High frequency trading strategies

September 5, 2021

Abstract. We examine the effect of high frequency trading on market quality from the perspective of a limit order trader. By competing with slower limit order traders, high frequency traders (HFT) impose a welfare externality by selectively crowding out the most profitable limit orders. The order book imbalance immediately before each order submission, cancelation and trade suggests that high frequency traders strategically 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.

Key words: High-frequency trading, institutional investors, retail investors

JEL: G14; G15

1 Introduction

High frequency trading represents one of the most significant changes to market structure in recent years (SEC, 2010). The rise in high frequency trading has triggered debate about the possibility of creating an unequal playing field between fast (i.e., high frequency traders (HFT)) and slow (i.e., institutional and retail) traders. Benefits or costs may also vary by order type. While some empirical literature finds that high frequency trading is beneficial from the perspective of a market order trader, the effects of HFT on limit order traders is less clear. For example, some suggest that HFT may increase price discovery and improve traditional measures of market quality, such as depth and bid ask spread, meaning that market order traders can execute at a lower cost and at a more efficient price. However, improvement in these metrics provides little insight on the costs or benefits to limit order traders.

Theory sheds some light on potential welfare effects and proposes that slower limit order traders face lower execution probabilities when faster traders are present, a type of 'crowding out effect' (Hoffmann (2014); Han, Khapko, and Kyle (2014)). Li, Wang, and Ye (2021) focus on one mechanism that can cause the crowding out effect. Specifically, when the bid-ask spread is binding at the minimum tick size, HFT dominate liquidity provision as they can use speed to compete for queue priority. However, while theory shows some potential effects, the existing empirical test have been limited, perhaps due to the paucity of appropriate measures and data. Moreover, while Li, Wang, and Ye (2021) identify the competition for queue priority as one mechanism driving the crowding out effect, it is possible that other mechanisms may also contribute to the crowding out effect.

We show that HFT selectively crowd out 'favorable' limit orders from the order book, an externality of strategic HFT behavior that tries to anticipate future price moves. Specifically, using a highly granular data set to track the evolution of limit orders in the order book, we show that fast traders (i.e., HFT) crowd out slower traders (i.e., institutional and retail traders) from the favorable executions in the limit order book, but not the unfavorable executions. Next, we uncover a new information channel driving the asymmetric crowding out effect. While all traders use information contained in the order book as one channel in their trading decisions, HFT are better able to exploit this channel, allowing them to 'piggyback' off the information of others, thereby crowding out slower limit order traders.

Our unique highly granular order book data contains every order submission, cancelation and amendment with an attached broker ID to classify brokers into three trader types (proprietary HFT firms, institutions and retail). Our full order book data from a largely unfragmented market allows us to derive new insights into trader behavior not yet investigated in the literature.[1] Using this high level of granularity, we track the lifespan of individual limit orders and their queue position over time, which allows us to examine how HFT use order book information in their order management decisions. By fully reconstructing the order book, we extend the existing literature investigating the crowding out effect (see Yao and Ye (2018)) to identify precisely the limit orders and market participants who are most affected by differential trading speeds.

Exploiting the granularity of the data, we propose two measures that directly capture the 'crowding out' risks HFT impose on limit order users: The first measure reflects the overall risk of non-execution, which we measure as the probability of execution, while the second measure, based on Handa and Schwartz (1996), captures the probability of missing out on the most profitable limit order fills, which we measure by decomposing the probability of execution into favorable and unfavorable fills depending on the stock price movement after the order executes. Next, we show that favorable limit orders can be detected *ex ante* using a simple measure, the order book depth imbalance. Because stock prices are likely to move in the direction of the depth imbalance,[2] 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. On the other hand, an unfavorable fill occurs when there is more depth on the opposite side of the book, as these limit orders are likely to face adverse selection. For identification, we use the introduction of a faster message transmission protocol known as ITCH on the Australian Securities Exchange (ASX), which further segmented

[1]Other datasets are typically limited to transaction-level data (e.g., van Kervel and Menkveld (2019), Menkveld (2013), Baron, Brogaard, Hagströmer, and Kirilenko (2019)) or data from highly fragmented markets (e.g., Brogaard, Hagströmer, Nordén, and Riordan (2015), Malinova, Park, and Riordan (2018), Malinova and Park (2020)).

[2]Cao, Hansch, and Wang (2009) and Cont, Kukanov, and Stoikov (2013) find strong evidence that order imbalances between the buy and sell schedules of the limit order book are significantly related to future stock returns.

the market into fast and slow traders.[3,4]

Initially, we confirm the results of Yao and Ye (2018) and show that when HFT gain a speed advantage, slower traders are 'crowded out' of the limit order book, reflected by an overall drop in their probability of limit order executions. Moreover, using the additional granularity in trader identification, we extend Yao and Ye (2018) and show that slower retail traders are crowded out more from the order book than institutional traders. Further, exploiting the cross sectional variation in our sample stocks, we demonstrate a stronger crowding out effect when speed competition for queue priority is most required. Specifically, we confirm the predictions of Li, Wang, and Ye (2021) and show that the crowding out effect is more pronounced when trading is tick constrained or for stocks that have a larger relative tick size.

Li, Wang, and Ye (2021) propose that when the mandated bid-ask spread is greater than the breakeven bid-ask spread, rents are extracted for providing liquidity, creating a speed competition for queue priority, which results in a crowding out effect. Based on this queuing channel, it is profitable to provide liquidity on *both* the bid and ask prices, and therefore, slower liquidity traders should be crowded out symmetrically from *both* sides of the book. Our findings reveal an asymmetric crowding out effect: We show that it is the probability of favorable executions (i.e., the orders resting on the thick side of the order book) that drops for slower traders, whereas the probability of an unfavorable execution is unaffected when HFT gain a speed advantage. This asymmetry suggests that HFT selectively crowd out slower limit order traders from the favorable side of the order book by predicting which side of the book to provide liquidity on. HFT's ability to predict the favorable side of the limit order book may be driven by their ability to piggyback off the information of other limit order traders, similar to the 'back runners' described in Yang and Zhu (2019).

We next investigate the piggybacking induced asymmetric crowding out effect by focusing on a simple measure, the order book depth imbalance, which encompasses the broad information

[3] ASX ITCH is an ultra-low latency protocol for accessing ASX market information for a monthly fee, which was designed to meet the requirements of speed sensitive traders.

[4] While several studies use speed changes to investigate the effect of HFT on overall market quality (see Brogaard, Hendershott, Hunt, and Ysusi (2014), Shkilko and Sokolov (2020), and Brogaard, Hagströmer, Nordén, and Riordan (2015)), we focus on the direct effect of faster trading speeds on different types of limit order traders.

contained in the order book.[5] We show that *all* traders trade in the direction of the order book imbalance: for all trader types, aggressive buying increases with the order book depth imbalance. However, we find that HFT are better at taking advantage of order book information under certain market conditions: when depth imbalances are large, the market is volatile and when HFT speed advantage increases. When depth imbalances are very large, suggesting large future price changes, HFT are more successful at trading in the direction of the imbalance relative to the other trader types. 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. Consistent with this view, HFT become even more successful at trading in the direction of the order book after the introduction of ITCH when they gain a larger speed advantage.[6] This finding provides support for theoretical predictions that faster traders pick off stale limit orders left by slower limit order traders.

Through strategic limit order management, HFT, on average, supply depth on the thick side of the order book and cancel from the thin side of the order book. Using multinomial logistic regression, we investigate HFT's order placement decisions and show that they 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 later 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 so at times when it is *least* needed, thereby crowding out other slower liquidity providers at the same price. Moreover, HFT remove liquidity quickly where it is *most* needed.

HFT appear to react to changes in the limit order book quickly. Of course, it is possible that

[5]Theoretical models by Goettler, Parlour, and Rajan (2009), Roşu (2020), Riccó, Rindi, and Seppi (2020) and Bhattacharya and Saar (2020) show that the state of the limit order book contributes to price discovery. Kwan, Philip, and Shkilko (2021) show empirically that the state of the order book is the most important conduit of price discovery.

[6]For example, Baron, Brogaard, Hagströmer, and Kirilenko (2019) show an improvement in HFT performance when their relative latency improves amongst other HFT firms.

they are instead reacting to external public news events (e.g., von Beschwitz, Keim, and Massa (2019) and Chakrabarty, Moulton, and Wang (2020)). We test this possibility by dividing our sample into high information days and low information days and show that the responses are invariant to the information content of the days. The limit order behavior of HFT shown here does not appear to be due to a faster response to external information. Instead, it appears HFT respond to information in the limit order book itself.

These results provide strong empirical evidence for the theoretical literature that focuses on speed differential between traders. First, we document a crowding out effect consistent with Li, Wang, and Ye (2021) who propose that tick size constraints and the time priority rule create a speed competition for queue priority. Second, our findings suggest that faster traders pick off stale limit orders left by slower traders, consistent with the predictions from Hoffmann (2014), Biais, Foucault, and Moinas (2015), Budish, Cramton, and Shim (2015), Foucault, Hombert, and Roşu (2016), and Foucault, Kozhan, and Tham (2016). Third, other theoretical models propose that HFT can anticipate future order flow of other traders (see Biais, Foucault, and Moinas (2015), Hoffmann (2014), Roşu (2019), and Yang and Zhu (2019)). We demonstrate that the order book depth imbalance is a channel that HFT use to anticipate future order flow. Fourth, the HFT behavior we document is consistent with the 'back-runners' from Yang and Zhu (2019) who exploit fundamental investors' information from their order flow. This 'back-running' behavior could explain why Brogaard, Hendershott, and Riordan (2019) find HFT limit orders submitted provide more price discovery than limit orders from other participants. While HFT limit orders contribute most to price discovery, they do so by 'piggybacking' off the fundamental information revealed by other limit order traders. Such behavior could have long term implications for financial markets. For example, slower fundamental traders could withdraw from lit exchanges or be discouraged from acquiring information. These negative externalities are consistent with a decrease in price informativeness in the longer term, consistent with Weller (2017) and Baldauf and Mollner (2020).

Our study also contributes to the growing body of literature on the possible externalities of HFT. Recently, Malceniece, Malcenieks, and Putniņš (2019) suggest that HFT lead to greater comovement in returns and liquidity, which could impact the cost of capital. Aquilina, Budish, and O'Neill (2020) suggest that eliminating latency arbitrage would reduce the cost of liquidity by 17%

and Boehmer, Li, and Saar (2018) report that an increase in competition between HFT firms that follow the same underlying strategy decreases stock volatility. Further, Kirilenko, Kyle, Samadi, and Tuzun (2017), Brogaard, Carrion, Moyaert, Riordan, Shkilko, and Sokolov (2018) and Shkilko and Sokolov (2020) report that HFT could exacerbate volatility as HFT withdraw liquidity from the market, at specific times, such as market crashes. Our contribution focuses on the systematic effects of HFT on limit order traders. Using trade and quote data, van Kervel and Menkveld (2019) and Korajczyk and Murphy (2018) report an increase in institutional trading costs when HFT compete in the same direction with institutional orders. In contrast, we have *order* data, so we know all orders, including cancelations, and not just trades or quotes. Exploiting the additional granularity of our order book data, we are able to show precisely how HFT's speed advantage impacts the welfare of retail and institutional limit order traders.

Finally, our results have important implications for studies on the market making role of HFT in equity markets. Using traditional measures of market depth aggregated across both bid and ask prices, Hendershott, Jones, and Menkveld (2011), Hasbrouck and Saar (2013), Brogaard, Hagströmer, Nordén, and Riordan (2015), and others find that HFT market making increases aggregate market depth. By examining the relative depth on each side of the market (one sided depth) separately and in relation to the other side, we find that HFT supply depth to the side of the book where it is not needed but demand depth from the side of the book where it is most needed. These asymmetric findings suggest that caution should be taken in interpreting results that aggregate market depth across both the bid and ask sides.

2 Data and sample selection

2.1 Data, limit order book reconstruction, and tracking orders

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 for the six months from January 3, 2012 to June 30, 2012, which spans the introduction of ASX ITCH. 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)).[7] 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. This sample period ensures that traders have sufficient time to implement strategies based on the ITCH protocol.

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 closing auction. 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 canceled. 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 reconstruction of the limit order book.

2.2 Broker IDs and trader identification

Data from the ASX offer several advantages over other exchanges. For example, in our dataset, broker identifiers are assigned into three trader categories based on the primary trading activity or

[7] Brogaard, Hendershott, and Riordan (2014) suggest that HFTs provide the most liquidity in larger stocks. 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.

client base of the broker: proprietary HFT firms (*HFT*), *Institutions*, and *Retail*.[8] Thus, we do not need to rely on HFT proxies such as message to trade ratios, which could be negatively correlated with the true measure of HFT activity (see Yao and Ye (2018), Li, Wang, and Ye (2021)).

