High-Frequency Market Making to Large Institutional Trades*

Robert A. Korajczyk
Northwestern University

Dermot Murphy
University of Illinois at Chicago

First Draft: November 30, 2014
This Draft: July 16, 2018

Forthcoming, *Review of Financial Studies*

Abstract

We study market-making high-frequency trader (HFT) dynamics around large institutional trades in Canadian equities markets using order-level data with masked trader identification. Following a regulatory change that negatively affected HFT order activity, we find that bid-ask spreads increased and price impact decreased for institutional trades. The decrease in price impact is strongest for informed institutional traders. During institutional trade executions, HFTs submit more same-direction orders and increase their inventory mean reversion rates. Our evidence indicates that high-frequency trading is associated with lower transaction costs for small, uninformed trades and higher transaction costs for large, informed trades. (*JEL* G1, G11, G18, G24)

*We thank the Investment Industry Regulatory Organization of Canada (IIROC) for providing us with access to the data used for this study and Victoria Pinnington and Helen Hogarth of IIROC for answering questions regarding the data and Canadian market structure details. We also thank Andrew Karolyi (the editor); two anonymous referees; Torben Andersen, Rob Battalio, Oleg Bondarenko, Jonathan Brogaard, Kevin Crotty, Robert Ferstenberg, Mark Grinblatt, Hans Heidle, Sean Kersey, Amy Kwan, Katya Malinova, Pamela Moulton, Ioanid Roşu, Gideon Saar, Andriy Shkilko, Brian Weller, Avi Wohl, and Mao Ye; and seminar participants at DePaul University, the European Finance Association Meetings, the FMA Asia/Pacific Conference, the IIROC/Capital Markets Institute High Frequency Trading Forum, the Inquire-Europe Conference, the Multinational Finance Society Conference, the Notre Dame Conference on Current Topics in Financial Regulation, Tel Aviv University, Texas A&M University, University of British Columbia, University of Geneva, University of Melbourne, University of Queensland, University of St. Gallen, the Western Finance Association Meetings, and Wilfrid Laurier University for helpful comments. This work was supported by IIROC. Send correspondence to Dermot Murphy, University of Illinois at Chicago, College of Business, Department of Finance, 601 South Morgan Street, Chicago, IL 60607-7121; telephone: (312) 355-4372. E-mail: murphyd@uic.edu.

High-frequency traders (HFTs) have largely assumed the market making role in modern equity markets. Extant research has shown that the presence of HFTs is associated with improvements in market quality, including lower bid-ask spreads and improved price efficiency (e.g., Menkveld 2013; Hagströmer and Norden 2013; Brogaard, Hendershott, and Riordan 2014; Jovanovic and Menkveld 2016). However, it is still unclear whether the presence of market-making HFTs leads to reduced transaction costs for large institutional traders, who frequently buy or sell millions of dollars worth of shares in short periods of time. In this paper, we examine the interaction of market-making HFTs and large institutional trades and how this affects the execution costs of these trades.

In the extensive literature on market making, a standard feature is that equilibrium price quotes are set by the market maker so that uninformed investors, those with noninformational reasons for trading, subsidize informed investors (e.g., Glosten and Milgrom 1985; Kyle 1985). That is, market makers lose to informed traders and make up those losses from uninformed traders. HFTs might provide increased liquidity for all traders if their speed advantage reduces the costs of market making for all trades. Alternatively, the speed advantage conferred by HFTs might allow them to detect informed traders and adjust quotes in a way that allows them to reduce their exposure to adverse selection. In this case, HFT activity will increase the trading costs for informed traders and decrease the subsidy from uninformed traders. This is touched on by Warren Buffett, Chairman and CEO of Berkshire Hathaway, who has stated that the speed advantage gained by HFTs has made the "big orders" more costly while acknowledging that small investors have "never had it so good" (Crippen 2014).

In theory, there are three primary mechanisms through which market-making HFTs may be increasing the execution costs of large institutional trades. The first mechanism is related to HFTs using their speed advantage to reprice standing limit orders in response to signals about incoming orders (Ait-Sahalia and Sağlam 2017a, 2017b). This behavior is sometimes described as "phantom liquidity," where limit orders are modified before slower traders are able to complete their trades. The second mechanism is related to the theoretical "back-running" model in Yang and Zhu (2017), where an informed institutional trader trades over time to reduce liquidity costs and an HFT attempts to infer information from the institutional order flow and trades in competition with the institutional order. The third mechanism is based on classic microstructure models of inventory management (Stoll 1978; Ho and Stoll 1981; Amihud and Mendelson 1980). A large institutional order will place opposite-direction pressure on HFT inventories. If HFTs proceed to revert their inventories, while the institutional trade is still underway, then they effectively will be competing with the institutional order. Although market maker inventory management is not new to the HFT environment, HFTs may be using their speed and information processing advantages to revert their inventories faster through the use of order modifications (Ait-Sahalia and Sağlam 2017a, 2017b) and same-direction trades (Yang and Zhu 2017).

We study the trading dynamics of market-making HFTs around a sample of approximately 1.2 million institutional trades on Canadian equities exchanges. The Investment Industry Regulatory Organization of Canada (IIROC) provides us with access to order-level data for all Canadian equities for the period from January 2012 to June 2013. For each of the approximately 60 billion messages, we are provided with a user ID, allowing us to track the order and trade activity for any of these IDs across time and in the cross-section of equities. In particular for this study, the information allows us to identify directional, institutional-sized orders and track their executions over time and also allows us to identify the market-making HFTs and how they interact with these orders.

We first provide evidence establishing an important link between HFT and the execution costs of large institutional trades. On April 1, 2012, IIROC introduced a new regulation called the "integrated fee model," in which traders would be charged on a pro rata basis for the messages they send to Canadian marketplaces. HFTs were especially affected by this regulatory change because their strategies typically involve very high message traffic, and we find that daily HFT message traffic decreased by about 20% following the change, a finding consistent with the evidence in Malinova, Park, and Riordan (2018). We show that the average price impact—the component of implementation shortfall related to trade size—for large institutional trades following the regulatory change decreased by about 15%, suggesting that the ability of HFTs to frequently and cheaply adjust their limit orders in the presence of large institutional trades was somewhat hindered by this new regulation. We also provide evidence that the average fixed cost for these trades increased by about three basis points following the regulatory change, suggesting that HFTs widened their bid-ask spreads to compensate for the new order submission fees. Taken together, this evidence indicates that execution costs increased for smaller-sized institutional trades and decreased for larger-sized institutional trades, and we find a trade size break-even point of approximately $2 million.

