High-frequency trading strategies

Michael Goldstein
Babson College
Babson Park, MA 02457-0310, USA
goldstein@babson.edu

Amy Kwan
University of Sydney
Sydney, NSW 2006, Australia
amy.kwan@sydney.edu.au

Richard Philip
University of Sydney
Sydney, NSW 2006, Australia
richard.philip@sydney.edu.au

Current version: December 8, 2016

Abstract

Using a unique, broker-level dataset, we document an important information channel driving high frequency trading strategies. High frequency traders (HFT) condition their strategies on order book depth imbalances, which are a strong predictor of future price movements. Examining the order book imbalance immediately before each order submission, cancelation and trade, we show HFT supply liquidity on the thick side of the order book and demand liquidity from the thin side. This strategic behavior is more pronounced during volatile periods and when trading speeds increase. However, by competing with non-HFT limit orders, HFT crowd out non-HFT limit orders.

The authors would like to thank the Centre for International Finance and Regulation which funded this research under CIFR grant T013.

High-frequency trading strategies

Abstract

Using a unique, broker-level dataset, we document an important information channel driving high frequency trading strategies. High frequency traders (HFT) condition their strategies on order book depth imbalances, which are a strong predictor of future price movements. Examining the order book imbalance immediately before each order submission, cancelation and trade, we show HFT supply liquidity on the thick side of the order book and demand liquidity from the thin side. This strategic behavior is more pronounced during volatile periods and when trading speeds increase. However, by competing with non-HFT limit orders, HFT crowd out non-HFT limit orders.

1.0 Introduction

High frequency traders (HFT) influence financial markets in many ways. For example, HFT reduces the bid-ask spread (Angel, Harris and Spatt, 2010; Jones, 2013; Harris, 2013; Hasbrouck and Saar, 2013; Brogaard, Hagstromer, Norden and Riordan, 2015; Malinova, Park and Riordan, 2016), increases price efficiency (Carrion, 2013; Brogaard, Hendershott and Riordan, 2014) and increases overall market depth (Hasbrouck and Saar, 2013). On the whole, HFT play a dominant role in providing liquidity (Hagstromer and Norden, 2013; Hasbrouck and Saar, 2013; Menkveld, 2013; Malinova, Park and Riordan, 2016, Conrad, Wahal and Xiang, 2015), but could also withdraw this liquidity during stressful periods (Kirilenko, Kyle, Samadi and Tuzun, 2016; Brogaard et al., 2016; van Kervel and Menkveld, 2016; Korajczk and Murphy, 2016). Similarly, other studies document that HFT can anticipate future order flow from other traders and predict future price movements (Brogaard, Hendershott and Riordan, 2014; Li, 2014; Hoffmann, 2014; Biais, Foucault and Moinas, 2015; Foucault, Hombert and Rosu, 2016; Hirschey, 2016; Rosu, 2016; Subrahmanyam and Zheng, 2016). While we understand much about how HFT affect financial markets, not much is known about the information channels that drive HFT behavior due to the proprietary nature of their business.

In this study, we examine HFT order placement strategies and document one such information channel, namely the order book depth imbalance. Similar to previous studies, we provide strong evidence that order book imbalances predict short term future price movements. While all traders (i.e., HFT, institutions and retail) attempt to trade in the direction of the imbalance, we find HFT are better at taking advantage of information contained in the order book, especially at times when the market is volatile and when HFTs become faster. Specifically, HFTs trade in the direction of the order book imbalance and cancel or amend orders when the order book imbalance moves against them.[1] As a result, HFT on average, supply liquidity on the thick side of the order book and demand liquidity from the thin side. By competing with non-HFT limit orders on the thick side of the order book, we find evidence that HFT crowd out non-HFT limit orders from the order book. Our study exposes one important mechanism through which HFT anticipate future order flow and returns.

For this investigation, we use a unique, broker-level dataset from the Australian Securities Exchange (ASX). Using this dataset, we reconstruct the full limit order book and measure the shape of the order book at the time of each order submission, cancelation, amendment, and trade. The dataset also classifies brokers into three trader types: proprietary HFT firms, institutions and retail. Our primary focus is on HFT trading behavior, relative to institutional and retail traders. Finally, the introduction of faster ITCH technology on the ASX provides a natural experiment for us to investigate the effects of increasing trading speeds.

We find strong evidence that the shape of the limit order book contains information about future price movements. In particular, when the volume available on the best five prices on the bid (ask) side exceeds the volume available on the best five ask (bid) prices, we show that prices are more likely to rise (fall) in the short-term.[2] We also find that non-HFTs want to act like HFTs, but are less successful when the market indicates it is most important to do so. We demonstrate that all trader types attempt to trade in the direction of the price change predicted by the depth imbalance. Specifically, for all trader types, the percentage of buyer initiated trade volume

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

[2] Cont, Kukanov and Stoikov (2013) also find that depth imbalances predicts future short term price changes, but only use depth imbalances at the best bid and ask prices. Using a sample period before the growth of HFT, Cao, Hansch and Wang (2008) find that order imbalances behind the best bid and offer contribute to approximately 22% of price discovery. We also show that depth on levels 2 to 5 contain additional information on future price changes in a high frequency world (see Appendix 1).

increases as the depth available on the bid prices grows relative to the depth on the ask prices. However, when depth imbalances are very large (i.e., predicting the largest price changes), we find that HFTs are more successful at trading in the direction of the imbalance, relative to the other trader types. This finding is consistent with the HFT's speed advantage, in that HFT monitor the limit order book more effectively than institutional and retail investors and react faster to large depth imbalances by submitting aggressive market orders.

For passive limit orders, we find that HFTs are also better at using information contained in the limit order book. HFTs submit limit orders to the order book primarily when there is a small favorable depth imbalance. As the depth imbalance becomes more favorable, the resting limit order executes in the same direction as the imbalance. On the other hand, if the depth imbalance becomes less favourable, HFT are quick to cancel or amend their orders, reducing adverse selection costs. Together, these strategies mean that HFT supply liquidity to the thick side of the order book and demand liquidity from the thin side, which could exacerbate future depth imbalances.

Next, we investigate HFT trading behavior around times of high stock volatility. We find HFT demand more liquidity when the market is volatile, in contrast to non-HFT. Further, for both aggressive and passive trades, we show that HFT are even more successful at trading in the direction of the order book during periods of high volatility, relative to non-HFT. In volatile markets, the HFTs use their speed advantage by strategically submitting market orders and cancelling limit orders. When markets are volatile, fast HFT are better able to use market orders to pick off stale limit orders from slower institutional and retail investors. Similarly, HFT use their speed advantage to cancel their own stale limit orders to reduce their adverse selection costs. These results imply that trading speed differentials are particularly advantageous during periods of high volatility.

The introduction of ITCH in 2012 offers a natural experiment to investigate the effects of a speed change on HFT trading behavior. Using a difference-in-difference framework, we demonstrate that HFT are more successful at trading in the direction of the order book when they gain a larger speed advantage. However, one externality of HFT trading behavior is that there is a crowding out effect on non-HFT limit orders. We find that the probability of execution for institutional and retail limit orders submitted to the best bid and ask prices decreases when HFT gain a larger speed advantage. Further, conditioning limit order executions on the shape of the order book, we show that it is the probability of favourable executions (i.e. non-HFT limit order trading in the direction of the order book imbalance), which falls.

While much is known about the effects of HFT, the literature is unclear on how HFTs trade to influence financial markets. While the theoretical literature predicts that fast traders can anticipate the order flow of slow traders (Biais, Foucault and Moinas, 2015; Li, 2014; Hoffmann, 2014; Rosu, 2016), the exact mechanism has not been previously documented in the theoretical or empirical literatures. Empirically, Hirschey (2016) documents that HFTs can anticipate order flow from other investors and Subrahmanyam and Zheng (2016) conclude that HFT manage their limit orders in anticipation of short-term price movements, but do not show how HFT predict future price movements. We demonstrate that the order book depth imbalance is one important mechanism through which HFT gain information on future price movements.

Second, several studies find that HFT increase price efficiency. Brogaard, Hendershott and Riordan (2014) demonstrate that HFT buy in the direction of permanent price changes through liquidity demanding orders. Carrion (2013) shows that HFT incorporate information from order flow and market-wide returns more efficiently.[3] But how do they do so? We uncover the information channel through which HFT incorporates information into price. Specifically, we find that HFT increase price efficiency by trading in the direction of the order book imbalance, which is a strong predictor of future price movements.

Third, our results have implications for studies on the market making role of HFT in equity markets, which find that HFT market making increases market depth (Hendershott, Jones and Menkveld, 2011; Hasbrouck and Saar, 2013; Brogaard, Hagstromer, Norden and Riordan, 2015). However, these findings are typically based on traditional measures of market depth, aggregated across both bid and ask prices.[4] Aggregated measures of market depth do not capture the amount of depth available on the side of the limit order book where it is most needed. For example, a trader submitting a buy market order is more concerned about the depth available on the ask side of the limit order book, rather than aggregated depth over both bid and ask prices. Our results show that HFT on average supply depth on the thick side of the order book but demand depth from the thin side of the order book, i.e., they may add depth, but not on the side that is thin. Similarly, HFT cancel limit orders from the thin side of the order book, which face larger adverse selection costs.

Finally, several studies in the literature examine HFT liquidity provision during stressful times. Brogaard et al. (2016) examine the stability of liquidity supply by high frequency traders, who do not have the obligation to supply liquidity during stressful periods. They find that HFTs supply liquidity to non-HFTs during extreme price moves in a single security but demand liquidity when several stocks experience simultaneous extreme price moves. However, their analysis

[3] Several studies, including Chaboud Chiquoine, Hjalmarsson and Vega (2014), Hendershott, Jones and Menkveld (2011), Boehmer, Fong and Wu (2015) find that algorithmic trading, of which HFT is a subset, improves informational efficiency of prices.

[4] In a fragmented market setting, Van Kervel (2015) shows that consolidated measures of liquidity could overestimate the actual amount of available. Observing a trade on one venue, Van Kervel (2015) shows that HFT market makers cancel outstanding limit orders on all other venues to reduce their adverse selection costs.

focuses exclusively on the liquidity demanded and supplied through the Nasdaq exchange, which represents only 30-40% of all trading activity of the sample stocks. Thus, it is possible that HFT are supplying liquidity on Nasdaq while demanding liquidity from other trading venues. By analysing a mostly consolidated market, we provide further insights on HFT trading activity over the whole market. Unlike Brogaard et al. (2016), our results show that HFT are net demanders of liquidity and become even more aggressive at times of high market volatility. Further, we find that they demand more liquidity from the thin side of the order book, which could exacerbate future volatility.[5]

In a related group of studies, van Kervel and Menkveld (2016) and Korajcyzk and Murphy (2016) study liquidity provision to institutional trades and show that HFTs initially trade 'against the wind' but eventually trade 'with the wind' as a large institutional trade progresses. By trading in the same direction as the institutional trader, HFTs increase the implementation shortfall of institutional orders. Our study reveals the mechanism through which HFT detect institutional orders. Specifically, we show that HFT can detect institutional demand through the order book depth imbalance and trade in the same direction of the imbalance before the predicted price rise.