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.[9] Following Upson, McInish, and Johnson (2021), 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. 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. As such, we are one of the few studies to

[8]Appendix A.1 provides more details on the classification procedure. 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. We acknowledge some smaller proprietary HFT firms could trade through institutional brokers and thus, *Institutions* may also contain some proprietary HFT activity. 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. 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. Last, ASIC has banned payment for order flow services and thus, order routing of retail order flow and/or institutional flow does not play a major role on the ASX.

[9]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 (2018)), 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.

analyze HFT in a largely unfragmented market.

2.3 Introduction of ASX ITCH

ASX ITCH, introduced to the Australian Securities Exchange (ASX) on April 2, 2012, 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). 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.[10]

We test this assumption in Internet Appendix A.2.1 and analyze whether HFT speeds increase, relative to other traders after the implementation of ITCH. Using a measure of response time based on Baron, Brogaard, Hagströmer, and Kirilenko (2019), we report an almost two-fold increase in HFT response times while the response time of slower participants remain unchanged. This finding shows that 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 study.[11]

2.4 Summary statistics

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 of volume executed is \$27.8 AUD million and the average number of trades per day is 2,264.

[10] Aquilina, Budish, and O'Neill (2020) highlight the importance of speed in winning latency-arbitrage races.

[11] While the response times of HFT improve post-ITCH, in Appendix A.2.2 we show that the order management strategies of slower traders do not change immediately after the adoption of ITCH.

[Insert Table 1]

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 actively. Relative to *Institutions* and *Retail*, *HFT* have a higher cancel to trade ratio, a lower trade to order volume ratio, and their median submission to cancel time is significantly lower. $EODInventory\%$, which is a broker's net inventory as a percentage of the broker's trading volume is also lower for *HFT* relative to the other trader types. Last, we also find that HFT firms have the greatest number of switches between long and short positions in a stock day: on average, HFT firms cross over 2.57 times across all stocks while *Retail* cross over 1.3 times. When we look at the maximum number of cross overs per stock, $InventoryCrossesZero$ increases to 9.23 for *HFT*, indicating that some HFT firms in our sample frequently change the direction of their inventory positions.

3 Probability of limit order executions post-ITCH

We start our empirical investigation by examining the effects of a speed change on limit order executions for our trader types. The use of limit orders involves two risks: the risk of non-execution and the risk of undesirable execution (Handa and Schwartz (1996)).

3.1 Non-Execution risk; probability of fill

Hoffmann (2014), Han, Khapko, and Kyle (2014) and Li, Wang, and Ye (2021) predict that the risk of non-execution increases for slower traders in the presence of faster traders. To test this prediction, we measure the risk of non-execution by computing the probability of limit order executions for HFT (i.e., faster traders) and non-HFT (i.e., slower traders) limit orders as follows:

$$P(Fill)^T = \frac{\sum TradeVolume^T}{\sum SubmitVolume^T} \tag{1}$$

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.[12] For each stock, we measure $P(Fill)^T$ on a daily basis for each trader type.

Figure 1 shows the evolution of $P(Fill)^T$ for *HFT* (gray line) and *Retail* (black line) around the introduction of ITCH.[13] The plot shows a clear reduction in $P(Fill)^T$ for *Retail* after the implementation of ITCH (black line) while we do not observe significant changes in $P(Fill)^T$ for *HFT*. Additionally, we do not see significant divergences in *HFT* and *Retail* $P(Fill)^T$, providing support that parallel trends hold between the trader types in the pre-ITCH period.[14]

[Insert Figure 1]

To formally test whether $P(Fill)^T$ changes around the introduction of ITCH, we use a difference-in-difference framework and estimate the following regression model:

$$\begin{aligned} P(Fill)^T = \quad & \beta_0 \quad + \beta_1 I(Non-HFT)^T \times I(Post) + \beta_2 I(Non-HFT)^T \\ + \quad & \beta_3(Post) + \beta_4 Volatility + \beta_5 Volume + \beta_6 Price + \beta_7 Spread + \epsilon^T \end{aligned} \tag{2}$$

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.

[12] For all our analyses, we treat marketable limit orders the same as market orders. Therefore, marketable limit orders are not included in our $P(Fill)$ calculation.

[13] For clarity, we only plot $P(Fill)$ for *HFT* and *Retail*. However, we perform formal statistical tests on $P(Fill)$ for all trader types in Table 2.

[14] In Internet Appendix A.2.3 we use a falsification test to confirm that the parallel trends assumption holds between HFT and non-HFT trading activity in the pre-ITCH period.

Our main variable of interest is $I(Non-HFT)^T \times I(Post)$. If fast traders increase the risk of non-execution of slow traders, we expect a negative coefficient for $I(Non-HFT)^T \times I(Post)$, indicating that the probability of a non-HFT order receiving execution decreases when HFT become faster.

[Insert Table 2]

Consistent with predictions, Table 2, Column 1 shows that the probability of limit order execution falls for non-HFT after the introduction of ITCH, as the coefficient on $I(Non-HFT)^T \times I(Post)$ is -0.037. In Table 2, 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 statistically significant indicating that both *Institutions* and *Retail* are crowded out from the order book by HFT. Further, given retail investors are less sophisticated and slower than institutional investors, we anticipate that the crowding out effect is more pronounced for *Retail*. Our findings are consistent with this hypothesis, with a significantly more negative coefficient for the interactions with *Retail* (-0.058) than *Institutions* (-0.021).

3.2 Crowding out by stock characteristics

To gain an insight into the underlying channel driving the crowding out effect, we exploit the cross sectional variation among stocks in our sample. In the first test, we provide robustness to our initial findings and examine whether the crowding out effect is stronger for stocks in which HFT are most active. Because the introduction of ITCH creates a larger competitive advantage for HFT over other traders, we expect the effect of the speed change to be more pronounced in stocks with higher HFT participation.

Focusing on the subsample of stocks with the highest HFT participation, our second test investigates a potential channel driving the crowding out effect. Li, Wang, and Ye (2021) propose that when the one-tick mandated bid-ask spread is greater than the breakeven bid-ask spread, rents are extracted for providing liquidity. HFT exploit their faster trading speeds to compete on queue priority to extract these rents. Thus, we expect to find a stronger crowding out effect for highly tick constrained stocks, in which the mandated bid-ask spread is typically binding. Similarly, the crowding out effect is likely to be more pronounced in stocks with a large relative tick size. Since a uniform tick size constrains price competition, traders must then compete on speed for queue priority, as described in Yao and Ye (2018). For this reason, speed is likely to be more advantageous for HFT in stocks that have higher relative tick sizes.

To test whether the crowding out effect is more pronounced in stocks with higher HFT participation, we divide the sample of firms into *High HFT* (top tercile) and *Low HFT* (bottom tercile) based on the level of HFT participation in the pre-ITCH period. For *High HFT*, the results in Table 2, Column 3 confirm that *P(Fill)* falls for non-HFT after the adoption of ITCH (-0.032). Consistent with the notion that retail investors are less sophisticated and slower than institutional investors, Table 2, Column 4 shows that the decline in *P(Fill)* is driven by *Retail*. In contrast to the results for *High HFT*, we do not find a fall in *P(Fill)* for non-HFT limit orders for stocks with low HFT participation. The interaction terms in Table 2, Columns 5 and 6 are statistically insignificant indicating that *P(Fill)* remains unchanged for non-HFT traders after the introduction of ITCH for *Low HFT* stocks.

Next, using the *High HFT* subsample, we investigate the predictions of Li, Wang, and Ye (2021), who propose that competition for queue priority is the underlying mechanism driving the crowding out effect. We create two subsamples based on whether trading in a stock is tick constrained or unconstrained.[15] In support of the competition for queue priority channel, we find that *P(Fill)* decreases for non-HFT firms after HTF gain a speed advantage when trading is constrained by the minimum tick size (Table 3, Column 1). In contrast, for stocks that are unconstrained by the minimum tick size in Table 3, Column 2, our results do not show a statistically significant decrease in *P(Fill)* for non-HFT limit orders.

[Insert Table 3]

For further robustness, we divide the *High HFT* sample into *High relative tick* and *Low relative*

[15] Specifically, we designate a stock as constrained or unconstrained based on the average bid-ask spread in the pre-ITCH period.

tick based on the tercile of relative tick sizes in Table 3, Columns 3 and 4.[16] Consistent with our findings above and those in Yao and Ye (2018), we show that the crowding out effect is most pronounced in high relative tick stocks, providing further support for the predictions of Li, Wang, and Ye (2021). As before, we do not find a decline in *P(Fill)* for non-HFT limit orders in stocks with low relative tick sizes.

Together, these results support the queuing channel of Li, Wang, and Ye (2021). The crowding out effect is strongest for stocks in which speed is most critical, namely stocks that are tick constrained or stocks with higher relative tick sizes.

3.3 Asymmetry in the crowding out effect

The results in the previous section show that faster trading speeds crowd out slower traders from the limit order book. One underlying channel is the competition for queue priority channel proposed by Li, Wang, and Ye (2021); when the mandated bid-ask spread is greater than the breakeven bid-ask spread, a speed competition ensues to extract rents for providing liquidity, thereby crowding out slower traders. Under this scenario, it should be profitable for HFT to provide liquidity on *both* sides of the order book and the crowding out effect should be symmetrical. However, theory also predicts that slower traders face a higher adverse selection risk while faster traders face a lower adverse selection risk. For example, Han, Khapko, and Kyle (2014) predict that slower traders' quotes have a lower probability of trade execution when public information makes a trade more profitable. In this section, we ask whether HFT crowd out slower traders equally from both sides of the order book, or do they selectively crowd out limit orders with certain characteristics?

To test these predictions, we decompose $P(Fill)$ into $P(Favorablefill)$ and $P(Unfavorablefill)$ based on the price reaction after the limit order executes. Specifically, a favorable fill is a buy (sell) limit order, which executes prior to a price rise (fall) over the following 10 trades, and vice versa for unfavorable fills. The results reported in Table 4 show an asymmetry in the crowding out effect of slower non-HFT traders: Column 1 shows that $P(Favorablefill)$ decreases after the implementation of ITCH (-0.011) while Column 3 shows that $P(Unfavorablefill)$ does not change. These

[16] Specifically, we form terciles based on the stock's average relative tick size over the pre-ITCH period.

results are confirmed in Columns 2 and 4, when we divide non-HFT into *Institutions* and *Retail*. $P(Favorablefill)$ declines while $P(Unfavorablefill)$ remains at pre-ITCH levels for both *Institutions* and *Retail*. This finding suggests that HFT are able to predict the side of the book that is favorable for liquidity provision, and predominantly compete for liquidity provision on the favorable side.

[Insert Table 4]

Given that HFT appear to selectively crowd out the favorable fills *ex post*, it is natural to ask how HFT determine which is the favorable side of the order book, *ex ante*. To answer this question, we must identify an appropriate proxy for the *ex ante* adverse selection risk. Chordia, Roll, and Subrahmanyam (2002) and Chordia and Subrahmanyam (2004) document a strong relation between trade imbalances and future returns. Using more granular limit order book data, Cao, Hansch, and Wang (2009) and Cont, Kukanov, and Stoikov (2013) find evidence that order imbalances between the buy and sell schedules of the limit order book are significantly related to future stock returns. Ranaldo (2004) examines how the state of the limit order book can affect a trader's order submission strategy, and Cont, Kukanov, and Stoikov (2013) 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).

Collectively, these studies suggest that the information contained in the state of the limit order book is a good proxy for *ex ante* adverse selection risk. We therefore use the depth on both sides of the limit order book as our proxy of *ex ante* adverse selection risk for any given order. Specifically, a passive limit order that executes while resting on the thin side of the order book is likely to be adversely selected because we expect the future price to move against the limit order, *ex ante*.

To measure the shape of the limit order book immediately prior to a trade execution, we calculate depth imbalance (DI) as the difference between the depth available at the best bid and ask prices, as a proportion of the total depth available at the best bid and ask prices. For each limit order execution, we determine:

$$DI_t = \frac{\sum_{i=1}^{n} DepthBid_{i,t} - \sum_{i=1}^{n} DepthAsk_{i,t}}{\sum_{i=1}^{n} DepthBid_{i,t} + \sum_{i=1}^{n} DepthAsk_{i,t}}, \tag{3}$$

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 a trade hits the limit order book, at time t.[17] We compute DI immediately before the time of the trade to avoid capturing the volume of the trade itself. For ease of exposition, we calculate everything in terms of buy imbalances, so sell imbalances are negative numbers.

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.[18] 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. Because DI is not based on stock returns, our measure of DI is an *ex ante* measure of adverse selection: all information can be observed prior to the limit order execution.[19]

Now, we decompose $P(Fill)$ into *ex ante* favorable and unfavorable fills based on the DI at the time of the trade. We define a favorable fill for a limit order as a limit order that executes while resting on the side of the order book with more depth immediately prior to the trade (i.e., a positive DI for limit buy executions and a negative DI for limit sell executions).