Choi, Larsen, and Seppi (Forthcoming) predict that portfolio-rebalancing institutional traders engage in predictable, "sunshine" trading (Admati and Pfleiderer 1991) to differentiate themselves from informed institutional traders, who pool their trades with non-informational order flow to avoid being detected by market-making HFTs. Motivated by this theory, we separate institutional traders into "informed" groups based on the profitability of their past trades and examine the differential effect of the regulatory change on group-level transaction costs. First, we confirm that traders from the high-informed group profitably trade out-of-sample, indicating that these traders exhibit some degree of skill. Importantly, we find that the regulatory change has the strongest effect on the high-informed group, with a reduction in price impact of about 28%. In contrast, the price impact reductions for the remaining types are low and have weaker statistical significance. Thus, our results suggest that trading costs for informed traders are higher when HFTs can cheaply modify their orders. From a price efficiency standpoint, this may be problematic if HFT deters other traders from acquiring costly information. Indeed, Weller (Forthcoming) shows that algorithmic trading, which encompasses HFT, is associated with decreased price efficiency, suggesting that information acquisition is lower when algorithmic trading activity is higher.

How do HFTs dynamically interact with large institutional trades? Theoretically, HFT inventory dynamics nonlinearly depend on their traders current inventory level (Amihud and Mendelson 1980) and on their information about the fundamental value of the security, which HFTs partially infer from the institutional order flow. Using the empirical strategy in Hansch, Naik, and Viswanathan (1998) (HNV), who examine specialist inventory dynamics on the London Stock Exchange, we first test how HFT inventory dynamics relate to their current inventory levels. For nonextreme inventory positions, HFTs mean-revert 27.6% of that position in the following 15-minute period, while for extreme inventory positions, HFTs mean-revert 37.7% of that position in the following period. These reversion rates correspond to inventory half-lives of 32.2 minutes and 22.0 minutes, respectively, and are consistent with the nonlinearities reported in HNV. We then examine how HFT inventory reversion rates differ in the presence of large institutional trades. If an institutional order places opposite-direction pressure on HFT inventory levels and HFTs also infer information from the order flow, then we would expect inventory reversion to occur at a faster rate. For nonextreme inventory positions, we find that the inventory mean reversion rate is 32.4% when an institutional trade is underway, compared to only 21.6% when one is not. We find similar differences in inventory mean reversion rates for extreme HFT inventory levels. Our evidence implies that one-third of HFT inventory reversion during an institutional trade execution $(1 - 21.6/32.4)$ can be attributed to information that HFTs infer from that trade.

We separately examine how abnormal buying and selling activity by HFTs change in the presence of large institutional trades. If HFTs are more inclined to back-run institutional orders like in Yang and Zhu (2017), then we would expect higher buying (selling) activity in the presence of a large buy (sell) order. Alternatively, if HFTs become less inclined to provide liquidity to institutional orders, then we would expect lower selling (buying) activity in the presence of a large buy (sell) order. We find a significant increase in same-direction abnormal trading activity by HFTs relative to their opposite-direction abnormal trading activity when an institutional trade is being executed. We also find similar evidence for abnormal limit-order submission activity by HFTs. Overall, this evidence indicates that the same-direction inventory changes of HFTs in the presence of large institutional trades are more likely to be driven by an increase in same-direction trades and limit orders than a decrease in opposite-direction trades and limit orders.

Finally, we provide suggestive evidence of the information that HFTs use to effectively compete with large institutional orders. Using a probit regression framework, we first identify predictors of large institutional trades and find that past returns, trade imbalances, and limit-order imbalances are all strong predictors. Using these publicly observable variables, we calculate the predicted component of large institutional trades and find that it has a strong relationship with same-direction inventory changes by HFTs. These inventory changes are even more pronounced in the presence of aggressive institutional trades that consist of a relatively high proportion of liquidity-demanding orders and that are more likely to move prices and be motivated by short-lived private information. Overall, our evidence suggests that HFTs use information from publicly observable variables to infer that an institutional trade is underway.

Several studies are closely related to ours. Malinova, Park, and Riordan (2018) study the effect of the integrated fee model on the execution costs of both retail and institutional traders. They find that the effective spread for retail traders increases significantly by 0.9 basis points. They also find that implementation shortfall significantly increases by 4.9 basis points for institutional parent orders that only use marketable orders. For institutional orders that use both marketable and nonmarketable orders, implementation shortfall increases by a statistically insignificant 1.2 basis points. When we study average implementation shortfall, unconditional on trade size, we obtain results that are in line with those in Malinova, Park, and Riordan (2018). We break our implementation shortfall into a spread component and a price impact component and find the former increases and the latter decreases upon implementation of the integrated fee model. While the average trade has higher implementation shortfall, we find that the largest trades have lower implementation shortfall. Van Kervel and Menkveld (Forthcoming) study the interaction of high-frequency traders on NASDAQ and 6,000 institutional trades in Swedish stocks from four large Swedish institutional investors. They show that HFTs initially trade in the opposite direction of institutional orders, but later trade in the same direction, leading to higher execution costs for these orders. Their HFTs seem to accommodate the institutional trade for a longer time than our HFTs. Using unique data that allow us to identify specific HFTs and track their inventory levels, we show that HFTs compete with large institutional orders through inventory management and order anticipation channels, while also providing additional causal evidence linking HFTs to the execution costs of large institutional trades. Anand and Venkataraman (2016) examine whether stock exchanges should impose market maker obligations. Using a similar transaction-level data set with masked trader identity from the TSX for the year 2006, they find that "endogenous liquidity providers" provide different levels of liquidity based on their trading profits, inventory risks, and capital commitments and based on different market conditions, such as large-price-movement days and high-volatility days. We focus on how market-making HFTs dynamically interact with large institutional trades and what this ultimately means for the costs of these large trades.

Our study adds to a growing literature about the potentially negative effects of HFT on market quality metrics. Hirschey (2018) provides evidence that HFTs trade in anticipation of future order flow by non-HFTs on the NASDAQ stock market. Breckenfelder (2013) finds that when HFTs compete for trades, liquidity deteriorates and short-term volatility rises. Carrion (2013) finds that HFTs tend to take more liquidity when it is plentiful and supply liquidity when it is scarce. Kirilenko et al. (2017) find that the HFTs in their sample engage in "stale quote sniping," by disproportionately trading in the direction of a subsequent price movement and then offering much of the liquidity available at that new price. Our study adds to this literature by using comprehensive, order-level data to show how HFTs dynamically interact with large institutional traders and ultimately affect the costs of institutional trades, many of which are likely to originate from a large cross-section of mutual and pension funds.