2.0 Institutional details

The Australian Securities Exchange (ASX) is the dominant stock exchange for Australian equities, with over 90% market share of on-market traded volume in 2012.[6] The ASX operates as a continuous limit order book between approximately 10:00 am and 4:00 pm, matching orders based on price and time priority, with a randomized open and a randomized close. Each stock

[5] Using transaction level data around the 2010 flash crash, Kirilenko, Kyle, Samadi and Tuzun (2016) find that the trading pattern of non-HFT did not change when prices fell during the crash. We study an additional dimension, namely the order submission and cancellation behaviors of HFT, around times of large price movements and find that HFTs cancel orders that are at a high risk of being adversely selected.

[6] See Aitken, Chen and Foley (2013). Comerton-Forde and Putnins (2016) report that off-exchange dark and block trades make up approximately 18% of total dollar volume over the 2008 to 2011 period.

opens with an opening auction at a random time between 10:00 and 10:10 am depending on the starting letter of their ASX code. Similarly, the closing price is determined via a closing price auction that takes place between 4:10 pm and 4:12 pm. While trading on the ASX has been anonymous since the removal of real-time broker identifiers in November 2005, the broker identification behind all executed trades is available to all market participants on a t+3 basis.

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

3.0 Data and variable construction

3.1 Data and sample selection

We obtain full order book and trade data for stocks in the S&P/ASX 100 index from the AusEquities database provided by the Securities Industry Research Centre of Asia Pacific.[7] The securities contained in our dataset are the most liquid and actively traded among HFT and institutional investors on the ASX. We analyze one year of order level data for the period January 1, 2012 to December 31, 2012, which incorporates the introduction of ITCH on April 2, 2012. To avoid the randomized open and close, we include only trades and orders entered between 10:10:00 and 16:00:00 to ensure that our sample is not contaminated by the opening and closing call

[7] 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.

auctions. We assume that all outstanding orders remaining in the limit order book at the end of the trading day are cancelled.

Data from the ASX offer several advantages over other exchanges. For each order, the data contain the stock symbol, date and time of order entry to the millisecond level, order size and price, order identification number and an identifier for the submitting broker. In our dataset, broker identifiers are assigned into three trader categories: proprietary HFT firms (*HFT*), *Institutions*, and *Retail*.[8] We refer to orders originating from *Institutions* and *Retail* collectively as *non-HFT*. Additionally, we use the order identification number to trace subsequent amendments, executions or cancelation back to the original order entry, allowing for a full reconstruction of the limit order book and the tracking of the order and its queue position through time. We rely on the granularity of the data to compute the depth imbalance proxy for trading strategies. Furthermore, because we can replay the full order book, we do not have to rely on trade classification algorithms, such as Lee and Ready (1991),[9] to determine whether a trade is buyer or seller initiated. Following Upson, Johnson and McInish (2015), we aggregate all trade reports at the same price, in the same trade direction, from the same broker, and reported in the same millisecond timestamp into one marketable order. Finally, in comparison to U.S. and European equity markets, the ASX is less fragmented, operating as a virtual monopoly in Australian equities during our period with over 90% of the daily trading volume.

[Insert Table 1]

Table 1, Panel A reports the summary statistics for the 94 stocks, which appear in the S&P/ASX 100 index over the full sample period. *Market capitalization* is measured on January 3,

[8] We classify HFT firms based on van Kervel and Menkveld (2016). We note that some smaller proprietary HFT firms could trade through institutional brokers and thus, *Institutions* could also contain some proprietary HFT activity.
[9] Ellis, Michaely and O'Hara (2000) report that the Lee and Ready (1991) rule misclassifies approximately 20% of all trades.

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 $11.67. The average daily dollar volume is 25.5 AUD million and the average number of trades is 2,176. Given that the minimum pricing increment on the ASX is $0.01 for stocks priced above $2.00, an average daily time-weighted quoted spread (*Spread*) of 1.04 cents indicates that many stocks in the sample are likely to be spread constrained.[10]

Table 1, Panel B reports the summary statistics for *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 percentage of order cancelations, and the median submission to cancel time is significantly lower for *HFT*. The average number of active and passive trades is approximately equal for *HFT*, whereas *Institutions* and *Retail* are predominantly limit order traders.[11] In the following sections, we investigate how *HFT* incorporate information contained in the order book to their trading strategies.

3.2 Depth imbalance

Previous studies document a strong relationship between trade imbalances and future returns (Chordia, Roll and Subrahmanyam, 2002; Chordia and Subrahmanyam, 2004). Using more granular limit order book data, recent studies also find strong evidence that order imbalances between the buy and sell schedules of the limit order book are significantly related to future stock returns (Cao, Hansch and Wang, 2008; Cont, Kukanov and Stoikov, 2014). Cont, Kukanov and

[10] Twelve stocks in our sample have an average stock price under $2. Our results are robust to removing these stocks from the sample.

[11] As described above, we aggregate all trade reports at the same price, in the same trade direction, and reported in the same millisecond timestamp into one marketable order. As a result, the total number of passive trades exceeds the total number of aggressive trades.

Stoikov (2014) show that high-frequency price changes are mainly driven by imbalances between supply and demand at the best bid and ask prices. Specifically, large buying (selling) pressure on the bid (ask) price predicts future price rises (falls). Further, Ranaldo (2004) examines how the state of the limit order book can affect a trader's order submission strategy. We use the information contained in the state of the limit order book to proxy for strategic trading.

To measure the shape of the limit order book at the time of order submission, we calculate depth imbalance (*DI*) as the difference between the volume available at the best bid and ask prices, as a proportion of the total volume available at the best bid and ask prices.[12] Specifically, for each order book event (i.e., submission, trade, amendment or cancelation) we determine:

$$DI_t = \frac{\sum_{i=1}^{n} VolBid_{it} - \sum_{i=1}^{n} VolAsk_{it}}{\sum_{i=1}^{n} VolBid_{it} + \sum_{i=1}^{n} VolAsk_{it}}$$

where $\sum_{i=1}^{n} VolBid_t$ ($\sum_{i=1}^{n} VolAsk_t$) is the volume available at the top *n* bid (ask) price levels immediately before the order book event, *t*.[13] For our main results, we calculate *DI* based on the volume available at the top five bid and ask prices (n = 5).[14] 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) prices is much larger than the bid (ask) depth.

For some tests, for which trade direction is not important, we multiply *DI* by an indicator for whether the order or trade is a buy or sell to remove the effects of trade direction. We refer to this directionally adjusted *DI* measure as *Adjusted DI*. When *Adjusted DI* is positive, the trade or order event occurs in the direction of the depth imbalance (e.g., a buy trade executes when the bid depth exceeds the ask depth).

[12] In contrast to Naes and Skjeltorp (2006), we are more interested in the liquidity available at bid and ask prices rather than the slope of the order book.
[13] We compute *DI* immediately before the time of event to avoid capturing the volume of order book event itself.
[14] For robustness, we also test our results using one and three price levels.

Table 1, Panel C summarizes the average *Adjusted DI* immediately before active trades, passive trades, order submissions, amendments and cancelations for each trader type. For active trades, we find that all traders trade in the direction of the imbalance, indicating that traders are more likely to submit a market buy (sell) order when the bid (ask) depth is much larger than the ask (bid) depth.

Comparing the magnitude of *Adjusted DI*, *HFT* submit market orders when *Adjusted DI* is much larger (0.148) compared to 0.024 for both *Institutions* and *Retail*. *Adjusted DI* at the time of passive trades is 0.083, -0.030, and -0.012, for *HFT*, *Institutions*, and *Retail*, respectively. A negative *Adjusted DI* indicates that *Institutions* and *Retail* limit orders are picked off the thin side of the limit order book. In contrast, we find that *HFT* submit orders when there is a moderate *Adjusted DI* (*Adjusted DI* = 0.059) but 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*. We formally investigate *HFT* trading strategies and their potential impact on *non-HFT* trading in the next section.

4.0 Empirical Results

4.1 Depth imbalance, future stock prices and aggregate trading volumes

First, we need to establish the information content of order book depth imbalances for our sample stocks. To investigate whether depth imbalances contain information about the future stock price, we rank trades into deciles based on the depth imbalance immediately before the trade for each stock-day. For each transaction, we also calculate future returns by comparing the midpoint of the best bid and ask prices at the time of the trade with the bid-ask midpoint 10 trades in the future. Figure 1, Panel A presents the average future return for trades from each depth imbalance decile. We observe a strong positive relationship 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.[15] Specifically, when there are a lot of buyers in the order book, relative to the number of sellers, we observe a rise in future stock price.

[Insert Figure 1]

Next, we examine how the market responds to order book depth imbalances. For each depth imbalance decile, we calculate the percentage of total volume that is buyer or seller initiated. Given that depth imbalances predict future returns, we expect strategic traders to trade in the direction of the order book imbalance. Specifically, we expect more aggressive buying (i.e., more buyer initiated trades) when a large positive depth imbalance exists and more aggressive selling when there are large negative imbalances. Consistent with strategic trading, our full market results in Figure 1, Panel A confirm a strong positive (negative) relationship between the size of the depth imbalance and the percentage of buyer (seller) initiated trade volume. In the next section, we investigate whether some trader types trade more strategically than other traders based on order book depth imbalances.

4.2 Depth imbalance and trading volumes by trader type

Our previous analysis show that in aggregate, traders buy aggressively when there is a large positive depth imbalance and sell aggressively when a large negative imbalance exists. To investigate whether the relation between depth imbalance and trading volumes differs by trader, for each trader type, we calculate the amount of buyer and seller initiated volumes as a percentage of total market volume. Figure 2, Panels A to C presents the results separately for *HFT*, *Institutions*, and *Retail*, respectively. Consistent with the full sample results from Figure 1, Panel A, we observe

[15] Cao, Hansch and Wang (2008) and Cont, Kukanov and Stoikov (2014) also document that limit order book imbalances contribute to price discovery.

a general positive (negative) relationship between depth imbalance and aggressive buying (selling) for all trader types, indicating that all traders trade in the direction of the depth imbalance.

[Insert Figure 2]

Comparing between the panels, *HFT* are more successful than *Institutions* and *Retail* when depth imbalances are very positive or very negative. Figure 2, Panel A shows that *HFT* buy (sell) most aggressively when depth imbalance is the most positive (negative). For *Institutions* (Panel B) and *Retail* (Panel C), the percentage of buyer (seller) initiated trades increases with the size of the positive (negative) depth imbalance for moderate levels of imbalances. However, in the extremes (i.e., when depth imbalance is very positive or very negative), both *Institutions* and *Retail* are less successful at trading in the direction of the imbalance.

To further assess whether *HFT* are more successful at trading on information contained in the depth imbalance, we calculate the executed volume imbalance for each *DI* decile. Specifically, for each trader type, we calculate the volume imbalance as:

$$Volume\ imbalance_j = \frac{\sum_{j=1}^{n} BuyVolume_j - \sum_{i=1}^{n} SellVolume_j}{\sum_{j=1}^{n} BuyVolume_j + \sum_{i=1}^{n} SellVolume_j}$$

where $\sum_{j=1}^{n} BuyVolume_j(\sum_{j=1}^{n} SellVolume_j)$ is the total aggressive buying (selling) volume for depth imbalance decile, *j*.

