correspond to times of large depth imbalances.

To further assess whether *HFT* are more successful at trading on information contained in the order book, for each stock day, we calculate the executed volume imbalance that occurs at each *DI* decile, *j,* for each trader type, *T*. Specifically we calculate the volume imbalance as:

$$Volume\ imbalance_j^T = \frac{\sum_{k=1}^{n} BuyVolume_{k,j}^T - \sum_{k=1}^{n} SellVolume_{k,j}^T}{\sum_{k=1}^{n} BuyVolume_{k,j}^T + \sum_{k=1}^{n} SellVolume_{k,j}^T}$$

where $\sum_{k=1}^{n} BuyVolume_{k,j}^T (\sum_{k=1}^{n} SellVolume_{k,j}^T)$ is the total aggressive buying (selling) volume, i.e., volume from the submission of market or marketable buy (sell) limit orders, for depth imbalance decile, *j,* initiated by trader type, *T*.

Figure 2, Panel D, shows the relation between *Volume imbalance* and *DI* for our three trader types. Given that Figure 1, Panel A established that the size of *DI* predicts future returns, a steeper slope between *Volume imbalance* and *DI* indicates a trader is more focused on trading with the order book *DI*, ahead of future predicted price changes. Comparing the slopes for *HFT*, *Institutions* and *Retail*, our results show *HFT Volume imbalance* is most sensitive to *DI*, indicating that *HFT* are most successful at buying aggressively before an expected price rise and selling aggressively before an expected price fall, as predicted by *DI*. [19]

In Table 2, Panel B we test whether a statistically significant difference in *Volume imbalance* exists between our trader types for each *DI* decile. When the average *DI* is the most negative (*DI* = -0.367; Depth imbalance decile 0), the volume imbalance for *HFT* is -61.4%,

[19] In Appendix 5, we standardize the depth imbalance by average trade size and report similar findings.

while *Volume imbalance* for *Institutions* and *Retail* is only -19.7% and -5.8%, respectively. For the most positive *DI* decile (*DI* = 0.382; Depth imbalance decile 9), we observe positive volume imbalances for *HFT* (61.7%), *Institutions* (19.3%) and *Retail* (7.0%). Finally, when the order book is balanced, such that the bid depth is approximately equal to the ask depth, the difference in the *Volume imbalance* is less severe. For example, when *DI* is only 0.035 (decile 5), the volume imbalances range from -1% (*Retail*) to only 6.1% (*HFT*), consistent with our finding in Table 2, Panel A that HFT aggressive market share % is lower when the order book is moderately flat.

Importantly, we find that *HFT Volume imbalance* is always significantly below the institutional and retail *Volume imbalance* when a negative depth imbalance exists (i.e., there is selling pressure in the limit order book). In contrast, when buying pressure exists in the limit order book, volume imbalances are significantly larger for *HFT*, relative to *Institutions* and *Retail.* This result indicates that HFT are more successful at buying when the order book is predicting a future price rise and selling before expected future price declines. Further, comparing between *Institutions* and *Retail*, we find that *Institutions* are more strategic than *Retail* in trading with the imbalance in eight of the ten depth imbalance deciles.

Brogaard, Hendershott and Riordan (2014) show that HFTs demand liquidity in the direction of the limit order book imbalance.[20] Our results show this behavior is not unique to HFT and that all broker types attempt to trade in the direction of a stock's depth imbalance. This trading behavior is also consistent with the theoretical predictions of Parlour (1998). However, *HFT* are more successful at trading on information contained in the depth imbalance than the other trader types, and the difference is even more severe at extreme levels of order book

[20] Using a measure of trade imbalance, Malinova and Park (2016) also find that high frequency market makers demand liquidity in the direction of the imbalance.

imbalances. Arguably, this could be because *HFT* have a different trading objective function to *non-HFT*. While *non-HFT* are typically longer term investors, *HFTs* generally trade to make short term gains and thus, *DI* should be a more important component of *HFT* trading decisions.

We estimate the welfare effects of this behavior and find that HFT generate over 40% of their total profits through market order executions when the absolute size of the order book imbalance is in the most extreme decile. One implication for our results is that *HFT* could be crowding out *non-HFT* limit orders, especially when large depth imbalances exist, which we later test formally in Section 3.5.

In Table 3, we formally test the sensitivity of volume imbalances to depth imbalances for our trader types after controlling for trading volumes, stock and day fixed effects. Based on the findings from Anand and Venkataraman (2016), Kirilenko et al. (2016), Brogaard et al. (2017), van Kervel and Menkveld (2017) and Korajczk and Murphy (2016) we also examine stressful periods, by separately analyzing the results on high and low volatility days. Using all stock day observations, we estimate the following regression:

$$\begin{aligned} Volume\ imbalance_j^T = {} & \beta_0 + \beta_1 I(HFT)^T \times DI_j + \beta_2 I(Institutions)^T \times DI_j \\ & + \beta_3 I(HFT)^T + \beta_4 I(Institutions)^T \\ & + \beta_5 DI_j + \beta_6 Volume_j + \varepsilon_j^T \qquad (1) \end{aligned}$$

where $I(HFT)^T (I(Institutions)^T)$ is 1 if trader type, *T,* is *HFT* (*Institutions*), and 0 otherwise. DI_j is the average depth imbalance for the trades in the depth imbalance decile, *j,* and $Volume_j$ is the natural log of the total traded volume in the decile. We also include controls for stock and day fixed effects.

[Insert Table 3]

Table 3, Column 1 presents the results for all trading days in the sample. The main variables of interest are the interaction terms between the trader type and *DI*. A positive and significant coefficient implies that a trader's *Volume imbalance* is more sensitive to the level of *DI* in the order book. Consistent with our earlier results, we find that $I(HFT) \times DI$ is positive and significant indicating that relative to the other broker categories, *HFT* are more likely to submit buyer initiated trades when *DI* is larger. In contrast, $I(Institutions) \times DI$ is insignificant and *DI* is negative and significant indicating that *Institutions* and *Retail* trade less on order book information than *HFT*.

To investigate the effects of stock volatility on HFT trading behavior, for each stock we rank trading days into terciles (low, medium, and high) based on the daily stock volatility.[21] Table 3, Columns 2 and 3, present the results separately for low and high volatility days, respectively. For both low and high volatility days, we find that *HFT* use more order book information in their trading strategies than *Institutions* and *Retail*. Further, *DI* is negative and significant for all samples, indicating that *Retail* are less successful at trading in the direction of the expected future price movements. To test for differences in high and low volatility effects, in untabulated results we use a three way interaction between *I(HFT), DI*, and an indicator variable for high volatility days, and find that *HFT* volume imbalances are more sensitive to *DI* on volatile days (*p-value* = 0.004).

It is possible that some smaller proprietary HFT firms trade through institutional brokers. While it is difficult for these smaller HFT traders to change overall institutional volume imbalances, due to their small trade size, these traders could influence trading imbalances based

[21] We calculate daily volatility as the difference between the log of the intraday high ask price and the log of the intraday low bid price. In robustness tests, we calculate volatility as the standard deviation of 30-minute bid-ask midpoint returns and the results remain the same.

on the number of trades. To investigate this possibility, in Table 3, Columns 4-6, we replace the dependent variable from Equation (1) with trade imbalance, which is based on the number of buyer and seller initiated trades, rather than the volume of buyer and seller initiated trades. For *HFT*, our results are largely consistent with our findings based on volume imbalances. For *Institutions*, in contrast to our results based on volume imbalances, we also find that $I(Institutions) \times DI$ is positive and significant for the full sample (Column 4) and for the subsample of high volatility days (Column 6). This result reveals that some institutional investors behave like HFT and submit aggressive orders that capture information contained in the order book. In contrast, our volume results in columns 1-3 suggest that large trades from institutional investors are less likely to execute in the direction of the imbalance. This result could be driven by small HFT firms executing their strategies through larger institutional brokers, especially when the markets are highly volatile. Collectively, these results show that relative to non-HFT traders, HFT submit more buyer (seller) initiated orders when there are already low levels of liquidity available on the sell (buy) side of the order book.

3.3 Endogeneity

In the previous subsection, we show that HFT respond more strategically to the limit order book imbalance than institutional and retail investors. One possible explanation for our results is that HFT are simply better at responding to public news events. For example, a public news event could trigger fast traders to withdraw or trade against stale limit orders, resulting in the large depth imbalances we observe prior to market orders executing against the thin side of the order book. Thus, rather than HFT responding to the limit order book imbalance, they could be simply trading on public news events.

To test this possibility, for each stock, we divide stock days into information terciles based on the number of daily news articles obtained from Bloomberg. Our aim is to re-examine our full sample results from Figure 2, Panel D and Table 3 using these high and low information subsamples. If HFT were simply better at responding to public news, we would only observe a steeper slope for HFT, versus other trader types, for the high news day subsample. Figure 3 shows that our results are almost identical between the low news day and high news day subsamples. Comparing between the results for low news days (Panel A) and high news days (Panel B), we find that *HFT* act more strategically than *Institution*s and *Retail* regardless of the level of public news arrival.

[Insert Figure 3]

Next, we formally test the relationship between *Volume imbalance* (or *Trade imbalance*) and *DI* for each of our trader types. Table 4 shows the regressions results for our high and low information subsamples. Consistent with the full sample results from Table 3, we find that $I(HFT) \times DI$ is positive and significant indicating that *HFT Volume imbalance* (or *Trade imbalance*) is more sensitive to *DI*, relative to the other broker categories. Importantly, our results show that HFT respond to the limit order book imbalance, irrespective of whether the depth imbalance is triggered by a public news event.

[Insert Table 4]

3.4 Order submission strategies

In this section, we analyze how order submission strategies differ between investor categories. Specifically, we measure the adjusted depth imbalance, *Adjusted DI,* immediately before a trader submits, amends or cancels an order.[22] Since these orders include market orders and marketable limit orders, we are therefore also measuring *Adjusted DI* immediately before a trade execution takes place (i.e., when a market or marketable limit order is submitted or a limit order is executed against). We call each trade execution or the arrival of an order submission, amendment, or cancellation an order book "event", as it will change the limit order book, either by causing an execution to occur or by otherwise adding to or altering the limit order book.

These events are signed, in the sense that an event is classified as a "buy" if:

1) a market or marketable limit order is submitted to remove volume from the ask side ("buy aggressive execution"),
2) a limit order on the bid side is executed against by an incoming market or marketable limit order ("buy passive execution"),
3) a limit order is submitted to the bid side ("buy submission"), or
4) in the case of cancelations and amendments, volume is subtracted from the bid side (and so is a "buy cancelation" or "buy amendment").

We apply a similar classification for "sell" events.

As discussed earlier, to remove the effects of trade direction, we multiply *DI* by an indicator for whether the order or trade is a buy or sell, resulting in the *Adjusted DI* measure, which allows purchases and sales to be interpreted together. An *Adjusted DI* value of 0 indicates that the order book is balanced at the time of the event, while a high positive (negative) *Adjusted DI* value indicates a large depth imbalance in the same (opposite) direction as the order book

[22] We capture the depth imbalance before an event (i.e., submission, amendment, execution, or cancelation) to avoid capturing the depth effects of the event itself, but it represents the state of the order book at the time of the arrival of the new order. More generally, we can think of *DI* occurring at time t and the event occurring at $t + \varepsilon$, where $\varepsilon \to 0$. As such, we can interpret *DI* and the event as occurring at the same time. This is analogous to comparing trades to the quotes that existed immediately prior to the arrival of a market order; in fact, for n=1 depth, it is exactly analogous to looking at the depth-weighted bid-ask spread when an order arrives.

event.[23] For example, a negative *Adjusted DI* at the time of a buy trade indicates that a trader is buying when the ask depth exceeds the bid depth. Since a larger ask depth, relative to the bid depth, is associated with an expected future price fall, a negative *Adjusted DI* at the time of a buy trade indicates that a trader is buying before an expected future price fall. Thus, a strategic trader who uses information contained in the order book should execute trades when *Adjusted DI* is highly positive and cancel or amend orders when *Adjusted DI* is low or negative.

To investigate the order submission behavior of HFT, we use a multinomial logistic regression model to assess the probability of each order book event based on the prevailing market conditions in the limit order book. For each trader type, *T,* we estimate the following regression controlling for stock and day fixed effects:

$$\begin{aligned} OrderBookEvent_E^T = \ & \beta_0^T + \beta_1^T Adjusted\ DI_E^T + \beta_2^T Volatility \\ & + \beta_3^T Volume + \beta_4^T Price + \beta_5^T Spread + \varepsilon_E^T \end{aligned} \qquad (2)$$

where $OrderBookEvent_E^T$ is the dependent variable indicating one of five order book events, *E*: Aggressive execution, passive execution, limit order submission, amendment or cancelation. $Adjusted\ DI_E^T$ is the adjusted depth imbalance immediately before the order book event. All stock control variables are measured at the daily level. *Volatility* is the standard deviation of 30-minute mid-quote returns, *Volume* is the daily dollar volume, *Price* is the value-weighted average price and *Spread* is the time-weighted quoted spread. We estimate the model with limit order submission as the baseline category. Due to the extremely large number of order book events in our full sample, we select a subsample of 12 trading days for our multinomial logistic

[23] In the case of a buy execution (aggressive or passive), a positive *Adjusted DI* indicates that the bid depth exceeds the ask depth at the time of trade. For a limit order cancelation, a positive *Adjusted DI* indicates that the trader is cancelling an order from the thick side of the order book.

analysis. To form the sample trading days, for each of the six months in the sample, we select the first trading day of the month and the 15th day of the month (if this day falls on a non-trading day, we select the next trading day), resulting in a dataset of over 35 million observations.

[Insert Table 5]

Table 5, Panel A presents the results for *HFT*. Consistent with our strategic trading hypothesis, we find a strong positive association between HFT aggressive executions and large depth imbalances. In contrast, HFT cancelations or amendments occur when the order book imbalance is much smaller. Table 5, Columns 1 and 2, show that *Adjusted DI* is 1.075 and 0.467 for aggressive and passive executions, respectively, indicating that on average, trades take place when the depth imbalance is larger than the depth imbalance at the time of a limit order submission, which is the baseline category for our logistic regressions. Additionally, the difference between the coefficients on the aggressive and passive executions is statistically different at the 1% level and suggests that HFTs submit market orders when the order imbalance is larger and future prices will move in the direction of their trade, while their resting limit orders, which still execute under favorable order imbalances, do so when the order imbalance is smaller. Further, we find that for amendments and cancelations, *Adjusted DI* is negative and significant (-0.103 and -0.398, respectively), indicating HFT are quick to amend or cancel orders when the depth imbalance becomes less favorable to trade. In doing so, HFT remove stale limit orders before these orders can be picked off the order book by other traders.

