Revealing High-Frequency Trading Provision of Liquidity with Visualization

Michael Hirsch
Monash University
Wellington Rd, Clayton VIC 3800, Australia
michaelpaulhirsch@gmail.com

Dianne Cook
Monash University
Wellington Rd, Clayton VIC 3800 Australia
Dicook@monash.edu

Paul Lajbcygier
Monash University
Wellington Rd, Clayton VIC 3800, Australia
paul.lajbcygier@monash.edu

Rob Hyndman
Monash University
Wellington Rd, Clayton VIC 3800 Australia
rob.hyndman@monash.edu

ABSTRACT

Liquidity is crucial for successful financial markets. It ensures that all investors are able to buy and sell assets quickly at a fair price. High Frequency Traders (HFTs) utilize sophisticated algorithms operating with extreme speed and are frequently cited as liquidity providers. The objective of this paper is to investigate the liquidity provision of a number of HFTs to determine their effects on aggregate marketplace liquidity. We consider a large data set collected from the Australian Securities Exchange throughout 2013, providing a near complete picture of all trading activity. Our method is to consider temporal bar charts, association scatterplots, faceted plots and network diagrams to provide visualizations that yield both novel and conventional insights into how HFTs are operating in the market, specifically with respect to liquidity provision. Consistent with HFTs avoiding adverse selection, our results show that on aggregate, HFTs often consume rather than provide liquidity. Furthermore, liquidity consumption often occurs very quickly over intra-millisecond time periods. We conclude that HFTs are not exclusively focused on liquidity provision.

CCS Concepts

• Cross-computing tools and techniques → Empirical studies
• Computing methodologies→Modeling and simulation→ Visual analytics

Keywords

High Frequency Trading; Tick Data; Large Data; Data Visualization

ICSIM 2019, January 10–13, 2019, Bali, Indonesia

ACM ISBN 978-1-4503-6642-7/19/01…$15.00
https://doi.org/10.1145/3305160.3305214

1. INTRODUCTION

This paper describes a visual exploration of the liquidity provision by high frequency traders (HFTs) on financial exchanges. Financial exchanges have morphed from ‘open outcry’ pits, physically populated by traders shouting orders to each other, to their current form comprising warehouses of networked computer servers, silently matching orders generated by different traders’ computer algorithms. These fast electronic financial exchanges have transformed the business of trading dramatically to the point that algorithmic trading now accounts for the majority of market turnover on the world’s equity exchanges [7]. HFTs are traders who utilize sophisticated algorithms operating with extreme speed, and who typically trade in high volume for small profits per trade. By maintaining privileged access to an exchange via Co-located servers in order to reduce the travel time of messages between themselves and the market, HFTs have established themselves as the most nimble of algorithmic traders.

In order to study liquidity provision this study uses a method common in the finance literature, examining the passive and aggressive sides of a trade. The difference in price between the best quoted ask and the best bid is known as the ‘spread’. Should a broker submit an order to buy at the price of the best ask then the arriving buy order is said to have ‘crossed the spread’ and will transact with the resting sell order in the order book. That is, the buyer has demonstrated that they would prefer to pay the seller’s asking price for that share immediately rather than wait for the price to change. In this example, the best ask has been waiting in the order book for some time already, available to any prospective buyers. The buyer has initiated the trade by crossing the spread and accepting the offer price, while the seller made no action to move toward the buyer. We say in this case that the buyer acted ‘aggressively’ and ‘took liquidity’, and that the seller acted ‘passively’ and ‘provided liquidity’. This behavior is symmetric for buyers and sellers (i.e. buyers need not always be aggressive, nor sellers passive). In this way, aggressive orders take liquidity and passive orders make liquidity. The techniques in this study measure liquidity accordingly.

The provision of liquidity has long been thought a ‘public good’ and specialized traders, known as market makers, have been incentivized in various ways to provide liquidity to other traders. Market makers play an important role in financial markets. They provide liquidity and immediacy to all market participants. In this way, all market participants are able to buy or sell their stock holdings at a fair price whenever they need, irrespective of any temporary lack of supply or demand for their stock. To facilitate this immediacy and liquidity provision market makers must hold inventories of stock. They sell (thus decreasing) or buy (thus increasing) their stock inventories to/from market participants and thereby smooth the supply and/or demand of the stock for market participants. In turn, the market makers make a profit by charging a spread, commensurate with the cost and risk associated with providing this liquidity. These spread profits should compensate market makers for the risk of holding their inventory (amongst other costs). This risk is known as 'adverse selection': adverse selection occurs because informed traders will either sell overvalued stock to market makers or buy undervalued stock from market makers, resulting in inventory losses for market makers. Thus, managing a stock inventory, which is necessary for market makers to supply liquidity to market participants, may lead to a market maker making losses when they trade against informed traders.

In the USA, but never in Australia, liquidity was provided by designated market-makers who enjoyed privileged market access in return for fulfilling their market-making responsibilities which consisted of always being ready to make liquidity. HFTs are well-suited to performing a similar role—that of liquidity provision—because they possess the ability to change their offerings to reflect new information about the true price of a stock very quickly. Trading speed is essential for the success of modern market makers: If liquidity providers (known as high frequency market makers) are too slow, the prices they offer will be 'stale' and they may be 'adversely selected' by better informed (and faster) traders (known as high frequency bandits) which means they will likely incur trading losses.

However, while HFTs are not required to be liquidity-makers, they nonetheless share the characteristic of having privileged market access (e.g, Colocation) with designated market-makers. Thus, it is interesting to examine if HFTs are performing a similar function.

An examination of the most recent academic research suggests that HFTs do increase liquidity [4, 8, 9]. However these studies have been undertaken on US markets and are often faced with data quality issues including:

• *Untagged trades*: Agents on each side of a trade are anonymous.

• *Fragmented markets*: There are many exchanges in the US (NYSE, NASDAQ, BATS, etc.) which makes it difficult to see the full picture of activity in a stock.

The methods used in this paper are the result of an iterative process of experimentation with visualization methods. The vast amount of stock market tick data means that visualizations can help to understand whether HFTs provide liquidity.