1 Related Theory

We study institutional-size directional trades that result from a large trader either building or exiting a position. The optimal execution strategies derived in Bertsimas and Lo (1998) and Collin-Dufresne and Fos (2016) imply that institutional order flow is autocorrelated, which is consistent with the data (Campbell, Ramadorai, and Schwartz 2009). This institutional order flow can potentially be exploited by sophisticated traders who have the technological capabilities to detect that a large institutional trade is in the process of being executed. Ait-Sahalia and Sağlam (2017a, 2017b) model high-speed market makers who receive a signal about the direction and impatience of the next low-frequency trade. This allows HFTs to cancel existing limit orders and replace them with orders at less favorable prices if they anticipate an impatient low-frequency trader. This corresponds to the notion of phantom liquidity: liquidity that disappears before the order from a low-frequency trader can execute against it. Ait-Sahalia and Sağlam (2017a) analyze their model's implications for the effect of a hypothetical fee change, similar to the integrated fee model analyzed in Section 3, on liquidity provision. They find that the fee leads to smaller quoted spreads in low volatility regimes but higher quoted spreads in high volatility regimes.

Yang and Zhu (2017) present a model based on the standard Kyle (1985) setting in which a strategic informed trader can split their order across two periods. In the second period, an HFT obtains a noisy signal about the informed trader's order in the first period; the authors argue that HFTs use sophisticated algorithms that detect order flow "footprints" left by the institutional trader, thus providing justification for this signal. HFTs also observe the stock price from the first period and use this information to infer the aggregate net order flow. In the second period, HFTs use their signal and past order flow to partially infer the fundamental value of the security. Consequently, HFTs will "back-run" the informed trader's order by trading in the same direction as the order in the second period. Bessembinder et al. (2016) show that predictable trading can lead to worse liquidity outcomes if strategic liquidity traders can time their trades ahead of the order with a high enough level of precision, even if markets are highly resilient. To the extent that HFTs represent these strategic liquidity traders, this model also implies that HFTs can back-run institutional orders, leading to higher execution costs for those institutions. In a similar vein, in Section 4, we will analyze the extent to which HFTs compete with institutional trades.

The empirical analysis of HFT inventory dynamics in Section 4 is also motivated by models of inventory management by liquidity suppliers. Market-making HFTs provide liquidity to other market participants by posting bids and offers to the limit-order book. Inventory management is an important component of the market-making operation, because HFTs want to minimize their exposure to adverse price movements when their position deviates from zero. Stoll (1978) and Ho and Stoll (1981) assume a risk-averse market maker and show that the price quoted by the market maker is a function of their inventory level. Amihud and Mendelson (1980) assume a risk-neutral market maker with exogenously specified upper and lower bounds on their inventory levels. These bounds can be thought of as capital constraints on the market maker or, in the context of Stoll (1978), maximum risk tolerance levels. When market makers inventory deviates from their optimal position (for market-making HFTs, this is typically zero), they raise prices to encourage incoming sells or lower prices to encourage incoming buys, thus reverting their position to the optimum. Importantly, the impact of the market maker's inventory on pricing is nonlinear once these capital constraints are taken into account, meaning that the market maker's pricing strategy becomes increasingly aggressive as their inventory nears the upper or lower bound. We use the methodology in Hansch, Naik, and Viswanathan (1998), which is also motivated by these theoretical models of inventory management, to analyze HFT inventory dynamics during the executions of large institutional trades.

2 Data and Classification Methodology

We have been provided access to detailed order-level data by the Investment Industry Regulatory Organization of Canada (IIROC), a Canadian national self-regulatory organization that regulates securities dealers in Canada's equity markets. IIROC carries out its regulatory responsibilities through setting and enforcing rules regarding the proficiency, business, and financial conduct of dealer firms and their registered employees, and through setting and enforcing market integrity rules regarding trading activity on Canadian equity marketplaces.[1] The closest equivalent in the United States would be the Financial Industry Regulatory Authority (FINRA), a self-regulatory, nongovernmental organization that regulates brokerage firms and exchange markets.

Through the monitoring of the Canadian equities markets, IIROC collects detailed records on all orders submitted to Canadian exchanges. IIROC provides us with access to a data set that contains all trades, orders, order cancellations, and order modifications for the period from January 1, 2012 to June 30, 2013. Each record contains an "event" field that allows us to determine whether that observation is an order, trade, order cancellation, or order modification. We are also provided with the security ID and the price, quantity, date, and time associated with each record, where the time is reported at the millisecond level. More importantly, each record contains a masked identification for the trader submitting an order, allowing us to track the activity of any user ID over time, along with the direction (buy or sell) of that order. For trades, we are provided with masked identification for both the buyer and the seller. Finally, for each trade, we are provided with an "active/passive" indicator that identifies the party submitting the marketable limit order, thus making trade-direction inference algorithms, such as the one used in Lee and Ready (1991), unnecessary. Altogether, the data set comprises approximately 60 billion observations.

The IIROC data set contains a high level of detail that allows us to classify traders

[1] This information and additional details can be found at `www.iiroc.ca/about`.

as market-making HFTs and identify large, institutional-sized trades.[2] In the next two subsections, we outline the methodology for identifying market-making HFTs and large institutional trades. The initial sample of stocks used for these identifications is based on IIROC's definition of "highly liquid securities," which are stocks that have an average of at least 100 trades and $10 million in dollar volume per day in Canadian marketplaces. IIROC exempts highly liquid securities from certain restrictions and prohibitions governing trading activity in securities with low liquidity. The low-liquidity securities are mostly traded in dealer-driven markets that have very little HFT activity and thus are not applicable to our study of market-making HFTs and large institutional trades. Of the approximately 4,200 publicly traded equity securities traded during the sample period, 295 stocks meet the criteria for highly liquid securities. This represents about 7% of the publicly traded equity securities. On a dollar volume basis, the highly liquid securities comprise approximately $5 trillion during the sample period, whereas the remaining publicly traded equity securities comprise approximately $700 billion, indicating that the highly liquid securities represent a large majority of the dollar volume on Canadian exchanges (about 88%).[3]

2.1 Classifying market-making HFTs

Following the methodologies in Comerton-Forde, Malinova, and Park (2018) and Malinova, Park, and Riordan (2018), who use a similar data set from IIROC, we classify HFTs based on their operating speeds. For each user ID in our sample, we calculate the median time between submitting an order and cancelling it. Neuroscience research suggests that the median reaction time of humans to external stimuli is approximately 250 milliseconds (Laming 1968); we primarily classify a user ID as an HFT if the median order-to-cancel time is below this 250-millisecond threshold. We also follow those studies by classifying a user ID as an HFT if the corresponding trader submitted at least 1,000 orders within the

[2]Henceforth, we refer to large, institutional-sized trades as "institutional trades," although it is possible that some of these trades are coming from high net worth individuals.