We test the assertion that a favorable (unfavorable) fill occurs for orders resting on the thick (thin) side of the order book in Internet Appendix A.3 and confirm the results in Cao, Hansch, and Wang (2009), Cont, Kukanov, and Stoikov (2013), and others. Specifically, Internet Appendix A.3 shows that a limit order trader benefits when an order executes with a lot of depth on the same

[17] In Section 4.2, we also determine the depth imbalance before other order book events (i.e., submission, amendment, execution, or cancelation).

[18] 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).

[19] Based on Cao, Hansch, and Wang (2009) and Cont, Kukanov, and Stoikov (2013), our measure could also be interpreted as a relative buying interest index, where values close to 1 (-1) reflect high buying (selling) pressure.

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).

Consistent with our results based on *ex post* measures of favorable and unfavorable fill, Table 4, Columns 5 to 8 reveals that the decrease in *P(Fill)* in the post-period for non-HFT is driven by a fall in the favorable limit order executions, measured *ex ante*, rather than a fall in the unfavorable limit order executions. While Column 5 shows that the overall probability of a favorable fill increases in the post-ITCH period as the coefficient on *I(Post-ITCH)* is both positive and significant (0.036), this increase is not shared equally across players but primarily accrues to the HFT. Comparing between Columns 5 and 7, we find that *P(Favorable fill)* falls for non-HFT after the implementation of ITCH, as $I(Non-HFT) \times I(Post-ITCH)$ is -0.04, while *P(Unfavorable fill)* remains unchanged after the implementation of ITCH. Separating non-HFT into *Institutions* and *Retail*, we document similar findings: For *P(Favorable fill)*, the interaction terms for *I(Institutions)* and *I(Retail)* with *I(Post)* is negative and significant at -0.029 and -0.051 respectively, in Column 6 and insignificant in Column 8. Thus, the likelihood of receiving a favorable limit order execution falls for both *Institutions* and *Retail* when HFT gain a speed advantage.

Previous studies show that faster trading speeds improve market quality for the market order trader (e.g., through tighter spreads, increased depth and improved price efficiency). We provide new evidence that non-HFT limit orders face higher non-execution risk when trading becomes faster, which supports existing theory.

4 Strategic order placement strategies

In this section, we investigate one information channel that drives trader behavior to uncover how speed differentials could impact on the probability of non-HFT limit order executions. As we are concerned with the information channel overall, we start by using data from the entire period, and then examine the effects on trader behavior as the speed differential increases. We investigate strategic market and limit order management in Sections 4.1 and 4.2, respectively.

4.1 Market orders

To understand how faster data feeds may be differentially affecting HFT and non-HFT trading behavior, we first need to gain an understanding of the information channels traders use. In Yang and Zhu (2019)'s model, 'back-runners' exploit order flow information in their trading decisions. For reasons outlined in Section 3.3, we use the order book depth imbalance (DI) to proxy for the information contained in the order flow and examine how market participants respond to DI.

We start by ranking trades into deciles based on DI immediately before the trade for each stock-day. As we are interested in the information channel that affects trading strategies, we use data from the entire period. Since depth imbalances predict future returns, we expect strategic traders to trade in the direction of the order book imbalance. That is, more aggressive buying (i.e., more buyer initiated trades) when a large positive depth imbalance exists and more aggressive selling when a large negative depth imbalance exits.

Figure 2, Panels A to C present the results separately for *HFT*, *Institutions*, and *Retail*, respectively. Consistent with strategic trading, we observe a general positive (negative) relation 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 1, 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). This result is consistent with the notion that *HFT* are more likely to trade opportunistically on order book information compared to *Institutions* and *Retail*, who tend to trade for reasons exogenous to the order book. Thus, we expect *HFT* to trade most aggressively when large, profitable opportunities exist, which correspond to times of large depth imbalances. This behavior is consistent with Parlour (1998), who proposes that it is optimal for a trader to submit market orders, instead of limit orders, when trading in the direction of a large depth imbalance.

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:

$$VolumeImbalance_j^T = \frac{\sum_{k=1}^{n} BuyVolume_{k,j}^T - \sum_{k=1}^{n} SellVolume_{k,j}^T}{\sum_{k=1}^{n} BuyVolume_{k,j}^T + \sum_{k=1}^{n} SellVolume_{k,j}^T} \tag{4}$$

where $\sum_{k=1}^{n} BuyVolume_{k,j}$ ($\sum_{k=1}^{n} SellVolume_{k,j}^T$) is the total aggressive buying (selling) volume, i.e., volume from the submission of market or marketable buy (sell) limit orders, that occurs during 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 the size of DI predicts future returns (see Internet Appendix A.3), a steeper slope between *Volume imbalance* and *DI* signals that 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*.[20]

Brogaard, Hendershott, and Riordan (2014) show that HFT demand liquidity in the direction of the limit order book imbalance. Our results reveal that this behavior is not unique to HFT and that all broker types attempt to trade in the direction of a stock's depth imbalance. 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 imbalances. Arguably, this could be because HFT have a different trading objective function to non-HFT. While

[20] We arrive at similar findings if we standardize the depth imbalance by average trade size.

non-HFT are typically longer term investors, HFT generally trade to make short term gains and thus, *DI* should be a more important component of HFT trading decisions.

In Table 5, we test the sensitivity of volume imbalances to depth imbalances for our trader types after controlling for trading volumes, stock and day fixed effects over the entire period using the following regression:

$$\begin{aligned} VolumeImbalance_j^T = \quad & \beta_0 \quad + \beta_1 I(HFT)^T \times DI_j + \beta_2 I(Institutions)^T \times DI_j + \beta_3 I(HFT)^T \\ + \quad & \beta_4 I(Institutions)^T + \beta_5 DI_j + \beta_6 Volume_j + \epsilon^T \end{aligned} \tag{5}$$

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. Further, Anand and Venkataraman (2016), Kirilenko, Kyle, Samadi, and Tuzun (2017), Brogaard, Carrion, Moyaert, Riordan, Shkilko, and Sokolov (2018), van Kervel and Menkveld (2019) and Korajczyk and Murphy (2018) find HFT withdraw liquidity during periods of stress. We therefore also examine stressful periods, by separately estimating Equation 5 on subsamples of high and low volatility days.

[Insert Table 5]

Table 5, 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 *VolumeImbalance* is more sensitive to the level of *DI* in the order book, relative to the baseline.[21] Consistent with our earlier results from Figure 2, we find that the coefficient on $I(HFT) \times DI$ is positive and significant (1.046) indicating that relative to the other trader types, HFT are more likely to submit buyer initiated trades when *DI* is larger.

To investigate the effects of stock volatility on HFT trading behavior, for each stock, we rank

[21]In this regression, *Retail* is the omitted or baseline group and their sensitivity to *DI* is captured by the coefficient on *DI*.

trading days into terciles (low, medium, and high) based on the daily stock volatility.[22] Table 5, Columns 2 and 3, presents 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*.[23] 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 ($p - value = 0.004$). As will be shown later in Section 5, these results are not due to HFT responding faster to external public announcements, but instead, appear to be reactions to the limit order book itself.

Due to the limitations of our data, it is possible that some smaller proprietary HFT firms trade through institutional brokers and their order flow has been misclassified as *Institutions*. While it is difficult for these smaller, misclassified HFT traders to change overall volume imbalances of their large institutional brokers, due to their small trade size, these misclassified HFT traders could influence trading imbalances of their institutional brokers based on the number of trades, as HFT typically make many small but frequent trades. As robustness, to investigate this possibility, in Table 5, Columns 4-6, we replace the $VolumeImbalance$ dependent variable from Equation 5 with a trade imbalance dependent variable, which is based on the *number* of buyer and seller initiated trades, rather than the *volume* of buyer and seller initiated trades. Using the trade imbalance dependent variable, we find that $I(Institutions) \times DI$ is now positive and significant for the full sample (Column 4) and for the subsample of high volatility days (Column 6), which is qualitatively similar to our findings for HFT. These results suggest that smaller HFT firms who execute their order flow via an institutional broker, have similar behavior to the larger HFT firms who have their own exchange membership that we can accurately classify. In contrast, and consistent with expectations, our results for *HFT* and *Retail* remain unchanged regardless of whether we use volume imbalances or trade imbalances as the dependent variable.

[22]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.

[23]For further robustness, we reproduce the results in Table 5 using another order book signal, the weighted midpoint price (see Hagströmer (2021)). The weighted midpoint price uses information contained in the best bid and ask depths while our measure of *DI* uses information from the top 5 levels of the order book. Results using the weighted midpoint price are slightly weaker than results using *DI* suggesting that the order book contains information beyond that of the best bid and ask.

Large depth imbalances, which are strong predictors of future price movements, should not persist in the order book for extended periods of time. Figure 3, Panel A shows the inverted U-shaped relation between the size of a depth imbalance and the median time the limit order book remains within the depth imbalance decile. In these graphs, a balanced book is in the middle and books with extreme depth imbalances are at either ends. For this analysis, depth imbalance deciles are calculated based on all order book events for the entire period. Thus, depth imbalances can change due to order submissions, amendments, and cancelations, as well as trade executions.

Overall, 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 (milliseconds) in the state of the order book at the extremes, it is likely that only HFT are able to participate effectively in the extreme depth imbalance deciles. Consistent with this intuition, Figure 3, 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 3, Panels C and D), resulting in a strong inverted U-shaped pattern.

[Insert Figure 3]

The stark contrast between the graph for *HFT* (inverted U-shaped) and the reverse shape for *Retail* or *Institutions* (both U-shaped) provide further support for the notion that HFT pick off stale non-HFT limit orders (Hoffmann (2014), Biais, Foucault, and Moinas (2015), Budish, Cramton, and Shim (2015), Foucault, Hombert, and Roşu (2016), Foucault, Kozhan, and Tham (2016)). Market share aggressively increases for *HFT* in the extreme *DI* deciles, precisely the deciles that contain the most information on future price movements.

4.1.1 Depth imbalance and volume imbalances by trader type post-ITCH

Now that we have established that all traders respond to the shape of the limit order book, we investigate the effects of further increasing the relative speed differentials. Because HFT strategies are most likely to benefit from even faster trading speeds, we expect that *HFT Volume imbalance* is more sensitive to the level of *DI* after switching to ITCH, when their speed advantage increased. On the other hand, we expect a lesser effect on the slope of the relation between *Volume imbalance* and *DI* for *Retail* and *Institutions*, who are less speed sensitive and are unable to capture short-lived information in extreme depth imbalances. To assess whether ITCH affects trading behavior, we use a difference-in-difference framework and re-estimate Equation 5 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:

$$\begin{aligned} VolumeImbalance_j^T = \quad & \beta_0 \quad + I(Pre)\Big(\beta_1 I(HFT)^T \times DI_j + \beta_2 I(Institutions)^T \times DI_j + \beta_3 DI_j\Big) \\ & + \quad I(Post)\Big(\beta_4 I(HFT)^T \times DI_j + \beta_5 I(Institutions)^T \times DI_j + \beta_6 DI_j\Big) \\ & + \quad \beta_7 I(HFT)^T + \beta_8 I(Institutions)^T + \beta_9 Volume_j + \epsilon^T \end{aligned} \tag{6}$$

For ease of exposition, we calculate everything in terms of buy imbalances, so sell imbalances are included as negative numbers. Table 6, Columns 1 and 3, report the two sets of coefficients $\{\beta_1, \beta_2, \beta_3\}$ and $\{\beta_4, \beta_5, \beta_6\}$. Table 6, Column 4 reports the test of equality between the coefficients on the interaction between HFT indicator variables and the depth imbalance, 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 6, Column 4 also reports the test of equality between β_2 and β_5, and β_3 and β_6, which similarly 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 6, Column 2.

[Insert Table 6]

Table 6, 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 submit market buy orders when *DI* is positive and market sell orders 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 *Volume imbalance* against *DI* is steeper in the post-ITCH period. Collectively, as their market access speeds increase, HFT trade more strategically on information contained in the limit order book. In contrast, for *Institutions* and *Retail*, we cannot reject the null hypothesis that $\beta_2 = \beta_5$ and $\beta_3 = \beta_6$. Thus, as expected, we do not find evidence that non-HFT, who are less speed sensitive, trade more strategically conditional on the shape of the book after the adoption of ITCH.