Table 5, Panels B and C show that *Institutions* and *Retail* are less strategic in their order placement strategies. While *Institutions* are more likely to submit aggressive market orders when the *Adjusted DI* is large (coefficient of 0.511), both *Institutions* and *Retail* are more likely to receive a limit order execution when *Adjusted DI* is lower (coefficients of -0.381 and -0.360, respectively), relative to *Adjusted DI* at the time of order submission. Together, these results indicate that institutions and retail fail to cancel their resting limit orders when the depth imbalance moves in an unfavorable direction, meaning that their stale orders are more likely to be picked off the limit order book. These results are supported by the results in Appendix 4, where we show that HFT successfully trade on information contained in the limit order book, while institutions and retails do not. By comparing the bid-ask midpoint price 10 trades into the future with the current bid-ask midpoint, we show that HFTs consistently buy before future price rises and sell before future price falls. On the other hand, institutional and retail limit orders are picked off the limit order book.

So far, our findings indicate that HFT actively monitor the order book and thus large depth imbalances, which are strong predictors of future price movements, should not persist in the order book for extended periods of time. Figure 4, Panel A shows the inverted U-shaped relationship between the size of a depth imbalance and the median time the limit order book remains within the depth imbalance decile. For this analysis, depth imbalance deciles are calculated based on all order book events. Thus, depth imbalances can change due to order submissions, amendments, and cancelations, as well as trade executions. We find that large imbalances (i.e., deciles 0 and 9) persist for a much shorter period of time than more balanced order books (i.e., deciles 4 and 5). Our results indicate that large buy (sell) imbalances remain for a median time of only 34 milliseconds (121 milliseconds), while an order book that is relatively balanced remains in the same decile for a median time of approximately 1.4 seconds, more than 40 times longer. Given the rapid changes in the state of the order book, it is likely that only HFT are able to participate effectively in the extreme depth imbalance deciles. Consistent with this intuition, Figure 4, Panel B show that HFT are more active in the extreme depth imbalance deciles, relative to when the order book is more balanced, resulting in a strong U-shaped pattern. In contrast, *Institutions* and *Retail*, who are less able to compete on speed, reduce their activity when large imbalances exist in the order book (Figure 4, Panels C and D), resulting in a strong inverted U-shaped pattern.

[Insert Figure 4]

3.5 Volatility and HFT strategies

An important question to regulators is whether HFTs (and other market participants) act differently during times of stress, and in particular, whether the liquidity provided by HFT during normal periods may disappear during stressful periods when it is needed. Brogaard et al. (2017) examine HFT liquidity supply during extreme price moves, and find liquidity provision varies as to whether the large price movement is in just one stock or multiple stocks simultaneously, although data limitations in a highly fragmented market make these results unclear if they hold for the larger market as a whole. Anand and Venkataraman (2016) find that market makers reduce their participation on days with large order imbalances but increase their participation on days with high stock volatility. We test this further by investigating both active and passive HFT participation during times of high market volatility and examine the market conditions at the exact time of trade execution.

For this investigation, we divide each trading day into 30 minute intervals and for each interval, measure its volatility by taking the natural log of the high price divided by the low price during the period. For each stock, we then rank the 30 minute intervals into deciles based on its volatility. Decile 0 (9) contains the least (most) volatile periods. For each trader type, we also determine the amount of aggressive volume (i.e., due to market or marketable limit order submissions) and passive volume (i.e., limit order executions resulting from an incoming market or marketable limit order), as a percentage of their total volume, in each decile.

Figure 5, Panels A to C present the graphs of volatility against aggressive and passive volumes for *HFT*, *Institutions*, and *Retail*, respectively.

[Figure 5]

If HFT trade to decrease volatility, we expect them to supply more passive volume during times of high market uncertainty. In contrast, Figure 5, Panel A shows that HFT aggressive volume (black circles) increases while their passive volume (gray triangles) decreases as the market becomes more volatile. For *Institutions* in Figure 5, Panel B, we observe a fall in aggressive volumes (black circles are low and dropping) as volatility increases, from about 48.6% for the lowest volatility decile to about 46.6% for the highest, while HFTs jump from about 66.7% to 74.1%, which shows that *Institutions* are less active in the market in periods of high uncertainty. For *Retail*, with higher stock volatility, we observe a sharp decrease in aggressive volume (Figure 5, Panel C). For both *Institutional* and *Retail*, passive volume increases when the market becomes more volatile. This trading pattern could indicate that stale institutional and retail limit orders are picked off the limit order book by aggressive HFT orders during volatile periods and that retail and institutional brokers supply liquidity in times of need.

We test the relationships observed in Figure 5 more formally using the following regression model, controlling for stock and day fixed effects:

$$Aggressive\ volume\ \%_I^T = \beta_0 + \beta_1 Volatility_I + \beta_{10} Volume_I + \varepsilon \tag{3}$$

where the dependent variable, $Aggressive\ volume\ \%_I^T$, is the aggressive volume as a percentage of total aggressive and passive volume for each broker type, *T*, that is executed in the

30 minute interval, *I*. For each time interval, *I*, *Volatility* is the natural log of the high price divided by the low price and *Volume* is the total number of shares traded. We estimate Equation (3) separately for each trader type.

Table 6, Panel A shows that HFT *Aggressive volume %* increases with stock volatility (Column 1, coefficient of 447.6), consistent with our observations from Figure 4. This finding is similar across both large stocks and small stocks in Columns 2 and 3. In contrast, for institutional and retail traders in Table 6, Panels B and C, we find that *Aggressive volume %* generally decreases when stock volatility increases (coefficients of -84.7 and -174.3, respectively). These results support Boehmer, Li and Saar (2017), who find that short-horizon directional HFT strategies are more active when a thin order book exists and markets are volatile. Taken together, our results show that HFT trade more aggressively in times of high market volatility.

Our results contrast Brogaard et al. (2017), who find that HFTs supply liquidity during extreme price movements for a single stock, using a subsample of Nasdaq trades. One possible explanation for their findings is that HFT could be providing liquidity on Nasdaq, which has a maker/taker pricing schedule, while taking liquidity from other markets.[24] Our results further reveal that it is important to investigate consolidated trading volumes when analyzing overall HFT behavior.

[Insert Table 6]

We continue this examination of differential actions by examining strategic behavior during high and low stress periods. We test whether HFT are more strategic in times of market volatility by investigating the relationship between *Adjusted DI* and stock volatility for each category (HFT, Institution, or Retail). As discussed earlier, *Adjusted DI* measures a trader's ability to condition their trades on information contained in the order book. Specifically, a

[24] Unlike U.S. market design, the Australian market does not have maker-taker pricing.

positive *Adjusted DI* indicates that a trade executes in the direction of a favorable imbalance. Thus, *Adjusted DI* is more positive for traders better able to trade in the direction of the imbalance, or capture predicted future price movements based on the shape of the order book. For each market volatility decile, we calculate the average *Adjusted DI* for both the aggressive and passive executions in the decile.

[Insert Figure 6]

Figure 6, Panels A and B, present the average *Adjusted DI* for aggressive and passive executions, respectively. Comparing between the trader types, we observe a large difference in trading behaviors. Notably, for both passive and aggressive executions, we observe a sharp increase in *Adjusted DI* for *HFT* as stock volatility increases.

In contrast, for *Institutions* and *Retail*, *Adjusted DI* is relatively flat across the volatility deciles for their aggressive executions (Figure 6, Panel A) while for their passive executions (Figure 6, Panel B), *Adjusted DI* decreases with rising market volatility. This finding supports our hypothesis that non-HFT limit orders are picked off the thin side of the order book, especially in times of high market uncertainty. When the market is volatile, limit orders from slower traders could potentially become stale, leaving more opportunities for faster, more sophisticated traders.

We test the relationship between *Adjusted DI* and stock volatility more formally using a regression framework, after controlling for stock and day fixed effects.

$$Adjusted\ DI_I^T = \beta_0 + \beta_1 I(HFT)^T \times Volatility_I + \beta_2 I(Institutions)^T \times Volatility_I + \beta_3 I(HFT)^T + \beta_4 I(Institutions)^T + \beta_5 Volatility_I + \beta_6 Volume_I + \varepsilon \quad (4)$$

where $Adjusted\ DI_I^T$ is the average *Adjusted DI* for trades by trader *T* in the 30 minute interval, *I*. $I(HFT)^T(I(Institutions)^T)$ is 1 if trader type, *T,* is *HFT* (*Institutions*) and 0 otherwise and $Volatility_I$ and $Volume_I$ are defined as in Equation (3).

We estimate the regression separately for aggressive and passive executions. Table 7, Column 1 presents the results for aggressive executions based on the full sample of stocks. As expected, we find for *HFT*, *Adjusted DI* is larger when the market is more volatile, indicating that HFT are more strategic when uncertainty exists. Comparing the results for large and small stocks (Table 7, Columns 2 and 3), we show that the relationship between *Adjusted DI* and volatility is driven by the large stock sample. This finding supports the notion that HFT are likely to benefit more from trading speed differentials when trading in larger, more liquid stocks, in which trading is typically faster.

[Insert Table 7]

Table 7, Columns 4 to 6 present the regression coefficients for passive executions. Consistent with our earlier results, we find that HFT are able to successfully implement limit order strategies, especially when the market is volatile. Relative to *Retail* passive orders, which is captured in the intercept coefficient, *HFT* passive orders execute with a larger *Adjusted DI*, when more stock volatility exists (the coefficient on $I(HFT) \times Volatility$ is 5.224). While *Institutions* exhibit similar trading behaviors to *HFT*, the magnitudes of the coefficients are significantly lower (the coefficient on $I(Institutions) \times Volatility$ is 1.621). This finding supports our hypothesis that HFT are more successful at monitoring their limit orders, in particular, when there is high stock volatility. One further implication of our results is that *HFT* aggressively pick off stale orders from the thin side of the order book, particularly when the market is already highly volatile.

3.6 Introduction of ITCH

Using a difference-in-difference framework, we exploit a natural experiment to investigate whether an increase in trading speed improves HFT skill. On April 2, 2012, the ASX implemented ASX ITCH, which increased market information access speeds for a monthly subscription fee. While subscribing to ASX ITCH is voluntary, and the identity of subscribing brokers is confidential, it is reasonable to assume that traders who are most speed sensitive will be the first to subscribe to the faster data feed. To leave sufficient time for implementation, the pre-ITCH period is the one-month period prior to April 2, 2012 (i.e., March 2, 2012 to March 30, 2012) and the post-ITCH period begins one week after the introduction of ITCH and ranges from April 9, 2012 to May 9, 2012.[25]

Given that HFT strategies are most likely to benefit from the faster trading speeds, we expect that *HFT Volume imbalance* is more sensitive to the level of *DI* after switching to ITCH. On the other hand, the slope of the relationship between *Volume imbalance* and *DI* is less affected for *Retail* and *Institutions*, who are less speed sensitive. To assess empirically whether ITCH affects trading behavior, we use a difference-in-difference framework and re-estimate Equation (1) after including two interaction terms, *I(Pre)* and *I(Post)*, which are indicator variables indicating whether the observation occurs on a trading day before or after the introduction of ITCH. The regression specification is now:

$$\begin{aligned} Volume\ imbalance_j^T = {} & \beta_0 + I(Pre)[\beta_1 I(HFT)^T \times DI_j + \beta_2 I(Institutions)^T \times DI_j + \beta_3 DI_j] \\ & + I(Post)[\beta_4 I(HFT)^T \times DI_j + \beta_5 I(Institutions)^T \times DI_j + \beta_6 DI_j] \\ & + \beta_7 I(HFT)^T + \beta_8 I(Insto)^T + \beta_9 Volume_j + \varepsilon_j^T \qquad (5) \end{aligned}$$

[25] In unreported results, we find a 3.4% increase in the order to trade ratio after the introduction of ITCH (p-value = 0.01).

We run the tests in Table 8 on buy volume imbalances and buy trade imbalances. Table 8, Columns 1 and 3, report the two sets of coefficients $\{\beta_1, \beta_2, \beta_3\}$ and $\{\beta_4, \beta_5, \beta_6\}$. Table 8, Column 4 reports the test of equality between β_1 and β_4 (i.e., $\beta_4 - \beta_1 = 0$), which indicates whether HFT strategies capture more information contained in the depth imbalance after the implementation of ITCH. Similarly, Table 8, Column 4 also reports the test of equality between β_2 and β_5, and β_3 and β_6, which tests whether institutional and retail strategies change as a result of ITCH, respectively. Lastly, β_7 to β_9 and the intercept term are reported in Table 8, Column 2.

[Insert Table 8]

Table 8, Columns 1 and 3 show that β_1 and β_4 are both positive and significant (0.942 and 1.064, respectively) indicating that *HFT* buy when *DI* is positive and sell when *DI* is negative in both the pre- and post-ITCH periods. Importantly, the estimate of β_4 is larger than β_1 and the *F*-test in Column 4 (*F*-test = 5.35) shows that the difference is statistically significant at the 5% level. This result indicates that for *HFT*, the slope of *Buy Volume imbalance* against *DI* is steeper in the post-ITCH period. As a group, *HFT* trade more strategically on information contained in the limit order book as their market information access speeds increase. In contrast, we cannot reject the null hypothesis that $\beta_2 = \beta_5$ and $\beta_3 = \beta_6$. Consistent with our expectations, we do not find evidence that *non-HFT*, who are less speed sensitive, trade more strategically after the adoption of ITCH.

In Table 8, Columns 5 to 7, we replace the dependent variable with buy trade imbalance. Similar to our results based on buy volume imbalances, we find that *HFT* trade more strategically post-ITCH (*F*-test = 7.71). For *Institutions*, we also find that their trading is more sensitive to *DI* after the implementation of ITCH (*F*-test = 5.27). Consistent with our earlier results, it is possible that some more speed sensitive institutional brokers also subscribed to ITCH, on behalf of their smaller HFT clients, to take advantage of the faster speeds. While it is difficult for a small HFT trader to influence average volume imbalances for the broker, we find that their number of buy and sell orders capture more information contained in depth imbalances after trading becomes faster. Based on trade imbalances, we find that *Retail*, who are less likely to compete on speed, trade less strategically post-ITCH. Specifically, we find that β_6 is negative and significant at -0.121 in the post-ITCH period.