[3]We should also note that although all dollar figures in this paper are reported in Canadian dollars, the exchange rate between U.S. and Canadian dollars was close to parity during our sample period.

first 500 milliseconds after 3:40 p.m. EST; this is when regularly scheduled announcements are made by the TSX about net order imbalances for closing call auctions in TSX-listed securities. These announcements contain valuable information about closing prices in TSX-listed stocks that fast traders can exploit by acting on it before other traders. Thus, traders who consistently submit orders immediately following information releases at 3:40 p.m. EST also can be classified as HFTs. Using these classification schemes, we identify 103 user IDs as HFTs. This number is fairly close to the numbers reported in Comerton-Forde, Malinova, and Park (2018) and an IIROC research report that examines trading groups (Devani et al. 2014).

Several of the stocks in our sample are marginally classified as highly liquid securities, in that either the average number of trades per day is just above the cutoff of 100 or the average dollar volume per day is just above the cutoff of $10 million. For many of these stocks, there is little to no HFT presence, making them unsuitable for our analysis of HFT dynamics around large institutional trades. Using a randomly selected sample of stocks provided by NASDAQ, Brogaard, Hendershott, and Riordan (2014) report that HFTs represent 42% of volume in large stocks and 18% in small stocks. Motivated by this, we exclude any stock from our sample if HFTs represent less than 15% of trading volume in that stock.[4] Post-removal, we are left with a final sample of 181 stocks. This represents approximately 77% of the total dollar volume of the publicly traded equity securities during the sample period.

Next, we identify the HFTs who act as market makers, because the purpose of our study is to examine how market-making HFTs dynamically interact with large institutional orders. The primary function of a market maker is to provide liquidity to the market by posting bids and offers to the limit-order book, profiting from the bid-ask spread and liquidity rebates.[5] Thus, the share volume of buy and sell orders posted to the limit-order book by

[4] We also test the robustness of our main results to alternative cutoffs of 5% and 25% and find qualitatively similar results.

[5] Liquidity rebates are monetary incentives offered by many exchanges for posting liquidity to the exchange's limit-order book. In 2012, for example, TSX provided a rebate of $0.0031 per share for posted limit

a market maker should be fairly balanced. Following Comerton-Forde, Malinova, and Park (2018), we calculate the "market maker index" (MMI) for each HFT (i) as follows:

$$\mathrm{MMI}_{i,j,d} = \left| \frac{\text{Passive Buy Order Volume}_{i,j,d} - \text{Passive Sell Order Volume}_{i,j,d}}{\text{Passive Buy Order Volume}_{i,j,d} + \text{Passive Sell Order Volume}_{i,j,d}} \right|, \tag{1}$$

where i is the HFT, j is the stock, and d is the date. By construction, this index is bound between zero and one. For each HFT i, we calculate the median value of MMI across all stocks and days in which that HFT was present. A median MMI close to zero indicates that the HFT consistently submits similar share quantities of buy and sell orders, suggesting that the HFT is a market maker. Like in Comerton-Forde, Malinova, and Park (2018), we find a structural break in the median MMI at about 0.20, and classify an HFT as a market-making HFT if their median MMI is below this threshold. Using this methodology, we classify 68 HFTs as market-making HFTs.

Using all stock-days in our sample, we calculate summary statistics for our market-making HFTs. These statistics are reported in Table 1. We find that market-making HFTs are responsible for 31.6% of total daily trading volume, on average. For 5% of the stock-days in our sample, market-making HFTs are responsible for at least 53.4% of daily trading volume. These numbers suggest that market-making HFTs are significantly present in our sample of stocks, implying that large institutional traders will be trading with market-making HFTs fairly frequently. Furthermore, we find that market-making HFTs are responsible for 55.4% of total limit-order submission volume, on average, which is almost twice as high as their average percentage of trading volume. This is unsurprising: frequent limit-order submissions, cancellations, and modifications are hallmarks of HFT market-making strategies, because market-making HFTs frequently adjust their limit orders to avoid the adverse-selection risk associated with trading with informed counterparties. The high average order-to-trade ratio of 33.1 in Table 1 further supports our claim that market-making HFTs frequently cancel their limit orders, and implies that market-making HFTs only ex-

orders that executed against incoming marketable limit orders.

ecute about 3% of the orders that they post to the limit-order book. We also find that market-making HFTs execute approximately 27.8% of their trades using marketable limit orders, suggesting that liquidity provision is not the only component of their market-making strategies. Market-making HFTs may be using these orders to quickly revert extreme inventory positions and compete with large institutional trades, which we will analyze later in this paper. Finally, we find that the median order size for a trade involving a market-making HFT is 147 shares, with a median value of $2,685. The median order size of 147 shares suggests that most trades involving a market-making HFT are for one or two round lots of 100 shares.

2.2 Classifying large institutional trades

To minimize execution costs, institutions commonly execute their large orders by executing a series of smaller-sized orders over time. Doing so allows the institution to "hide" in the noise trader order flow, reducing the ability of other traders to detect the information content contained in the institutional order flow. The IIROC database provides us with masked identification for every trader, allowing us to track these orders. Large institutional orders are characterized by their large trade sizes aggregated across one or more transactions. We classify a series of orders as a "large institutional trade" if the same user ID is on the same side of one or more transactions across 1 or more consecutive days and the total dollar volume of these transactions is at least $100,000.[6] The $100,000 cutoff is based on institutional trade dollar size statistics reported in other studies. Chan and Lakonishok (1995), for example, report a median institutional trade dollar size of under $200 thousand, whereas Cready, Kumas, and Subasi (2014) report average institutional trade sizes that range from approximately $100,000 to $500,000, depending on the size of the institutional investor. The $100,000 cutoff used in our study is conservative and allows for a greater cross-section of institutional investors. Using this methodology, we identify 1,173,482 large institutional

[6] We link an institutional trade across days if its execution involves at least one trade in both the last half-hour of day t and the first half-hour of day $t + 1$.

trades in our sample. In our main tests later in the paper, we experiment with different cutoffs to ensure our results are robust to alternative specifications for large institutional trades.