2. DESCRIPTION OF DATASET

The initial dataset presented here contains all trades on the Australian Securities Exchange (ASX) for 253 days in 2013 across five stocks from a single sector with varying levels of turnover. This dataset consists of 6 million trades and represents nearly A$28 billion of market activity. A second, smaller dataset contains six highly liquid stocks from the Australian banking sector, a crucially important sector to Australian markets representing over 27% of value on the ASX in 2013. This dataset has been extracted from important news announcement days in this sector, where liquidity provision is expected to be particularly relevant due to increased activity. Contained in this dataset are 1 million trades (A$7.5 billion of activity) with Broker IDs and the (anonymous) orders that transpired each day on the lit (normal) market. The orders in the second dataset allow one to determine the best quotes at any time and consequently when price changes occur. For privacy, traders are de-identified in this paper.

In order to process the Australian SIRCA data-set we have built a 5TB database. To create and query this database we have utilized large IT infrastructure comprising a 64 core processor machine with 1TB RAM. We have developed the expertise required to divide any database query into parallel queries and run each parallel query on one of the separate processors provided, resulting in considerable speedup when compared to implementing on a single processor and hence making the research computationally feasible.

These datasets are well-suited to investigate the liquidity provision of HFTs, for several reasons.

1. Australian markets are not heavily-fragmented and thus the vast majority of all trading activity is contained in a single marketplace, from which our data is extracted.
2. Many traded companies in the dataset are either traded exclusively in Australia or only traded in Australia in the hours that ASX is open for trading.

3. All trades are appended with two coded identifiers: the Broker ID for the buyer and the Broker ID for the seller. When examining multiple trades, the Broker ID can verify whether a trader who was buying stock in the morning is selling stock that afternoon, for example.

The first two attributes of the data provide assurances that there is little to no trading activity occurring outside of the dataset. The third attribute allows analysis of a broker's trading when it spans multiple orders and hence multiple hours, days, or months.

3. METHODS

The liquidity provided by HFTs is examined using a selection of visualization methods chosen to address limitations of working with large, dynamic, time-sensitive, partially anonymous data. We have restricted our study to trades and do not consider the effect of orders on liquidity.

We construct three metrics that can be used to help visually measure the efficacy of a broker's performance as a liquidity-maker. We determined the broker identification tags of eleven HFTs, nine of which appear in this dataset from 2013. We profile eight of these HFTs in detail, discarding the ninth due to very infrequent appearance. The HFTs have been labelled A through H in ascending order of contribution to market share across the examined five stocks, which themselves have been ordered 1–5 in decreasing order of turnover.

3.1 Metric 1: Measuring Liquidity

As described earlier in the example using the hypothetical XYZ stock, for each trade we can determine the broker on the aggressive side, the passive side, and the dollar volume of that trade (number of shares × price per share). By the market's close each day, an agent's total activity can be divided into groups of liquidity-making (passive) or liquidity-taking (aggressive), and then compared against other agents and the total amount of liquidity provided and consumed that day. Market share of liquidity provision action, $MS.Liq_j$, is calculated for a broker j, for each date and stock according to the following:

$$\overrightarrow{Actions_j} = \{\text{passive sells; passive buys; aggressive sells; aggressive buys}\} = \left\{ \begin{array}{l} \left(\sum_{i\in B;i\neq j}\sum_{k=1}^{M}\left(\mathbb{I}_{BI,k}P_{j\to i,k}\times V_{j\to i,k}\right)\right); \\ \\ \left(\sum_{i\in B;i\neq j}\sum_{k=1}^{L}\left(\mathbb{I}_{SI,k}P_{i\to j,k}\times V_{i\to j,k}\right)\right); \\ \left(\sum_{i\in B;i\neq j}\sum_{k=1}^{L}\left(\mathbb{I}_{BI,k}P_{i\to j,k}\times V_{i\to j,k}\right)\right); \\ \left(\sum_{i\in B;i\neq j}\sum_{k=1}^{M}\left(\mathbb{I}_{SI,k}P_{j\to i,k}\times V_{j\to i,k}\right)\right) \end{array} \right\} \quad (1)$$

$$\overrightarrow{MS.Liq_j} = \frac{\overrightarrow{Actions_j}}{\sum_{i\in B;j\in B;i\neq j;k<\infty}\left(\mathbb{I}_{BI,k}\left[P_{j\to i,k}\times V_{j\to i,k}\right]+\mathbb{I}_{SI,k}\left[P_{i\to j,k}\times V_{i\to j,k}\right]\right)} \quad (2)$$

Where B is the set of Brokers who traded that day, M is the number of trades in which broker j sold to i, L is the number of trades in which broker i sold to j, BI indicates Buyer-Initated and SI indicates Seller-Initiated, $P_{x\to y,k}$ is the price per share of the k[th] trade where broker x sold stock to broker y, $V_{x\to y,k}$ is the volume of shares traded in the k[th] trade from x to y, and $\mathbb{I}_{SI,k}$ is the indicator function that returns 1 when the k[th] trade is initiated by the seller and zero otherwise (and similarly for buyer initiated trades). $Actions_j$ is a vector containing broker j's aggressive and passive, buy and sell actions and $MS.Liq_j$ is broker j's actions measured as market share of the underlying stock.