Figure 2, Panel D, shows the relation between *Volume imbalance* and a *DI* for our three trader types. Given the size of *DI* predicts future returns, a steeper slope indicates a trader is more focused on trading with the order book *DI*, ahead of future predicted price changes. Comparing the slopes for *HFT*, *Institutions* and *Retail*, our results show *HFT Volume imbalance* is most sensitive to an order book depth imbalance, 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*.

In Table 2, we test whether a statistically significant difference in *Volume imbalance* exists between our trader types for each *DI* decile. When the average *DI* is the most negative (*DI* = -0.375), the volume imbalance for *HFT* is -61.8%, while *Volume imbalance* for *Institutions* and *Retail* is only -16.8% and -6.2%, respectively. For the most positive *DI* decile (*DI* = 0.380), we observe positive volume imbalances for *HFT* (62.6%), *Institutions* (17.1%) and *Retail* (6.1%). Finally, when the order book is balanced, such that the bid depth is approximately equal to the ask depth, the difference in the *Volume imbalance* is less severe. For example, when *DI* is only 0.028 (decile 5), the volume imbalances range from 0% (*Retail*) to only 6.4% (*HFT*).

Importantly, we find that *HFT Volume imbalance* is always significantly below the institutional and retail *Volume imbalance* when a negative depth imbalance exists (i.e., there is selling pressure in the limit order book). In contrast, when buying pressure exists in the limit order book, volume imbalances are significantly larger for *HFT*, relative to *Institutions* and *Retail.* This result indicates that HFT are more successful at buying when the order book is predicting a future price rise and selling before expected future price declines. Further, comparing between *Institutions* and *Retail*, we find that *Institutions* are more strategic than *Retail* in trading with the imbalance in 9 of the depth imbalance deciles.

[Insert Table 2]

Taken together, these results indicate that all broker types attempt to trade in the direction of a stock's depth imbalance. However, *HFT* are more likely to trade 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. One implication for our results is that *HFT* could be crowding out *non-HFT* limit orders, especially when large depth imbalances exist.

In Table 3, we formally test the sensitivity of volume imbalances to depth imbalances for our trader types after controlling for trading volumes, stock and day fixed effects. Based on the volume imbalance for each broker and depth imbalance decile, we estimate the following regression:

$$Volume\ imbalance\ \% = \beta_0 + \beta_1 I(HFT) \times DI + \beta_2 I(Institutions) \times DI + \beta_3 I(HFT) + \beta_4 I(Institutions) + \beta_5 DI + \beta_6 Volume + \varepsilon \quad (1)$$

where $I(HFT)$ and $I(Institutions)$ are indicator variables for HFT and institutional brokers. DI is the average depth imbalance for the trades in the depth imbalance decile and $Volume$ is the natural log of the total traded volume in the decile. We also include controls for stock and day fixed effects.

[Insert Table 3]

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

To investigate the effects of stock volatility on HFT trading behavior, for each stock we rank trading days into terciles based on the daily stock volatility.[16] Table 3, Columns 2 and 3,

[16] 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.

present the results separately for low and high volatility days, respectively. For both low and high volatility days, we find that *HFT* use more order book information in their trading strategies than *Institutions* and *Retail*.[17] Further, $I(Institutions) \times DI$ is negative and significant on high volatility days indicating that *Institutions* are less successful at trading in the direction of the expected future price movement when the market is highly volatile.

It is possible that some smaller proprietary HFT firms trade through institutional brokers. While these HFT traders could influence trading imbalances based on the number of trades, it is much more difficult to change overall volume imbalances. To investigate this possibility, in Table 3, Columns 4-6, we replace the dependent variable from Equation (1) with trade imbalance, which is based on the number of buyer and seller initiated trades, rather than the volume of buyer and seller initiated trades. For *HFT*, our results are largely consistent with our findings based on volume imbalances. For *Institutions*, in contrast to our results based on volume imbalances, we also find that $I(Institutions) \times DI$ is positive and significant for all samples. This result reveals that institutional investors submit small trades that capture information contained in the order book while their large trades are less likely to execute in the direction of the imbalance. This result could be driven by small HFT firms executing their strategies through larger institutional brokers.

So far, our results show that relative to non-HFT traders, HFT submit more buyer (seller) initiated orders when there are already low levels of liquidity available on the sell (buy) side of the order book. One implication of these results is that HFT strategies could exacerbate future order imbalances, especially when the market is volatile.

4.3 Order submission strategies

[17] In unreported tests, 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) .

In this section, we analyze how order submission strategies differ between our investor categories. Specifically, we measure the order book depth imbalance immediately before a trader submits, amends or cancels an order. For each stock, we estimate the daily average *DI* for each order book event: aggressive trade (i.e., market or marketable limit order), passive trade, order submission, amendment, and cancelation.

As discussed earlier, to remove the effects of trade direction, we multiply *DI* by an indicator for whether the order or trade is a buy or sell so that purchases and sales can be interpreted together (*Adjusted DI*). An *Adjusted DI* value of 0 indicates that the order book is balanced and a high positive *Adjusted DI* value indicates a large depth imbalance in the direction of the order book event.[18] A negative *Adjusted DI* indicates that the order book is moving against the trader. Specifically, a negative *Adjusted DI* at the time of a buy trade indicates that a trader is buying when the ask depth exceeds the bid depth, before an expected future price fall. Thus, a strategic trader who uses information contained in the order book should execute trades when *Adjusted DI* is highly positive and cancel or amend orders when *Adjusted DI* is low or negative. We estimate the following regression:

$$Adjusted\ DI = \beta_0 + \beta_1 I(Aggressive\ trade) + \beta_2 I(Passive\ trade) + \beta_3 I(Amend) + \beta_4 I(Cancel) + \beta_5 Volatility + \beta_6 Volume + \beta_7 Price + \beta_8 Spread + \varepsilon \quad (2)$$

where $I(Aggressive\ trade)$ is an indicator variable equal to 1 for a market or marketable limit order and 0 otherwise. $I(Passive\ trade)$ is an indicator variable equal to 1 for a passive trade and 0 otherwise. $I(Amend)$ is an indicator variable equal to 1 for an order amendment and 0 otherwise. $I(Cancelation)$ is an indicator variable equal to 1 for an order cancelation and 0 otherwise. All

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

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 *Qspread* is the time-weighted quoted spread. We also control for stock and day fixed effects.

Consistent with our strategic trading hypothesis, we find that HFT trade aggressively when a large depth imbalance exists in the order book, and cancel or amend orders when the order book imbalance moves against them. Table 4, Column 1, shows that *I(Aggressive trade)* and *I(Passive trade)* is positive and significant indicating that on average, trades take place when the depth imbalance is larger than the depth imbalance at the time of a limit order submission, which is captured in the constant term. Comparing between aggressive and passive trades, we find HFTs submit market orders only when the order imbalance is highly favorable (*p-value* < 0.01). Further, we find that *I(Amend)* and *I(Cancel)* are negative and significant indicating HFT are quick to amend or cancel orders when the depth imbalance becomes less favorable to trade. In doing so, HFT remove stale limit orders before these orders can be picked off the order book by other traders.

[Insert Table 4]

We find institutions and retail investors are generally less strategic than HFT on trading on information contained in the limit order book. Similar to *HFT*, institutions submit aggressive orders when the depth imbalance is in the same direction (Table 4, Column 2). However, for their limit order strategies, *I(Passive trade)* is negative and significant while *I(Amend)* and *I(Cancel)* are both positive and significant. Together, these results indicate that institutions 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.

For *Retail*, Table 4, Column 3 reveals that *I(Passive trade)* is negative and significant indicating that retail limit orders are also picked off the limit order book. Further, the negative and significant coefficient for *I(Aggressive trade)* suggests that retail investors are unable to strategically time their market orders. Taken together, these result shows that retail investors buy (either through passive or aggressive orders) before an expected fall in the stock price and sell before an expected price rise.

In Table 4, Columns 4 to 6, we replace the dependent variable with *Adjusted DI* based on the depth available at the best bid and ask prices (i.e., 1 level of the order book). This test allows us to compare whether some traders are only trading on information contained in the top level of the order book. Comparing with Table 4, Columns 1 and 3, we find that for *HFT*, the coefficients are similar in sign and significance using either 5 levels or 1 level of order book information. One exception is *I(Passive Trade),* which is positive and significant when *Adjusted DI* is calculated using 5 levels of the order book but insignificant when the dependent variable is based on the top level of the order book. The insignificant coefficient on *I(Passive trade)* indicates that HFT use information contained in the order book, beyond the best bid and ask levels, in their order placement strategies.

In Table 4, Column 5, the negative coefficients on *I(Amend)* and *I(Cancel)* show that *Institutions* are more likely to cancel and amend orders when the *Adjusted DI* for the top level of the order book is highly unfavorable. However, *I(Passive trade)* remains negative and significant, indicating that institutional limit orders are picked off the limit book. For *Retail,* the results in Column 6 are similar to the results reported earlier using 5 levels of the order book. One exception is *I(Aggressive trade)*, which is now positive and significant, indicating that retail traders cross the spread when there is a large favourable depth imbalance based on the top level of the order book.

Thus, when a retail investor wishes to buy (sell), and a large limit order queue exists on the best bid (ask) price, they are more likely to demand immediacy by submitting an aggressive market order. Overall, our results provide further support for the conclusion that HFT are more successful at monitoring the limit order book than other trader types. To avoid stale limit orders, HFT cancel or amend their resting limit orders when the order book depth imbalance moves in an unfavourable direction.

To further investigate the order submission behavior of HFT, we use a multinomial logistic regression model to assess the probability of each order book event based on the prevailing market conditions in the limit order book. For each broker type, we estimate the following regression:

$$\Pr(OrderBookEvent) = \beta_0 + \beta_1 Adjusted\ DI + + \beta_2 Volatility + \beta_3 Volume + \beta_4 Price + \beta_5 Spread + \varepsilon \quad (3)$$

where *OrderBookEvent* is the dependent variable indicating one of five order book events: Aggressive trade, passive trade, limit order submission, amendment or cancelation. *Adjusted DI*, *Volatility*, *Volume*, *Price* and *Spread* are defined as in Equation (2). We estimate the model with limit order submission as the baseline category.

[Insert Table 5]

Table 5, Panel A presents the regression coefficients for *HFT*. Consistent with our expectations, we find that *Adjusted DI* is positive for aggressive trades. Thus, when *Adjusted DI* is large, HFT are more likely to submit a market (or marketable limit) order, than a less aggressive limit order. Similarly, *Adjusted DI* is positive for passive trades, meaning that limit order trade executions are more likely than a limit order submission when *Adjusted DI* is large. In contrast, *Adjusted DI* is negative for amendments and cancelations. When *Adjusted DI* is small, HFT are more likely to amend or cancel an order than to submit a limit order. These results are consistent with strategic HFT order placement strategies. HFT trade when there is a large favourable order book depth imbalance and cancel or amend their resting limit orders when the imbalance moves in an unfavourable direction.