In addition to dollar trade size, we identify other important attributes of each institutional trade that could potentially influence the dynamics of how HFTs interact with that trade. "Aggressiveness" is defined as the percentage of the institutional trade that is executed using marketable limit orders, which are typically used by liquidity-demanding traders who must execute a trade quickly. "Time to Completion" is defined as the number of trading hours it takes to execute the large trade; a lower value is likely to reflect an institutional trade that is based on short-lived private information. Finally, we calculate the implementation shortfall (IS) of each institutional trade. This is based on the "implicit cost of interacting with the market" from Perold (1988) and calculated as follows:

$$IS_{i,t} = \frac{\sum_{n=1}^{N} p_n x_{i,n} - p_0 x_{i,N}}{p_0 x_{i,N}} \times (\mathbf{1}_B - \mathbf{1}_S), \tag{2}$$

where i represents the institutional trade, t represents the date that the trade was initiated, p_n and x_n are the price and volume of trade n within the parent order, p_0 is the bid-ask midpoint at the initiation of the parent order, $x_N \equiv \sum_{n=1}^{N} x_n$ is the total number of shares executed in the institutional parent order, and $\mathbf{1}_B$ ($\mathbf{1}_S$) is an indicator variable that equals one if the institutional trade is a buy (sell). This is a standard approach used in other studies of institutional trading, such as Keim and Madhavan (1997) and Anand et al. (2012). The implementation shortfall variable will be particularly important for the tests in the following section, where we examine the relationship between HFT and institutional trading costs.

Summary statistics for the sample of large institutional trades are reported in Table 2. The average dollar size of a large institutional trade is about \$720,000, with the upper 5% of these trades exceeding \$2.5 million (the upper 1% exceeds \$6.8 million). On average, an institutional trade is executed using 118 smaller trades and 234 limit orders. The discrepancy between the average number of trades and limit orders is the result of unexecuted limit orders that are cancelled by the institution. The average and median order-to-trade ratios for an institutional trade are 4.9 and 1.0, respectively, further reflecting the use of order cancellations by some institutions. About 57% of the shares in an institutional order are bought or sold using marketable limit orders, indicating that the child orders composing an average institutional trade are about evenly split between liquidity-demanding and liquidity-supplying limit orders. It takes an average of 3.0 hours to execute an institutional trade, and 95% of all institutional trades are executed within 6.5 hours, the number of hours in a single trading day. Related, we find that 6.3% of the institutional trades in our sample span 2 or more days. Finally, we find that the average implementation shortfall of a large institutional trade equals 7.1 basis points. The large interquartile range of 31.8 basis points further suggests sizable variation in execution quality.

3 Institutional Trading Costs and HFT

The first major step in this study is to establish a link between institutional trading costs and high-frequency trading activity. Theory suggests that HFTs can use their speed advantage to profitably modify their orders in the presence of a large institutional trade, thereby increasing the execution costs associated with that trade (Ait-Sahalia and Sağlam 2017a, 2017b; Yang and Zhu 2017). An exogenous event that improves or impairs the ability of HFTs to modify their orders in the presence of large institutional trades would be ideal for establishing a link between HFT and institutional execution costs. We utilize a regulatory change that was implemented by IIROC on April 1, 2012, called the "integrated fee model." IIROC recognized that message traffic from Canadian exchanges was steadily increasing over time, increasing the burden on IIROC to monitor the traders on these exchanges. As a result, IIROC implemented a fee model in which traders would be charged on a pro rata basis for both the trades they execute and the messages they send to Canadian marketplaces, as opposed to the prior fee model in which traders were only charged for their trades. In developing the new fee model, IIROC stated that traders with a "greater share of messages or trades compared to their share of shares traded will incur higher fees under the proposed model compared to the current model" (IIROC Notice 10-0316). High-frequency traders were strongly affected by the new message fees because their strategies typically involve high message activity and high order-to-trade ratios, as noted in Table 1. In contrast, the new message fees for our sample of institutional traders were largely trivial due to their low message activity relative to their trades, and we do not observe any significant changes in the arrival rates of institutional orders after the regulation. Pro-HFT commenters noted that the proposed model would extend "an apparent bias against HFTs," further noting that "taxing message traffic will disproportionately hurt HFTs." In their responses, IIROC stated that they "developed the proposed fee model to be as neutral as possible between liquidity providers and liquidity takers."[7] The goal of this section is to provide a clearer picture of how this regulatory change affected the spreads and market depth of the institutional trades in our sample.

As a preliminary test, we examine HFT trade, order, and cancellation statistics surrounding the fee change and do indeed see significant changes. For each stock, we calculate the average daily number of trades, orders, and cancellations during the 3-month periods immediately before and after the fee change, and we also calculate the percentage change in these averages after the fee change. Table 3 reports these summary statistics. Before the fee change, HFTs were involved in an average of 5,220 trades per stock-day, while after the fee change, they were involved in only 4,451 trades per stock-day (with no significant increase in the number of shares per trade). This represents a decrease of 14.7%, which is statistically significant at the 1% level. Orders and cancellations were also significantly affected. Before the fee change, HFTs submitted 116,783 orders and 112,611 cancellations per stock-day, while, after the fee change, they submitted 91,778 orders and 88,250 cancellations per day,

[7]For more information about the integrated fee model and feedback from marketplace members about the proposed model, see IIROC Notices 11-0125 and 12-0043.

representing decreases of 21.4% and 21.6%, respectively. Both changes are also significant at the 1% level. These results corroborate the findings in Malinova, Park, and Riordan (2018), who provide a detailed examination of market quality metrics around the integrated fee model. In particular, they show that, following the fee change, market-making HFTs reduced their relative presence at the most competitive prices by posting limit orders less frequently, clearly indicating that the fee change had a significant effect on HFT behavior. In the following two subsections, we examine the implications of these changes on the transaction costs for large institutional trades.

3.1 Baseline results

We examine the effect of the integrated fee model on the execution costs of large institutional trades by testing the following ordinary least squares (OLS) regression model:

$$\begin{aligned} IS_{i,j,t} &= \beta_1 \cdot \ln(TSize_{i,j,t}) + \beta_2 \cdot Fee_t + \beta_3 \cdot (Fee_t \times \ln(TSize_{i,j,t})) + \\ &\quad \gamma \cdot X_{i,j,t} + \delta_j + \varepsilon_{i,j,t}. \end{aligned} \tag{3}$$

In this specification, i represents the institutional trade, j represents the stock, t represents the date that the trade was initiated, and δ_j denotes stock fixed effects. We also double cluster the standard errors by stock and date. Fee is an indicator variable that equals one if the trading day is on or after April 1, 2012, the date that the integrated fee model went into effect, and zero otherwise. $\ln(TSize)$ is the natural log of the standardized dollar value of the institutional order, where the standardized dollar value is calculated as the number of shares in the institutional order multiplied by the bid-ask midpoint at the initiation of the order and then divided by the trade size minimum of \$100,000. The β_2 coefficient on Fee represents the change in the average fixed cost (which we refer to as the "spread") for institutional trades following the fee change, while the β_3 coefficient on $Fee \times \ln(TSize)$ represents the change in the average price impact for institutional trades following the fee change. If the fee change makes it costly for HFTs to quickly modify orders in response to institutional order flow, then we would expect a negative β_3 and a positive β_2 coefficient. Finally, X is a vector of the following market-level control variables during the execution of the large institutional trade: (1) $Mret$, the contemporaneous S&P/TSX 60 market return multiplied by the direction of the large institutional trade, which is meant to control for the fact that the average execution price is also driven by market-wide price movements unrelated to this particular trade, and (2) $|Mret|$, the absolute value of $Mret$, which is meant to control for contemporaneous market-wide volatility, like in Chordia, Roll, and Subrahmanyam (2002) and Ait-Sahalia and Sağlam (2017a, 2017b).