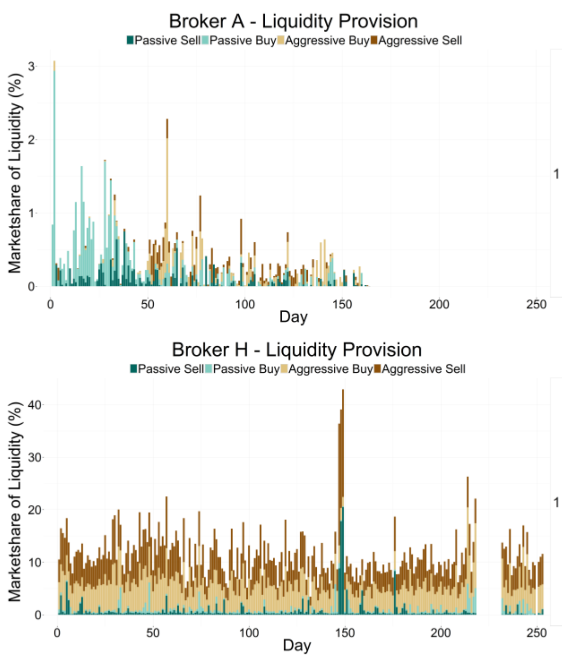

Figure 1. Market share of liquidity consumed/provided against trading day by broker A (top) and H (bottom) for one stock. Color indicates type of action: passive behavior in green shades and aggressive behavior in brown shades. Broker A is purely a liquidity-provider initially (more green) before changing behaviors and finally exiting the market. Broker H consistently consumes significant amounts of liquidity, on the order of 10% of all market liquidity daily

This is computed only on trades that occur on the market during normal operating hours and only for trades where a broker is not on both sides simultaneously (which dilutes the results of other brokers while meaninglessly increasing the former's). Note that the sum over all agents of Market Share of Liquidity adds up to 200%. This is because 100% of liquidity is made and 100% is taken.

3.2 Metric 2: Social Network Analysis

Liquidity-makers facilitate flow of stock volume and stock value between other brokers. Thus, one measure of the efficacy of a liquidity-maker is the dollar volume (*price* × *volume*) of stock that passes through a broker en route to other brokers. This can be rendered visually in a network diagram with the more central positions indicating a liquidity-making broker. In order to compute the association measures on which the network is computed we calculate the sum of dollar volume traded between broker i and broker j. Let A be an $N \times N$ association matrix where $A_{j,j} := 0$ and

$$A_{i,j} := \sum_{k=1}^{L}\left(P_{i\to j,k}\times V_{i\to j,k}\right) + \sum_{k=1}^{M}\left(P_{j\to i,k}\times V_{i\to j,k}\right)$$

for all $i \neq j$; N being the total number of brokers, i and j correspond to the i[th] and j[th] brokers, $P_{x\to y,k}$ is the price per share of the k[th] trade where broker x sold stock to broker y, $V_{x\to y,k}$ is the volume of shares traded in the kth trade from x to y, L is the number of trades in which broker j bought from broker i, and M is the number of trades in which broker j sold to broker i.

3.3 Metric 3: Broker speed relative to informative events

As HFTs are dependent on speed, examining what constitutes an HFT's stimuli (i.e., what events cause a trader to take action) and how quickly traders respond to them can help us understand their effect on marketplace liquidity.

Within this section we consider two types of events to which we expect brokers to react:

- price changes
- other trades

To calculate price changes, we require orders in addition to trades. To this end, we employ a smaller dataset that contains orders. For example, a buy order that arrives with a price higher than the previous best bid (and still lower than the best ask) has shifted the price upward slightly. All orders in the order book must be tracked simultaneously to determine their time of entry and exit and at what prices they were posted at different times. This constructs a picture of the hitherto untransacted orders in the market at any time, known as market depth (i.e., How many bid and ask orders are present and at what prices and volumes).

Having locating instances where the best bid or the best ask has changed, we then associate each trade with its most recent prior price change and its earliest following price change. The trade is then appended with the time interval since the last price change. A trade is also appended with the time interval before the next price change. In the case that a trade causes a price change, this number will be zero. The (x, y) position of a trade in the scatterplot is determined as follows in the price-change measure:

$$\left(x_i, y_j\right) := \left(t_i - \max(p_k; p_k < t_i), \min(p_k, p_k \geq t_i) - t_i\right) \quad (3)$$

where x_i and y_i refer to the i^{th} trade's position in the scatterplot, t_i is the time in milliseconds of the i^{th} trade, and p_k is the time of the k^{th} price change of that day, on-market, during normal trading hours.

Trade events are calculated differently using the interval between trades on the lit (normal) market. Not every trade will cause a price shift, some are too small. If there are 605 shares to be sold at \$70, buying 600 will not change the price. Nonetheless, that trade is a significant event and a broker may make a decision about their next action on seeing it. A trade such as this will change the market depth (specifically, the number of shares offered at the best bid and ask), and hence will impact on what the market looks like to all participants. Thus, traders may act on changes in market depth as well as changes in price.

Similar to price changes, every trade is preceded and followed by another trade by some interval. We assume, in both measures in this section, that the timing of each trade is deliberate. It is only the initiating agent that causes a trade to transact at the precise time that it does and hence the timing should only be relevant for the initiator. Furthermore, the agent initiating a trade is reasonably expected to have acted intentionally, whereas the same cannot be said regarding the passive agent. The passive side may have traded unwillingly because of a resting limit order that was not updated quickly enough to reflect new information before the initiating order was transacted.

Using intervals between trades amounts to a difference equation with additional categorical information for the broker identity. The following interval is plotted against the previous interval to reveal flurries and saunters of the succession of events, as well as broker reaction times and contexts in which initiating agents take liquidity. The next/last trade interval measure can be calculated according to the equation:

$$(x_i, y_j) \coloneqq (t_i - t_{i-1}, t_{i+1} - t_i) \quad \forall\, i \notin \{1, I\} \tag{4}$$

where x_i and y_i refer to the i^{th} trade's position in the scatterplot, t_i is the time in milliseconds of the i^{th} trade and I is the number of on-market trades during normal trading hours.

In this metric, a strong presence does not imply that an agent is only taking liquidity, rather it elucidates the contexts in which the agent takes liquidity. An ideal liquidity-maker is expected to have some presence in this metric despite the fact that it only examines aggressive actions. This is because a liquidity-maker may acquire more inventory than is preferred through long stretches of passive trading and then must act to get rid of it. They may be expected to 'cross the spread' to relinquish (or acquire) enough stock to adjust their inventory and return to a comfortable position.

4. RESULTS AND DISCUSSION

4.1 Market Share of Liquidity-Making and Liquidity-Taking

This section shows how metric 1 is utilized to explore liquidity-making. Trade-events are aggregated, per broker, into totals per day of the categories: passive sells, passive buys, aggressive buys, and aggressive sells. Restricting attention to trades only, as done in this paper, ignores liquidity-making that does not result in a trade. A broker may make large contributions to liquidity by posting many competitive orders (making it easier for others to trade), however if these orders do not transact successfully, they will not be seen in this metric.