Table 5, Panels B and C show that *Institutions* and *Retail* are less strategic in their order placement strategies. While *Institutions* are more likely to submit aggressive market orders when the *Adjusted DI* is large, both *Institutions* and *Retail* are more likely to receive a limit order execution when *Adjusted DI* is lower, relative to *Adjusted DI* at the time of order submission. Consistent with our earlier findings, this result indicates that *Institutions* and *Retail* are less successful at monitoring the limit order book and are more likely to face picking-off risk due to stale orders resting in the book.

4.4 Volatility and HFT strategies

In this section, we further investigate HFT trading behavior in times of high market volatility. We divide each trading day into 30 minute intervals and for each interval, measure its volatility by taking the natural log of the high price divided by the low price during the period. For each stock, we then rank the 30 minute intervals into 10 deciles based on its volatility. Decile 0 (9) contains the least (most) volatile periods. For each decile, we also determine the amount of aggressive volume and passive volume as a percentage of total market volume for each broker type. Figure 3, Panels A to C present the graphs of volatility against aggressive and passive volumes for *HFT*, *Institutions*, and *Retail*, respectively.

[Figure 3]

If HFT trade to decrease market volatility, we expect them to supply more passive volume in times of high market uncertainty. In contrast, Figure 3, Panel A shows that HFT aggressive volume increases while their passive volume decreases as the market becomes more volatile. For

Institutions in Figure 3, Panel B, we observe a fall in both aggressive and passive volumes as volatility increases, which is consistent with *Institutions* withdrawing from the market in periods of high uncertainty. For *Retail*, with higher stock volatility, we observe a sharp decrease in aggressive volume (Figure 3, Panel C). *Retail* passive volume at first decreases and then increases when the market becomes more volatile. This trading pattern could indicate that stale retail limit orders are picked off the limit order book by aggressive HFT orders during volatile periods.

We test the relationships observed in Figure 3 more formally using the following regression model, controlling for stock and day fixed effects:

$$\begin{aligned} Aggressive\ volume\ \% &= \beta_0 + \beta_1 I(HFT) \times I(Low\ volatility) + \beta_2 I(HFT) \times I(High\ volatility) \\ &+ \beta_3 I(Institutions) \times I(Low\ volatility) + \beta_4 I(Institutions) \\ &\times I(High\ volatility) + \beta_5 I(Low\ volatility) + \beta_6 I(High\ volatility) \\ &+ \beta_7 I(HFT) + \beta_8 I(Institutions) + \beta_9 Volatility + \beta_{10} Volume + \varepsilon \qquad (4) \end{aligned}$$

where the dependent variable, *Aggressive volume %*, is the aggressive volume as a percentage of total aggressive and passive volume for each broker type and volatility decile. *I(HFT)* and *I(Institutions)* are indicator variables for HFT and institutional brokers, respectively. *I(Low volatility)* (*I(High volatility)*) is an indicator variable equal to 1 for the lowest (highest) volatility decile, and 0 otherwise. For each volatility decile, *Volatility* is the natural log of the high price divided by the low price and *Volume* is the total number of shares traded.

Consistent with our observations from Figure 3, Table 6, Column 1 shows that HFT *Aggressive volume %* is larger (smaller) when volatility is high (low). This finding is similar across both large stocks and small stocks in Columns 2 and 3. In Table 6, Columns 4 to 6, we replace the dependent variable with *Aggressive Trade %*. Again, we find HFT trade more aggressively in times of high market volatility, which could potentially exacerbate future volatility especially if HFT are taking liquidity from the thin side of the order book.

[Insert Table 6]

To test whether HFT are more strategic in times of market volatility, we investigate the relationship between *Adjusted DI* and stock volatility for each broker category. As discussed earlier, *Adjusted DI* measures a trader's ability to condition their trades on information contained in the order book. Specifically, a positive *Adjusted DI* indicates that a trade executes in the direction of a favourable imbalance. Thus, *Adjusted DI* is more positive for traders better able to trade in the direction of the imbalance, or capture predicted future price movements based on the shape of the order book. For each market volatility decile, we calculate the average *Adjusted DI* for both the aggressive and passive trades in the decile.

[Insert Figure 4]

Figure 4, Panels A and B, present the average *Adjusted DI* for aggressive and passive trades, respectively. Comparing between the broker categories, we observe a large difference in trading behaviors. Notably, for both passive and aggressive trades, we observe a sharp increase in *Adjusted DI* for *HFT* as stock volatility increases.

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

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

$$
\begin{aligned}
Adjusted\ DI = \\
&= \beta_0 + \beta_1 I(HFT) \times I(Low\ volatility) + \beta_2 I(HFT) \times I(High\ volatility) \\
&+ \beta_3 I(Institutions) \times I(Low\ volatility) + \beta_4 I(Institutions) \\
&\times I(High\ volatility) + \beta_5 I(Low\ volatility) + \beta_6 I(High\ volatility) \\
&+ \beta_7 I(HFT) + \beta_8 I(Institutions) + \beta_9 Volatility + \beta_{10} Volume + \varepsilon \qquad (5)
\end{aligned}
$$

We estimate the regression separately for aggressive trades and passive trades. Table 7, Column 1 presents the results for aggressive trades based on the full sample of stocks. As expected, we find for *HFT*, *Adjusted DI* is larger when the market is more volatile, indicating that HFT are more strategic when uncertainty exists. The findings are consistent across both the large stock and small stock subsamples (Columns 2 and 3, respectively). While *Institutions* exhibit similar trading behaviors to *HFT*, the magnitudes of the coefficients are significantly lower.

[Insert Table 7]

Table 7, Columns 4 to 6 present the regression coefficients for passive trades. Consistent with our earlier results, we find that HFT are able to successfully implement limit order strategies, especially when the market is volatile. Relative to *Retail* passive orders, which is captured in the intercept coefficient, *HFT* passive orders execute with a larger *Adjusted DI*, when more stock volatility exists. This finding supports our hypothesis that HFT are more successful at monitoring their limit orders, in particular, when there is high stock volatility. In contrast, *I(High volatility)* is negative and significant across all stock samples indicating that retail limit orders are picked off more when the market is volatile. One further implication of our results is that *HFT* aggressively pick off stale orders from the thin side of the order book, which potentially exacerbates future stock volatility when the market is already highly volatile.

4.5 Introduction of ITCH

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

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

$$Volume\ imbalance\ \% = \beta_0$$
$$+I(Pre - ITCH) \times [\beta_1 I(HFT) \times DI + \ \beta_2 I(Institutions) \times DI + \beta_3 DI]$$
$$+I(Post - ITCH) \times [\beta_4 I(HFT) \times DI + \ \beta_5 I(Institutions) \times DI + \beta_6 DI]$$
$$+\beta_7 I(HFT) + \ \beta_8 I(Institutions) + \beta_9 Volume + \varepsilon \qquad (6)$$

Table 8, Columns 1 and 2, reports the two sets of coefficients $\{\beta_1, \beta_2, \beta_3\}$ and $\{\beta_4, \beta_5, \beta_6\}$. The test of equality between β_1 and β_4 (i.e., $\beta_4 - \beta_1 = 0$) indicates whether HFT strategies capture more information contained in the order book depth imbalance after the implementation of ITCH. Similarly, the test of equality between β_2 and β_5, and β_3 and β_6 tests whether institutional and retail strategies change as a result of ITCH, respectively.

[Insert Table 8]

Table 8, Columns 1 and 2 show that β_1 and β_4 are both positive and significant indicating that *HFT* buy when *DI* is positive and sell when *DI* is negative in both the pre- and post-ITCH periods. Importantly, the estimate of β_4 is larger than β_1 and the F-test in Column 3 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. As a group, *HFT* trade more strategically on information contained in the limit order book as their market information access speeds increase. In contrast, we cannot reject the null hypothesis that $\beta_2 = \beta_5$ and $\beta_3 = \beta_6$. Consistent with our expectations, we do not find evidence that *non-HFT*, who are less speed sensitive, trade more strategically after the adoption of ITCH.

In Table 8, Columns 4 and 5, we replace the dependent variable with trade imbalance. Similar to our results based on volume imbalances, we find that *HFT* trade more strategically post-ITCH (F-test = 7.71). For *Institutions*, we also find that their trading is more sensitive to *DI* after the implementation of ITCH (F-test = 5.27). It is possible that some more speed sensitive institutional brokers also subscribed to ITCH to take advantage of the faster speeds. While it is difficult for an individual broker to influence average volume imbalances, we find that their number of buy and sell orders capture more information contained in order book depth imbalances after trading becomes faster. Based on trade imbalances, we find that *Retail*, who are less likely to compete on speed, trade less strategically post-ITCH. Specifically, we find that β_6 is negative and significant in the post-ITCH period.

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

$$P(fill) = \frac{\sum TradeVolume}{\sum SubmitVolume}$$

Where $\sum SubmitVolume$ is the total daily volume submitted to the top level of the limit order book and $\sum TradeVolume$ is the total volume that is executed. We measure $P(fill)$ on a daily basis for each broker category. We then estimate the following regression model:

$$\begin{aligned} P(Fill) = {} & \beta_0 + \beta_1 I(Non-HFT) \times I(Post-ITCH) + \beta_2 I(Non-HFT) \\ & + \beta_3 I(Post-ITCH) + \beta_4 Volatility + \beta_5 Volume + \beta_6 Price + \beta_7 Spread + \varepsilon \end{aligned} \tag{6}$$

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

[Insert Table 9]

Our main variable of interest is $I(Non - HFT) \times I(Post - ITCH)$. If HFT are crowding out non-HFT limit orders from the order book, we expect a negative coefficient for $I(Non - HFT) \times I(Post - ITCH)$, indicating that the probability of a non-HFT order receiving execution decreases when HFT become faster. Consistent with our expectations, Table 8, Column 1 shows that the probability of limit order execution falls for non-HFT after the introduction of ITCH. In Table 9, Column 2, we separate non-HFT traders into *Institutions* and *Retail*. For both *Institutions* and *Retail*, we find the interaction term with *I(Post-ITCH)* is negative and significant, indicating that both *Institutions* and *Retail* are crowded out from the order book by HFT.

From our earlier results, we show that a limit order trader benefits when an order executes with a lot of depth on the same side of the order book. 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. In Table 9, Columns 3 to 6, we separate *P(Fill)* into favorable and unfavorable fills. We define a favourable 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. For example, if the limit order is sitting on the best bid price, a favourable fill occurs when there is more depth on the bid side of the order book, relative to the ask side.

Our results show that the decrease in *P(Fill)* is driven by a fall in the volume of favorable executions. Comparing between Columns 3 and 5, we find that *P(Favorable fill)* falls for non-HFT while *P(Unfavorable fill)* remains unchanged after the implementation of ITCH. Separating non-HFT into *Institutions* and *Retail* in Columns 4 and 6, we document similar findings. Specifically, the interaction terms with *I(Post-ITCH)* is negative and significant in Column 4 and insignificant in Column 6. Thus, the likelihood of receiving a favourable limit order execution falls for both institutional and retail traders when HFT gain a speed advantage. Taken together, these results provide evidence that HFT are crowding out non-HFT limit orders from the order book, especially when trading becomes faster.