The results from the regression in Equation (3) are reported in Column 1 of Table 4. In this test, we only focus on institutional trades executed during the 6-month period surrounding the fee change. Importantly, we find a significant decrease in the price impact coefficient following the fee change. Specifically, the coefficient on $\ln(TSize)$ decreases by 14.5% (0.98/6.74), which provides support for our hypothesis that the fee change led to lower price impact for institutional trades because it impaired the ability of HFTs to cheaply modify their orders in the presence of these trades. We also find that the average spread for institutional trades increased by 3.0 basis points following the fee change; our interpretation is that market-making HFTs widened their bid-ask spreads to offset the costs associated with the regulation. Further, this latter result corroborates evidence in Malinova, Park, and Riordan (2018) showing that average execution costs for retail trades and institutional trades using all marketable limit orders increased after the fee change. For sufficiently large institutional trades, the reduced transaction costs from the deeper market can outweigh the increased transaction costs from the wider spread, leading to lower overall execution costs. We calculate the "break-even" trade size by setting the marginal cost from the increased spread equal to the marginal savings from the increased depth and find a break-even point of about \$2.1 million ($\exp(3.0/0.98) \times \$100,000$). The top 7% of institutional trades by size exceed this \$2.1 million threshold and account for approximately 45% of total trading volume from our sample of institutional trades.

Columns 2 to 5 of Table 4 demonstrate that the results from Column 1 are robust to alternative model specifications. In Column 2, we include additional trade and stock-level control variables: (1) the aggressiveness of the institutional trade (Agg), (2) the number of hours it takes to complete the institutional trade ($Time$), and (3) the natural log of stock dollar volume, expressed in millions of dollars, on the day that the institutional trade was initiated ($\ln(Dvol)$). We find that more aggressive trades have higher implementation shortfall, reflecting the cost of taking liquidity from the limit-order book. Trades with shorter time spans also have higher implementation shortfall; our interpretation is that these traders are trading on short-lived information and do not have sufficient time to strategically spread out their trades in a way that makes it difficult for market makers to distinguish these trades from noise trader activity, like in Kyle (1985) and Collin-Dufresne and Fos (2016). We also find that implementation shortfall is lower on days with higher trading volume, as informed traders can more easily "hide" within the noise when executing their trades (Kyle 1985).[8] In Column 3, we use all of the trading days in the sample as opposed to the 6 months surrounding the fee change. In Columns 4 and 5, we redefine institutional trades as having a minimum institutional trade size of \$500,000 and \$1 million and standardize the $Tsize$ variable using the minimum institutional trade size in that subsample. For the tests in Columns 3 to 5, our baseline results are similar, with spreads increasing and price impact decreasing for large institutional trades following the fee change.

3.2 Price impact analysis by institutional trader type

It is possible that our results in the previous subsection are being driven by changes in overall market conditions or other unobserved outcomes around the time of the fee change.

[8] We should note that the fee change also affected these control variables, with average aggressiveness decreasing by 2.5 percentage points (as shown in Malinova, Park, and Riordan 2018), time to execution increasing by 0.20 hours, and average stock dollar volume decreasing by about 7%. Combined, we find that these post-fee changes do not have a significant effect on average implementation shortfall. Further, the test in Column 1 of Table 4 indicates that our baseline results are robust to the exclusion of these control variables.

We address this concern by exploiting heterogeneities in the trading motivations of institutional traders. Some institutions allocate resources toward obtaining an informational advantage, and thus are motivated to trade on their information. Other institutions, such as pension funds, often trade for reasons unrelated to information about fundamental security values, such as a need for portfolio rebalancing. Theory suggests that HFTs compete with informed institutional traders, who trade with greater urgency because of their often short-lived information advantage. If the integrated fee model makes it more costly for HFTs to compete with these trades, as our results in the previous subsection suggest, then we would expect a stronger reduction in price impact for informed institutional traders compared to uninformed institutional traders. In addition, changes in overall market conditions or other unobserved outcomes around the time of the fee change are unlikely to have a differential effect on the trading costs of informed versus uninformed traders, further strengthening the link between the fee change and institutional execution costs.

Our methodology for classifying informed institutional traders is as follows. First, we calculate the return for each institutional trade. For institutional buy (sell) orders, the return is calculated as (the negative of) the percentage difference between the closing price 5 days after the trade has been executed and the size-weighted average price of the trade. The 5-day window for calculating returns for institutional trades is also used in Chan and Lakonishok (1995). This represents a balance between a shorter window, which is more likely to produce return estimates that are biased by the transitory price impact of the trade, and a longer window, which is more likely to produce noisier return estimates of institutional trade performance. Then we calculate the average return for each institutional trader using their moving history of institutional trade returns. Calculations are done on a monthly basis. Finally, institutions are placed into terciles based on their average historical returns up to the end of the previous month, where the terciles are denoted by $g \in \{H, M, L\}$ (high, medium, low). Institutions in the highest tercile are considered most likely to be trading on private information.

In a preliminary test, we examine whether informed institutional traders are trading ahead of same-direction price movements out-of-sample. For each institutional buy (sell) order, we calculate the (negative of the) return for various time windows relative to the price at the beginning of the trade. The time windows we consider are the end of the trade at day t and the close of days $t+1$, $t+5$, and $t+20$. For example, if an institution initiated a buy order when the price of the security was \$100.00 and the closing price 5 days later was \$105.00, then the return equals 5%. If institutions in the high-informed group are truly informed, then we should expect these institutions to incur positive and significant average returns in future time periods. Figure 1 displays the cumulative returns for institutional trades originating from each of the three informed groups. For the highly informed group, we find that the mean 5-day return relative to the price at the beginning of the institutional trade is about 17 basis points, which is significantly higher than the 5-day returns for the low and medium informed groups. The results are even stronger for mean 20-day returns, with the highly informed group averaging about 26 basis points and the low and medium informed groups averaging about 10 basis points. If we factor in the transaction costs from the implementation shortfall of the trade, then we find that the positive 20-day returns are only significant for the highly informed group. Overall, our evidence indicates that institutions in the highly informed group trade profitably out of sample.