To provide additional robustness and to address concerns that some smaller proprietary HFT firms could trade through institutional brokers with their order flow misclassified as *Institutions*, in Table 6, Columns 5 to 7, we replace the dependent variable with trade imbalance, which is computed using the *number* of buyer and seller initiated trades, rather than the *volume* of buyer and seller initiated trades. While it is unlikely for these small, misclassified HFT traders to change the overall volume imbalances of their larger institutional brokers, these HFT traders could influence institutional trade imbalances, as HFT typically trade with small order size but high frequency. This adjustment to the dependent variable aims to capture this phenomenon. Using the trade imbalance dependent variable, our results for *Institutions* become qualitatively similar to our results for *HFT*: we now find *Institutions* are more sensitive to *DI* after the implementation of ITCH (F-test = 5.27). This result is consistent with our earlier results for HFT in Section 4.1, and suggest 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. In contrast, for *HFT* and *Retail*, where potential misclassification is not a concern, we obtain the same results regardless of whether the dependent variable is measured using volume or trade imbalances.

4.2 Limit order strategies

In Section 4.1, we investigate strategic trading through the relation between the order book depth imbalance and volume imbalances, which are based on *market* order executions. In this section, we analyze strategic trading more comprehensively by investigating how *limit* order strategies (i.e., passive executions, order submissions, amendments or cancelations) differ between investor categories.

We call each passive execution or the arrival of an order submission, amendment, or cancelation an order book 'event' and measure *DI* immediately before every event, as described in Section 4.1. To remove the effects of order direction, we multiply *DI* by an indicator for whether the order rests on the bid (+1) or offer (-1), resulting in the *Adjusted DI* measure, which allows bids and offers to be interpreted together.

An *Adjusted DI* value of 0 denotes 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 event.[24] For example, a negative *Adjusted DI* at the time of a passive bid execution shows the ask depth exceeds the bid depth at the time the limit order executes from the bid price. Since a larger ask depth, relative to the bid depth, predicts a future price fall, a negative *Adjusted DI* indicates that the order is picked off from the thin side of the order book. That is, the limit order trader is buying before an expected future price fall.

Table 7 reports the average *Adjusted DI* for each event (i.e., passive execution, submission, amendment or cancelation) for our three trader types. Notably, not all limit orders face the same picking off risk: The average *Adjusted DI* for passive executions for *HFT*, *Institutions* and *Retail* is 0.083, -0.029 and -0.009 respectively. These values highlight that HFT passive order executions typically occur on the thick side of the book (i.e., positive *Adjusted DI*), whereas non-HFT passive executions are, on average, picked off from the thin side of the order book (i.e., negative *Adjusted DI*).

[24] In the case of a passive buy execution, 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.

[Insert Table 7]

The findings from Table 7 reveal a pattern of strategic trading by HFT: 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 emphasize that a trader does not have full control of the *Adjusted DI* for passive executions, as these executions rely on an incoming market order from another market participant. However, if fast enough, the trader can manage their limit order up until execution to ensure their limit orders do not become stale.

To investigate the order management behavior of HFT in a more formal setting, 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. We estimate the following regression on HFT orders, controlling for stock and day fixed effects:[25]

$$\begin{aligned} OrderBookEvent_E = \quad & \beta_0 \quad + \beta_1 AdjustedDI_E + \beta_2 Volatility + \beta_3 Volume \\ + \quad & \beta_4 Price + \beta_5 Spread + \epsilon_E \end{aligned} \tag{7}$$

where $OrderBookEvent_E$ is the dependent variable indicating one of four order book events, E: Passive execution, limit order submission, amendment or cancelation. $AdjustedDI_E$ 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

[25]For robustness, in unreported tests, we run the regression for each stock separately and find similar results.

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 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).

[Insert Table 8]

Table 8 reports a strong positive association between HFT passive executions and the depth imbalance. Confirming our earlier findings, *Adjusted DI* is 0.467 and statistically significant at the 1% level for passive executions, revealing 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. 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 than it was at the time of the original limit order submission. In other words, as the market conditions worsen, HFT cancel stale limit orders before these orders can be picked off the order book by other traders.[26]

This strategic behavior provides a mechanism through which HFT crowd out slower market participants from the limit order book. HFT infer information from the order flow of slower traders, behavior consistent with the 'back-runners' from Yang and Zhu (2019). By using order book information in their limit order strategies, HFT limit orders contribute to short term price discovery as in Brogaard, Hendershott, and Riordan (2019). However, by piggy-backing off the fundamental information of slower traders, this improvement in short term price informativeness could come at a cost in the longer term as slower market participants have fewer incentives to acquire information. (Weller (2017), Baldauf and Mollner (2020)).

[26]For further robustness and to assess the probability of fill more directly, we replace $OrderBookEvent_E$ with a 1 or 0 dependent variable, where filled limit orders are equal to 1 and unfilled limit orders (i.e., those that are subsequently amended or canceled) are equal to 0. In unreported results, we also find that *HFT* limit orders execute in the direction of the order book imbalance while *Institutions* and *Retail* receive passive executions when the order book moves against them.

5 Endogeneity

In the previous section, 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 5 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 4 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 *Institutions* and *Retail* regardless of the level of public news arrival.

[Insert Figure 4]

Next, we formally test the relation between *Volume imbalance* (or *Trade imbalance*) and *DI* for each of our trader types. Table 9 shows the regressions results for our high and low information subsamples. Consistent with the full sample results from Table 5, 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 9]

6 Conclusion

Recent theory models suggest that HFTs use their speed advantage to crowd out slower traders from the limit order book. For example, Hoffmann (2014) and Han, Khapko, and Kyle (2014) suggest that HFTs use market orders to pick off stale limit orders, leading to higher adverse selection costs for slower traders. In Li, Wang, and Ye (2021), faster traders use speed to compete for queue priority, which results in a crowding out effect. However, empirical research testing these predictions is sparse due predominantly to the lack of highly granular data. Further, while Li, Wang, and Ye (2021) identify one channel that drives the crowding out effect, little is known about other potential information channels that HFT use to crowd out other traders from the order book. In this study, we use detailed order book data to analyze the impact of HFT on limit order traders, namely institutional and retail traders.

We document three main findings. First, confirming the crowding our effect empirically documented in Yao and Ye (2018), we find that the probability of limit order execution decreases for slower traders when market access speeds increase. Exploiting the additional granularity in our data, we show a larger deterioration in limit order executions for slower retail traders than institutional traders. The crowding out effect is also stronger for tick constrained firms and for firms with a larger relative tick size, which is consistent with the predictions of Li, Wang, and Ye (2021). Importantly, we also document an asymmetry in the crowding out effect: the decrease in execution probabilities is driven by a decrease in the probability of favorable executions, which corresponds to an increase in adverse selection risk for slower traders. To the best of our knowledge, this result documenting the impact of HFT on the limit order trader is novel to the literature.

Second, we show that the limit order book is an important information channel for HFT. While Brogaard, Hendershott, and Riordan (2014) show that HFT respond to limit order book imbalances, we show that all traders attempt to trade in the direction of the order book but HFT are more successful at doing so, especially at times of high market volatility. Because large depth imbalances only persist in the order book for a fraction of a second, speed is critical to strategies that take advantage of order book information: HFT become even more successful at trading in the direction of the order book when they gain a larger speed advantage. Theoretical models show that

HFT use market orders to pick off stale limit orders (Biais, Foucault, and Moinas (2015), Budish, Cramton, and Shim (2015), Foucault, Hombert, and Roşu (2016), Foucault, Kozhan, and Tham (2016)). We show that HFT use the order book depth imbalance to pick off stale limit orders from the thin side of the order book.

Third, we contribute to the growing literature on limit order price discovery. Brogaard, Hendershott, and Riordan (2019) show limit orders, in particular limit orders submitted by HFT, provide price discovery. However, little is known about the information channels HFT used in their limit order strategies: We show that order book information is one important channel. Thus, while HFT limit orders add to price discovery, they do so by 'piggybacking' off the fundamental information of other limit order traders, which increases the adverse selection risk for slower traders and decreases their execution probabilities. While this behavior improves price efficiency in the short-term, Weller (2017) and Baldauf and Mollner (2020) argue that longer-term price efficiency could deteriorate as faster traders erode information rents of information acquirers.

By using the order book depth imbalance in their limit order submission and cancelation decisions, HFT on average supply liquidity on the thick side of the order book but demand liqudiity from the thin side of the order book. These asymmetric findings suggest that caution should be taken in interpreting results that aggregate market depth across both the bid and ask sides.

A Internet Appendix

This Internet Appendix, which comprises five sections, provides additional details and robustness to the results reported in the main body of the paper. Section A.1 provides additional details on the trader classification process. Section A.2 provides additional robustness for the ITCH experiment. Sections A.2.1 and A.2.2 investigate changes to trader reaction times and cancelation rates after the adoption of ITCH, respectively, and Section A.2.3 presents the results of a falsification test. Section A.3 confirms the positive relation between depth imbalance and future stock returns as documented in Cao, Hansch, and Wang (2009) and Cont, Kukanov, and Stoikov (2013). Section A.4 investigates the profitability of the depth imbalance signal. Section A.5 corresponds to our multinomial logit analysis in Section 4.2 and provides additional results for the order management behavior of non-HFT trader types.

A.1 Broker identification

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*, *Institutions* 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.[27] Over our sample period, there are approximately 100 registered market participants with unique broker identifiers.

We classify these 100 registered market participants based on their primary trading activity or client base. We identify HFT firms that have exchange membership, many of which are identified in van Kervel and Menkveld (2019), if their primary business is high frequency proprietary trading (e.g., Virtu and Getco). Thus, similar to Nasdaq HFT data, our HFT category captures the activity

[27] 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.

of pure HFT firms. Similarly, we are able to identify brokerage firms that cater predominately to retail investors (e.g., E-Trade and CommSec), which we classify as *Retail*. Last, we identify large institutional brokers (e.g., Goldman Sachs and Deutsche Bank) if their main client base is institutional investors. 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 *Institutions*.

A.2 Robustness tests for ITCH experiment

In this Appendix, we provide additional robustness tests for the ITCH analysis.

A.2.1 Response times by broker type

To test whether HFT become faster after the implementation of ITCH, we follow Baron, Brogaard, Hagströmer, and Kirilenko (2019) and compute a measure of response time, which captures the reaction time of an HFT firm, with an adjustment to account for our more granular order level data. In contrast to Baron, Brogaard, Hagströmer, and Kirilenko (2019), who analyze time differences between a passive trade and an aggressive trade by the same firm, in the same stock and at the same trading venue, we exploit the order book level data and analyze the difference in time stamps from a passive trade to any subsequent order book event (i.e., passive or aggressive order submission, amendment or cancelation) by the same firm in the same stock. For each day, we identify all events where a passive trade is followed by a subsequent order book event by the same broker in the same stock within one second. Next, to capture the fastest possible reaction times, we identify the daily percentage of these events that occur within a 1 millisecond time stamp for each broker.[28] We refer to this measure as $FastResponse\%$.

To determine whether HFT respond more quickly, we compare $FastResponse\%$ before and after the implementation of ITCH in a difference-in-difference framework. Because *Institutions* could contain some HFT transactions (see Footnote 8), we compare HFT response times against that

[28]We are unable to use the 0.1% quantile threshold of Baron, Brogaard, Hagströmer, and Kirilenko (2019) as this threshold frequently falls within the one millisecond granularity of our data.

of retail response times. Table AI, Panel A presents the mean $FastResponse\%$ for HFT and Retail brokers before and after the implementation of ITCH. The percentage of events occurring within the 1 millisecond time stamp approximately doubles after ITCH for *HFT* ($FastResponse\%$ increases from 1.62% to 2.97%) while $FastResponse\%$ remains unchanged for *Retail*. Table AI, Panel B, Column 1 shows that the interaction term $I(HFT) \times I(Post - ITCH)$ is positive and significant confirming that HFTs respond more quickly after they connect to the faster data feed. In contrast, our placebo tests based on sub-periods in the pre-event window as discussed in Section A.2.3 do not indicate a significant change in HFT response times in the month before the implementation of ITCH (Table AI, Panel B, Column 2).

Table AI
Reaction time

Table AI analyzes the reaction time for HFT 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). For the placebo experiment, we analyze the periods March 1, 2012 to March 14, 2012 (pre-period) and March 16, 2012 to March 30, 2012 (post-period). To measure reaction time, we compute the variable *FastResponse%*. For each day, we identify all events where a passive trade is followed by an order book event by the same firm in the same stock within one second. Of these events, we compute the daily percentage of these events that occur within a 1 millisecond time stamp for each firm. Panel A reports the mean *FastResponse%* for each broker type. Panel B reports the regression results for the following regression model:

$$FastResponse\%^T = \beta_0 \quad + \beta_1 I(HFT)^T \times I(Post) + \beta_2 I(HFT)^T + \beta_3 I(Post) + \epsilon^T$$

where $I(HFT)^T$ is an indicator variable equal to 1 for HFT and zero for *Retail*. $I(Post)$ is an indicator variable equal to 1 if the trading day falls in the post-ITCH period and zero for the pre-ITCH period.***, ** and * indicate significance levels of 1%, 5% and 10%, respectively.