Figure 1 (top) shows a stacked bargraph of the trading activity of the smallest HFT in our study, an HFT we will call A, in the most liquid of the stocks in our dataset. Broker A shows similar liquidity making trends in other stocks, but is only sporadically active in those stocks. Thus, we focus on stock 1. The dataset contains, for each trade, a qualifier describing the type of trade that occurred. In the context of this metric, we examine only those trades where the qualifier contains information about which agent (buyer or seller) initiated the trade—this omits less than 5% of trading activity for which the qualifier does not indicate an initiating party. Stacked vertical bars represent the day's proportion of liquidity-making (passive) and liquidity-taking (aggressive) activity attributable to that broker. Each of the four categories is coded to a color, the color scheme has been chosen primarily to represent liquidity provisions (brown and green) and secondarily for buying and selling (light and dark, respectively). Thus liquidity-making behavior can be identified easily as a majority in green or brown for the inverse.

Broker A is predominantly on the liquidity-making side of all trades (as seen by the prominence of green in Figure 1) (top) for several weeks, providing about 0.3% of all liquidity in this stock for the first 50 days. From day 50 onwards, a consistent portion of broker A's activity is liquidity-taking and by day 170, after alternating between liquidity-making and taking behaviors, broker A ceases trading in stock 1 and others. If one interprets that the first 50 days are representative of this trader's normal behavior, then the remainder displays broker A employing a new strategy before ultimately deciding to exit the market altogether. Broker A is the only pure liquidity-maker of all traders isolated in the data.

By contrast, broker H (Figure 1, bottom plot) is almost exclusively taking liquidity in the same stock. This HFT has a much higher turnover than broker A and consumes on average 10.1% of all liquidity (while providing 1.4%) in stock 1. This broker has a sporadic presence in the remaining four stocks, but similar liquidity-taking behavior in five (of six) other stocks from the banking dataset which contains only high-turnover, highly-liquid stock (similar to stock one in this dataset). A commonality throughout is the tendency for this trader to buy and sell in equal amounts as can be recognized by the near equal proportions of light and dark colors. This demonstrates that this broker is averse to the risk of holding inventory overnight. A similar property can be seen in many of the HFTs examined, which aligns well with characteristics for identifying HFTs provided by ASIC (Australian Securities and Investments Commission) [2] and the SEC (Securities Exchange Commission) in the US [1].

Broker E (Figure 2) displays similar behavior in all stocks of our larger dataset, as well as the six bank stocks from the smaller dataset. Broker E does not provide liquidity before day 165. After day 182 however, this broker returns after 2 weeks' absence with a new liquidity-making strategy that appears to ramp up in market share over the last few months. Toward the end, broker E is generating a net decrease in liquidity of 1.6%, 0.1%, and 0.5% in stocks 1, 2 and 3 respectively, but a net increase of 0.7% and 0.3% in stocks 4 and 5. Stocks are arranged in order of decreasing overall turnover, thus broker E can be said to be providing a slight net increase in liquidity to the low-turnover, illiquid stocks of our dataset, where it is more needed.

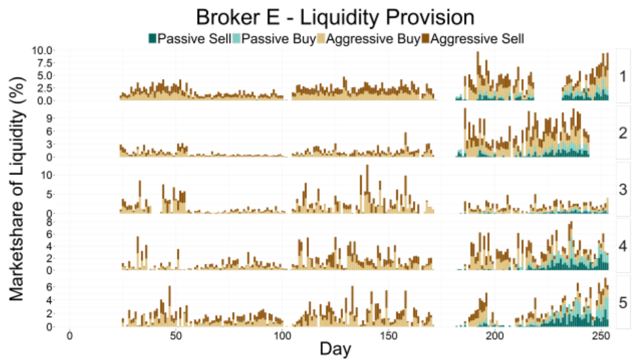

Figure 2. Market share of liquidity consumed/provided by broker E. Broker E appears to begin a new liquidity-making behavior from day 182 onwards, in some cases generating more liquidity overall than broker A had done.

Ultimately it can be argued that brokers A, C, D, E and F (shown in Figure 3) are making liquidity, though only C and F do so consistently. While F generates a net increase in liquidity of 2.4%, 0.3%, and 1.1% in stocks 3, 4, and 5 respectively, it is apparent that the the HFTs with heavy liquidity-taking strategies are removing more liquidity overall than is provided by this broker. Brokers G and H each commonly generate a net decrease in liquidity of between 1.5–8.7% (excluding H's liquidity-neutral behavior in stocks 2 and 4).

Combining the marginal effect of the smaller brokers and the large reduction in liquidity from larger brokers as well as broker F, Figure 4 shows that overall the HFTs that we have identified generate a net decrease in liquidity across most stocks, while stock 3 appears to have no major change.

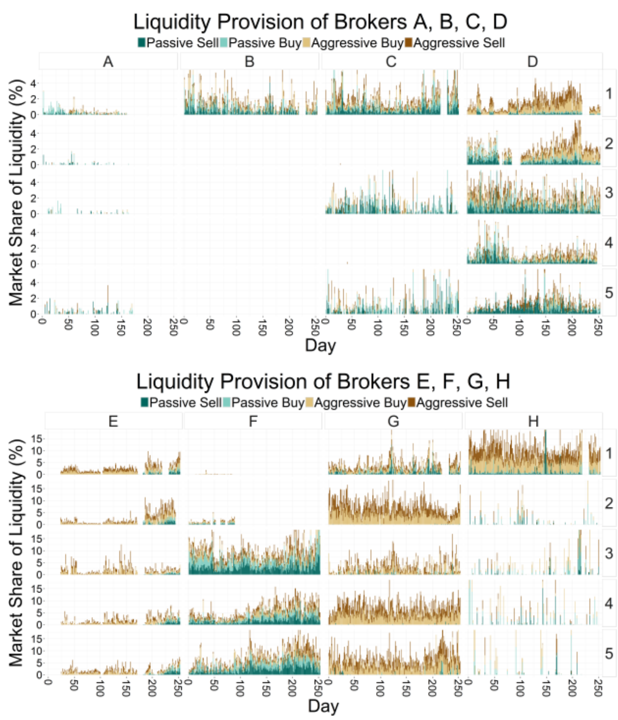

Figure 3. Market share of liquidity consumed/provided by the smallest HFTs (A, B, C, D) by turnover (top) and largest brokers (D, E, F, G) (bottom) over five stocks in order of decreasing turnover. Values above 5.5% (top) and 18% (bottom) market share are cropped. Besides broker F (and partially broker E), these larger HFTs consume much more liquidity than they provide. Broker F generates a net increase in liquidity in stocks 3, 4, and 5.