We examine the differential effect of the integrated fee model on the execution costs for the three informed institutional trader groups. For each informed group $g \in \{H, M, L\}$, we test the following regression model:

$$\begin{aligned} IS_{i,g,j,t} &= \beta_{1,g} \cdot \ln(TSize_{i,g,j,t}) + \beta_{2,g} \cdot Fee_t + \beta_{3,g} \cdot (Fee_t \times \ln(TSize_{i,g,j,t})) + \\ &\quad \gamma_g \cdot X_{i,g,j,t} + \delta_j + \varepsilon_{i,g,j,t}, \end{aligned} \tag{4}$$

where all variables are defined as before and the standard errors are double clustered by stock and date. The vector X contains the same control variables as the test in Column 2 of Table 4, although our results are not affected if we exclude Agg, $Time$, and $\ln(Dvol)$. If our hypothesis is correct about the integrated fee model having the largest decrease in price impact for the highly informed group, then we would expect a more negative β_3 for group H compared to the other two groups (M and L).

The results from the above regression model are reported in Columns 1 to 3 of Table 5. Institutional trades from the high-informed group are analyzed in Column 1. We find that, following the fee change, the price impact of large institutional trades from the high-informed group decreases by approximately 27.5% (2.69/9.79); this change is statistically significant at the 1% level. In contrast, Columns 2 and (3) indicate a 10.1% decrease (0.89/8.80) in price impact for the medium-informed group (significant at the 10% level) and no statistically significant decrease in price impact for the low-informed group. The regression in Column 4 formally tests the change in price impact for the high-informed group relative to the changes in price impact for the other two groups. Specifically, we use a pooled sample of institutional trades from all three groups and construct indicator variables representing the medium-informed and low-informed groups ($\mathbf{1}_M$ and $\mathbf{1}_L$) and interaction terms between each of these indicator variables and the key dependent variables (Fee and $Fee \times \ln(TSize)$). Similar to the tests in Columns 1 to 3, we find a much stronger negative effect on the price impact of large institutional trades for the highly informed group than for the other two groups. The differences are Statistically significant.[9] We also provide graphical evidence of the implementation shortfall changes for all traders and each of the informed groups in Figure 2 by calculating the average IS for each week. Panels A through C indicate that the mean IS for the pooled sample and the low and medium informed groups increased, and panel D indicates that the mean IS for the highly informed group decreased, a finding consistent with the evidence from our regressions. Further, these graphs indicate that our results are not being driven by general trends in institutional execution costs. Over-

[9] We also find evidence that the spread effect of the fee change, represented by the coefficient on the Fee variable, is positive and significant for only the low-informed and medium-informed groups, suggesting that smaller-sized institutional trades from the high-informed group also slightly benefit from the decrease in price impact following the fee change.

all, the results from the regressions and figures indicate that the integrated fee model led to higher informational rents for informed institutional traders through the price impact channel. Furthermore, this evidence suggests that some traders could be deterred from acquiring costly information in the presence of HFT, which is consistent with the evidence in Weller (Forthcoming) showing that algorithmic trading can lead to reduced price efficiency because it deters information acquisition.

4 HFT Dynamics around Large Institutional Orders

Microstructure theory suggests that fast traders can use their superior technological capabilities to infer information from and compete with large institutional trades. While the evidence in the previous section indicates an important link between HFT and the execution costs of large institutional trades, particularly those originating from informed institutional traders, it is still unclear how HFTs dynamically interact with these trades. In this section, we address this question by examining how the inventory dynamics and limit-order submission activities of HFTs change in the presence of large institutional trades.

4.1 HFT inventory and quote dynamics

We begin by examining how HFT inventory dynamics depend on current HFT inventory levels and institutional order flow. Theoretical models of market maker inventory dynamics suggest that market makers become increasingly aggressive about reverting their inventory levels as they deviate further from zero. When market makers inventory is highly positive, for example, they can revert this position faster by posting more competitive ask prices and submitting more marketable limit orders to quickly sell at available bid prices. Hansch, Naik, and Viswanathan (1998) (HNV) examine inventory dynamics for dealers on the London Stock Exchange in 1991 and 1992 and show that inventory mean reversion rates are highly nonlinear and increasing in their absolute inventory levels. In this section, we adapt the HNV methodology to the HFT market-making environment and examine how institutional trades affect these mean-reversion rates.

First, we outline our methodology for examining HFT inventory mean reversion rates as a function of their current absolute inventory levels. One of the main variables used in HNV is $I_{m,j,t}$, which denotes the standardized inventory level of dealer m in stock j during time period t. $I_{m,j,t}$ is calculated as the accumulated net share position for dealer m in stock j up to time period t minus the previous-months time-series mean of the net accumulated net share position. Demeaning this variable removes any unobserved inventory positions for dealer m at the beginning of the sample. The demeaned variable is then divided by the previous-months standard deviation of the inventory position, thus controlling for differences in dealer risk aversions and making dealer inventories comparable in the cross-section. By construction, $I_{m,j,t}$ has a mean of zero and a standard deviation of 1. The main difference between our study and HNV when calculating $I_{m,j,t}$ is the time index t, which represents 15-minute periods in our study and 1-day periods in HNV. This is a consequence of the high-frequency trading environment used in this study, as opposed to the relatively slower-moving dealer environment from the early 1990s explored in HNV.

We do not focus on the HFT market makers m in aggregate, as they trade with both other trader types and each other, complicating aggregate inventory dynamics. Instead, we focus on inventory dynamics for individual HFTs. For our subsequent tests, we focus on the primary HFT market maker in each stock, which is the HFT market maker with the highest average trading volume for that stock across the sample period, averaging about 28.4% of HFT trading volume. The main advantage of this approach is that the primary market maker ends the day with a flat position for almost all of the stock-days in our sample. In particular, we find that the primary HFT ends the day with an average absolute inventory position of 3.3% of their trading volume for that day, with 75.5% of those stock-days ending below 1%. The nonprimary HFTs, on the other hand, end the day with an average absolute inventory position of 45.8% of their trading volume for that day, with only 10.6% of those stock-days ending below 1%. The primary HFTs also typically end the day flat if they are trading in other stocks in which they are not the primary HFTs, with an average end-of-day inventory position of 2.6%. For these stocks, the primary HFT averages about 13.7% of HFT trading volume. For cross-listed securities, these statistics suggest that if the HFT is trading in both the Canadian and U.S. markets, then the HFT inventory management is within country. Otherwise, it would be possible that the HFT closes a particular stock-day with a nonzero position of x shares in the Canadian market and $-x$ shares in the U.S. market. Another advantage of our approach is that because these HFT market makers tend to end the day with a zero position, they are unlikely to be trading with multiple user IDs. Finally, we should note that the results in this section are robust to defining the primary market maker on a daily basis or examining HFT inventory dynamics for all HFTs who have been identified as primary in any stock, which are those with generally well-behaved inventory dynamics.