	ITCH		Placebo	
Panel A: Summary statistics				
	Pre	Post	Pre	Post
HFT	1.624	2.972***	1.646	1.602
Retail	0.013	0.018	0.009	0.016
Panel B: Regression analysis				
I(HFT) × I(Post-ITCH)	1.343	***	-0.051	
	(3.58)		-(0.41)	
I(HFT)	1.611	***	1.637	***
	(6.07)		(18.57)	
I(Post-ITCH)	0.005		0.007	
	(0.02)		(0.08)	
Constant	0.013		0.009	
	(0.07)		(0.15)	
Obs.	88		44	
Adj. R-square	0.662		0.939	

A.2.2 Cancelation rates by broker type

While the faster data feed improves the response times of HFT, it is unlikely for non-HFT to intentionally change their order placement strategies *immediately* after the implementation of ITCH for several reasons. First, it is unlikely that large institutions have the capability to significantly change all their algorithms in the few days after the implementation of ITCH. Substantial changes in trading strategies is also unlikely for *Retail*, who are typically manual traders on the Australian Securities Exchange. To confirm this intuition, we compare $Cancel\%$, the number of cancels as a percentage of all submissions, for the trader types around the time of ITCH. In Table AII, we do not reject the hypothesis that there were no significant changes to $Cancel\%$ for our trader types after the implementation to ITCH.

Table AII
Cancelation rates before and after the implementation of ITCH

Table AII analyzes the cancelation rates 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). For each trader type T, we calculate the daily *Cancel%* by dividing the number of order cancelations by the total number of orders submitted. In Column 1, we estimate the following regression:

$$\begin{aligned} Cancel\%^T = & \quad \beta_0 \quad + \beta_1 I(Non-HFT)^T \times I(Post) + \beta_2 I(Non-HFT)^T \\ + & \quad \beta_3(Post) + \beta_4 Volatility + \beta_5 Volume + \beta_6 Price + \beta_7 Spread + \epsilon^T \end{aligned}$$

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. 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.

	Cancel%	
	(1)	(2)
I(Non-HFT) × I(Post-ITCH)	-0.556	
	(-0.54)	
I(Non-HFT)	-13.022***	
	(-12.82)	
I(Institutions) × I(Post-ITCH)		-0.790
		(-0.67)
I(Institutions)		-4.257***
		(-4.04)
I(Retail) × I(Post-ITCH)		-0.211
		(-0.20)
I(Retail)		-23.260***
		(-22.13)
I(Post-ITCH)	1.903**	1.933**
	(2.61)	(2.68)
Volatility	-16.963	-11.279
	(-1.68)	(-1.18)
Volume	-1.269***	-1.026***
	(-4.53)	(-3.84)
Price	-2.624***	-2.565***
	(-3.70)	(-4.09)
Qspread	17.476	-1.917
	(0.24)	(-0.02)
Constant	56.847***	53.074***
	(11.86)	(11.53)
Obs.	10,435	10,435
Adj. R-square	0.287	0.595

A.2.3 Placebo test

We perform a placebo test in our pre-event period to test whether the parallel trends assumption holds between HFT and non-HFT trading activity in the pre-ITCH period. For this test, we falsely assume that the event occurs on March 15, 2012, which is in the middle of our pre-event window. Thus, we evaluate Equation 2 with a pre-event window from March 1 to 14, 2012 and a post-event window from March 16 to 31, 2012. Table AIII shows that the effect is isolated to the period after the implementation of ITCH: $I(Non-HFT) \times I(Post-ITCH)$ is indistinguishable from zero in all specifications of the placebo experiment. Similarly, the interaction terms of $I(Institutions)$ and $I(Retail)$ with $I(Post-ITCH)$ are indistinguishable from zero.

Table AIII
Placebo

Table AIII provides a falsification test of the ITCH experiment. We analyze trade and quote data for the periods March 1, 2012 to March 14, 2012 (pre-period) and March 16, 2012 to March 30, 2012 (post-period). The 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:

$$\begin{aligned} P(Fill)^T = \quad & \beta_0 \quad + \beta_1 I(Non-HFT)^T \times I(Post) + \beta_2 I(Non-HFT)^T \\ + \quad & \beta_3 I(Post) + \beta_4 Volatility + \beta_5 Volume + \beta_6 Price + \beta_7 Spread + \epsilon^T \end{aligned}$$

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 Columns 3 and 4 (5 and 6), we replace the dependent value with $P(Favorablefill)^T$ ($P(Unfavorblefill)^T$). 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. In Columns 2, 4 and 6, 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. 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)	(4)	(5)
I(Non-HFT) × I(Post-ITCH)	0.004		0.015		-0.005	
	(0.37)		(1.25)		(-0.75)	
I(Non-HFT)	0.107***		-0.038*		0.137***	
	(5.17)		(-1.97)		(18.41)	
I(Institutions) × I(Post-ITCH)		0.008		0.015		-0.001
		(0.75)		(1.44)		(-0.18)
I(Institutions)		-0.075***		-0.133***		0.053***
		(-3.66)		(-7.08)		(8.25)
I(Retail) × I(Post-ITCH)		0.001		0.014		-0.010
		(0.05)		(0.96)		(-0.98)
I(Retail)		0.299***		0.071***		0.230***
		(13.70)		(3.47)		(23.60)
I(Post-ITCH)	0.000	-0.003	-0.002	-0.006	0.003	0.002
	(0.01)	(-0.18)	(-0.15)	(-0.43)	(0.23)	(0.19)
Volatility	-0.506**	-0.554**	-0.200	-0.396	-0.121	-0.238
	(-2.19)	(-2.38)	(-0.79)	(-1.58)	(-0.68)	(-1.37)
Volume	0.077***	0.075***	0.046***	0.045***	0.024***	0.023***
	(16.83)	(16.76)	(9.74)	(9.14)	(6.24)	(5.91)
Price	-0.048	-0.044	-0.162*	-0.184**	0.061	0.047
	(-0.67)	(-0.61)	(-1.98)	(-2.16)	(1.05)	(0.79)
Qspread	0.881	0.537	2.659*	2.055	-1.912	-1.877
	(0.51)	(0.32)	(1.74)	(1.43)	(-1.48)	(-1.44)
Constant	-0.577***	-0.543***	-0.204*	-0.143	-0.237***	-0.199**
	(-5.44)	(-5.10)	(-1.85)	(-1.25)	(-2.82)	(-2.41)
Obs.	5622	5622	5131	5131	5063	5063
Adj. R-square	0.190	0.617	0.129	0.373	0.210	0.466

A.3 Depth imbalance and future returns

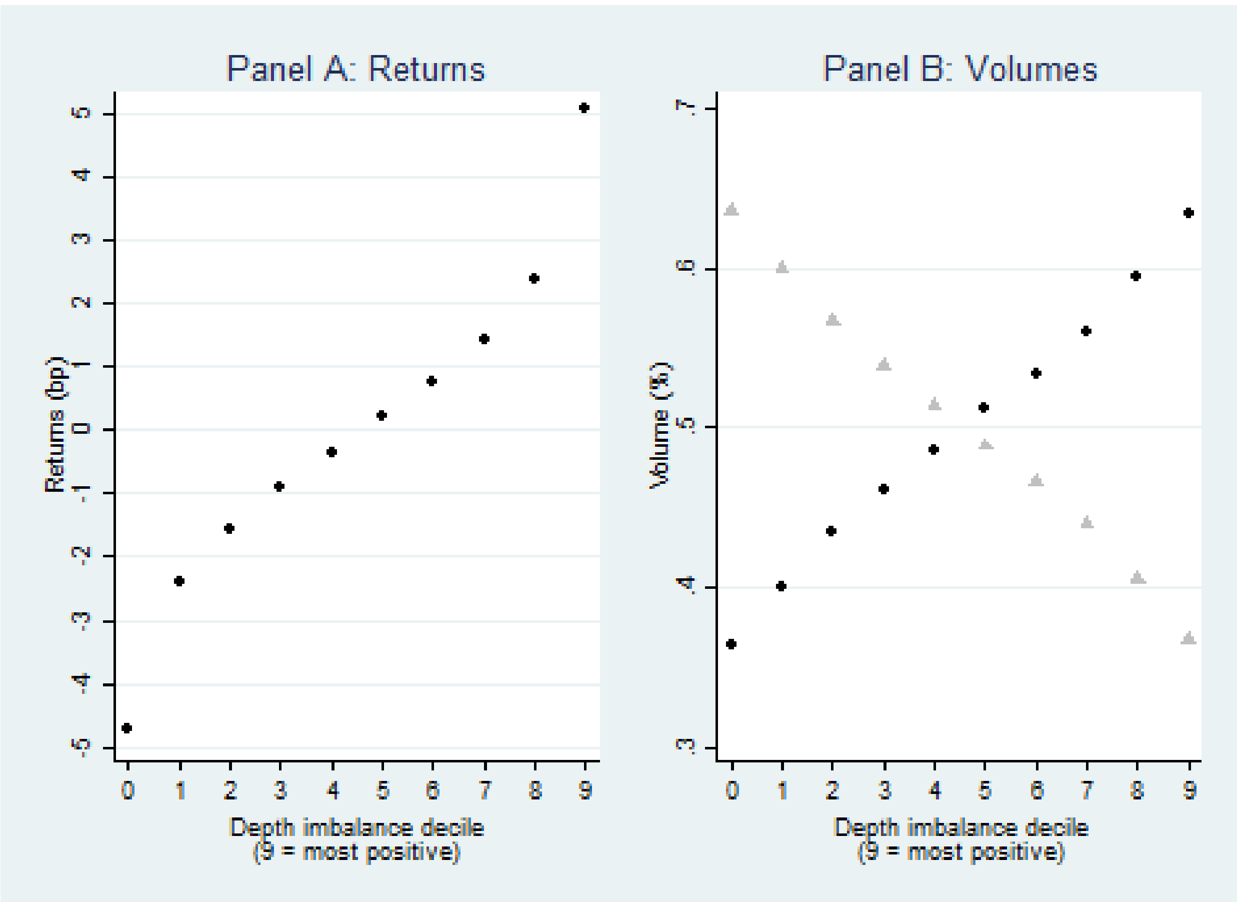
In this Appendix, we confirm the results of Cao, Hansch, and Wang (2009) and Cont, Kukanov, and Stoikov (2013), who find strong evidence that buy and sell order book depth imbalances are significantly related to future stock returns. To investigate whether depth imbalances contain information about the 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 A.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.[29] Specifically, as the number of buyers relative to sellers in the limit order book increase, the relative level of stock prices in the future also increases.

Based on the results from Ranaldo (2004), who examines how the state of the limit order book can affect a trader's order submission strategy, for each depth imbalance decile, we also 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 A.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.

[29] Cont, Kukanov, and Stoikov (2013) 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.

Figure A.1. Depth imbalance and future returns

Fig. A.1 shows the relations 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.



A.4 Profitability of Depth Imbalance signal

In our main analysis, we show that HFTs make extensive use of the DI signal in their order placement strategies. Here, we investigate the profitability of the DI signal using a measure of gross trading profits based on Menkveld (2013) and Brogaard, Hendershott, and Riordan (2014):

$$PROFIT_t = \sum_{i=1}^{n} D_{i,t} \times PRICE_{i,t} \times QUANTITY_{i,t} \times -1 + INV_t \times REFPRICE, \qquad \text{(A.1)}$$

where D is equal to 1 for buys and -1, $PRICE$ is the match price and $QUANTITY$ is the number of shares transacted for i on day t. At the end of the trading day, t, INV is the number of shares outstanding. We assume that all inventory is closed out in the closing auction at the closing auction price, $REFPRICE$, as in Brogaard, Hendershott, and Riordan (2014).

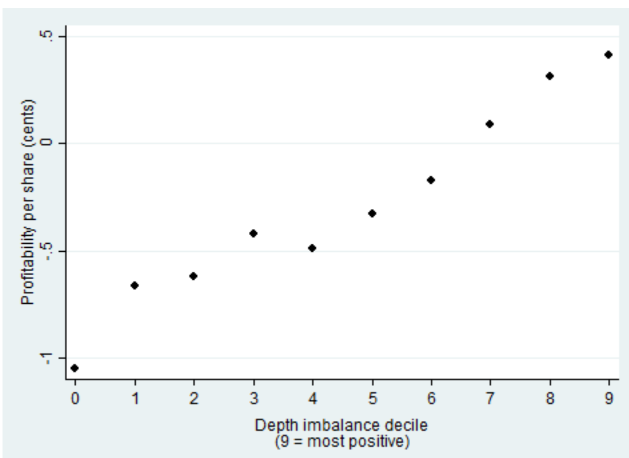
For each stock and depth imbalance decile, we compute $PROFIT_t$ for all transactions by the same broker. We then divide $PROFIT_t$ by the volume of shares traded by the broker. Figure A.2 presents the average gross profit generated by the DI signal for each share traded for HFT. We see a clear positive relation between the depth imbalance decile and $PROFIT_t$: HFT make approximately half a tick (i.e., 0.5 cents) per share when they buy in the direction of a large positive depth imbalance but lose approximately one tick per share when they buy in the direction of a large negative depth imbalance.