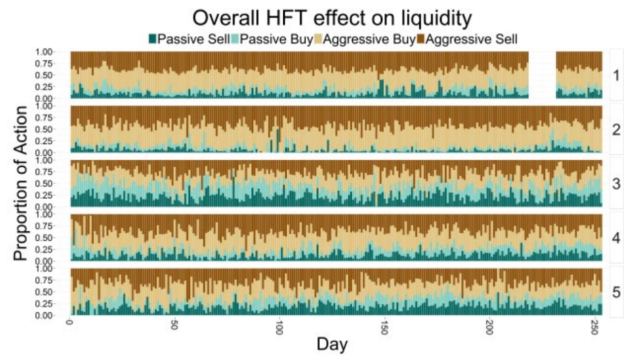

Figure 4. Market share of liquidity consumed/provided by HFTs overall. Examined as a group, HFTs consume significantly more liquidity in stocks 1, 2, 4, and 5.

4.2 Network Visualization of Stock Flow

Figure 5 hows an undirected, weighted network diagram of dollar volume flow as traded on the ASX during normal trading hours on day 125, in the most-liquid stock from our dataset. A broker's position in the graph is determined by the Fruchterman-Reingold algorithm [6] which repels all nodes from each other equally and then introduces an attractive force between nodes according to a given association measure described in Section 3. The attractive force between two brokers is proportional to the dollar volume turnover traded between them. Thus, if an agent W who trades $7 with agent L and $78 with an agent Z, the Fruchterman-Reingold algorithm would place W and Z closer to each other than it would W to L.

The diameter of a node is proportional to the total turnover of the trader that it represents and within each node is embedded a pie chart representing proportions of liquidity making (green) and taking (brown) as in Section 4.1. Edge opacities and widths are proportional to the dollar volume size of the shared turnover that it represents. Though trades (edges) from non-HFTs to other non-HFTs are hidden, they still inform the structure of the graph. HFTs have labels attached to the top-right of their nodes, while other brokers are not labelled at all.

Of the layout algorithms investigated, none had reliable repeatability and thus interpreting finer details from such a graph is not recommended. Furthermore, such a network diagram would only scale well for periods of time where there are no major changes in broker behavior. The results from Section 4.1 demonstrate this not to be the case. Nonetheless, Figure 5 can be used to interpret a broker's liquidity provision and turnover relative to its trading partners and the overall market simultaneously. Ultimately, Figure 5 gives a much more nuanced picture of the flow liquidity over a subset of the data.

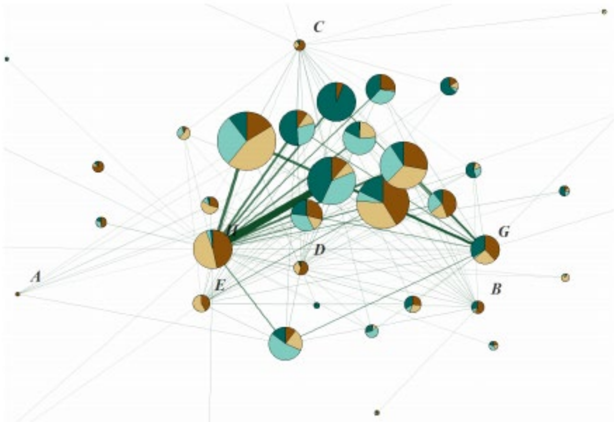

Figure 5. A cropped Network diagram of dollar volume flow in stock 1 on day 125. Trades (edges) that do not involve HFTs rendered invisibly. HFTs (with labels A–H; F not present) lie mostly on the periphery of the center, nonetheless brokers D, G and H occupy the most central position of HFTs. Few traders shown here have making/taking strategies as onesided as C, D, E, and H except for perhaps the mostly dark green nonHFT broker near the top of the central region. There are a few brokers which appear to perform well as central liquidity-makers in this example, however none are HFTs.

Many HFTs can be seen to occupy positions that are close to the periphery of the central region (the extreme boundaries of the graph have been cropped to improve visibility). One would expect ideal liquidity-makers, inasmuch as liquidity-makers connect unconnected brokers, to be close to the center of such a graph. In practice, a broker is more likely to occupy the central region by having traded a large amount, aggressively or passively, with many brokers.

Examining broker H, the largest HFT in this stock on this day, reveals those traders that are providing liquidity to them, typically medium-large non-HFT brokers. There are a few medium-sized non-HFT brokers that are providing liquidity on that day near the center of the network diagram. The broker closest to the absolute center is a promising candidate for a liquidity provider due to the mostly one-sided, passive proportion and as they have not acquired (or relinquished) significant inventory overall (near equal dark and light proportions). Interestingly, high proportions of liquidity-taking is relatively rare outside of those exhibited by Broker C, D, E, and H on this day.

As a further point of interest, it appears somewhat common for non-HFTs to exhibit strong liquidity-making behaviors (green). Five medium-large brokers display a majority of three-quarters or greater in the passive direction, whereas two medium-large brokers display the same tendency in the aggressive direction, one of which is broker H. Investigating different measures of centrality, including a turnover discounted centrality, did not yield useful results.

4.3 Speed and Liquidity Consumption

This section examines the timing of a broker's trades when they cross the spread relative to the time of an event that contains information. This section provides an indication of how HFTs behave in a marketplace at small time-scales and to determine whether their trading activity has other downstream effects. Informative events in this section are defined either as the next and previous trade (Figure 6, left) or as price changes (Figure 6, right). For both figures, points close to the y-axis occur very soon *after* the last event and points close to the x-axis occur very soon *before* the next event. Similarly, points that are not close to the axes occur a long time before (or after) the next (or previous) event.