5.0 Conclusion

The information channels through which HFT trade are relatively unknown. We present strong evidence of one such channel, which provides an explanation for many of the findings documented in the prior literature. We show that order book depth imbalances are strong predictors of future prices. HFT are highly sophisticated in monitoring order book imbalances, which allows them to trade ahead of these predicted price changes. At the same time, when the order book imbalance moves in an unfavourable direction, they are quick to cancel or amend orders that are at high risk of being picked off by other traders.

HFT order placement strategies based on order book imbalances are particularly successful when the market is volatile. During times of high market volatility, the chance of an institutional or retail limit order becoming stale increases. We find that HFT trade more aggressively and are more successful at picking off stale orders from institutional and retail investors when the market is volatile. However, by demanding liquidity from the thin side of the order book, one consequence is that HFT could potentially exacerbate future limit order book imbalances.

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

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

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

References

Aitken, M., Chen, H. and Foley, S. (2016) The impact of fragmentation, exchange fees and liquidity provision on market quality. *Journal of Empirical Finance,* Forthcoming.

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

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

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

Brogaard, J., Carrion, A., Moyaert, T., Riordan, R., Shkilko, A. and Sokolov, K. (2016) High-frequency trading and extreme price movements. *SSRN Working Paper*.

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

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

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

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

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

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

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

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

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

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

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

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

Hagströmer, B. and Nordén, L. 2013. The diversity of high-frequency traders. Stockholm University Working Paper.

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

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

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

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

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

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

Kirilenko, A., Kyle, A., Samadi, M. and Tuzun, T. (2016) The flash crash: High frequency trading in an electronic market. *Journal of Finance,* Forthcoming.

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

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

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

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

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

Naes, R. and Skjeltorp, J. (2006) Order book characteristics and the volume-volatility relation: Empirical evidence from a limit order market. *SSRN Working Paper*.

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

Rosu, I. (2016) Fast and slow informed trading. *SSRN Working Paper*.

Subrahmanyam, A. and Zheng, H. (2016) Limit order placement by high-frequency traders. *SSRN Working Paper*.

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

Van Kervel, V. and Menkveld, A. J. (2016) High-frequency trading around large institutional orders. *SSRN Working Paper*.

Appendix 1

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

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

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

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

Table A1 summarizes the mean adjusted R-square for the restricted and unrestricted models. If $DI_{Levels2-5}$ adds incremental information about the future price movements, we expect a higher adjusted R-square for the unrestricted model, relative to the restricted model. For over 85% of our regressions, the F-test is significant at the 1% level, indicating that $DI_{Levels2-5}$ adds additional explanatory power. This result indicates that limit orders behind the best bid and offer also contains information on future stock returns.

Table A1. Adjusted R-square for restricted and unrestricted model.

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

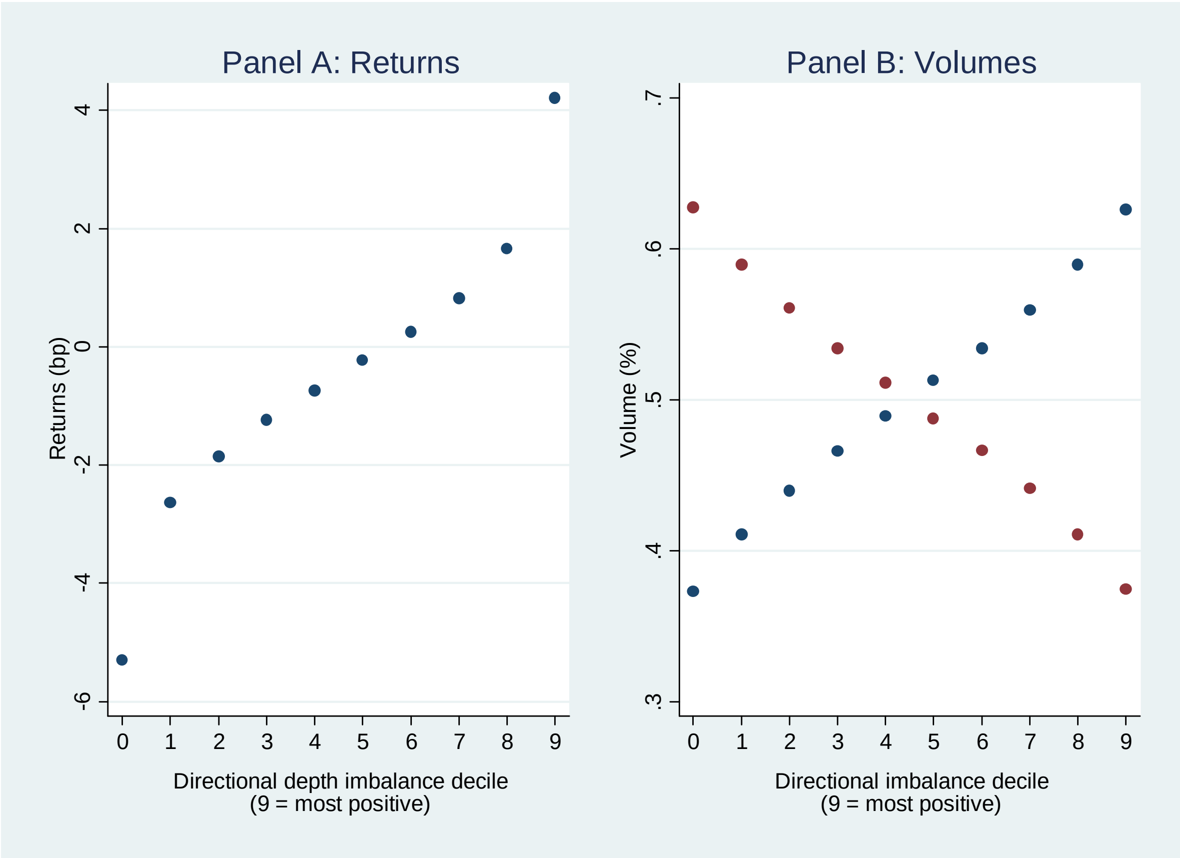

Fig. 1. Fig. 1 shows the relationship between depth imbalance, returns (Panel A) and volumes (Panel B). Directional depth imbalance is calculated as the difference between the depths available at the five best bid and ask prices, scaled by total depth available at these price levels, immediately before each trade. For each stock day, we rank trades into 10 depth imbalance deciles. Trades with the most negative depth imbalances (i.e., ask depth > bid 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 blue (red) dots represent the average percentage of buyer (seller) initiated volume, relative to total trade volume, for each depth imbalance decile.

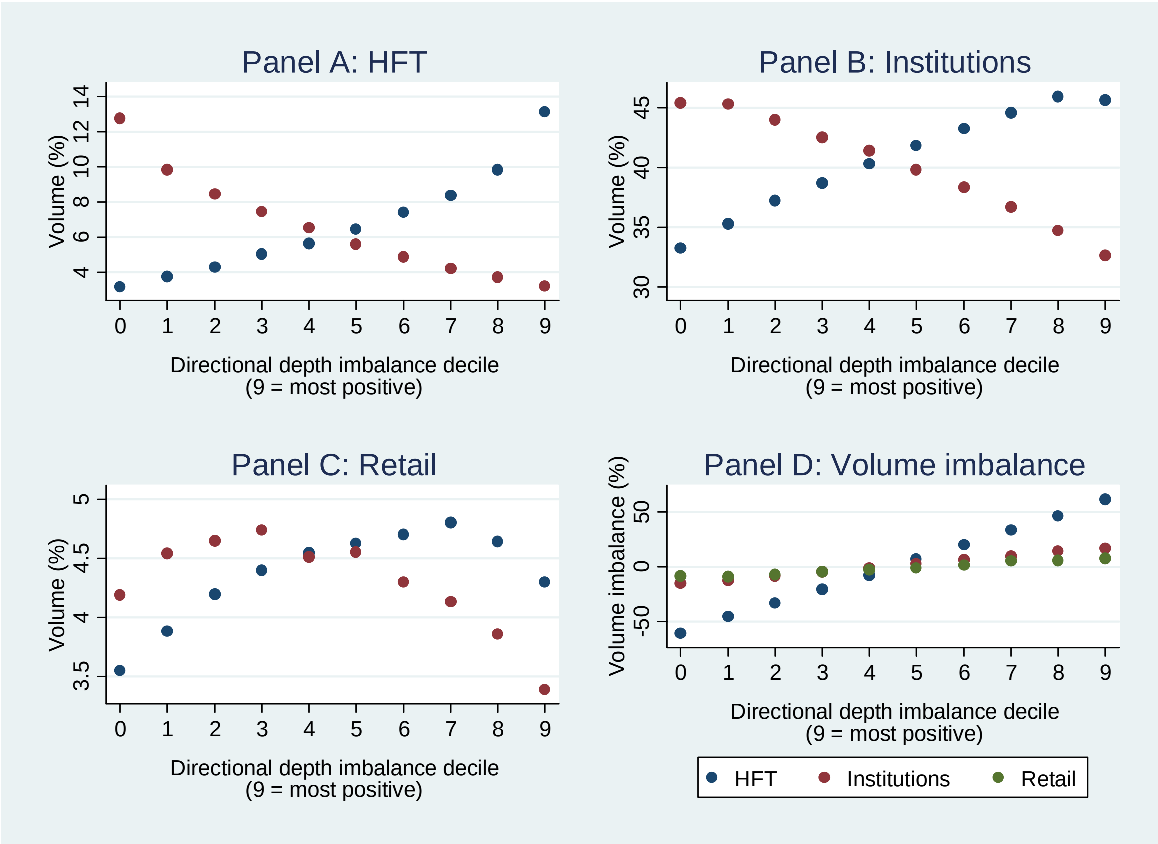

Fig 2. Fig. 2 shows the relationship between depth imbalance and trading volumes for each broker category. Directional depth imbalance is calculated as the difference between the depths available at the five best bid and ask prices, scaled by total depth available at these price levels, immediately before each trade. For each stock day, we rank trades into 10 depth imbalance deciles. Trades with the most negative depth imbalances (i.e., ask depth > bid 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 blue (red) dots represent the average percentage of buyer (seller) initiated volume, relative to total trade volume, for each depth imbalance decile and broker type. Panel D shows the volume imbalance (i.e., (Buys-Sells)/(Buys + Sells)) for each broker type and depth imbalance decile.