Figure A.2. Depth imbalance and HFT profitability

Fig. A.2 shows the relation between depth imbalance and the average profitability per share traded by HFT.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 compute gross trading profits as follows:

$$PROFIT_t = \sum_{i=1}^{n} D_{i,t} \times PRICE_{i,t} \times QUANTITY_{i,t} \times -1 + INV_t \times REFPRICE$$

where D is equal to 1 for buys and -1, $PRICE$ is the match price and $QUANTITY$ is the number of shares transacted for i on day t. At the end of the trading day, t,INV is the number of shares outstanding and $REFPRICE$ is the closing auction price. We compute $PROFIT_t$ for all transactions by the same broker and divide by the volume of shares traded by the broker. The figure reports the average $PROFIT_t$ for all HFT.



A.5 Order management behavior of non-HFT

In this appendix, we extend our analysis of HFT order management behavior to non-HFT. Specifically, we apply the multinomial logistic regression model from Equation 7 to *Institutions* and *Retail* and report the results in Table AIV. The results show that *Institutions* and *Retail* are less strategic in their order placement strategies. In Section 4.2, we show that HFT receive passive executions when *Adjusted DI* is positive, reflecting strategic order management strategies (i.e., HFT cancel orders before they are adversely selected). Table AIV shows that both *Institutions* (Panel A) and *Retail* (Panel B) 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.

Table AIV

Multinomial logistic regressions for limit order placement strategies for non-HFT

Table AIV assesses the probability of each order book event based on prevailing market conditions. We present the coefficient estimates for the following multinomial logistic regression on institutional (Panel A) and retail (Panel B) orders:

$$
\begin{aligned}
OrderBookEvent_E = \quad & \beta_0 \quad + \beta_1 AdjustedDI_E + \beta_2 Volatility + \beta_3 Volume \\
+ \quad & \beta_4 Price + \beta_5 Spread + \epsilon_E
\end{aligned}
$$

where $OrderBookEvent_E$ is the dependent variable indicating one of four order book events, E: 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:

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

where $\sum_{i=1}^{n} DepthBid_{i,t}$ ($\sum_{i=1}^{n} DepthAsk_{i,t}$) 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. All regressions control for stock and day fixed effects. t-statistics are reported in parentheses. ***, ** and * indicate significance levels of 1%, 5% and 10%, respectively.

	Passive execution		Amendment		Cancelation	
Panel A: Institutions						
Adjusted DI	-0.381	***	-0.003	*	0.069	***
	(0.002)		(0.002)		(0.002)	
Volatility	2.755	***	0.224	***	-2.455	***
	(0.058)		(0.049)		(0.054)	
Volume	-0.020	***	0.010	***	0.002	***
	(0.001)		(0.000)		(0.001)	
Price	0.010	***	0.046	***	-0.063	***
	(0.001)		(0.001)		(0.001)	
Spread	-10.702	***	5.238	***	-0.160	
	(0.224)		(0.170)		(0.200)	
Constant	-0.987	***	-1.212	***	-0.962	***
	(0.010)		(0.008)		(0.009)	
Obs.			33,808,108			
Pseudo R-square			0.003			
Panel B: Retail						
Adjusted DI	-0.360	***	-0.266	***	-0.134	***
	(0.010)		(0.009)		(0.015)	
Volatility	4.248	***	-7.229	***	9.362	***
	(0.270)		(0.281)		(0.390)	
Volume	-0.156	***	0.187	***	-0.157	***
	(0.003)		(0.003)		(0.005)	
Price	0.008		0.005		-0.068	***
	(0.005)		(0.005)		(0.008)	
Spread	4.949	***	-13.234	***	28.592	***
	(1.425)		(1.445)		(2.040)	
Constant	1.660	***	-3.813	***	0.394	***
	(0.052)		(0.052)		(0.079)	
Obs.			998,682			
Pseudo R-square			0.014			

References

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

Anand, A., and K. Venkataraman, 2016, "Market conditions, fragility, and the economics of market making," *Journal of Financial Economics*, 121(2), 327 – 349.

Aquilina, M., E. Budish, and O'Neill, 2020, "Quantifying the High-Frequency Trading 'Arms Race': A Simple New Methodology and Estimates," *Working Paper*, University of Chicago and the Financial Conduct Authority.

Baldauf, M., and J. Mollner, 2020, "High-Frequency Trading and Market Performance," *The Journal of Finance*, 75(3), 1495–1526.

Baron, M., J. Brogaard, B. Hagströmer, and A. Kirilenko, 2019, "Risk and Return in High-Frequency Trading," *Journal of Financial and Quantitative Analysis*, 54(3), 993–1024.

Bhattacharya, A., and G. Saar, 2020, "Information and the Limit Order Market," *Working Paper*, Cornell University.

Biais, B., T. Foucault, and S. Moinas, 2015, "Equilibrium fast trading," *Journal of Financial Economics*, 116(2), 292 – 313.

Boehmer, E., D. Li, and G. Saar, 2018, "The Competitive Landscape of High-Frequency Trading Firms," *The Review of Financial Studies*, 31(6), 2227–2276.

Brogaard, J., A. Carrion, T. Moyaert, R. Riordan, A. Shkilko, and K. Sokolov, 2018, "High frequency trading and extreme price movements," *Journal of Financial Economics*, 128(2), 253 – 265.

Brogaard, J., B. Hagströmer, L. Nordén, and R. Riordan, 2015, "Trading Fast and Slow: Colocation and Liquidity," *Review of Financial Studies*, 28(12), 3407–3443.

Brogaard, J., T. Hendershott, S. Hunt, and C. Ysusi, 2014, "High-Frequency Trading and the Execution Costs of Institutional Investors," *Financial Review*, 49(2), 345–369.

Brogaard, J., T. Hendershott, and R. Riordan, 2014, "High-Frequency Trading and Price Discovery," *The Review of Financial Studies*, 27(8), 2267–2306.

Brogaard, J., T. Hendershott, and R. Riordan, 2019, "Price Discovery without Trading: Evidence from Limit Orders," *The Journal of Finance*, 74(4), 1621–1658.

Budish, E., P. Cramton, and J. Shim, 2015, " The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response," *The Quarterly Journal of Economics*, 130(4), 1547–1621.

Cao, C., O. Hansch, and X. Wang, 2009, "The information content of an open limit-order book," *Journal of Futures Markets*, 29(1), 16–41.

Chakrabarty, B., P. Moulton, and X. F. Wang, 2020, "Limited Attention and Market Pricing of Earnings in a High Frequency World," *Working Paper*, St Louis University.

Chordia, T., R. Roll, and A. Subrahmanyam, 2002, "Order imbalance, liquidity, and market returns," *Journal of Financial Economics*, 65(1), 111–130.

Chordia, T., and A. Subrahmanyam, 2004, "Order imbalance and individual stock returns: Theory and evidence," *Journal of Financial Economics*, 72(3), 485–518.

Cont, R., A. Kukanov, and S. Stoikov, 2013, "The Price Impact of Order Book Events," *Journal of Financial Econometrics*, 12(1), 47–88.

Ellis, K., R. Michaely, and M. O'Hara, 2000, "When the Underwriter Is the Market Maker: An Examination of Trading in the IPO Aftermarket," *The Journal of Finance*, 55(3), 1039–1074.

Foucault, T., J. Hombert, and I. Roşu, 2016, "News Trading and Speed," *The Journal of Finance*, 71(1), 335–382.

Foucault, T., R. Kozhan, and W. W. Tham, 2016, "Toxic Arbitrage," *The Review of Financial Studies*, 30(4), 1053–1094.

Goettler, R. L., C. A. Parlour, and U. Rajan, 2009, "Informed traders and limit order markets," *Journal of Financial Economics*, 93(1), 67–87.

Hagströmer, B., 2021, "Bias in the effective bid-ask spread," *Journal of Financial Economics*, 142(1), 314–337.

Han, J., M. Khapko, and A. Kyle, 2014, "Liquidity with High-Frequency Market Making," *Working paper*, University of Maryland.

Handa, P., and R. A. Schwartz, 1996, "Limit Order Trading," *The Journal of Finance*, 51(5), 1835–1861.

Hasbrouck, J., and G. Saar, 2013, "Low-latency trading," *Journal of Financial Markets*, 16(4), 646 – 679.

Hendershott, T., C. M. Jones, and A. J. Menkveld, 2011, "Does Algorithmic Trading Improve Liquidity?," *The Journal of Finance*, 66(1), 1–33.

Hoffmann, P., 2014, "A dynamic limit order market with fast and slow traders," *Journal of Financial Economics*, 113(1), 156 – 169.

Kirilenko, A., A. S. Kyle, M. Samadi, and T. Tuzun, 2017, "The Flash Crash: High-Frequency Trading in an Electronic Market," *The Journal of Finance*, 72(3), 967–998.

Korajczyk, R. A., and D. Murphy, 2018, "High-Frequency Market Making to Large Institutional Trades," *The Review of Financial Studies*, 32(3), 1034–1067.

Kwan, A., R. Philip, and A. Shkilko, 2021, "The conduits of price discovery: A machine learning approach," *Working Paper*, University of Sydney.

Lee, C. M. C., and M. J. Ready, 1991, "Inferring Trade Direction from Intraday Data," *The Journal of Finance*, 46(2), 733–746.

Li, S., X. Wang, and M. Ye, 2021, "Who Provides Liquidity, and When?," *Journal of Financial Economics*, 141(3), 968–980.

Malceniece, L., K. Malcenieks, and T. J. Putniņš, 2019, "High frequency trading and comovement in financial markets," *Journal of Financial Economics*, 134(2), 381 – 399.

Malinova, K., and A. Park, 2020, "'Sniping' in Fragmented Markets," *Working Paper*, McMaster University and University of Toronto.

Malinova, K., A. Park, and R. Riordan, 2018, “Do Retail Investors Suffer from High Frequency Traders?,” *Working Paper*, McMaster University, University of Toronto and Queen’s University.

Menkveld, A. J., 2013, “High frequency trading and the new market makers,” *Journal of Financial Markets*, 16(4), 712–740.

Parlour, C. A., 1998, “Price Dynamics in Limit Order Markets,” *The Review of Financial Studies*, 11(4), 789–816.

Ranaldo, A., 2004, “Order aggressiveness in limit order book markets,” *Journal of Financial Markets*, 7(1), 53 – 74.

Riccó, R., B. Rindi, and D. J. Seppi, 2020, “Information, Liquidity, and Dynamic Limit Order Markets,” *Working Paper*, Bocconi University.

Roşu, I., 2019, “Fast and slow informed trading,” *Journal of Financial Markets*, 43, 1 – 30.

——— , 2020, “Liquidity and Information in Limit Order Markets,” *Journal of Financial and Quantitative Analysis*, 55(6), 1792–1839.

Shkilko, A., and K. Sokolov, 2020, “Every Cloud Has a Silver Lining: Fast Trading, Microwave Connectivity, and Trading Costs,” *The Journal of Finance*, 75(6), 2899–2927.

Upson, J., T. McInish, and H. Johnson, 2021, “Order based versus level book trade reporting: An empirical analysis,” *Journal of Banking and Finance*, 125.

van Kervel, V., 2015, “Competition for Order Flow with Fast and Slow Traders,” *The Review of Financial Studies*, 28(7), 2094–2127.

van Kervel, V., and A. J. Menkveld, 2019, “High-Frequency Trading around Large Institutional Orders,” *The Journal of Finance*, 74(3), 1091–1137.

von Beschwitz, B., D. B. Keim, and M. Massa, 2019, “First to ‘Read’ the News: News Analytics and Algorithmic Trading,” *The Review of Asset Pricing Studies*, 10(1), 122–178.

Weller, B. M., 2017, “Does Algorithmic Trading Reduce Information Acquisition?,” *The Review of Financial Studies*, 31(6), 2184–2226.

Yang, L., and H. Zhu, 2019, “Back-Running: Seeking and Hiding Fundamental Information in Order Flows,” *The Review of Financial Studies*, 33(4), 1484–1533.

Yao, C., and M. Ye, 2018, “Why Trading Speed Matters: A Tale of Queue Rationing under Price Controls,” *The Review of Financial Studies*, 31(6), 2157–2183.

Figure 1. Probability of limit order executions around the implementation of ITCH

Fig. 1 shows the probability of limit order executions ($P(Fill)$) for *HFT* and *Retail* around the implementation of ITCH on April 2, 2012. $P(Fill)$ is defined 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 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. We present daily measures of $P(Fill)$ for *HFT* and *Retail*, averaged across all stocks. The plot also indicates the pre-ITCH (March 2, 2012 to March 30, 2012) and post-ITCH (April 9, 2012 to May 9, 2012) periods for our analysis.