So far, our results show that HFT are more successful at trading on information contained in the order book imbalance when their trading becomes faster. One implication for these results is that non-HFT traders could be crowded out of the order book as HFT compete to trade in the same direction as non-HFT. We investigate whether HFT have a crowding out effect by measuring the probability of execution, $P(Fill)$, for HFT and non-HFT limit orders.

$$P(Fill)^T = \frac{\sum TradeVolume^T}{\sum SubmitVolume^T}$$

where $\sum SubmitVolume^T$ is the total daily volume submitted to the top level of the limit order book by trader type, *T,* and $\sum TradeVolume^T$ is the total volume of orders submitted to the top of the order book by trader type, *T*, which is executed. For each stock, we measure $P(Fill)^T$ on a daily basis for each trader type. We then estimate the following regression model:

$$P(Fill)^T = \beta_0 + \beta_1 I(Non - HFT)^T \times I(Post) + \beta_2 I(Non - HFT)^T + \beta_3 I(Post) + \beta_4 Volatility + \beta_5 Volume + \beta_6 Price + \beta_7 Spread + \varepsilon^T \quad (6)$$

where $I(Non - HFT)^T$ is an indicator variable equal to 1 if trader type *T* is an institutional or retail trader, and 0 for an HFT trader and *I(Post)* is an indicator variable equal to 1 for the post-ITCH period, and 0 for the pre-ITCH period. All other control variables are measured on a daily basis. *Volatility* is the standard deviation of 30-minute mid-quote returns, *Volume* is the daily dollar volume, *Price* is the value-weighted average price and *Spread* is the time-weighted quoted spread. We also control for stock and day fixed effects.

[Insert Table 9]

Our main variable of interest is $I(Non - HFT)^T \times I(Post)$. If HFT are crowding out non-HFT limit orders from the order book, we expect a negative coefficient for $I(Non - HFT) \times I(Post)$, indicating that the probability of a non-HFT order receiving execution decreases when HFT become faster. Consistent with our expectations, Table 9, Column 1 shows that the probability of limit order execution falls for non-HFT after the introduction of ITCH, as the coefficient on $I(non - HFT) \times I(Post)$ is -0.037. In Table 9, Column 2, we separate non-HFT traders into *Institutions* and *Retail*. For both *Institutions* and *Retail*, we find the interaction term with *I(Post)* is negative and significant (-0.021 and -0.058, respectively), indicating that both *Institutions* and *Retail* are crowded out from the order book by HFT.

To further investigate the crowding out effect on non-HFT traders, in Table 9, Columns 3 to 6, we decompose *P(Fill)* into favorable and unfavorable fills. We define a favorable fill as an order execution when the limit order rests on the side of the order book with more depth immediately prior to the trade. From our earlier results, we show that a limit order trader benefits when an order executes with a lot of depth on the same side of the order book (i.e., favorable fill). On the other hand, if depth builds up on the opposite side of the order book, a limit order is likely to face adverse selection (i.e., unfavorable fill).

Our results show that the decrease in *P(Fill)* is driven by a fall in the volume of favorable executions. Comparing between Columns 3 and 5, we find that *P(Favorable fill)* falls for non-HFT after the implementation of ITCH, as *I(Non-HFT)xI(Post)* is -0.04, while *P(Unfavorable fill)* remains unchanged after the implementation of ITCH. Separating non-HFT into *Institutions* and *Retail* in Columns 4 and 6, we document similar findings. Specifically, the interaction terms with *I(Post)* is negative and significant at -0.029 and -0.051 for *Institutions* and *Retail*, respectively, in Column 4 and insignificant in Column 6. Thus, the likelihood of receiving a favorable limit order execution falls for both institutional and retail traders when HFT gain a speed advantage. Taken together, these results provide evidence that HFT are crowding out non-HFT limit orders from the order book, especially when trading becomes faster. These findings are consistent with Van Kervel and Menkveld (2017) and Korajczk and Murphy (2016), who report an increase in institutional trading costs increase when HFTs compete in the same direction with institutional orders.

4.0 Conclusion

The information channels through which HFT trade are relatively unknown. We present strong evidence of one such channel, which provides an explanation for many of the findings documented in the prior literature. We show that order book depth imbalances are strong predictors of future prices. HFT are highly sophisticated in monitoring order book imbalances, which allows them to trade ahead of these predicted price changes. At the same time, when the order book imbalance moves in an unfavorable direction, they are quick to cancel or amend orders that are at high risk of being picked off by other traders. However, this strategic HFT behavior has potential negative welfare effects on slower traders. A strategy using fast market orders to pick of stale limit orders when the size of the order imbalance is in the largest size decile contributes over 40% to overall HFT trading revenues.

HFT order placement strategies based on order book imbalances are particularly successful when the market is volatile. During times of high market volatility, the chance of an institutional or retail limit order becoming stale increases. We find that HFT trade more aggressively and are more successful at picking off stale orders from institutional and retail investors when the market is volatile.

Using the introduction of ITCH as a natural experiment, we find that HFT become even better at acting on information contained in the order book when their trading becomes faster. However, by competing for favorable trade execution in the same direction as the buying or selling pressure, HFT have a crowding out effect on non-HFT limit orders, which potentially increases non-HFT limit order execution costs.

Our results on HFT trading behavior have implications for market quality. Several studies show that HFT enhance market quality by improving the informational efficiency of stock prices (Carrion, 2013; Brogaard, Hendershott and Riordan, 2014; Brogaard, Hendershott and Riordan, 2017). On the other hand, HFT could increase transaction costs when they trade in the same direction of the institutional order flow (van Kervel and Menkveld, 2017; Korajcyzk and Murphy, 2016). We show that HFT increase price efficiency by trading in the direction of the order book imbalance, which is a strong predictor of future price movements. However, HFT can also use the information contained in the order book imbalance to detect institutional buying or selling pressure.

While earlier studies show that overall depth improves (Hasbrouck and Saar, 2013; Hendershott, Jones and Menkveld, 2011), when analyzing directional depth, we show that HFT supply liquidity to the order book, but only to the side where there is a lot of existing depth. In contrast, we find that HFT demand liquidity from the thin side of the order book, which is more prominent in times of high market volatility.

References

Aitken, M., Chen, H. and Foley, S. (2017) The impact of fragmentation, exchange fees and liquidity provision on market quality. *Journal of Empirical Finance,* 41**,** 140-160.

Anand, A. and Venkataraman, K. (2016) Market conditions, fragility, and the economics of market making. *Journal of Financial Economics,* 121**,** 327-349.

Angel, J., Harris, L. and Spatt, C. (2010) Equity trading in the 21st century. Marshall School of Business Working Paper No. FBE 09-10.

Biais, B., Foucault, T. and Moinas, S. (2015) Equilibrium fast trading. *Journal of Financial Economics,* 116**,** 292-313.

Boehmer, E., Fong, K. and Wu, J. (2015) International evidence on algorithmic trading. *SSRN Working Paper*.

Boehmer, E., Li, D. and Saar, G. (2017) The competitive landscape of high-frequency trading Firms. *SSRN Working Paper*.

Brogaard, J., Carrion, A., Moyaert, T., Riordan, R., Shkilko, A. and Sokolov, K. (2017) High-frequency trading and extreme price movements. *Journal of Financial Economics,* Forthcoming.

Brogaard, J., Hagströmer, B., Nordén, L. and Riordan, R. (2015) Trading fast and slow: Colocation and liquidity. *Review of Financial Studies,* 28**,** 3407-3443.

Brogaard, J., Hendershott, T. and Riordan, R. (2014) High-frequency trading and price discovery. *Review of Financial Studies,* 27**,** 2267-2306.

Brogaard, J., Hendershott, T. and Riordan, R. (2017) Price discovery without trading: Evidence from limit orders. *SSRN Working Paper*.

Cao, C., Hansch, O. and Wang, X. (2009) The information content of an open limit-order book. *Journal of Futures Markets,* 29**,** 16-41.

Carrion, A. (2013) Very fast money: High-frequency trading on the Nasdaq. *Journal of Financial Markets,* 16**,** 680-711.

Cespa, G. and Vives, X. (2017) High frequency trading and fragility. CESifo Working Paper Series No. 6279.

Chaboud, A. P., Chiquoine, B., Hjalmarsson, E. and Vega, C. (2014) Rise of the machines: Algorithmic trading in the foreign exchange market. *The Journal of Finance,* 69**,** 2045-2084.

Chordia, T., Roll, R. and Subrahmanyam, A. (2002) Order imbalance, liquidity, and market returns. *Journal of Financial Economics,* 65**,** 111-130.

Chordia, T. and Subrahmanyam, A. (2004) Order imbalance and individual stock returns: Theory and evidence. *Journal of Financial Economics,* 72**,** 485-518.

Comerton-Forde, C. and Putniņš, T. J. (2015) Dark trading and price discovery. *Journal of Financial Economics,* 118**,** 70-92.

Conrad, J., Wahal, S. and Xiang, J. (2015) High-frequency quoting, trading, and the efficiency of prices. *Journal of Financial Economics,* 116**,** 271-291.

Cont, R., Kukanov, A. and Stoikov, S. (2014) The price impact of order book events. *Journal of Financial Econometrics,* 12**,** 47-88.

Ellis, K., Michaely, R. and O'hara, M. (2000) The accuracy of trade classification rules: Evidence from nasdaq. *Journal of Financial and Quantitative Analysis,* 35**,** 529-551.

Foucault, T., Hombert, J. and Roşu, I. (2016) News trading and speed. *The Journal of Finance,* 71**,** 335-382.

Grossman, S., and Stiglitz, J. (1980) On the impossibility of informationally efficient markets. *The American Economic Review,* 70(3)**,** 393-408.

Hagströmer, B. and Nordén, L. (2013) The diversity of high-frequency traders. *Journal of Financial Markets,* 16**,** 741-770.

Harris, L. (2013) What to do about high-frequency trading. *Financial Analysts Journal,* 69**,** 6-9.

Hasbrouck, J. and Saar, G. (2013) Low-latency trading. *Journal of Financial Markets,* 16**,** 646-679.

Hendershott, T., Jones, C. M. and Menkveld, A. J. (2011) Does algorithmic trading improve liquidity? *The Journal of Finance,* 66**,** 1-33.

Hirschey, N. (2016) Do high-frequency traders anticipate buying and selling pressure? London Business School Working Paper.

Hoffmann, P. (2014) A dynamic limit order market with fast and slow traders. *Journal of Financial Economics,* 113**,** 156-169.

Jones, C. (2013) What do we know about high-frequency trading? Columbia Business School Research Paper No. 13-11.

Kirilenko, A., Kyle, A. S., Samadi, M. and Tuzun, T. (2017) The flash crash: High-frequency trading in an electronic market. *The Journal of Finance,* 72**,** 967-998.

Korajczyk, R. and Murphy, D. (2016) High frequency market making to large institutional trades. *SSRN Working Paper*.

Lee, C. M. C. and Ready, M. J. (1991) Inferring trade direction from intraday data. *The Journal of Finance,* 46**,** 733-746.

Li, W. (2014) High frequency trading with speed hierarchies. University of Maryland Working Paper.

Malinova, K. and Park, A. (2016) "Modern" market makers. University of Toronto Working Paper.

Malinova, K., Park, A. and Riordan, R. (2016) Taxing high frequency market making: Who pays the bill? University of Toronto Working Paper.

Menkveld, A. J. (2013) High frequency trading and the new market makers. *Journal of Financial Markets,* 16**,** 712-740.

Naes, R. and Skjeltorp, J. (2006) Order book characteristics and the volume-volatility relation: Empirical evidence from a limit order market. *Journal of Financial Markets,* 9**,** 408-432.

Parlour, C. (1998) Price Dynamics in Limit Order Markets. *The Review of Financial Studies*, 11, 789–816.

Ranaldo, A. (2004) Order aggressiveness in limit order book markets. *Journal of Financial Markets,* 7**,** 53-74.

Rosu, I. (2016) Fast and slow informed trading. HEC Paris Research Paper No. FIN-2015-1123.

Subrahmanyam, A. and Zheng, H. (2016) Limit order placement by high-frequency traders. *Borsa Istanbul Review,* 16**,** 185-209.

Upson, J., Johnson, H. and Mcinish, T. H. (2015) Order versus trades on the consolidated tape. *SSRN Working Paper*.

Van Kervel, V. (2015) Competition for order flow with fast and slow traders. *Review of Financial Studies,* 28**,** 2094-2127.

Van Kervel, V. and Menkveld, A. J. (2017) High-frequency trading around large institutional orders. *SSRN Working Paper*.

Yang, L. and Zhu, H. (2017) Back-running: Seeking and hiding fundamental information in order flows. Rotman School of Management Working Paper No. 2583915.

Yao, C. and Ye, M. (2017) Why trading speed matters: A tale of queue rationing under price controls. *Review of Financial Studies,* Forthcoming.

Ye, M. (2017) Who provides liquidity and when: An analysis of price vs. Speed competition on liquidity and welfare. University of Illinois at Urbana-Champaign Working Paper.

Appendix 1

To trade on the ASX, traders must submit their orders through an ASX market participant. Each ASX market participant has their own unique broker identifier, which we use to classify brokers as *HFT*, *Institutional* or *Retail*. Large, speed-sensitive proprietary HFT firms are typically registered ASX market participants with their own broker identifier. However, smaller proprietary HFT firms, who may not be able to absorb the costs associated with becoming an ASX market participant, could trade through institutional brokers and thus, *Institutions* could also contain some HFT activity.[26] Over our sample period, there are approximately 100 registered market participants with unique broker identifiers.

We use a three-step classification process to assign brokers, based on their primary trading activity. We start with the set of global HFT firms identified in van Kervel and Menkveld (2017); HFT firms from their study that also operate in the Australian market are classified as *HFT* in our sample. Next, we search through newspaper articles and the websites of individual brokers to identify the main trading activity or client base. Finally, we confirm our broker classifications through extensive conversations with regulators and industry participants from HFT firms, institutional and retail brokerage houses. Using this classification system, we assign approximately 10% of market participants as *HFT* and 20% of market participants as *Retail*. The remaining brokers, which could also contain some proprietary HFT or retail trading, are classified as *Institutional*.

[26] These costs include application fees, annual fees for ASX membership, and monthly fees for connection to ASX trading platforms. ASX market participants also have additional regulatory, compliance and risk management obligations.