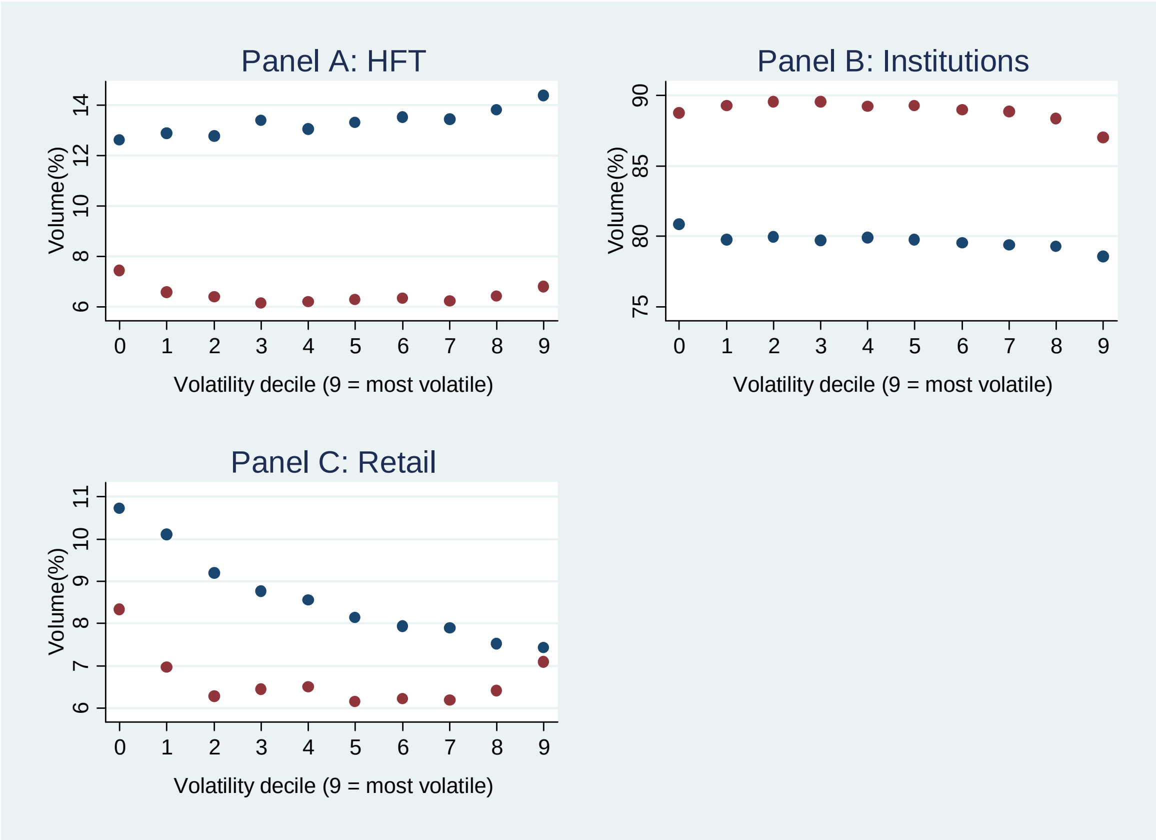

Fig. 3. Fig. 3 shows the relationship between volatility and trading volume for each broker category. Volatility is calculated as the log of ratio of the high to the low price over each 30 minute trading interval. Trading intervals are then ranked into volatility deciles. Panels A-C present the results for *HFT*, *Institutions*, and *Retail*, respectively. For each broker category, we calculate the percentage of aggressive (blue) and passive (red) trading volume, relative to total trading volume, for each volatility decile.

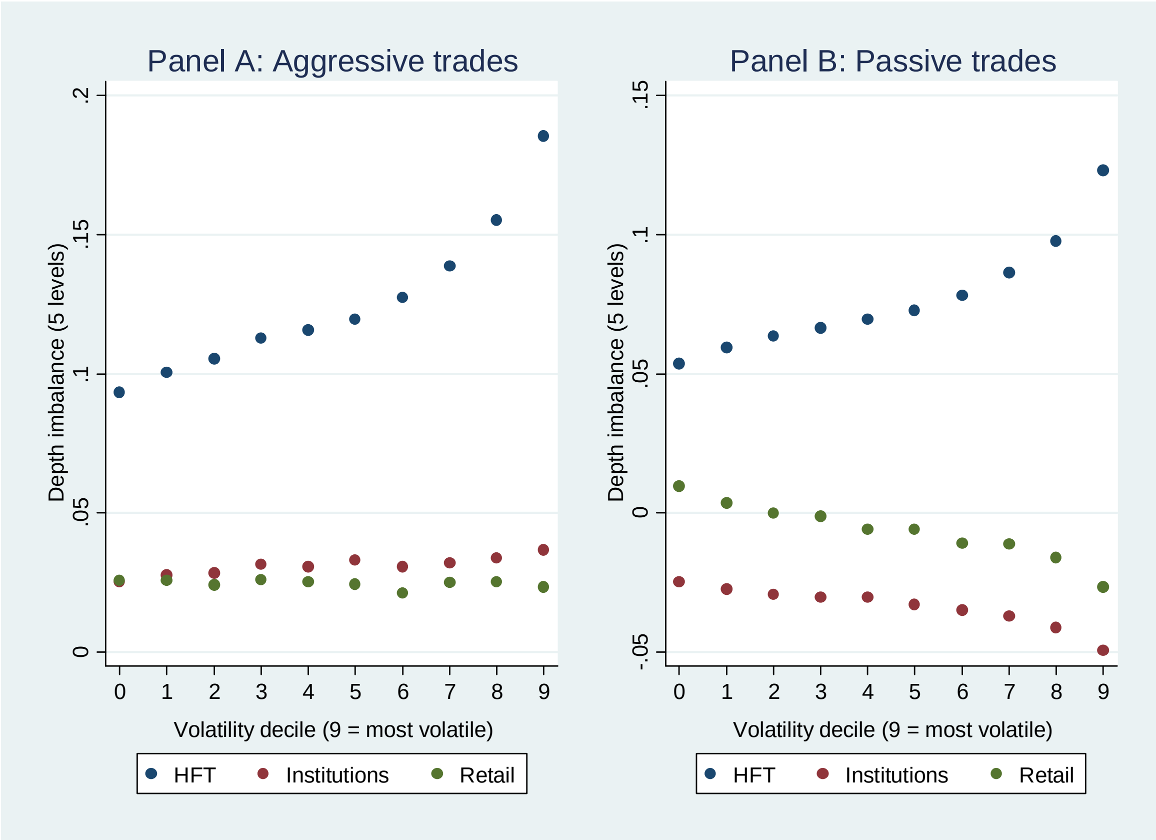

Fig. 4. Fig. 4 shows the relationship between volatility and depth imbalance for each broker category. Volatility is calculated as the log of ratio of the high to the low price over each 30 minute trading interval. Trading intervals are then ranked into volatility deciles. For each broker category, we calculate the depth imbalance immediately before each aggressive (Panel A) or passive (Panel B) trade execution. Depth imbalance is calculated as the difference between the depths available at the five best bid and ask prices, scaled by total depth available at these price levels, immediately before each trade. We multiply depth imbalance by a buy or sell indicator so that buys and sells can be interpreted together. Depth imbalances are then averaged over each volatility decile by broker category.

Table 1

Summary statistics

Table 1, Panel A reports statistics for the 94 stocks that remain in the ASX 100 index for the period January 3, 2012 to December 31, 2012. *Market capitalization* is the stock's market capitalization on January 3, 2012. *Dollar volume* is the average daily dollar volume in AUD. *Number of trades* is the average daily number of transactions. *Price* is the average trade price in AUD. *Volatility* is the difference between the log of the intraday high ask price and the log of the intraday low bid price. *Spread* is the time weighted average difference between the best bid and offer prices in AUD cents. The broker associated with each order book event is classified into three types: proprietary HFT (*HFT*), institutional (*Institutions*), or retail (*Retail*). Panel B reports the trading characteristics for each broker type. Panel C reports the average adjusted depth imbalance (*Adjusted DI*) for each trader type. For each order book event, *Adjusted DI* is calculated as:

$$Adjusted\ DI_t = q \times \frac{\sum_{i=1}^{5} VolBid_{it} - \sum_{i=1}^{5} VolAsk_{it}}{\sum_{i=1}^{5} VolBid_{it} + \sum_{i=1}^{5} VolAsk_{it}}$$

where $\sum_{i=1}^{5} VolBid_t$ ($\sum_{i=1}^{5} VolAsk_t$) is the volume 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.

Panel A: Stock characteristics					
	Mean	Std.dev.	Q1	Median	Q3
Market capitalization (AUD billions)	13.52	22.77	2.844	10.00	114.8
Dollar volume (AUD millions)	25.54	43.67	5.179	10.51	23.44
Number of trades	2,176	1,718	1,088	1,633	2,614
Price (AUD)	11.67	13.18	3.052	6.431	15.05
Volatility (%)	2.026	1.215	1.280	1.756	2.443
Spread (cents)	1.037	0.369	0.956	1.014	1.119

	HFT	Institutions	Retail
Panel B: Trader characteristics			
Average daily submissions	839.5	12,781	525.6
Average daily cancelations	375.9	4,309	58.79
Average daily trades (aggressive)	241.9	1,463	98.90
Average daily trades (passive)	279.4	3,529	167.7
Median trade size	1,681	926.5	2,187
Median submission to cancel time	128.7	246.8	3,034
Panel C: Adjusted depth imbalance			
Trades (aggressive)	0.148	0.024	0.024
Trades (passive)	0.083	-0.030	-0.012
Submissions	0.059	-0.005	0.030
Amendments	0.043	-0.003	0.014
Cancelations	0.017	0.003	0.027

Table 2

Volume and depth imbalances by trader type

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

$$DI_t = \frac{\sum_{i=1}^{5} VolBid_{it} - \sum_{i=1}^{5} VolAsk_{it}}{\sum_{i=1}^{5} VolBid_{it} + \sum_{i=1}^{5} VolAsk_{it}}$$

where $\sum_{i=1}^{5} VolBid_t$ ($\sum_{i=1}^{5} VolAsk_t$) is the volume available at the top 5 bid (ask) price levels immediately before the order book event, *t*. Then, for each *DI* decile, we calculate *Volume imbalance* as:

$$Volume\ imbalance_j = \frac{\sum_{j=1}^{n} BuyVolume_j - \sum_{i=1}^{n} SellVolume_j}{\sum_{j=1}^{n} BuyVolume_j + \sum_{i=1}^{n} SellVolume_j}$$

where $\sum_{j=1}^{n} BuyVolume_j$($\sum_{j=1}^{n} SellVolume_j$) is the total aggressive buying (selling) volume for depth imbalance decile, *j*. Column 2 reports the average *DI*. Columns 3-5 report the average *Volume imbalance* for *HFT, Institutions* and *Retail*. Columns 6-8 report the differences in *Volume imbalance* means between the trader types as indicated in the column headings based on a *t*-test. ***, ** and * indicate significance levels of 1%, 5% and 10%, respectively.

(1) Depth imbalance decile	(2) Depth imbalance	(3) HFT	(4) Institutions	(5) Retail	(6) HFT vs. Institutions		(7) HFT vs. Retail		(8) Institutions vs. Retail	
0 (most negative)	-0.375	-61.8	-16.8	-6.2	-44.9	***	-55.6	***	-55.6	***
1	-0.219	-46.8	-12.8	-7.4	-34.0	***	-39.5	***	-39.5	***
2	-0.141	-34.2	-8.6	-5.6	-25.6	***	-28.6	***	-28.6	***
3	-0.080	-21.3	-4.6	-4.1	-16.7	***	-17.2	***	-17.2	
4	-0.025	-6.9	-1.0	-1.8	-5.9	***	-5.0	***	-5.0	*
5	0.028	6.4	2.8	0.0	3.5	***	6.3	***	6.3	***
6	0.084	20.2	6.0	1.3	14.3	***	18.9	***	18.9	***
7	0.146	33.3	9.8	3.7	23.5	***	29.6	***	29.6	***
8	0.225	47.1	14.2	4.4	32.9	***	42.7	***	42.7	***
9 (most positive)	0.380	62.6	17.1	6.1	45.5	***	56.5	***	56.5	***

Table 3

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

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

$$Volume\ imbalance_j = \frac{\sum_{j=1}^{n} BuyVolume_j - \sum_{i=1}^{n} SellVolume_j}{\sum_{j=1}^{n} BuyVolume_j + \sum_{i=1}^{n} SellVolume_j}$$

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

$$Volume\ imbalance\ \% = \beta_0 + \beta_1 I(HFT) \times DI + \beta_2 I(Institutions) \times DI + \beta_3 I(HFT) + \beta_4 I(Institutions) + \beta_5 DI + \beta_6 Volume + \varepsilon$$

where *I(HFT)* and *I(Institutions)* are indicator variable for *HFT* and *Institutions*, respectively. *DI* is the average depth imbalance for the decile and *Volume* is the natural log of the total share volume traded in the decile. In Columns 4-6, we replace the dependent variable with *Trade imbalance*, which is calculated based on the number, rather than the volume, of aggressive trades. 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. We report heteroskedasticity-robust standard errors double clustered by stock and day in parentheses. ***, ** and * indicate significance levels of 1%, 5% and 10%, respectively.