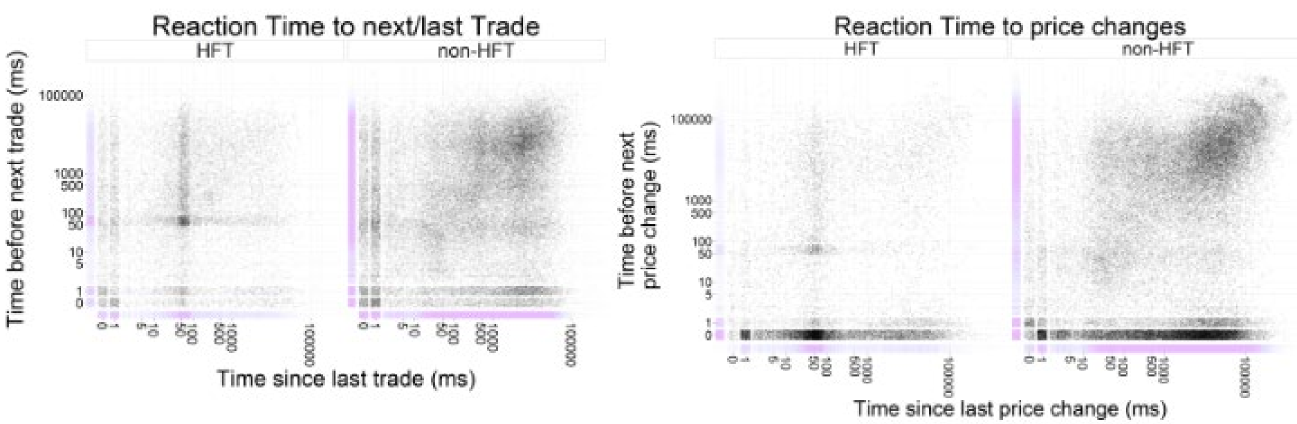

Figure 6. Timing of initiated trades with respect to next/last trades (left) and price changes (right). Overall, HFTs rarely act after more than 100ms of either type of event (both figures). With respect to price changes (right), HFTs appear to cause price changes relatively more frequently per trade than non-HFTs as shown by the higher tendency for an HFT's trades to occur zero milliseconds before a price change.

Figure 6 shows a jittered, alpha-blended scatterplot (and jittered rug plot) of all initiating broker's reaction times to trade events coerced into a log-log scale (with the value 0 mapped to 0.5), where each point represents a trade. For each trade we identify the initiating agent and then facet by whether that agent is one of the HFTs being studied in this paper. Recognizing the speed of HFTs, it is no surprise that HFTs do not have a strong presence past the 100ms mark on the x-axis, time since the previous informative event. We draw close attention to a comparison between the horizontal rug plots of each facet as an indicator of the significant speed discrepancy in the marketplace.

The dataset used to produce the visualizations in Sections 4.1 and 4.2 contains trades exclusively. Thus it is opaque to scrutiny about the timing of price changes that are caused by orders. For this reason, we focus on a second dataset which does contain orders here. This dataset examines interest rate announcement days in six liquid bank stocks. It is expected that liquidity is particularly relevant on these days given the importance of these announcements.

To ensure that the behaviors observed with regard to liquidity provision from section 4.1 can be analysed together with the price-change timing behaviors investigated in this section, we investigated the relationship between a trade's timing relative to next and last trades (which can be extracted from the both datasets) and a trade's timing relative to price changes (second dataset only) empirically. This is done in order to confirm whether a broker's stimuli (the events that a broker waits for before they act) is independent of a stock and their strategy in it. We examine the liquidity provision of the smaller dataset using the technique in 4.1 and find near-identical behaviors to those seen above for each broker in stock 1. Additionally, timing of a broker's trades do not appear to vary across stocks and thus we use these points as justification for a bridge between the datasets to compare liquidity provision with the timing of a broker's trades. As can be seen in Figure 7, patterns resulting from using the timing relative to other trades do not vary significantly between stocks or dates, only between brokers and level of activity in the underlying stock. In the case of broker H, the pattern in stocks 2–5 is not similar to the pattern in stock 1 due to H's sporadic activity in those stocks. Nonetheless the patterns appear to share from a subset of H's 'slower' behaviors in stock 1. Broker F does not appear strongly in Figure 8, due largely to the similarity between the bank dataset and stock 1 in which broker F is scarcely present.

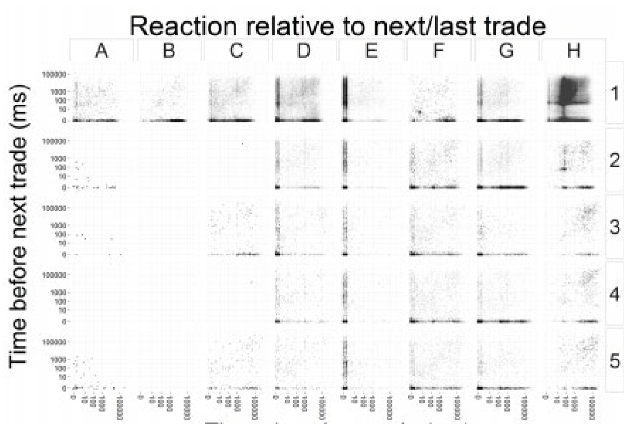

Figure 7. Timing of an initiating broker's trades relative to each trade's preceding and following informative event (other trades in this case) as computed on the original dataset, faceted by HFT (A–H) and by stock (1–5). Patterns do not differ dramatically in different stock (vertically) for the same broker except for broker H for whom the pattern in facets 2–5 is not clear due to inactivity. These similar patterns are present across different stock despite the contrasting behavior observed by brokers D, F, and G in between the stocks in Figure 3.

Thus, since a broker's stimuli appear to be obdurate to changes in stocks and liquidity provision strategies, comparing broker behaviors across data sets should not result in a distorted picture of timing behaviors relative to liquidity provision.