Appendix 2

In this Appendix, we show examples of how *DI* is constructed. Figure A1 shows two stylized diagrams of the limit order book. In the first example, the bid depth exceeds the ask depth and in the second example, the ask depth exceeds the bid depth. In Table A1, we show the calculation for *DI*, which is defined as:

$$DI_t = \frac{\sum_{i=1}^{n} DepthBid_{i,t} - \sum_{i=1}^{n} DepthAsk_{i,t}}{\sum_{i=1}^{n} DepthBid_{i,t} + \sum_{i=1}^{n} DepthAsk_{i,t}}$$

where $\sum_{i=1}^{n} DepthBid_{i,t}$ ($\sum_{i=1}^{n} DepthAsk_{i,t}$) is the depth available at the top *n* bid (ask) price levels immediately before the order book event, *t*.

Fig. A1. Limit order book examples
Each black rectangle represents a limit buy order of 100 shares and each grey rectangle represents a limit sell order of 100 shares.

Table A1. Calculation of depth imbalance (*DI*) based on examples in Figure A1.

	Example 1	Example 2
DI (5 levels, *n* = 5)	$\frac{1100-500}{1100+500} = 0.375$	$\frac{500-1100}{500+1100} = -0.375$
DI (1 level, *n* = 1)	$\frac{300-100}{300+100} = 0.5$	$\frac{100-300}{100+300} = -0.5$

Appendix 3

In this Appendix, we investigate the information content of resting limit orders behind the best bid and offer prices. To determine the incremental information content of resting limit orders at levels 2 to 5 of the order book, we estimate a restricted model, which only contains the depth imbalance for the best bid and offer, and an unrestricted model, which contains the *DI* for the best bid and offer and for levels 2 to 5 of the limit order book. For each stock and day, we perform the following regressions:

Restricted model: $Return = \beta_0 + \beta_1 DI_{TopLevel} + \varepsilon$

Unrestricted model: $Return = \beta_0 + \beta_1 DI_{TopLevel} + \beta_2 DI_{Levels2-5} + \varepsilon$

Return is calculated as the log of the difference between the bid-ask midpoint 10 trades in the future and the midpoint price just prior to the trade. *DI* is the depth imbalance immediately before the trade, which is calculated as the difference between the volumes available at the bid and ask prices as a proportion of the total volume available at the bid and ask prices. We calculate DI for the top level of the order book ($DI_{TopLevel}$) as well as for levels 2 to 5 of the limit order book ($DI_{Levels2-5}$).

Table A2 summarizes the adjusted R-square for the restricted and unrestricted models. If $DI_{Levels2-5}$ adds incremental information about the future price movements, we expect a higher adjusted R-square for the unrestricted model, relative to the restricted model. For over 85% of our regressions, the F-test is significant at the 1% level, indicating that $DI_{Levels2-5}$ adds additional explanatory power. This result indicates that limit orders behind the best bid and offer also contains information on future stock returns.

Table A2. Adjusted R-square for restricted and unrestricted model.

	Adjusted R-square		% with F-test significant at 1%
	Restricted model	Unrestricted model	
Mean	12.02%	13.48%	85.64%
Median	10.96%	12.45%	

A. Depth imbalance and stock returns over time

Figure A2 plots the difference in average returns between the trades with the most positive depth imbalances and trades with the most negative depth imbalances. We rank trades into deciles based on the depth imbalance immediately before the trade for each stock day. For each transaction, we also calculate future returns by comparing the midpoint of the best bid and ask prices at the time of the trade with the bid-ask midpoint 5, 10, 20, 50, 100, or 200 trades in the future. In Figure A2, each observation represents the return differential (*Differential*), which is the difference between average return of trades in the most positive depth imbalance decile (decile 9) and the average return of trades in the most negative depth imbalance decile (decile 0), for each return interval. Thus, a positive *Differential* indicates that the depth imbalance has predictive power over future returns while a *Differential* of 0 indicates that depth imbalance has no predictive power.

Figure A2 shows that the depth imbalance has predictive power over all time horizons. The *Differential* is 8.3 bps after 5 trades and increases monotonically to 11.9 bps after 200 trades, indicating that there is a permanent shift in midpoint price. Further, the information contained in the depth imbalance is fully reflected in prices after approximately 100 trades.

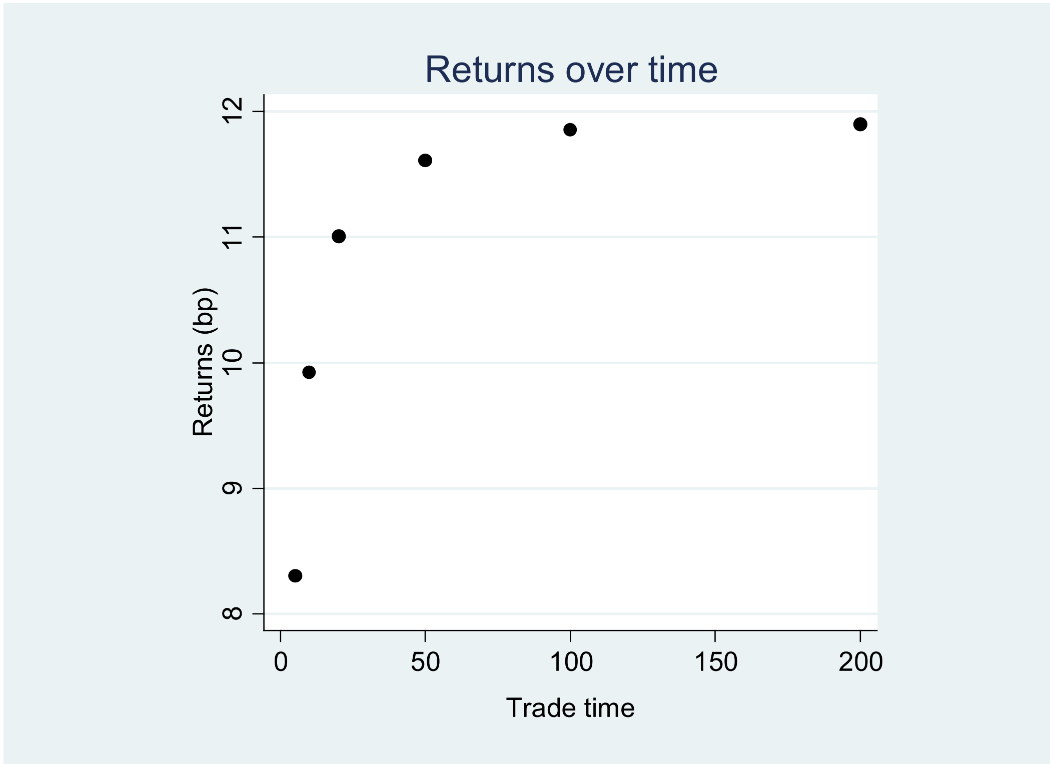

Fig. A2. Depth imbalance and stock returns over time

Fig. A2 shows the relation between depth imbalance and stock returns over time. Depth imbalance is calculated as the difference between the depths available at the five best bid and ask prices, scaled by total depth available at these price levels, immediately before each trade. For each stock day, we rank trades into 10 depth imbalance deciles. Trades with the most negative depth imbalances (i.e., bid depth << ask depth) are categorized as decile 0 and trades with the most positive depth imbalances (i.e., bid depth >> ask depth) are in decile 9. We calculate returns by comparing the current midpoint of the best bid and ask prices with the midpoint price 5, 10, 20, 50, 100, or 200 trades in the future. Each observation represents the difference in average returns between trades in decile 9 and trades in decile 0.

Appendix 4

In this Appendix, we investigate whether *HFT* successfully use their order placement strategies to buy before future price rises and sell before future price falls. Specifically, for each active (i.e., market or marketable limit order) and passive execution, we calculate the price impact and realized spread as:

$$Price\ Impact = q_t \times \frac{m_{t+10} - m_t}{m_t}$$

$$Realized\ spread = q_t \times \frac{p_t - m_{t+10}}{m_t}$$

where m_t is the bid ask midpoint at the time when the current trade takes place, m_{t+10} is the bid ask midpoint in 10 trades' time, and p_t is the trade price. Because each execution consists of both an active and passive side, we calculate the price impact and realized spread for both the active and passive parties to each trade. For active executions, the buy sell indicator, q_t, equals 1 (-1) if the trade is buyer (seller) initiated. For passive executions, q_t equals 1 (-1) if the existing limit order is resting on the bid (ask) side of the limit order book.

For active executions, we find that HFT outperform both institutional and retail market orders. The equally weighted average price impact for HFT is 7.38 basis points while the average price impact for institutional and retail is only 3.05 and 3.54 basis points, respectively. Consistent with this result, we also find that realized spreads for HFT active executions are larger than *Institutions* and *Retail*.[27]

Importantly, our results show that HFT are able to successfully time their passive executions. On average HFTs earn 0.46 basis points per limit order execution, based on comparing the current bid-ask midpoint price with the bid-ask midpoint 10 trades into the future.

[27] Value weighted realized spreads for active executions are positive for *HFT* and *Institutions* (0.67 and 0.25 basis points, respectively) but remain negative for *Retail* (-3.10 basis points).

This return increases to 1.04 basis points when we include the transaction price (i.e., a buy limit order transacts at the bid price, rather than at the bid-ask midpoint). In contrast, the price impact and realized spreads of institutional and retail limit orders are consistently negative, indicating that their orders are picked off the limit order book. Taken together, these results indicate that *HFT* have superior market timing abilities, relative to *Institutions* and *Retail*.

Table A3. Price impact and realized spreads by trader type

	Active executions (bps)		Passive executions (bps)	
	Price impact	Realized spread	Price impact	Realized spread
HFT	7.383	-0.768	0.460	1.038
Institutions	3.049	-5.120	-2.665	-4.211
Retail	3.539	-6.114	-1.670	-20.76
HFT - Institutions	4.334 ***	4.352 ***	3.125 ***	5.249 ***
HFT - Retail	3.844 ***	5.346 ***	2.131 ***	21.80 ***
Institutions - Retail	-0.490 ***	0.995 ***	-0.995 ***	16.55 ***

Appendix 5

DI is designed to capture the buying or selling pressure contained in the limit order book. *DI* compares the depth on the bid side *relative* to the ask side and accordingly, is bounded between 1 and -1. A value close to 1 (-1) reflects that the bid (ask) depth is substantially greater than the ask (bid) depth, while a value of 0 reflects that the bid depth is the same as the ask depth. One potential shortcoming of this measure is that we are comparing the bid and ask depth *relative* to each other while ignoring their magnitude *relative* to incoming orders. Consequently, we could compute the same value of *DI* for two distinct scenarios. For example, in the first scenario, the bid and ask depth is 1 and 10, respectively, whereas in the second scenario the bid and ask depth is 100 and 1000, respectively. In both of these scenarios, *DI* is -0.82, indicating a strong negative depth imbalance. However, the two scenarios do not necessarily reflect the same selling pressure.

If the average trade size is 200 shares, then in scenario one, the bid or ask depths are small *relative* to the average incoming trade size. As such, there is relatively little buying or selling pressure on both sides of the book *relative* to trade size. In contrast, in scenario two, the depth on the ask side is 5 times the average trade size, whereas the depth on the bid side is half the average trade size. When considering the incoming trade size, there appears to be more selling pressure on the ask side of the book relative to the buying pressure on the bid side of the order book.

For robustness, we modify *DI* to capture the effect of trade size. Specifically, rather than standardizing the depth imbalance between the bid and ask sides by the total depth available, we standardize the depth imbalance by the average trade size of the day. As such, our modified *DI* is now defined as:

$$ModifiedDI_t = \frac{\sum_{i=1}^{n} DepthBid_{i,t} - \sum_{i=1}^{n} DepthAsk_{i,t}}{TradeSize}$$

where $\sum_{i=1}^{n} DepthBid_{i,t}$ ($\sum_{i=1}^{n} DepthAsk_{i,t}$) is the total depth available at the top n bid (ask) price levels immediately before the order book event, t and *TradeSize* is the average trade size on the day.

Figure A3 shows the relation between *Volume imbalance* and the *ModifiedDI* for our three trader types. We observe that *HFT* are most successful at buying aggressively before an expected price rise and selling aggressively before an expected price fall. Importantly, our results using *ModifiedDI* are indistinguishable from our main results in Figure 2, Panel D.

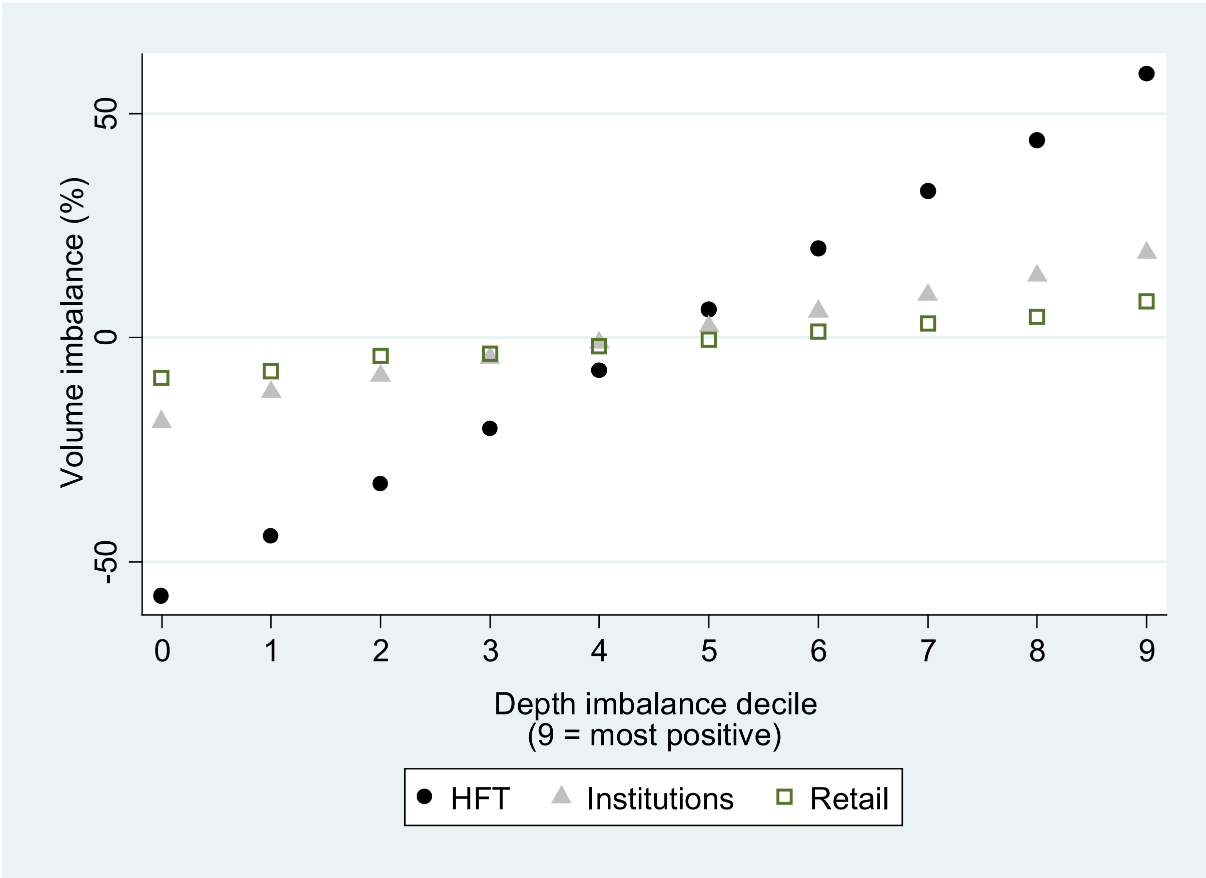

Fig. A3. Depth imbalance and volume imbalance by broker category

Fig. A3 shows the relationship between depth imbalance and trading volumes for each broker category. Depth imbalance is calculated as the difference between the depths available at the five best bid and ask prices, scaled by the average trade size on the trading day. For each stock day, we rank trades into 10 depth imbalance deciles. Trades with the most negative depth imbalances (i.e., bid depth << ask depth) are categorized as decile 0 and trades with the most positive depth imbalances (i.e., bid depth >> ask depth) are in decile 9. The graph shows the volume imbalance (i.e., (Buys-Sells)/(Buys + Sells)) for each broker type and depth imbalance decile.