	Volume imbalance%			Trade imbalance%		
	(1) All trading days	(2) Low volatility days	(3) High volatility days	(4) All trading days	(5) Low volatility days	(6) High volatility days
I(HFT) × DI	1.017***	0.980***	1.084***	0.921***	0.895***	0.962***
	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)
I(Institutions) × DI	-0.021	0.006	-0.036*	0.080***	0.081***	0.088***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
I(HFT)	0.015	0.024*	0.009	0.022*	0.028**	0.014
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
I(Institutions)	0.018*	0.024*	0.016	0.043***	0.047***	0.036***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
DI	-0.204***	-0.183***	-0.242***	-0.138***	-0.140***	-0.129***
	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)	(0.02)
Volume	0.011***	0.007**	0.008**	0.010***	0.006	0.009**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	-0.288***	-0.209***	-0.242***	-0.288***	-0.292***	-0.216***
	(0.03)	(0.04)	(0.05)	(0.03)	(0.05)	(0.05)
Obs.	503,990	150,376	166,644	503,990	150,376	166,644
Adj. R-square	0.183	0.175	0.198	0.254	0.242	0.283

Table 4
Limit order placement strategies

Table 4 compares the *Adjusted DI* immediately before order book events for *HFT*, *Institutions* and *Retail*. The dependent variable is *Adjusted DI,* which is the daily average *Adjusted DI* for each order book event and trader type:

$$Adjusted\ DI_t = q \times \frac{\sum_{i=1}^{n} VolBid_{it} - \sum_{i=1}^{n} VolAsk_{it}}{\sum_{i=1}^{n} VolBid_{it} + \sum_{i=1}^{n} VolAsk_{it}}$$

where $\sum_{i=1}^{n} VolBid_t$ ($\sum_{i=1}^{n} VolAsk_t$) is the volume available at the top *n* 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. We present the coefficient estimates for the following linear regression:

$$Adjusted\ DI = \beta_0 + \beta_1 I(Aggressive\ trade) + \beta_2 I(Passive\ trade) + \beta_3 I(Amend) + \beta_4 I(Cancel) + \beta_5 Volatility + \beta_6 Volume + \beta_7 Price + \beta_8 Spread + \varepsilon$$

where $I(\cdot)$ is an indicator variable equal to 1 for the order book event specified in the parentheses and 0 otherwise. All control variables are measured on a daily basis. *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. All regressions control for stock and day fixed effects. We report heteroskedasticity-robust standard errors double clustered by stock and day in parentheses. ***, ** and * indicate significance levels of 1%, 5% and 10%, respectively.

	Adjusted depth imbalance (5 levels)			Adjusted depth imbalance (1 level)		
	(1)	(2)	(3)	(4)	(5)	(6)
	HFT	Institutional	Retail	HFT	Institutional	Retail
I(Aggressive trade)	0.087***	0.029***	-0.006***	0.330***	0.221***	0.059***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)
I(Passive trade)	0.024***	-0.026***	-0.041***	0.002	-0.095***	-0.077***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
I(Amend)	-0.014***	0.001**	-0.015***	-0.026**	-0.061***	-0.030***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
I(Cancel)	-0.043***	0.008***	-0.003	-0.278***	-0.035***	-0.018***
	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)
Volatility	0.380***	0.105**	0.206**	0.101	0.070	0.046
	(0.10)	(0.04)	(0.08)	(0.08)	(0.04)	(0.07)
Volume	-0.001	0.004***	-0.000	0.001	0.002*	0.001
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Price	-0.007	-0.006**	0.001	0.025	-0.003	-0.005
	(0.01)	(0.00)	(0.01)	(0.02)	(0.00)	(0.01)
Spread	3.871***	0.545**	-0.496	1.753*	-1.009***	-0.483
	(0.64)	(0.26)	(0.71)	(0.99)	(0.27)	(0.58)
Constant	-0.004	-0.071***	0.040	-0.025	-0.025	0.033
	(0.04)	(0.02)	(0.03)	(0.05)	(0.02)	(0.03)
Obs.	109,351	111,417	110,409	109,351	111,417	110,409
Adj. R-square	0.265	0.132	0.042	0.531	0.712	0.091

Table 5

Multinomial logistic regressions for limit order placement strategies

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

$$\Pr(OrderBookEvent) = \beta_0 + \beta_1 Adjusted\ DI + + \beta_2 Volatility + \beta_3 Volume + \beta_4 Price + \beta_5 Spread + \varepsilon$$

here *OrderBookEvent* is the dependent variable indicating one of five order book events: Aggressive trade, passive trade, limit order submission, amendment or cancelation. *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 the daily average *Adjusted DI* for each order book event and trader type:

$$Adjusted\ DI_t = q \times \frac{\sum_{i=1}^{n} VolBid_{it} - \sum_{i=1}^{n} VolAsk_{it}}{\sum_{i=1}^{n} VolBid_{it} + \sum_{i=1}^{n} VolAsk_{it}}$$

where $\sum_{i=1}^{n} VolBid_t$ ($\sum_{i=1}^{n} VolAsk_t$) is the volume available at the top *n* 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. We estimate the model with limit order submission as the baseline category. Panels A to C present the results for *HFT*, *Institutions* and *Retail*, respectively. ***, ** and * indicate significance levels of 1%, 5% and 10%, respectively.

	Aggressive Trade	Passive Trade	Amendment	Cancelation
		Panel A: HFT		
Adjusted DI	8.062***	2.435***	-1.494***	-3.757***
	(0.10)	(0.10)	(0.10)	(0.09)
Volatility	-9.492***	-2.610***	0.092	3.607***
	(0.91)	(0.85)	(0.84)	(0.80)
Volume	0.225***	0.042***	0.069***	-0.008
	(0.01)	(0.01)	(0.01)	(0.01)
Price	-0.293***	-0.086***	-0.030*	0.106***
	(0.02)	(0.02)	(0.02)	(0.02)
Qspread	-18.047***	-3.990	4.127	4.982
	(4.37)	(4.19)	(4.38)	(4.27)
Constant	-3.530***	-0.605***	-1.110***	-0.049
	(0.18)	(0.17)	(0.17)	(0.16)
Obs.			109,351	
Pseudo R-square			0.0496	
		Panel B: Institutions		
Adjusted DI	13.432***	-10.111***	0.681***	3.897***
	(0.22)	(0.21)	(0.21)	(0.21)
Volatility	-0.816	1.775**	-0.069	-0.334
	(0.85)	(0.81)	(0.83)	(0.83)
Volume	-0.006	0.057***	-0.002	-0.010
	(0.01)	(0.01)	(0.01)	(0.01)
Price	0.071***	-0.012	0.002	0.011
	(0.02)	(0.02)	(0.02)	(0.02)
Qspread	-17.359***	-0.962	-0.291	-2.384
	(4.30)	(4.28)	(4.19)	(4.19)
Constant	0.053	-1.093***	0.039	0.179
	(0.17)	(0.17)	(0.16)	(0.16)
Obs.			111,417	
Pseudo R-square			0.0373	

Panel C: Retail				
Adjusted DI	-0.502***	-3.622***	-1.320***	-0.214**
	(0.09)	(0.09)	(0.09)	(0.09)
Volatility	-0.366	1.144	0.571	0.399
	(0.82)	(0.82)	(0.82)	(0.82)
Volume	0.007	0.005	0.016	0.008
	(0.01)	(0.01)	(0.01)	(0.01)
Price	0.015	0.017	0.008	0.004
	(0.02)	(0.02)	(0.02)	(0.02)
Qspread	-1.798	-2.801	0.915	0.513
	(4.16)	(4.22)	(4.19)	(4.17)
Constant	-0.092	-0.071	-0.302*	-0.163
	(0.16)	(0.16)	(0.16)	(0.16)
Obs.			110,409	
Pseudo R-square			0.0061	

Table 6

Relation between *Aggressive volume %*, *Aggressive trade %* and stock volatility

Table 5 presents the regression of *Aggressive volume %* or *Aggressive trade %* against volatility. The dependent variable is *Aggressive volume %*, which is the aggressive buying and selling volume as a percentage of total aggressive and passive volume for each broker type and volatility decile. The results are based on the following linear regression, which is based on 30-minute time intervals:

$$\begin{aligned} Aggressive\ volume\ \% &= \beta_0 + \beta_1 I(HFT) \times I(Low\ volatility) + \beta_2 I(HFT) \times I(High\ volatility) + \beta_3 I(Institutions) \times I(Low\ volatility) \\ &+ \beta_4 I(Institutions) \times I(High\ volatility) + \beta_5 I(Low\ volatility) + \beta_6 I(High\ volatility) + \beta_7 I(HFT) + \beta_8 I(Institutions) \\ &+ \beta_9 Volatility + \beta_{10} Volume + \varepsilon \end{aligned}$$

For each stock, we rank the 30 minute intervals into volatility deciles. *Volatility* is the difference between the log of the highest best ask price and the log of the lowest best bid price during the 30 minute interval. *I(Low volatility)* (*I(High volatility)*) is an indicator variable equal to 1 if the 30 minute interval is in the lowest (highest) decile based on stock volatility. *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 daily share volume during the 30 minute interval. *Large stocks* (*Small stocks*) refer to stocks contained in the largest (smallest) tercile of all sample stocks based on market capitalization. In Columns 4-6, we replace the dependent variable with *Aggressive trade %*, which is calculated based on the number, rather than the volume, of aggressive trades. All regressions control for stock and day fixed effects. We report heteroskedasticity-robust standard errors double clustered by stock and day in parentheses. ***, ** and * indicate significance levels of 1%, 5% and 10%, respectively.