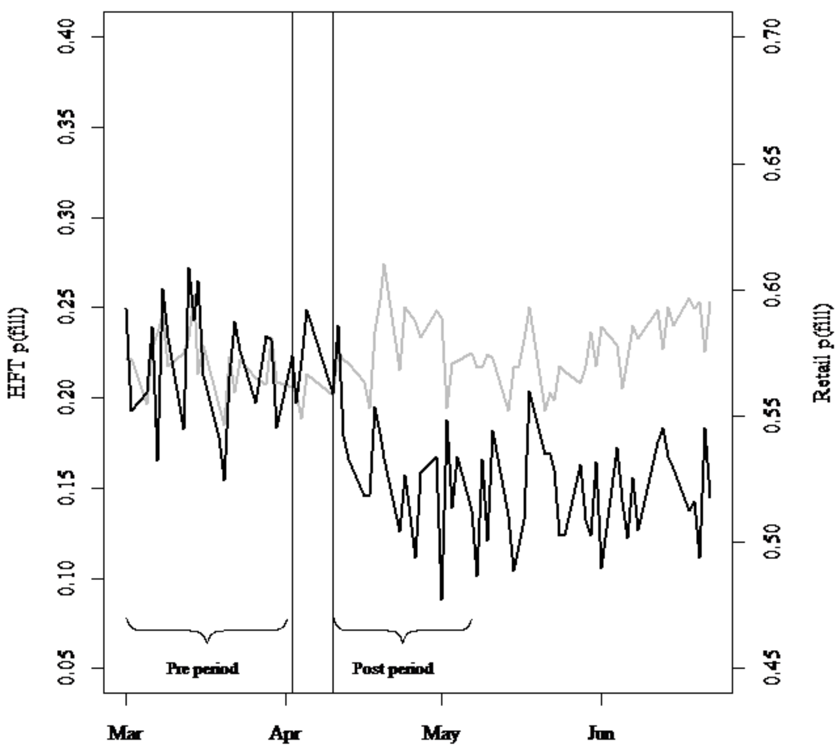


Figure 2. Depth imbalance and trading volumes

Fig. 2 shows the relation 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.

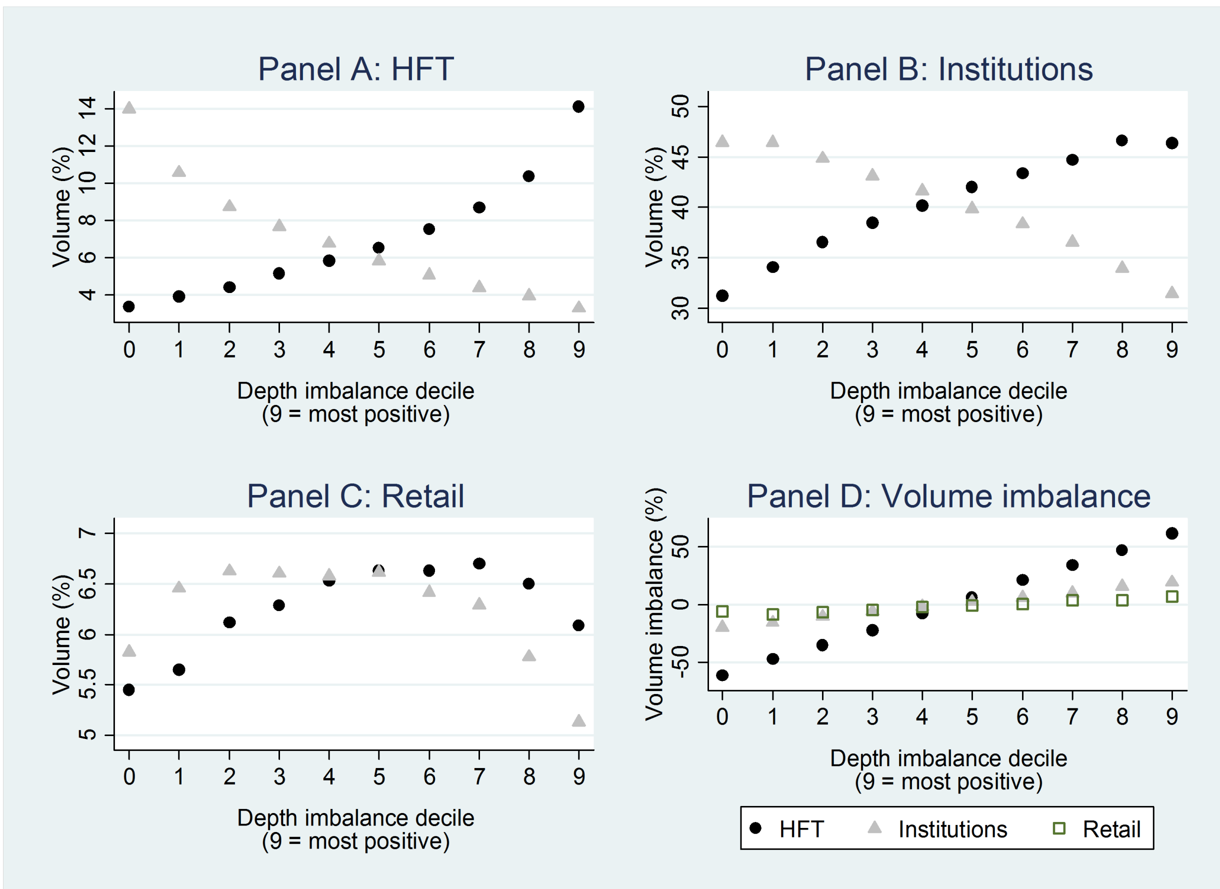


Figure 3. Depth imbalance and time in the order book

Fig. 3 shows the relation 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.

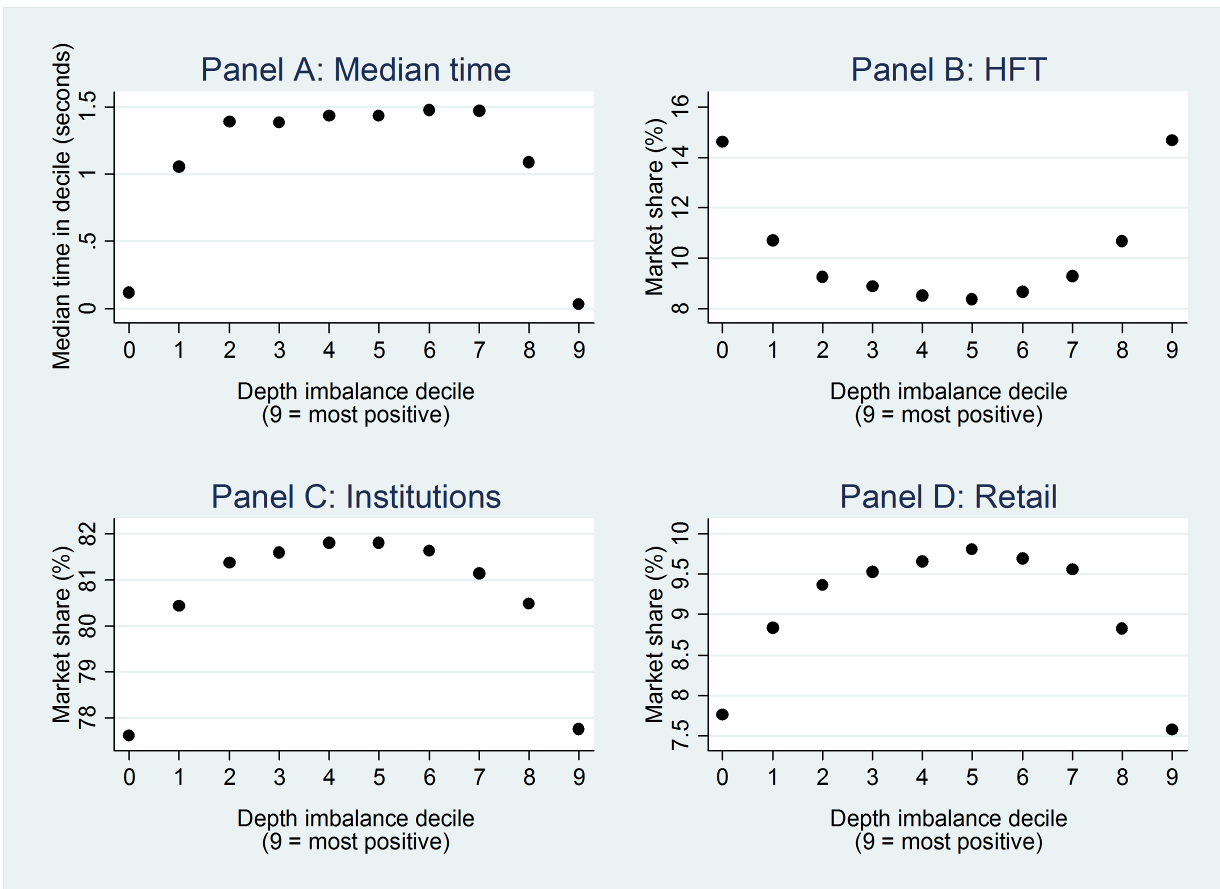


Figure 4. Depth imbalance and volume imbalance for low and high news days

Fig. 4 shows the relation 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) news tercile based on the number of daily Bloomberg news articles for each stock.

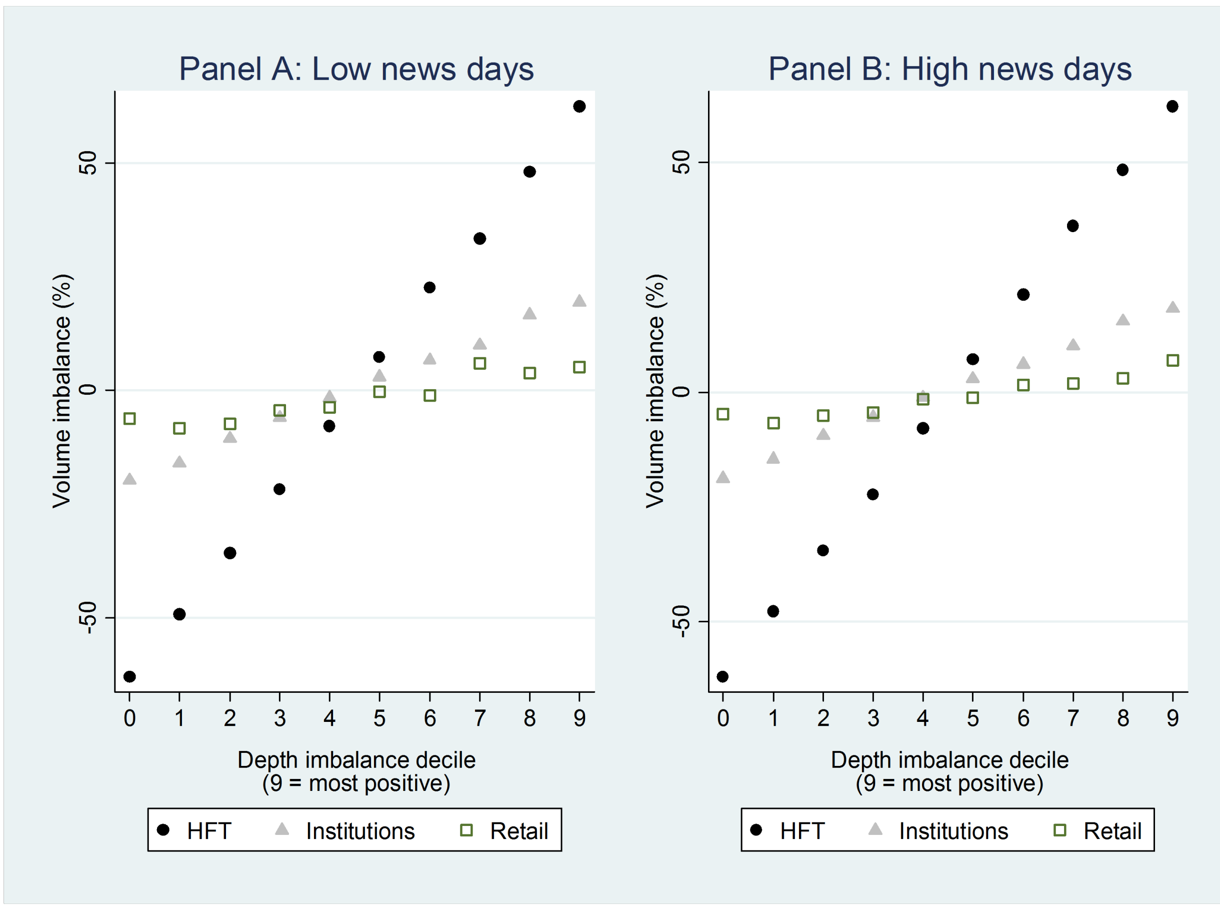


Table 1
Summary statistics

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. *Ntrades* 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 A: Stock characteristics

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

Panel B: Trader characteristics

	HFT	Institutions	Retail
Cancel to trade	3.85	2.12	0.60
Trade to order volume	0.14	0.31	0.29
Median trade size	1,611	915.4	2,308
Median submission to cancel time	188.3	233.3	3,200
EODInventory %	35	48	57
InventoryCrossesZero (Mean)	2.57	2.28	1.3
InventoryCrossesZero (Max)	9.23	4.57	2.81

Table 2
Probability of limit order executions before and after the implementation of ITCH

Table 2 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 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:

$$\begin{aligned} P(Fill)^T = \quad & \beta_0 \quad + \beta_1 I(Non-HFT)^T \times I(Post) + \beta_2 I(Non-HFT)^T \\ + \quad & \beta_3(Post) + \beta_4 Volatility + \beta_5 Volume + \beta_6 Price + \beta_7 Spread + \epsilon^T \end{aligned}$$

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 Columns 3 and 4 (5 and 6), we present results for stock subsamples based on the level of HFT participation. *High HFT* (*Low HFT*) contain stocks in the top (bottom) tercile of stocks based on the percentage of total volume executed by HFT during the pre-ITCH period. In Columns 2, 4 and 6, 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. 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.