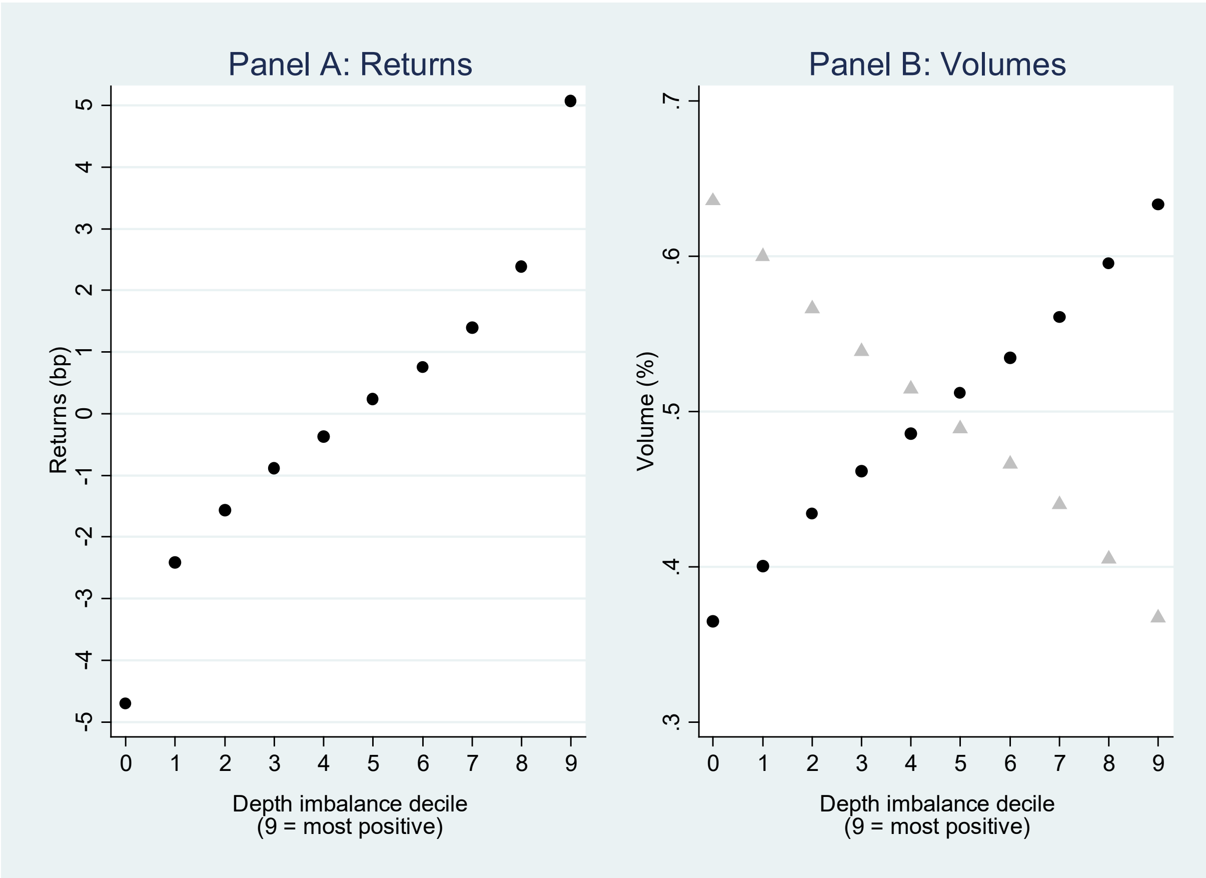

Fig. 1. Fig. 1 shows the relationship between depth imbalance, returns (Panel A) and volumes (Panel B). Depth imbalance is calculated as the difference between the depths available at the five best bid and ask prices, scaled by total depth available at these price levels, immediately before each trade. For each stock day, we rank trades into 10 depth imbalance deciles. Trades with the most negative depth imbalances (i.e., bid depth << ask depth) are categorized as decile 0 and trades with the most positive depth imbalances (i.e., bid depth >> ask depth) are in decile 9. In Panel A, we calculate returns by comparing the current midpoint of the best bid and ask prices with the midpoint price 10 trades in the future. In Panel B, the black circles (gray triangles) represent the average percentage of buyer (seller) initiated volume, relative to total trade volume, for each depth imbalance decile.

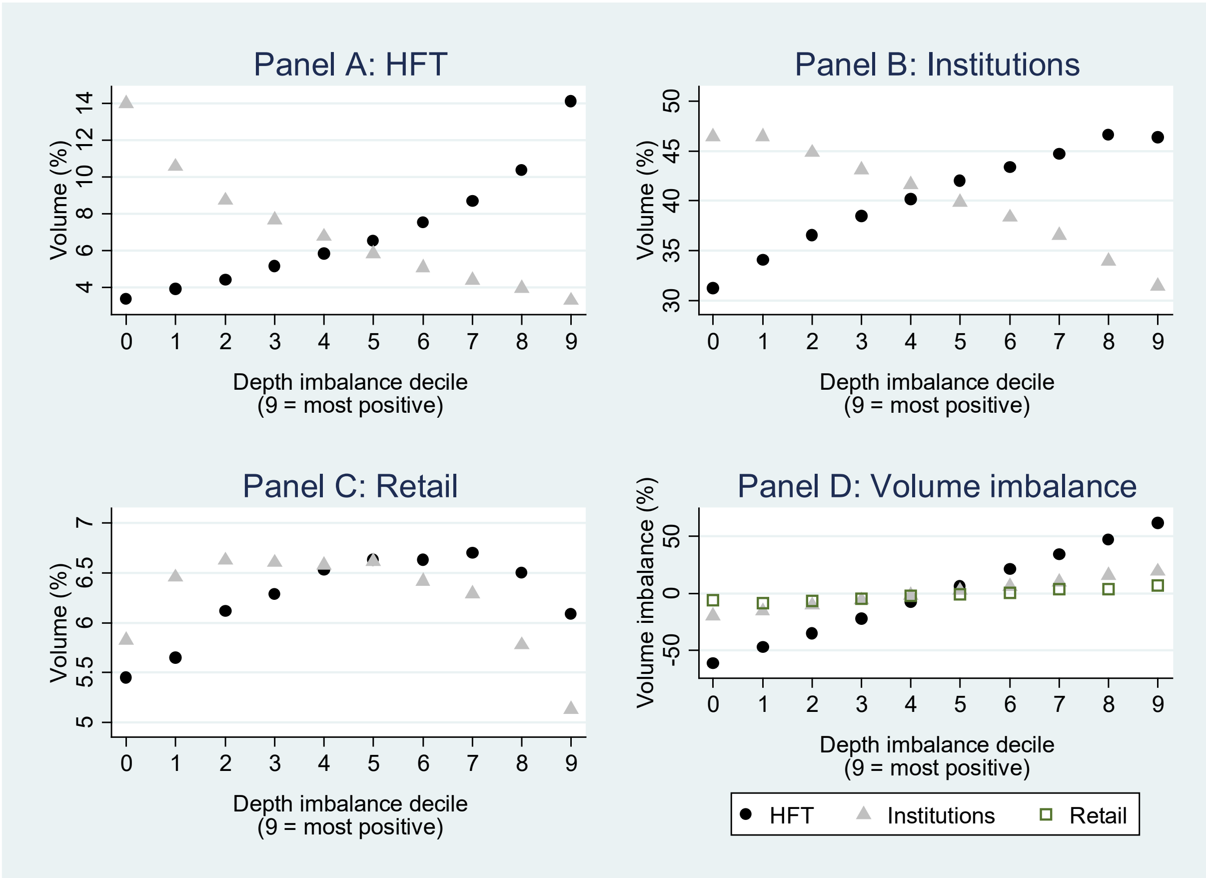

Fig 2. Fig. 2 shows the relationship between depth imbalance and trading volumes for each broker category. Depth imbalance is calculated as the difference between the depths available at the five best bid and ask prices, scaled by total depth available at these price levels, immediately before each trade. For each stock day, we rank trades into 10 depth imbalance deciles. Trades with the most negative depth imbalances (i.e., bid depth << ask depth) are categorized as decile 0 and trades with the most positive depth imbalances (i.e., bid depth >> ask depth) are in decile 9. Panels A-C present the results for *HFT*, *Institutions*, and *Retail*, respectively. The black circles (gray triangles) represent the average percentage of buyer (seller) initiated volume, relative to total trade volume, for each depth imbalance decile and broker type. Panel D shows the volume imbalance (i.e., (Buys-Sells)/(Buys + Sells)) for each broker type and depth imbalance decile.

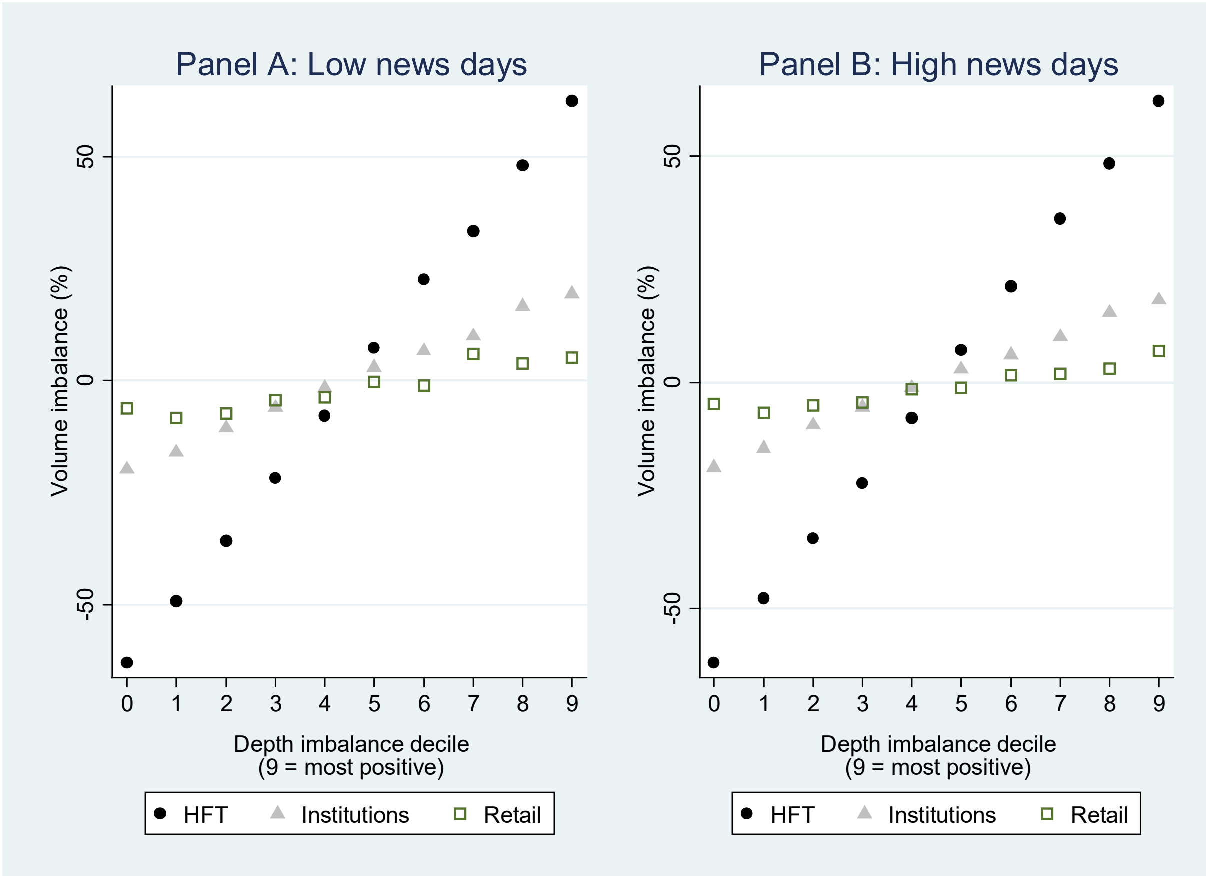

Fig. 3. Fig. 3 shows the relationship between depth imbalance and volume imbalance for each broker category based on subsamples of low and high news days. Depth imbalance is calculated as the difference between the depths available at the five best bid and ask prices, scaled by total depth available at these price levels, immediately before each trade. For each stock day, we rank trades into 10 depth imbalance deciles. Trades with the most negative depth imbalances (i.e., bid depth << ask depth) are categorized as decile 0 and trades with the most positive depth imbalances (i.e., bid depth >> ask depth) are in decile 9. Volume imbalance is calculated as (Buys-Sells)/(Buys + Sells) for each trader type. Panel A (Panel B) presents the results for the top (bottom) tercile based on the number of daily Bloomberg news articles for each stock.