The central region in the HFT facets of Figure 6, left and right) is populated primarily due to two high frequency traders, D and H, whose trade count obscures that of the smaller HFTs. Their appearance on this graph away from the axes suggests that these traders may be implementing slower strategies in addition to fast strategies.

It should be mentioned that not all fast traders are considered HFTs explicitly, thus there are fast traders in the non-HFT facet as well; however, these brokers do not meet the criteria for proprietary high frequency traders [2]. The speed discrepancy between HFTs and remaining traders is still apparent, nonetheless.

Broker E displays a very clear pattern (shared by D and G) of trading predominantly in the same millisecond as, or one millisecond after, another event (Figure 8). This implies that broker E's entire behavior can be well-documented by this visualization technique. It does not, however, necessarily imply that broker E is reacting to informative events in less than a millisecond.

To help explain this concept, consider three children racing each other to a jar of cookies. Even if all arrive exactly half a minute after they began, technically one has arrived marginally faster, the first runner-up arrived zero seconds after the winner, and the last child has arrived zero seconds after the first runner-up. In this situation we can understand how a flurry of intra-millisecond activity may show up in these scatterplots and why an agent, such as E, might have strong patterns along the axes. While this appears to be the case, the outcome would be the same even when traders do have sub-millisecond turnaround times on reacting to new events.

Interestingly, broker H does not have these vertical bands (Figure 8) despite being fairly similar to D, E, and G in terms of liquidity provision. Both broker D and H have horizontal bands near the x-axis, which implies that these brokers anticipate or cause other trades to occur. Alternatively, considering the above example with children racing, they are frequently the first trader in a flurry of trades, or perhaps follow their own trades with more trades. Further research shows that the traders D, E, G, H, appear to be involved in flurries of trades with each other quite frequently with broker H often being first. This would suggest that the HFTs D, E, G, H may have similar stimuli and hence, similar strategies in part. This is supported by comparing their behaviors in Figure 3 where they trade aggressively in stocks 1 and 2, (and similarly in the bank dataset).

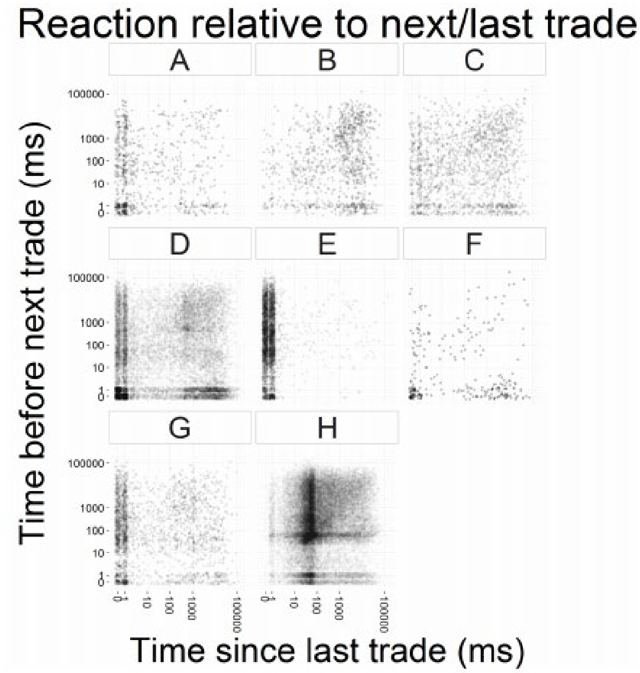

Figure 8. Reaction times for brokers A–H relative to next/last trade. Brokers D, E, and G all render with heavy vertical bands by the y-axis indicating that they often cross the spread 0ms or 1ms after another trade, but are themselves followed somewhat uniformly within 0–100ms (but with an additional strong preference for 0ms). D and H display horizontal bands at 0ms and 1ms suggesting that they anticipate trades, i.e. they are often the first trade in a flurry of trades. This suggests that these HFTs are likely to share similar stimuli and participate in a flurry of trades when these stimuli occur. Broker D and H also demonstrate a pattern away from the axes which shows that some of the features of Figure 6 are due to these brokers and are not necessarily common amongst HFTs.

For reference, individual non-HFT brokers when examined in this metric often have similar dense bands about the axes but show a consistently-strong pattern in the center in addition. This central pattern for non-HFTs is usually more dense than for HFTs and further from the axes as well.

Broker H displays an additional interesting pattern in the center of its facet of two intersecting lines, both around the 80ms mark. If not for this, broker H would appear similarly to most other HFTs. We have found that broker H's trades in the horizontal 80ms band are often being followed by the trades in the vertical 80ms band. This may suggest some kind of 'echoing' or periodic strategy. It is possible that some complicated behaviors that would render similarly to the pattern identified in the center of broker H's facet could be considered liquidity-making strategies, however, we do not investigate this further in this paper.

Finally, Figure 9 can be parsed quite simply. This plot uses the timing of price changes as informative events. All HFTs can be observed to have strong bands along the x-axis signifying trades that cause a price change. This is not entirely unexpected, as a price change will be caused more often by a trade than an order being deleted or an incoming order between the previous best bid and ask. Comparing the ratio of trades on the x-axis to trades elsewhere against the same ratio for non-HFTs in Figure 6 seems to suggest that a trade initiated by an HFT will more frequently cause a price change than may be expected from a random broker. Concretely, calculating the proportion of a broker's aggressive trades that cause price changes vs those broker's aggressive trades that don't affect the price reveals that the average value of this proportion over HFTs lies at 61% and for non-HFTs at 31%. Notable exceptions away from the average are brokers E, G, D, and F at 89%, 83%, 39%, and 37% respectively.