	Aggressive volume %			Aggressive trade %		
	(1)	(2)	(3)	(4)	(5)	(6)
	All stocks	Large stocks	Small stocks	All stocks	Large stocks	Small stocks
I(HFT) × I(Low volatility)	-4.058***	-3.879***	-4.376**	-2.707***	-4.521***	-3.298*
	(0.89)	(1.38)	(1.82)	(0.96)	(1.21)	(1.84)
I(HFT) × I(High volatility)	7.432***	6.517***	7.742***	7.381***	6.613***	7.790***
	(0.62)	(0.91)	(1.58)	(0.87)	(1.19)	(1.94)
I(Institutions) × I(Low volatility)	3.326***	-0.763	9.950***	4.912***	1.153	8.360***
	(0.81)	(0.83)	(1.54)	(0.78)	(0.97)	(1.37)
I(Institutions) × I(High volatility)	4.483***	6.948***	1.917	2.223***	5.015***	-1.644*
	(0.52)	(0.51)	(1.18)	(0.55)	(0.59)	(0.87)
I(Low volatility)	-1.432**	1.017	-6.787***	-2.378***	0.363	-5.427***
	(0.60)	(0.62)	(1.17)	(0.62)	(0.65)	(0.96)
I(High volatility)	-4.996***	-6.205***	-2.782***	-3.254***	-5.021***	-0.493
	(0.37)	(0.38)	(0.92)	(0.44)	(0.57)	(0.72)
I(HFT)	13.677***	10.162***	14.839***	16.521***	13.366***	16.784***
	(1.39)	(1.75)	(3.00)	(2.01)	(2.65)	(3.90)
I(Institutions)	-10.873***	-12.163***	-9.705***	2.624**	-5.138***	10.603***
	(0.72)	(1.04)	(1.65)	(1.15)	(1.52)	(1.51)
Volatility	88.122***	40.493	53.509**	131.281***	111.875*	116.592***
	(20.72)	(58.29)	(22.78)	(21.71)	(62.18)	(28.93)
Volume	-0.684***	-0.880***	-0.540***	-0.615***	-0.817***	-0.265
	(0.08)	(0.13)	(0.15)	(0.12)	(0.18)	(0.20)
Constant	61.237***	67.200***	59.557***	45.037***	55.510***	37.733***
	(1.23)	(2.54)	(2.11)	(1.76)	(2.96)	(2.18)
Obs.	449,556	175,152	118,303	449,556	175,152	118,303
Adj. R-square	0.241	0.239	0.207	0.200	0.245	0.162

Table 7

Relation between *Adjusted DI* and stock volatility

Table 5 presents the regression of *Adjusted DI* against volatility. The dependent variable is *Adjusted DI,* which is the average *Adjusted DI* for aggressive or passive trades during each 30 minute time interval, as indicated in the table heading. For each trade, *Adjusted DI* is calculated as:

$$Adjusted\ DI_t = q \times \frac{\sum_{i=1}^{5} VolBid_{it} - \sum_{i=1}^{5} VolAsk_{it}}{\sum_{i=1}^{5} VolBid_{it} + \sum_{i=1}^{5} VolAsk_{it}}$$

where $\sum_{i=1}^{5} VolBid_t$ ($\sum_{i=1}^{5} VolAsk_t$) is the volume available at the top *5* bid (ask) price levels immediately before the trade, *t*. The results are based on the following linear regression, which is based on 30-minute time intervals:

$$Adjusted\ DI = \beta_0 + \beta_1 I(HFT) \times I(Low\ volatility) + \beta_2 I(HFT) \times I(High\ volatility) + \beta_3 I(Institutions) \times I(Low\ volatility) + \beta_4 I(Institutions) \times I(High\ volatility) + \beta_5 I(Low\ volatility) + \beta_6 I(High\ volatility) + \beta_7 I(HFT) + \beta_8 I(Institutions) + \beta_9 Volatility + \beta_{10} Volume + \varepsilon$$

For each stock, we rank the 30 minute intervals into volatility deciles. *Volatility* is the difference between the log of the highest best ask price and the log of the lowest best bid price during the 30 minute interval. *I(Low volatility)* (*I(High volatility)*) is an indicator variable equal to 1 if the 30 minute interval is in the lowest (highest) decile based on volatility. *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 daily share volume during the 30 minute interval. *Large stocks* (*Small stocks*) refer to stocks contained in the largest (smallest) tercile of all sample stocks based on market capitalization. All regressions control for stock and day fixed effects. We report heteroskedasticity-robust standard errors double clustered by stock and day in parentheses. ***, ** and * indicate significance levels of 1%, 5% and 10%, respectively.

	Aggressive trades			Passive trades		
	(1) All stocks	(2) Large stocks	(3) Small stocks	(4) All stocks	(5) Large stocks	(6) Small stocks
I(HFT) × I(Low volatility)	-0.049***	-0.035***	-0.054***	-0.055***	-0.039***	-0.064***
	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)
I(HFT) × I(High volatility)	0.065***	0.064***	0.074***	0.080***	0.068***	0.092***
	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)
I(Institutions) × I(Low volatility)	-0.013***	-0.006*	-0.009	-0.008**	-0.008**	-0.004
	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)
I(Institutions) × I(High volatility)	0.010***	0.007**	0.010**	0.007***	0.003	0.005*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
I(Low volatility)	0.006*	0.008**	-0.000	0.018***	0.017***	0.014
	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)
I(High volatility)	-0.014***	-0.017***	-0.017***	-0.024***	-0.020***	-0.026***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
I(HFT)	0.115***	0.098***	0.121***	0.101***	0.082***	0.110***
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
I(Institutions)	0.002	0.007**	0.003	-0.022***	-0.027***	-0.019***
	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)
Volatility	2.476***	2.811***	3.047***	0.009	0.097	0.093
	(0.27)	(0.51)	(0.40)	(0.13)	(0.14)	(0.17)
Volume	-0.004***	-0.003***	-0.006***	-0.002***	-0.000	-0.003***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	0.051***	0.033***	0.055***	0.004	-0.010	0.043***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)
Obs.	519904	189054	168110	519670	191623	174017
Adj. R-square	0.157	0.185	0.155	0.148	0.189	0.149

Table 8

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

Table 7 reports the regression of *Volume imbalance %* or *Trade imbalance %* against *DI*. We analyse 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). Trades are sorted into deciles based on the size of the depth imbalance (*DI*) immediately before the trade. For each *DI* decile and trader type, we calculate *Volume imbalance* as:

$$Volume\ imbalance_j = \frac{\sum_{j=1}^{n} BuyVolume_j - \sum_{i=1}^{n} SellVolume_j}{\sum_{j=1}^{n} BuyVolume_j + \sum_{i=1}^{n} SellVolume_j}$$

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

$$\begin{aligned}&Volume\ imbalance\ \% = \beta_0\\&+I(Pre-ITCH)\times[\beta_1 I(HFT)\times DI + \ \beta_2 I(Institutions)\times DI + \beta_3 DI]\\&+I(Post-ITCH)\times[\beta_4 I(HFT)\times DI + \ \beta_5 I(Institutions)\times DI + \beta_6 DI]\\&+\beta_7 I(HFT) + \ \beta_8 I(Institutions) + \beta_9 Volume + \varepsilon\end{aligned}$$

I(Pre-ITCH) (*I(Post-ITCH)*) 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-ITCH)* and Column 2 presents the coefficients associated variables interacted with *I(Post-ITCH)*. We use a *F*-test to test for the equality of the coefficients interacted with *DI*. Column 3 presents the *F*-test and the associated *p*-value in parentheses. In Columns 4-6, we replace the dependent variable with *Trade imbalance* and perform the same analysis as the previous 3 columns. All regressions control for stock and day fixed effects. We report heteroskedasticity-robust standard errors double clustered by stock and day in parentheses. ***, ** and * indicate significance levels of 1%, 5% and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Volume imbalance %			Trade imbalance %		
	Pre-ITCH	Post-ITCH	F-Test	Pre-ITCH	Post-ITCH	F-Test
I(HFT) × DI	0.942***	1.064***	5.350**	0.938***	1.066***	7.71***
	(0.05)	(0.06)	(0.023)	(0.05)	(0.05)	(0.007)
I(Institutions) × DI	-0.030	-0.038	0.040	0.017	0.128***	5.27**
	(0.04)	(0.04)	(0.850)	(0.04)	(0.04)	(0.024)
DI	-0.095**	-0.118**	0.480	-0.028	-0.121***	6.36**
	(0.05)	(0.05)	(0.490)	(0.04)	(0.04)	(0.014)
I(HFT)	0.025			0.026		
	(0.02)			(0.02)		
I(Institutions)	0.033**			0.040**		
	(0.02)			(0.02)		
Volume	0.016***			0.019***		
	(0.00)			(0.01)		
Constant	-0.373***			-0.561***		
	(0.06)			(0.07)		
Obs.	80,666			80,666		
Adj. R-square	0.186			0.278		

Table 9

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

Table 8 analyzes the proabability of limit order executions for *HFT*, *Institutions* and *Retail* before and after the implementation of ITCH. We analyse trade and quote data for the periods March 2, 2012 to March 30, 2012 (pre-ITCH) and April 9, 2012 to May 9, 2012 (post-ITCH). The main dependent variable *P(Fill)* is calculated as

$$P(fill) = \frac{\sum TradeVolume}{\sum SubmitVolume}$$

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

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

where *I(Non-HFT)* is an indicator variable equal to 1 for *Institutions* and *Retail* and zero for *HFT*. *I(Post-ITCH)* 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)* with *I(Institutions)* and *I(Retail)*, which are indicator variables equal to 1 for the trader type specified in the parentheses and 0 otherwise. In Columns 3 and 4 (Columns 5 and 6), we replace the dependent variable with *P(Favorable fill)* (*P(Unfavorable fill)*).We define a favourable (unfavourable) fill as an order execution when the limit order rests on the side of the order book with more (less) depth immediately prior to the trade. All regressions control for stock and day fixed effects. We report heteroskedasticity-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)	-0.037***		-0.040***		-0.001	
	(0.01)		(0.01)		(0.01)	
I(Non-HFT)			-0.030*		0.134***	
			(0.02)		(0.01)	
I(Institutions) × I(Post-ITCH)		-0.021**		-0.029***		0.008
		(0.01)		(0.01)		(0.00)
I(Institutions)		-0.069***		-0.124***		0.053***
		(0.02)		(0.02)		(0.01)
I(Retail) × I(Post-ITCH)		-0.057***		-0.051***		-0.009
		(0.02)		(0.01)		(0.01)
I(Retail)		0.301***		0.078***		0.226***
		(0.02)		(0.02)		(0.01)
I(Post-ITCH)	0.028*	0.029*	0.036**	0.036**	0.006	0.008
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)
Volatility	0.087	0.072	0.054	-0.090	0.007	-0.097
	(0.17)	(0.17)	(0.16)	(0.16)	(0.10)	(0.09)
Volume	0.071***	0.069***	0.039***	0.036***	0.028***	0.026***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Price	0.031	0.033	0.015	0.009	-0.006	-0.014**
	(0.04)	(0.04)	(0.02)	(0.02)	(0.01)	(0.01)
Spread	-1.359	-1.328	-0.089	-0.411	-0.199	-0.365
	(2.37)	(2.38)	(1.40)	(1.30)	(1.62)	(1.59)
Constant	-0.564***	-0.534***	-0.291***	-0.231***	-0.249***	-0.205***
	(0.08)	(0.08)	(0.06)	(0.06)	(0.04)	(0.04)
Obs.	10646	10646	9718	9718	9574	9574
Adj. R-square	0.190	0.586	0.151	0.369	0.224	0.460