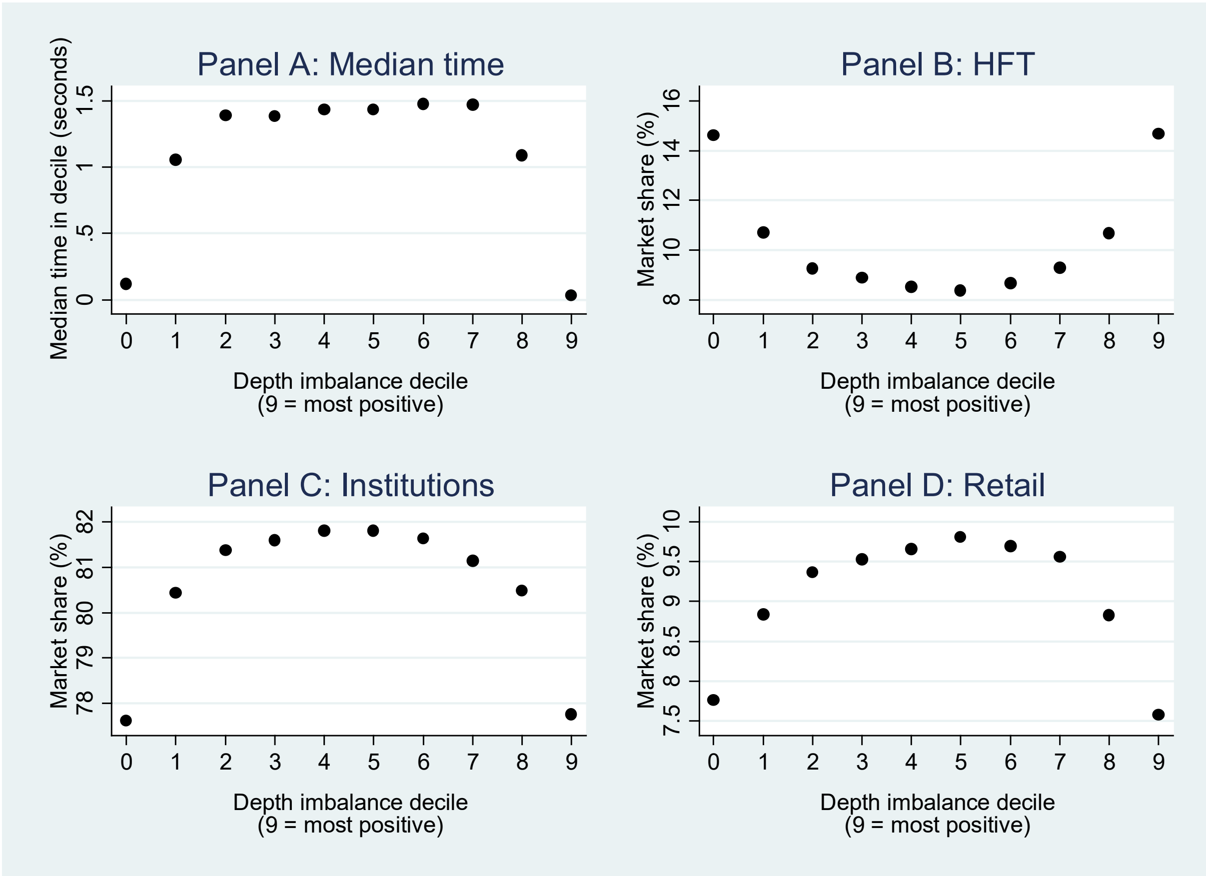

Fig. 4. Fig. 4 shows the relationship between depth imbalance and the median time the order book remains in each depth imbalance decile (Panel A) and market shares for *HFT*, *Institutions* and *Retail* (Panels B to D). In Panel A, we compute the depth imbalance for all order book events (submission, cancelation, amend, trade) and rank the depth imbalances into deciles. For all observations, we determine the time the order book remains within the same depth imbalance decile. Panel A plots the median time (seconds) the order book remains in each depth imbalance decile. For Panels B to D, we rank trades into 10 depth imbalance deciles for each stock day. Trades with the most negative depth imbalances (i.e., bid depth << ask depth) are categorized as decile 0 and trades with the most positive depth imbalances (i.e., bid depth >> ask depth) are in decile 9. Panels B to D plot the average daily market share per stock (based on aggressive orders) for each broker type and depth imbalance decile.

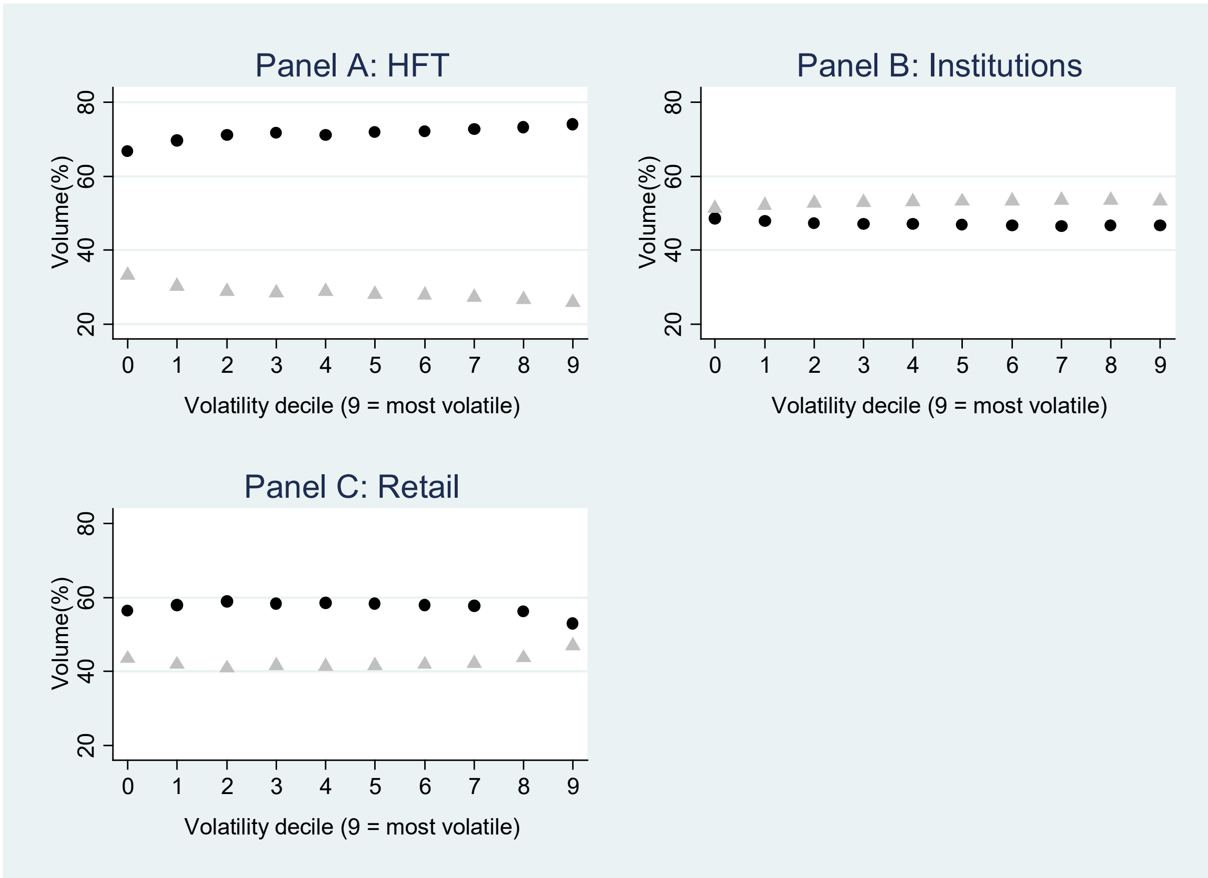

Fig. 5. Fig. 5 shows the relationship between volatility and trading volume for each broker category. Volatility is calculated as the log of ratio of the high to the low price over each 30-minute trading interval. For each stock, trading intervals are then ranked into volatility deciles. Panels A-C present the results for *HFT*, *Institutions*, and *Retail*, respectively. For each broker category, we calculate the percentage of aggressive (black circles) and passive (gray triangles) trading volume, relative to total aggressive and passive trading volume, for each volatility decile.

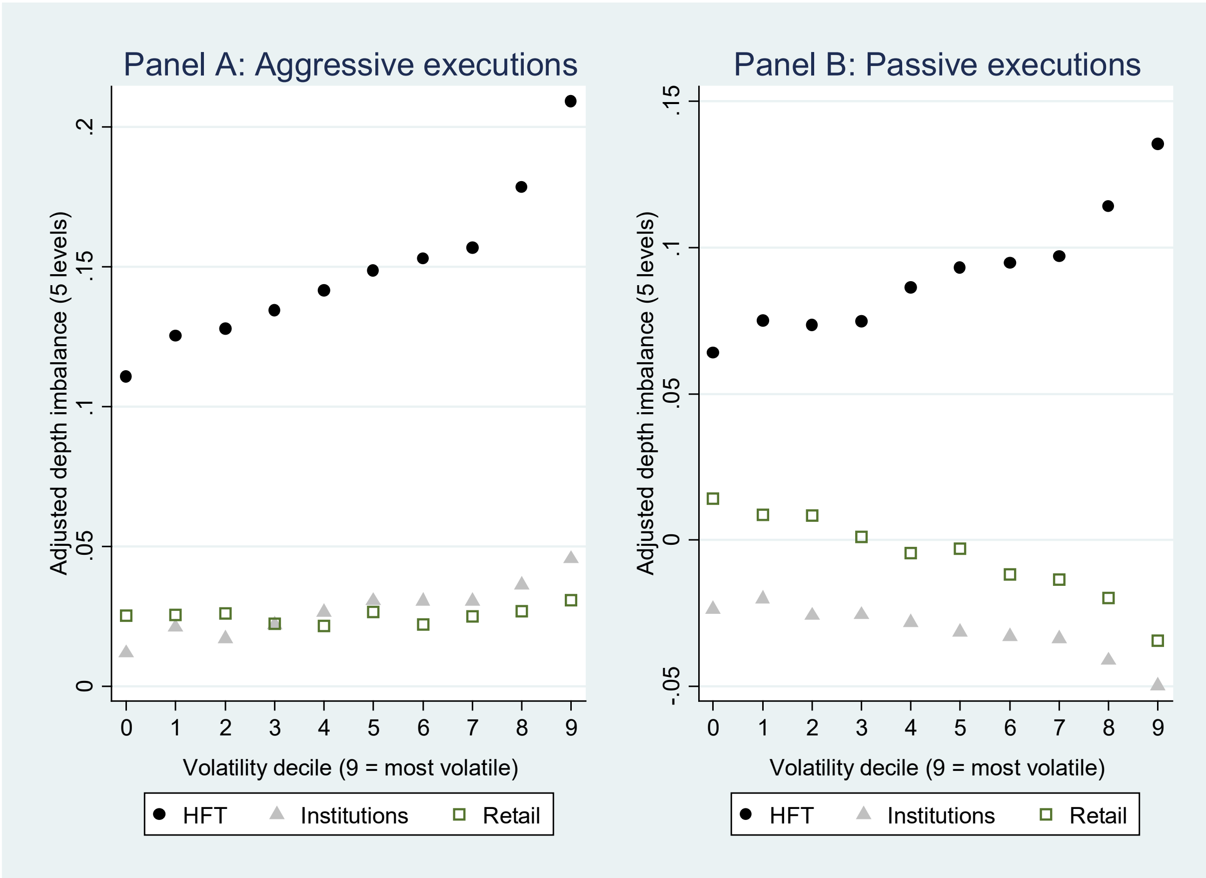

Fig. 6. Fig. 6 shows the relationship between volatility and adjusted depth imbalance for each broker category. Volatility is calculated as the log of ratio of the high to the low price over each 30-minute trading interval. Trading intervals are then ranked into volatility deciles. For each broker category, we calculate the depth imbalance immediately before each aggressive (Panel A) or passive (Panel B) trade execution. Depth imbalance is calculated as the difference between the depths available at the five best bid and ask prices, scaled by total depth available at these price levels, immediately before each trade. We multiply depth imbalance by a buy or sell indicator so that buys and sells can be interpreted together. Depth imbalances are then averaged over each volatility decile by broker category.

Table 1
Summary statistics

Table 1, Panel A reports statistics for the 94 stocks that remain in the ASX 100 index for the period January 3, 2012 to June 30, 2012. *Market capitalization* is the stock's market capitalization on January 3, 2012. *Dollar volume* is the average daily dollar volume in AUD. *Number of trades* is the average daily number of transactions. *Price* is the average trade price in AUD. *Volatility* is the difference between the log of the intraday high ask price and the log of the intraday low bid price. *Spread* is the time weighted average difference between the best bid and offer prices in AUD cents. The broker associated with each order book event is classified into three types: proprietary HFT (*HFT*), institutional (*Institutions*), or retail (*Retail*). Panel B reports the trading characteristics for each broker type. Panel C reports the average adjusted depth imbalance (*Adjusted DI*) for each trader type. For each order book event, *Adjusted DI* is calculated as:

$$Adjusted\ DI_t = q_t \times \frac{\sum_{i=1}^{n} DepthBid_{i,t} - \sum_{i=1}^{n} DepthAsk_{i,t}}{\sum_{i=1}^{n} DepthBid_{i,t} + \sum_{i=1}^{n} DepthAsk_{i,t}}$$

where $\sum_{i=1}^{5} DepthBid_{i,t}$ ($\sum_{i=1}^{5} DepthAsk_{i,t}$) is the depth available at the top *5* bid (ask) price levels immediately before the order book event, *t*. q_t is an indicator variable equal to 1 for buys and -1 for sells.