	All		High HFT		Low HFT	
	(1)	(2)	(3)	(4)	(5)	(6)
I(Non-HFT) × I(Post-ITCH)	-0.037***		-0.032***		-0.004	
	(-3.31)		(-2.93)		(-0.15)	
I(Non-HFT)	0.111***		0.103***		0.084**	
	(5.95)		(3.43)		(2.33)	
I(Institutions) × I(Post-ITCH)		-0.021**		-0.003		0.002
		(-2.04)		(-0.27)		(0.10)
I(Institutions)		-0.069***		-0.093***		-0.080**
		(-3.67)		(-3.05)		(-2.24)
I(Retail) × I(Post-ITCH)		-0.058***		-0.066***		-0.012
		(-3.77)		(-3.44)		(-0.43)
I(Retail)		0.301***		0.309***		0.262***
		(15.87)		(10.41)		(7.06)
I(Post-ITCH)	0.006	0.007	0.027	0.025	-0.060*	-0.057
	(0.39)	(0.49)	(1.60)	(1.47)	(-1.79)	(-1.70)
Volatility	0.076	0.062	-0.043	-0.091	0.750*	0.759*
	(0.46)	(0.36)	(-0.18)	(-0.35)	(1.87)	(1.89)
Volume	0.071***	0.069***	0.058***	0.057***	0.069***	0.066***
	(17.62)	(17.64)	(7.15)	(7.49)	(10.03)	(10.05)
Price	0.030	0.031	-0.060	-0.062	0.082***	0.087***
	(0.77)	(0.80)	(-0.72)	(-0.74)	(4.28)	(4.62)
Qspread	-1.334	-1.303	-4.214	-3.997	9.554**	9.940**
	(-0.56)	(-0.55)	(-1.56)	(-1.45)	(2.06)	(2.17)
Constant	-0.559***	-0.529***	-0.299	-0.274	-0.678***	-0.636***
	(-6.85)	(-6.61)	(-1.56)	(-1.48)	(-5.95)	(-5.70)
Obs.	10,646	10,646	3587	3587	3352	3352
Adj. R-square	0.188	0.586	0.162	0.647	0.177	0.525

Table 3
Probability of limit order executions conditional on the bid-ask spread

Table 3 analyzes the probability of limit order executions for HFT, Institutions and Retail before and after the implementation of ITCH for *High HFT*, conditional on the size of the bid-ask spread. 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 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. We estimate the following regression:

$$\begin{aligned} P(Fill)^T = \quad & \beta_0 \quad +\beta_1 I(Non-HFT)^T \times I(Post) + \beta_2 I(Non-HFT)^T \\ + \quad & \beta_3(Post) + \beta_4 Volatility + \beta_5 Volume + \beta_6 Price + \beta_7 Spread + \epsilon^T \end{aligned}$$

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. Columns 1 and 2 present the results for subsamples based on whether a stock is tick size constrained. *Constrained* (*Unconstrained*) contains stocks that are on average constrained (unconstrained) by the minimum tick size in the pre-ITCH period. Columns 3 and 4 present results for subsamples based on the stock's relative tick size. *High* (*Low*) contains stocks in the top (bottom) tercile based on the stock's average relative tick size in the pre-ITCH period. 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.

	Tick constraints		Relative tick size	
	Constrained	Unconstrained	High	Low
	(1)	(2)	(3)	(4)
I(Non-HFT) × I(Post-ITCH)	-0.059**	-0.021	-0.053***	-0.011
	(-3.28)	(-1.59)	(-3.87)	(-0.70)
I(Non-HFT)	0.182***	0.070*	0.092*	0.113**
	(4.10)	(1.89)	(2.02)	(2.76)
I(Post-ITCH)	-0.027	0.013	0.004	-0.005
	(-1.04)	(0.55)	(0.21)	(-0.24)
Volatility	0.179	-0.846*	-0.058	-0.017
	(0.63)	(-1.96)	(-0.22)	(-0.04)
Volume	0.066***	0.059***	0.061***	0.055***
	(4.04)	(7.85)	(5.48)	(6.98)
Price	-0.003	-0.069	-0.173	-0.021
	(-0.01)	(-1.37)	(-1.22)	(-0.28)
Qspread	7.932	-4.290	0.796	-5.951**
	(1.35)	(-1.52)	(0.18)	(-2.19)
Constant	-0.904	-0.062	-0.251	-0.345
	(-0.94)	(-0.29)	(-0.85)	(-0.83)
Obs.	1069	2518	1770	1817
Adj. R-square	0.257	0.067	0.207	0.093

Table 4
Probability of favorable and unfavorable limit order executions

Table 4 analyzes the probability of favorable and unfavorable 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). We divide $P(Fill)^T$ from Table 2 into $P(Favorable fill)^T$ and $P(Unfavorable fill)^T$. For *ex post* measures, a buy (sell) favorable fill is one in which the return is positive (negative) after the fill based on the midpoint price 10 trades in the future, and vice versa for unfavorable fills. For ex ante measures, a favorable (unfavorable) fill is 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.

	Ex-post measures				Ex-ante measures			
	P(Favorable fill)		P(Unfavorable fill)		P(Favorable fill)		P(Unfavorable fill)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
I(Non-HFT) × I(Post-ITCH)	-0.011**		-0.004		-0.040***		-0.001	
	(-2.21)		(-0.62)		(-4.18)		(-0.13)	
I(Non-HFT)	0.030***		0.092***		-0.030*		0.134***	
	(8.52)		(11.20)		(-1.80)		(22.11)	
I(Institutions) × I(Post-ITCH)		-0.009*		-0.001		-0.029***		0.008
		(-1.66)		(-0.20)		(-3.03)		(1.55)
I(Institutions)		0.004		0.025***		-0.124***		0.053***
		(1.04)		(3.35)		(-7.51)		(8.96)
I(Retail) × I(Post-ITCH)		-0.013**		-0.007		-0.051***		-0.009
		(-2.21)		(-0.80)		(-4.23)		(-1.00)
I(Retail)		0.056***		0.158***		0.078***		0.226***
		(13.86)		(15.72)		(4.52)		(29.49)
I(Post-ITCH)	0.024*	0.023*	-0.027	-0.028	0.036**	0.036**	0.006	0.008
	(1.68)	(1.70)	(-1.22)	(-1.23)	(2.24)	(2.33)	(0.57)	(0.68)
Volatility	2.980***	2.979***	1.520***	1.518***	0.054	-0.090	0.007	-0.097
	(20.14)	(20.42)	(4.19)	(4.20)	(0.33)	(-0.56)	(0.07)	(-1.04)
Volume	-0.004	-0.004	0.001	0.000	0.039***	0.036***	0.028***	0.026***
	(-1.45)	(-1.48)	(0.12)	(0.11)	(11.72)	(10.89)	(10.88)	(10.02)
Price	0.093***	0.093***	0.034**	0.034**	0.015	0.009	-0.006	-0.014**
	(7.92)	(8.02)	(2.32)	(2.29)	(0.81)	(0.62)	(-0.78)	(-2.01)
Qspread	-0.072	-0.046	0.042	0.109	-0.089	-0.411	-0.199	-0.365
	(-0.05)	(-0.04)	(0.03)	(0.07)	(-0.06)	(-0.32)	(-0.12)	(-0.23)
Constant	0.169***	0.169***	0.223***	0.225***	-0.291***	-0.231***	-0.249***	-0.205***
	(3.52)	(3.57)	(3.16)	(3.20)	(-4.81)	(-3.85)	(-5.73)	(-4.67)
Obs.	11027	11027	11027	11027	9,718	9,718	9,574	9,574
Adj. R-square	0.242	0.263	0.184	0.298	0.151	0.369	0.224	0.460

Table 5
Relation between Volume imbalance, Trade imbalance and Depth imbalance

Table 5 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 *VolumeImbalance* as:

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

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:

$$\begin{aligned} VolumeImbalance_j^T = \;& \beta_0 \quad + \beta_1 I(HFT)^T \times DI_j + \beta_2 I(Institutions)^T \times DI_j + \beta_3 I(HFT)^T \\ + \;& \beta_4 I(Institutions)^T + \beta_5 DI_j + \beta_6 Volume_j + \epsilon^T \end{aligned}$$

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 4-6, we replace the dependent variable with *TradeImbalance*, 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.

	VolumeImbalance %			TradeImbalance %		
	(1) All stocks	(2) Low volatility days	(3) High volatility days	(4) All stocks	(5) Low volatility days	(6) 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 6
Relation between Volume imbalance and DI before and after the implementation of ITCH

Table 6 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, j, and trader type, T, we calculate $VolumeImbalance_j^T$ using (4). For columns 1-3, we estimate the following linear regression, which is based on DI deciles:

$$\begin{aligned} VolumeImbalance_j^T = \quad & \beta_0 \quad + I(Pre)\Big(\beta_1 I(HFT)^T \times DI_j + \beta_2 I(Institutions)^T \times DI_j + \beta_3 DI_j\Big) \\ + \quad & I(Post)\Big(\beta_4 I(HFT)^T \times DI_j + \beta_5 I(Institutions)^T \times DI_j + \beta_6 DI_j\Big) \\ + \quad & \beta_7 I(HFT)^T + \beta_8 I(Institutions)^T + \beta_9 Volume_j + \epsilon^T. \end{aligned}$$

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
I(HFT) × DI	β_1 0.942***		β_4 1.064***	5.350**	β_1 0.938***		β_4 1.066***	7.71***
	(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)
DI	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 7
AdjustedDI for order book events by trader type

Table 7 shows the mean *AdjustedDI* for Passive trade executions, order submissions, order amendments and cancelations for each trader type. *AdjustedDI* is determined immediately before each order book event:

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

where $\sum_{i=1}^{n} DepthBid_{i,t}$ ($\sum_{i=1}^{n} DepthAsk_{i,t}$) 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.

	HFT	Institutions	Retail
Trade (passive)	0.083	-0.029	-0.009
Submission	0.063	-0.004	0.043
Amendment	0.040	-0.003	0.023
Cancelation	0.017	0.002	0.028

Table 8
Multinomial logistic regressions for limit order placement strategies

Table 8 assesses the probability of each order book event based on prevailing market conditions. We present the coefficient estimates for the following multinomial logistic regression on HFT orders:

$$\begin{aligned} OrderBookEvent_E = \quad & \beta_0 \quad +\beta_1 AdjustedDI_E + \beta_2 Volatility + \beta_3 Volume \\ + \quad & \beta_4 Price + \beta_5 Spread + \epsilon_E \end{aligned} \tag{F.2}$$

where $OrderBookEvent_E$ is the dependent variable indicating one of four order book events, E: 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:

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

where $\sum_{i=1}^{n} DepthBid_{i,t}$ ($\sum_{i=1}^{n} DepthAsk_{i,t}$) 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. All regressions control for stock and day fixed effects. t-statistics are reported in parentheses. ***, ** and * indicate significance levels of 1%, 5% and 10%, respectively.

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

Table 9
Relation between Volume imbalance, Trade imbalance and Depth imbalance for low and high information environments

Table 9 reports the regression of *Volume imbalance* or *Trade imbalance* against *DI*. 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:

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

where $\sum_{k=1}^{n} BuyVolume_{k,j}$ ($\sum_{k=1}^{n} SellVolume_{k,j}^T$) 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:

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

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.

	VolumeImbalance%		TradeImbalance%	
	(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