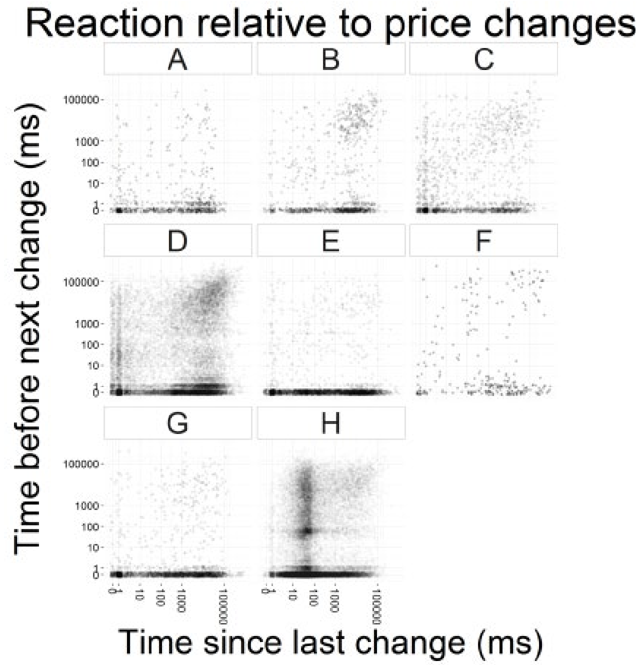

Figure 9. Reaction times for brokers A-H relative to price changes. The dense bands by the x-axis is almost universal here. This is not surprising as trades frequently cause price changes. However, the proportion of trades that cause price changes vs those that do not appears to be quite high for some brokers. Broker H has the same central pattern in all plots in this section. It appears to have an 'echoing' behavior where it trades about 80ms after an event, where this event is likely to have been caused by H.

This finding does not align HFTs well with the expectations of a liquidity-maker. Brokers that cause price changes can be considered 'informed traders' who know more about the true or future value of the stock than the rest of the market. This sketch of an informed trader is somewhat exclusive of the picture of a liquidity-maker that has been described so far. A liquidity-maker will trade aggressively when their inventory grows too large or too small, but one would not necessarily expect them to cause a disproportionately high number of price changes relative to the number of overall trades. Thus while a liquidity-maker would trade aggressively some of the time, they are not strongly associated with changing the price of a stock.

The clearest patterns present in both categories (Figures 8 and 9) of informative event simultaneously are E and G. This pattern suggests that for these brokers, almost all trades follow as a reaction to some prior trade less than 2 milliseconds earlier and these trades (moreso than is already common) cause price changes.

These HFTs have at least one instance of trading in a stock with a large majority of liquidity-taking. If one interprets this degree of liquidity-taking as suggestive of a radically different strategy to any strategy that would exhibit a weak majority of passive or aggressive trading, then it would seem that there is a liquidity-taking strategy employed by no insignificant subset of HFTs that not only takes the kind of liquidity discussed in this paper, but causes an above average number (relative to an average broker of the same size) of price fluctuations, seemingly increasing volatility. The effect that HFTs have on volatility is already disputed [3, 7], however these results seem to suggest that HFTs increase volatility in the stocks and dates examined here.

5. CONCLUSION

The liquidity provisions of HFTs in this dataset vary between HFT companies, stocks, and days. HFTs have been observed to vary dramatically in size (turnover) with the smallest broker usually close to 0.5% of total market share, and the largest hovering around 9% market share. Generally the smaller four of the eight HFTs examined appear to make liquidity marginally more often than they consume it. As for the larger four, these traders take liquidity significantly more often in the stocks and dates represented in the data. There are two clear exceptions to this, the third-largest broker makes significant liquidity contributions in one stock, generates a small net increase in two others, and exhibits little activity in the two most highly liquid stocks. Thus this broker appears to provide liquidity to low-liquidity stock. It has been observed that a liquidity-taking HFT changed behaviors and began making liquidity across the board toward the end of 2013. Thus, observations about this broker may not remain relevant in following years should its behavior continue to change.

Furthermore, having confirmed the speed of high frequency traders it appears that those HFTs that predominantly take liquidity, when they do so, often follow or precede each other in intra-millisecond trading flurries and cause price changes with a higher than average frequency. This characteristic further suggests these representative traders—contrary to the position generally cited in the literature—do not possess the characteristics of a liquidity-maker.

As a single group, high frequency traders consume significantly more liquidity than they provide in four of five stocks presented here. HFTs have been shown to have varied behaviors, however, and it is the pure-liquidity-consumption by some of the large brokers that outweighs the less pronounced liquidity provision by the rest. Thus, the claim that high frequency traders are liquidity-makers appears unlikely to be true given the data used in this study. Though some perform this role, the majority of HFTs do not appear to be concerned with liquidity provision as evidenced by the observation that they rarely bring about more than a slight increase in liquidity, and in many cases extreme net decreases.

We acknowledge that our study has a number of limitations, including: limited stocks and time periods, however despite these limitations we feel that we draw valid conclusions.

6. ACKNOWLEDGMENTS

The research was funded by the CSIRO-Monash Superannuation Research Cluster, a collaboration across institutions and industry, among stakeholders of the retirement system in the interest of better outcomes for all. We would like to thank Rohan Fletcher for outstanding research assistant work.

All calculations were done in R [10]. Plots were made with the ggplot2 package [11]. Network diagrams were created with the igraph package [5].

7. REFERENCES

[1] *Equity market structure literature review, part 2: High frequency trading.* U.S. Securities and Exchange Commission, March 2014.

[2] *Review of high-frequency trading and dark liquidity*. Australian Securities and Investments Commission, October 2015.

[3] E. Boehmer, K. Y. L. Fong, and J. J. Wu. *International Evidence on Algorithmic Trading.* Social Science Research Network Working Paper Series, Mar. 2012.

[4] J. Brogaard, T. Hendershott, and R. Riordan. *High frequency trading and price discovery*. European Central Bank, 2013.

[5] G. Csardi and T. Nepusz. *The igraph software package for complex network research*. InterJournal, 1695, 2006.

[6] T. M. J. Fruchterman and E. M. Reingold. *Graph drawing by force-directed placement*. Software Practice and Experience, 21(11):1129– 1164, 1991.

[7] B. Hagstrmer and L. Nordn. *The diversity of high-frequency traders*. Journal of Financial Markets, 16(4):741–770, 2013.

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

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

[10] R Core Team. R: *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016.

[11] H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. SpringerVerlag New York, 2009.