Panel A: Stock characteristics

	Mean	Std.dev.	Q1	Median	Q3
Market capitalization (bil.)	13.52	22.77	2.844	10.00	114.8
Dollar volume (mil.)	27.80	47.72	5.295	10.99	25.60
Ntrades	2,264	1,838	1,112	1,659	2,701
Price (dollars)	11.43	12.81	2.954	5.974	14.86
Volatility	2.105	1.188	1.360	1.861	2.541
Spread (cents)	1.021	0.376	0.926	0.998	1.115

	All	HFT	Institutions	Retail
Panel B: Trader characteristics				
Average daily submissions	5,556	1,025	15,185	440
Average daily cancelations	1,837	478	4,921	57
Average daily executions (active)	600	260	1,446	92
Average daily executions (passive)	1,427	319	3,796	162
Median submission to cancel time	413	188	233	3,200
Panel C: Adjusted depth imbalance				
Executions (active)	0.069	0.156	0.027	0.025
Executions (passive)	0.015	0.083	-0.029	-0.009
Submissions	0.034	0.063	-0.004	0.043
Amendments	0.019	0.040	-0.003	0.023
Cancelations	0.016	0.017	0.002	0.028

Table 2

Market shares, volume and depth imbalances by trader type

Table 2 reports the market share (Panel A) and volume imbalance (Panel B) for each depth imbalance (*DI*) decile by trader type. For each stock, we rank trades for each stock into *DI* deciles. For every trade, *DI* is calculated as:

$$DI_t = \frac{\sum_{i=1}^{n} DepthBid_{i,t} - \sum_{i=1}^{n} DepthAsk_{i,t}}{\sum_{i=1}^{n} DepthBid_{i,t} + \sum_{i=1}^{n} DepthAsk_{i,t}}$$

where $\sum_{i=1}^{n} DepthBid_{i,t}$ ($\sum_{i=1}^{n} DepthAsk_{i,t}$)is the depth available at the top *5* bid (ask) price levels immediately a trade, *t*. Panel A reports the average total volume, and for each trader type, the aggressive market share percentage across the *DI* deciles. In Panel B, for each *DI* decile, we calculate *Volume imbalance* as:

$$Volume\ imbalance_j^T = \frac{\sum_{k=1}^{n} BuyVolume_{k,j}^T - \sum_{k=1}^{n} SellVolume_{k,j}^T}{\sum_{k=1}^{n} BuyVolume_{k,j}^T + \sum_{k=1}^{n} SellVolume_{k,j}^T}$$

where $\sum_{k=1}^{n} BuyVolume_{k,j}^T$($\sum_{k=1}^{n} SellVolume_{k,j}^T$) is the total aggressive buying (selling) volume for depth imbalance decile, *j*. We also use a *t*-test to test for the differences in *Volume imbalance* means between the trader types. *** indicates a 1% significance level.

Depth imbalance decile	Depth imbalance	Avg. total volume	HFT %	Institutions %	Retail %
Panel A: Aggressive market share %					
0 (most negative)	-0.367	349,745	14.62	77.62	7.76
1	-0.210	286,940	10.73	80.43	8.84
2	-0.132	279,059	9.27	81.37	9.36
3	-0.071	273,599	8.87	81.60	9.53
4	-0.017	268,912	8.52	81.81	9.66
5	0.035	270,881	8.38	81.81	9.81
6	0.090	278,228	8.68	81.63	9.70
7	0.150	277,294	9.30	81.14	9.56
8	0.228	296,959	10.68	80.49	8.83
9 (most positive)	0.382	346,089	14.67	77.77	7.57

Depth imbalance decile	HFT	Institutions	Retail	HFT vs. Institutions		HFT vs. Retail		Institutions vs. Retail	
				Panel B: Volume imbalance %					
0 (most negative)	-61.4	-19.7	-5.8	-41.7	***	-55.6	***	-13.9	***
1	-47.3	-15.5	-8.3	-31.9	***	-39.1	***	-7.2	***
2	-35.0	-10.2	-6.6	-24.7	***	-28.4	***	-3.7	***
3	-21.9	-5.8	-4.8	-16.1	***	-17.1	***	-1.0	
4	-7.4	-1.7	-2.3	-5.7	***	-5.1	***	0.6	
5	6.1	2.6	-1.0	3.4	***	7.0	***	3.6	***
6	21.0	6.2	0.8	14.8	***	20.2	***	5.4	***
7	34.2	10.1	3.6	24.1	***	30.6	***	6.5	***
8	46.9	15.8	3.9	31.1	***	43.0	***	12.0	***
9 (most positive)	61.7	19.3	7.0	42.4	***	54.7	***	12.3	***

Table 3

Relation between *Volume imbalance*, *Trade imbalance* and *Depth imbalance*

Table 3 reports the regression of *Volume imbalance* or *Trade imbalance* against *Depth imbalance*. Trades are sorted into deciles based on the size of the depth imbalance (*DI*) immediately before the trade. For each *DI* decile and trader type, we calculate *Volume imbalance* as:

$$Volume\ imbalance_j^T = \frac{\sum_{k=1}^{n} BuyVolume_{k,j}^T - \sum_{k=1}^{n} SellVolume_{k,j}^T}{\sum_{k=1}^{n} BuyVolume_{k,j}^T + \sum_{k=1}^{n} SellVolume_{k,j}^T}$$

where $\sum_{k=1}^{n} BuyVolume_{k,j}^T(\sum_{k=1}^{n} SellVolume_{k,j}^T)$is the total aggressive buying (selling) volume for depth imbalance decile, *j*. For columns 1-3, we estimate the following linear regression, which is based on *DI* deciles:

$$Volume\ imbalance_j^T = \beta_0 + \beta_1 I(HFT)^T \times DI_j + \beta_2 I(Institutions)^T \times DI_j + \beta_3 I(HFT)^T + \beta_4 I(Institutions)^T + \beta_5 DI_j + \beta_6 Volume_j + \varepsilon_j^T$$

where $I(HFT)^T\,(I(Institutions)^T)$ is 1 if trader type, *T,* is *HFT* (*Institutions*) and 0 otherwise. *DI* is the average depth imbalance for the decile and *Volume* is the natural log of the total share volume traded in the decile. In Columns 4-6, we replace the dependent variable with *Trade imbalance*, which is calculated based on the number, rather than the volume, of aggressive executions. For each stock, low (high) volatility days represent the lowest (highest) tercile of trading days based on stock volatility, where volatility is the difference between the log of the intraday high ask price and the log of the intraday low bid price. All regressions control for stock and day fixed effects. Heteroscedastic-robust standard errors are double clustered by stock and day and *t*-statistics are reported in parentheses. ***, ** and * indicate significance levels of 1%, 5% and 10%, respectively.

	Buy volume imbalance%			Buy trade imbalance%		
	(1)	(2)	(3)	(4)	(5)	(6)
	All days	Low volatility days	High volatility days	All days	Low volatility days	High volatility days
I(HFT) × DI	1.046***	1.002***	1.117***	1.004***	0.997***	1.036***
	(25.95)	(19.73)	(28.90)	(24.58)	(18.78)	(26.18)
I(Institutions) × DI	0.025	-0.008	0.040	0.093***	0.049	0.129***
	(0.91)	(-0.20)	(1.33)	(3.37)	(1.15)	(4.79)
I(HFT)	0.011	0.025	0.007	0.022	0.041**	0.011
	(0.79)	(1.24)	(0.49)	(1.52)	(2.17)	(0.78)
I(Institutions)	0.016	0.032*	0.012	0.037**	0.053***	0.025*
	(1.26)	(1.81)	(0.82)	(2.43)	(2.72)	(1.66)
DI	-0.175***	-0.116***	-0.227***	-0.100***	-0.077**	-0.124***
	(-5.22)	(-2.65)	(-5.81)	(-4.10)	(-1.99)	(-4.77)
Volume	0.008**	0.002	0.006	0.011***	0.008*	0.011*
	(2.37)	(0.40)	(1.30)	(3.44)	(1.66)	(1.89)
Constant	-0.241***	-0.160***	-0.220***	-0.303***	-0.404***	-0.231***
	(-6.45)	(-2.76)	(-4.23)	(-7.29)	(-6.17)	(-3.48)
Obs.	245,382	64,203	93,143	245,382	64,203	93,143
Adj. R-square	0.186	0.186	0.201	0.275	0.274	0.302

Table 4

Relation between *Volume imbalance*, *Trade imbalance* and *Depth imbalance* for low and high information environments

Table 4 reports the regression of *Volume imbalance* or *Trade imbalance* against *Depth imbalance*. Trades are sorted into deciles based on the size of the depth imbalance (*DI*) immediately before the trade. For each *DI* decile and trader type, we calculate *Volume imbalance* as:

$$Volume\ imbalance_j^T = \frac{\sum_{k=1}^{n} BuyVolume_{k,j}^T - \sum_{k=1}^{n} SellVolume_{k,j}^T}{\sum_{k=1}^{n} BuyVolume_{k,j}^T + \sum_{k=1}^{n} SellVolume_{k,j}^T}$$

where $\sum_{k=1}^{n} BuyVolume_{k,j}^T(\sum_{k=1}^{n} SellVolume_{k,j}^T)$is the total aggressive buying (selling) volume for depth imbalance decile, *j*. For columns 1-3, we estimate the following linear regression, which is based on *DI* deciles:

$$\begin{aligned} Volume\ imbalance_j^T &= \beta_0 + \beta_1 I(HFT)^T \times DI_j + \beta_2 I(Institutions)^T \times DI_j + \beta_3 I(HFT)^T + \beta_4 I(Institutions)^T \\ &+ \beta_5 DI_j + \beta_6 Volume_j + \varepsilon_j^T \end{aligned}$$

where $I(HFT)^T(I(Institutions)^T)$ is 1 if trader type, *T,* is *HFT* (*Institutions*) and 0 otherwise. *DI* is the average depth imbalance for the decile and *Volume* is the natural log of the total share volume traded in the decile. In Columns 3-4, we replace the dependent variable with *Trade imbalance*, which is calculated based on the number, rather than the volume, of aggressive executions. For each stock, low (high) information days represent the lowest (highest) tercile of trading days based on the number of news articles. All regressions control for stock and day fixed effects. Heteroscedastic-robust standard errors are double clustered by stock and day and *t*-statistics are reported in parentheses. ***, ** and * indicate significance levels of 1%, 5% and 10%, respectively.

	Buy volume imbalance%		Buy trade imbalance%	
	(1) Low information	(2) High information	(3) Low information	(4) High information
I(HFT) × DI	1.100***	1.059***	1.038***	0.996***
	(24.81)	(23.48)	(22.09)	(22.93)
I(Institutions) × DI	0.047	0.031	0.097***	0.103***
	(1.48)	(1.03)	(2.90)	(3.69)
I(HFT)	0.011	0.007	0.028*	0.007
	(0.66)	(0.47)	(1.71)	(0.44)
I(Institutions)	0.021	0.015	0.045**	0.023
	(1.31)	(0.97)	(2.54)	(1.32)
DI	-0.204***	-0.199***	-0.136***	-0.092***
	(-5.57)	(-5.09)	(-4.41)	(-3.40)
Volume	0.010**	0.007*	0.014***	0.011***
	(2.05)	(1.93)	(2.96)	(2.60)
Constant	-0.278***	-0.249***	-0.329***	-0.300***
	(-5.16)	(-6.45)	(-5.52)	(-5.64)
Obs.	85,135	86,745	85,135	86,745
Adj. R-square	0.194	0.190	0.282	0.287

Table 5

Multinomial logistic regressions for limit order placement strategies

Table 5 assesses the probability of each order book event based on prevailing market conditions. We present the coefficient estimates for the following multinomial logistic regression:

$$OrderBookEvent_E^T = \beta_0^T + \beta_1^T Adjusted\ DI_E^T + \beta_2^T Volatility + \beta_3^T Volume + \beta_4^T Price + \beta_5^T Spread + \varepsilon_E^T$$

where $OrderBookEvent_E^T$ is the dependent variable indicating one of five order book events, *E,*: Aggressive execution, passive execution, limit order submission, amendment or cancelation. We estimate the model with limit order submission as the baseline category. *Volatility* is the difference between the log of the intraday high ask price and the log of the intraday low bid price. *Volume* is the natural log of the total daily share volume. *Price* is the average daily trade price. *Spread* is the time weighted average difference between the best bid and offer prices. The main independent variable is *Adjusted DI,* which is determined immediately before each order book event:

$$Adjusted\ DI_t = q_t \times \frac{\sum_{i=1}^{n} DepthBid_{i,t} - \sum_{i=1}^{n} DepthAsk_{i,t}}{\sum_{i=1}^{n} DepthBid_{i,t} + \sum_{i=1}^{n} DepthAsk_{i,t}}$$

where $\sum_{i=1}^{5} DepthBid_{i,t}$ ($\sum_{i=1}^{5} DepthAsk_{i,t}$) is the depth available at the top *5* bid (ask) price levels immediately before the order book event, *t*. *q* is an indicator variable equal to 1 for buys and -1 for sells. The sample consists of 12 trading days, formed by selecting the first and 15th day of each month in our sample period (if this day falls on a non-trading day, the next trading day is chosen). Panels A to C present the results for *HFT*, *Institutions* and *Retail*, respectively. *t*-statistics are reported in parentheses. ***, ** and * indicate significance levels of 1%, 5% and 10%, respectively.

	Aggressive execution		Passive execution		Amendment		Cancelation	
				Panel A: HFT				
Adjusted DI	1.075	***	0.467	***	-0.103	***	-0.398	***
	(0.007)		(0.007)		(0.008)		(0.006)	
Volatility	20.844	***	16.149	***	-26.768	***	0.026	
	(0.252)		(0.230)		(0.319)		(0.214)	
Volume	-0.475	***	-0.343	***	0.312	***	0.083	***
	(0.003)		(0.002)		(0.003)		(0.002)	
Price	0.761	***	0.439	***	-0.530	***	-0.018	***
	(0.005)		(0.005)		(0.005)		(0.004)	
Spread	-56.840	***	-35.202	***	46.945	***	-2.916	***
	(0.796)		(0.741)		(0.815)		(0.651)	
Constant	4.873	***	3.649	***	-5.572	***	-2.153	***
	(0.042)		(0.038)		(0.046)		(0.035)	
Obs.				2,479,277				
Pseudo R-square				0.023				

Panel B: Institutions								
Adjusted DI	0.511	***	-0.381	***	-0.003	*	0.069	***
	(0.003)		(0.002)		(0.002)		(0.002)	
Volatility	-1.833	***	2.755	***	0.224	***	-2.455	***
	(0.086)		(0.058)		(0.049)		(0.054)	
Volume	-0.021	***	-0.020	***	0.010	***	0.002	***
	(0.001)		(0.001)		(0.000)		(0.001)	
Price	-0.228	***	0.010	***	0.046	***	-0.063	***
	(0.001)		(0.001)		(0.001)		(0.001)	
Spread	10.057	***	-10.702	***	5.238	***	-0.160	
	(0.329)		(0.224)		(0.170)		(0.200)	
Constant	-1.480	***	-0.987	***	-1.212	***	-0.962	***
	(0.014)		(0.010)		(0.008)		(0.009)	
Obs.	33,808,108							
Pseudo R-square	0.003							
Panel C: Retail								
Adjusted DI	-0.005		-0.360	***	-0.266	***	-0.134	***
	(0.012)		(0.010)		(0.009)		(0.015)	
Volatility	-2.763	***	4.248	***	-7.229	***	9.362	***
	(0.343)		(0.270)		(0.281)		(0.390)	
Volume	-0.176	***	-0.156	***	0.187	***	-0.157	***
	(0.004)		(0.003)		(0.003)		(0.005)	
Price	-0.124	***	0.008		0.005		-0.068	***
	(0.006)		(0.005)		(0.005)		(0.008)	
Spread	14.729	***	4.949	***	-13.234	***	28.592	***
	(1.758)		(1.425)		(1.445)		(2.040)	
Constant	1.788	***	1.660	***	-3.813	***	0.394	***
	(0.063)		(0.052)		(0.052)		(0.079)	
Obs.	998,682							
Pseudo R-square	0.014							

Table 6

Relation between *Aggressive volume %* and stock volatility

Table 6 presents the regression of *Aggressive volume %* against volatility. The dependent variable is $Aggressive\ volume\ \%_I^T$, which is the aggressive buying and selling volume of trader type, *T*, as a percentage of their total aggressive and passive volume executed in a 30 minute time interval, *I*. The results are based on the following linear regression:

$$Aggressive\ volume\ \%_I^T = \beta_0 + \beta_1 Volatility_I + \beta_{10} Volume_I + \varepsilon$$

Volatility is the difference between the log of the highest best ask price and the log of the lowest best bid price during the 30 minute interval. *Volume* is the natural log of the total daily share volume during the 30 minute interval. *Large stocks* (*Small stocks*) refer to stocks contained in the largest (smallest) tercile of all sample stocks based on market capitalization. All regressions control for stock and day fixed effects. Panels A to C present the results for *HFT*, *Institutions* and *Retail*, respectively. Heteroscedastic-robust standard errors are double clustered by stock and day and *t*-statistics are reported in parentheses. ***, ** and * indicate significance levels of 1%, 5% and 10%, respectively.

	(1) All stocks	(2) Large stocks	(3) Small stocks
	Panel A: HFT		
Volatility	447.573***	211.525***	453.829***
	(20.35)	(5.06)	(11.25)
Volume	0.504***	0.368*	1.698***
	(3.53)	(1.69)	(5.09)
Constant	40.377***	42.203***	23.657***
	(18.36)	(13.81)	(5.13)
Obs.	58,789	28,891	10,750
Adj. R-square	0.304	0.272	0.332
	Panel B: Institutions		
Volatility	-84.866***	-10.890	-91.190***
	(-15.77)	(-1.21)	(-8.75)
Volume	-0.033	0.157***	-0.409***
	(-1.23)	(4.17)	(-7.06)
Constant	49.163***	48.914***	52.524***
	(119.37)	(84.32)	(62.17)
Obs.	95,102	34,025	28,243
Adj. R-square	0.118	0.143	0.122
	Panel C: Retail		
Volatility	-174.323***	-663.288***	7.451
	(-7.14)	(-13.31)	(0.19)
Volume	-0.220	-0.916***	0.642**
	(-1.43)	(-3.62)	(2.20)
Constant	67.915***	67.170***	61.621***
	(27.96)	(18.38)	(14.20)
Obs.	61,640	27,154	14,913
Adj. R-square	0.053	0.074	0.065

Table 7

Relation between *Adjusted DI* and stock volatility

Table 7 presents the regression of *Adjusted DI* against volatility. The dependent variable is $Adjusted\ DI_I^T$, which is the average *Adjusted DI* for aggressive or passive executions for each trader type, *T*, over a 30 minute interval, *I*. For each trade, *Adjusted DI* is calculated as:

$$Adjusted\ DI_t = q_t \times \frac{\sum_{i=1}^{n} DepthBid_{i,t} - \sum_{i=1}^{n} DepthAsk_{i,t}}{\sum_{i=1}^{n} DepthBid_{i,t} + \sum_{i=1}^{n} DepthAsk_{i,t}}$$

where $\sum_{i=1}^{5} DepthBid_{i,t}$ ($\sum_{i=1}^{5} DepthAsk_{i,t}$) is the depth available at the top *5* bid (ask) price levels immediately before the order book event, *t*. *q* is an indicator variable equal to 1 for buys and -1 for sells. The results are based on the following linear regression:

$$Adjusted\ DI_I^T = \beta_0 + \beta_1 I(HFT)^T \times Volatility_I + \beta_2 I(Institutions)^T \times Volatility_I + \beta_3 I(HFT)^T + \beta_4 I(Institutions)^T + \beta_5 Volatility_I + \beta_6 Volume_I + \varepsilon$$

$I(HFT)^T$ ($I(Institutions)^T$) is 1 if trader type, *T,* is *HFT* (*Institutions*) and 0 otherwise. *Volatility* is the difference between the log of the highest best ask price and the log of the lowest best bid price during the 30 minute interval. *Volume* is the natural log of the total daily share volume during the 30 minute interval. *Large stocks* (*Small stocks*) refer to stocks contained in the largest (smallest) tercile of all sample stocks based on market capitalization. All regressions control for stock and day fixed effects. Heteroscedastic-robust standard errors are double clustered by stock and day and *t*-statistics are reported in parentheses. ***, ** and * indicate significance levels of 1%, 5% and 10%, respectively.

	Aggressive executions			Passive executions		
	(1) All stocks	(2) Large stocks	(3) Small stocks	(4) All stocks	(5) Large stocks	(6) Small stocks
I(HFT) × Volatility	3.940***	5.012***	1.088	5.224***	5.473***	2.634*
	(5.51)	(3.74)	(1.10)	(5.75)	(3.19)	(1.96)
I(Institutions) × Volatility	0.323	0.569	0.313	1.621***	2.361***	1.102***
	(1.09)	(1.21)	(0.89)	(5.75)	(3.85)	(2.76)
I(HFT)	0.100***	0.080***	0.140***	0.068***	0.053***	0.103***
	(15.63)	(11.92)	(9.43)	(6.95)	(5.10)	(3.95)
I(Institutions)	0.002	0.009**	-0.006	-0.036***	-0.040***	-0.030***
	(0.55)	(2.25)	(-1.04)	(-11.23)	(-9.63)	(-6.70)
Volatility	1.753***	0.991**	2.440***	-2.521***	-3.134***	-1.526***
	(6.22)	(1.98)	(6.64)	(-8.22)	(-4.32)	(-3.91)
Volume	-0.003***	-0.002**	-0.004**	-0.001*	-0.000	-0.003*
	(-3.46)	(-2.04)	(-2.43)	(-1.90)	(-0.04)	(-1.79)
Constant	0.036***	0.034***	0.041**	0.017**	0.011	0.016
	(4.58)	(2.96)	(2.56)	(2.31)	(0.69)	(0.97)
Obs.	253146	97567	80562	255094	99025	83576
Adj. R-square	0.178	0.207	0.173	0.157	0.197	0.138

Table 8

Relation between *Volume imbalance* and *DI* before and after the implementation of ITCH

Table 8 reports the regression of *Volume imbalance* or *Trade imbalance* against *DI*. We analyze trade and quote data for the periods March 2, 2012 to March 30, 2012 (pre-ITCH) and April 9, 2012 to May 9, 2012 (post-ITCH). For each stock, trades are sorted into deciles based on the size of the depth imbalance (*DI*) immediately before the trade. For each *DI* decile and trader type, we calculate *Volume imbalance* as:

$$Volume\ imbalance_j^T = \frac{\sum_{k=1}^{n} BuyVolume_{k,j}^T - \sum_{k=1}^{n} SellVolume_{k,j}^T}{\sum_{k=1}^{n} BuyVolume_{k,j}^T + \sum_{k=1}^{n} SellVolume_{k,j}^T}$$

where $\sum_{k=1}^{n} BuyVolume_{k,j}^T (\sum_{k=1}^{n} SellVolume_{k,j}^T)$ is the total aggressive buying (selling) volume for depth imbalance decile, *j*. For columns 1-3, we estimate the following linear regression, which is based on *DI* deciles:

$$\begin{aligned} Volume\ imbalance_j^T &= \beta_0 + I(Pre)[\beta_1 I(HFT)^T \times DI_j + \beta_2 I(Insto)^T \times DI_j + \beta_3 DI_j] + I(Post)[\beta_4 I(HFT)^T \times DI_j + \beta_5 I(Insto)^T \times DI_j + \beta_6 DI_j] + \beta_7 I(HFT)^T \\ &+ \beta_8 I(Insto)^T + \beta_9 Volume_j + \varepsilon_j^T \end{aligned}$$

I(Pre) (*I(Post)*) is an indicator variable equal to 1 if the trading day falls in the pre-ITCH (post-ITCH) period and zero otherwise. *I(HFT)* and *I(Institutions)* are indicator variables equal to 1 for the trader type specified in the parentheses and 0 otherwise. *Volume* is the natural log of the total share volume traded in the decile. For ease of comparison, Column 1 reports the coefficients associated variables interacted with *I(Pre)*, Column 2 presents the coefficients for the control variables, and Column 3 presents the coefficients associated variables interacted with *I(Post)*. We use an *F*-test to test for the equality of the coefficients interacted with *DI*. Column 4 presents the *F*-test and the associated *p*-value in parentheses. In Columns 5-8, we replace the dependent variable with *Trade imbalance* and perform the same analysis as the previous 4 columns. All regressions control for stock and day fixed effects. Heteroscedastic-robust standard errors are double clustered by stock and day and *t*-statistics are reported in parentheses. ***, ** and * indicate significance levels of 1%, 5% and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Volume imbalance				Trade imbalance			
	Pre-ITCH		Post-ITCH	F-Test	Pre-ITCH		Post-ITCH	F-Test
				5.350*				7.71**
I(HFT) × DI	β_1 0.942***		β_4 1.064***		β_1 0.938***		β_4 1.066***	
	(18.43)		(18.79)	(0.023)	(17.24)		(22.25)	(0.007)
I(Institutional) × DI	β_2 -0.030		β_5 -0.038	0.040	β_2 0.017		β_5 0.128***	5.27**
	(-0.79)		(-0.88)	(0.850)	(0.37)		(3.22)	(0.024)
Depth imbalance	β_3 -0.095**		β_6 -0.118**	0.480	β_3 -0.028		β_6 -0.121***	6.36**
	(-2.09)		(-2.50)	(0.490)	(-0.73)		(-3.24)	(0.014)
I(HFT)		β_7 0.025				β_7 0.026		
		(1.33)				(1.49)		
I(Institutional)		β_8 0.033**				β_8 0.040**		
		(2.05)				(2.33)		
Volume		β_9 0.016***				β_9 0.019***		
		(3.36)				(3.44)		
Constant		β_0 -0.373***				β_0 -0.561***		
		(-6.76)				(-8.55)		
Obs.		80,666				80,666		
Adj. R-square		0.186				0.278		

Table 9

Probability of limit order executions before and after the implementation of ITCH

Table 9 analyzes the probability of limit order executions for *HFT*, *Institutions* and *Retail* before and after the implementation of ITCH. We analyze trade and quote data for the periods March 2, 2012 to March 30, 2012 (pre-ITCH) and April 9, 2012 to May 9, 2012 (post-ITCH). The main dependent variable $P(Fill)^T$ is calculated as:

$$P(Fill)^T = \frac{\sum TradeVolume^T}{\sum SubmitVolume^T}$$

where $\sum SubmitVolume^T$ is the total daily volume submitted to the top level of the limit order book and $\sum TradeVolume^T$ is the total volume that is successfully traded, for trader type, *T*. In Column 1, we estimate the following regression:

$$P(Fill)^T = \beta_0 + \beta_1 I(Non-HFT)^T \times I(Post) + \beta_2 I(Non-HFT)^T + \beta_3 I(Post) + \beta_4 Volatility + \beta_5 Volume + \beta_6 Price + \beta_7 Spread + \varepsilon^T$$

where $I(Non-HFT)^T$ is an indicator variable equal to 1 for *Institutions* and *Retail* and zero for *HFT*. *I(Post)* is an indicator variable equal to 1 if the trading day falls in the post-ITCH period and zero for the pre-ITCH period. *Volatility* is the difference between the log of the intraday high ask price and the log of the intraday low bid price. *Volume* is the natural log of the total daily share volume. *Price* is the average daily trade price. *Spread* is the time weighted average difference between the best bid and offer prices. In Column 2, we replace $I(Non-HFT)^T$ with $I(Institutions)^T$ and $I(Retail)^T$, which are indicator variables equal to 1 for the trader type specified in the parentheses, and zero otherwise. In Columns 3 and 4 (Columns 5 and 6), we replace the dependent variable with $P(Favorable\ fill)^T$ ($P(Unfavorable\ fill)^T$). We define a favorable (unfavorable) fill as an order execution when the limit order rests on the side of the order book with more (less) depth immediately prior to the trade. All regressions control for stock and day fixed effects. We report heteroskedastic-robust standard errors double clustered by stock and day in parentheses. ***, ** and * indicate significance levels of 1%, 5% and 10%, respectively.

	P(Fill)		P(Favorable fill)		P(Unfavorable fill)	
	(1)	(2)	(3)	(4)	(5)	(6)
I(Non-HFT) × I(Post)	-0.037***		-0.040***		-0.001	
	(-3.31)		(-4.18)		(-0.13)	
I(Non-HFT)	0.111***		-0.030*		0.134***	
	(5.95)		(-1.80)		(22.11)	
I(Institutions) × I(Post)		-0.021**		-0.029***		0.008
		(-2.04)		(-3.03)		(1.55)
I(Institutions)		-0.069***		-0.124***		0.053***
		(-3.67)		(-7.51)		(8.96)
I(Retail) × I(Post)		-0.058***		-0.051***		-0.009
		(-3.77)		(-4.23)		(-1.00)
I(Retail)		0.301***		0.078***		0.226***
		(15.87)		(4.52)		(29.49)
I(Post-ITCH)	0.006	0.007	0.036**	0.036**	0.006	0.008
	(0.39)	(0.49)	(2.24)	(2.33)	(0.57)	(0.68)
Volatility	0.076	0.062	0.054	-0.090	0.007	-0.097
	(0.46)	(0.36)	(0.33)	(-0.56)	(0.07)	(-1.04)
Volume	0.071***	0.069***	0.039***	0.036***	0.028***	0.026***
	(17.62)	(17.64)	(11.72)	(10.89)	(10.88)	(10.02)
Price	0.030	0.031	0.015	0.009	-0.006	-0.014**
	(0.77)	(0.80)	(0.81)	(0.62)	(-0.78)	(-2.01)
Spread	-1.334	-1.303	-0.089	-0.411	-0.199	-0.365
	(-0.56)	(-0.55)	(-0.06)	(-0.32)	(-0.12)	(-0.23)
Constant	-0.559***	-0.529***	-0.291***	-0.231***	-0.249***	-0.205***
	(-6.85)	(-6.61)	(-4.81)	(-3.85)	(-5.73)	(-4.67)
Obs.	10,646	10,646	9,718	9,718	9,574	9,574
Adj. R-square	0.188	0.586	0.151	0.369	0.224	0.460