In theory, changes in market maker inventories will be nonlinearly related to the current inventory level, in that market makers will be more aggressive about reversing inventory positions that are further from zero. To take these nonlinearities into account, we follow the methodology in HNV and interact $I_{m,j,t-1}$ with indicator variables representing how far the normalized inventory level has deviated from zero. Specifically, for each $k \in \{1, 2, 3\}$, we define D^k as an indicator variable that equals one if $k - 1 \leq |I_{m,j,t-1}| < k$, and define D^4 as an indicator variable that equals one if $|I_{m,j,t-1}| \geq 3$ (each of these indicator variables equals zero otherwise). D^1 represents a nonextreme inventory position (within one standard deviation of the mean inventory level) and D^4 represents the most extreme inventory position (greater than 3 standard deviations from the mean inventory level). Similar to HNV, we then estimate inventory mean reversion rates as a function of the absolute inventory levels using the following OLS regression model:

$$\Delta I_{m,j,t} = \alpha + \sum_{k=1}^{K} \beta_k D^k I_{m,j,t-1} + \varepsilon_{m,j,t-1}, \tag{5}$$

where $\Delta I_{m,j,t} \equiv I_{m,j,t} - I_{m,j,t-1}$. If market-making HFTs become increasingly aggressive about reversing their inventory positions as those positions deviate further from zero, then β_k will be increasingly negative for higher values of k.

The results from this regression test are reported in Column 1 of Table 6. Like in HNV, we find stronger rates of mean reversion in HFT inventory levels when the inventory position becomes more extreme. In particular, for smaller deviations in the HFT market maker inventory level (i.e., within 1 standard deviation of the mean HFT inventory level), the mean reversion coefficient equals -0.276, implying that 27.6% of a smaller inventory positions are reversed in the following 15-minute period. In contrast, when an inventory position is 1–2, 2–3, or more than 3 standard deviations outside the mean inventory level, the mean reversion coefficient equals -0.281, -0.312, and -0.377, respectively, indicating that the inventory reversion rate is increasing in the absolute inventory level. These coefficients are also significantly different from each other. It is useful to express each mean reversion coefficient (β_k) as the corresponding half-life of the HFT inventory position, that is, the number of minutes it takes for an HFT to revert half of its inventory position. The half-life (H_k) is derived and then calculated as follows:

$$\frac{1}{2} = (1+\beta_k)^{H_k/15} \Rightarrow H_k = \frac{\ln(1/2)}{\ln(1+\beta_k)} \times 15. \tag{6}$$

The half-lives for different inventory levels are also reported in Column 1 of Table 6. For smaller inventory positions ($k = 1$), it takes a market-making HFT 32.2 minutes to unwind half of that position, while for the most extreme positions ($k = 4$), it takes 22.0 minutes, which is about 31.7% lower than the half-life for smaller inventory positions. Similar to HNV, we find that HFTs move their inventories at faster rates when these inventory levels are further from zero.

We examine how inventory mean reversion rates differ in the presence of large institutional trades. If market-making HFTs absorb institutional order flow and also infer information from that order flow, then they will likely reverse their position at a faster rate to reduce the adverse-selection risk associated with their position. We retest the regression model in Column 1 of Table 6 for time periods in which an institutional buy or sell order is underway. We also include controls for characteristics of the institutional trade, including the aggressiveness of the trade (Agg), the number of hours until completion ($Time$), and the dollar size of the trade ($Tsize$). The results are reported in Columns 2 and 3. For all inventory levels, we find stronger mean reversion rates in the presence of large institutional buy or sell orders compared to the reversion rates from the pooled sample in Column 1. The regression model in Column 4 formally tests the differences in these inventory mean reversion rates. First, we construct a signed indicator variable L that equals one if an institutional buy is currently being executed, -1 if an institutional sell is currently being executed, and zero otherwise. Then, we retest the regression model in Column 1 with the inclusion of the interaction terms $I_{t-1} \times D^k \times |L_{t-1}|$ for $k \in \{1, 2, 3, 4\}$. (Our results are similar if we interact $I_{t-1} \times D^k$ with indicator variables for institutional buy orders and institutional sell orders separately.) These interaction terms are meant to capture the incremental effect of an institutional trade on inventory reversion rates. The results in Column 4 indicate that reversion rates are significantly higher in the presence of an institutional trade at all inventory levels, with reversion rates increasing by about 50.5% (0.109/0.216) for smaller inventory positions and about 19.9% for extreme inventory positions (0.069/0.347). Importantly, we also include L_{t-1} in this regression model. This is meant to capture how HFTs trade in the presence of an institutional buy or sell order, inventory levels notwithstanding. The positive and significant coefficient on L_{t-1} indicates that HFTs buy more shares in the presence of institutional buy orders and sell more shares in the presence of institutional sell orders, regardless of their current inventory level, which is consistent with the HFT implementing order anticipation strategies.[10]

[10]In the Internet Appendix, we show that average HFT profit per share is negative at the beginning of an institutional trade and increases as the trade progresses. This finding is consistent with that of Van Kervel and Menkveld (Forthcoming), who similarly show that HFT profits in the presence of institutional trades are initially negative but eventually become positive as the trade progresses.

The coefficients from Column 4 of Table 6 also can be used to provide a useful quantitative breakdown of the reasons for HFT market maker inventory reversion. Recall that for smaller inventory positions, the HFT reverses 21.6% of their inventory position in the subsequent period when no institutional trade is being executed. This increases by 10.8 percentage points to 32.4% if a large institutional trade is underway. The additional 10.8% provides a useful proxy for the component of the HFT inventory reversal that is attributable to information inferred from the institutional order. For smaller inventory positions (D^1), this implies that 33.5% (10.8/32.4) of inventory reversion in the presence of a large institutional trade can be attributed to the updated information set of the HFT. For extreme inventory positions (D^4), we find that 16.6% ($6.9/(6.9 + 34.7)$) of HFT inventory reversion can be attributed to their updated information set. This is lower than the 33.5% reported for smaller inventory positions because of the nonlinearities in the inventory reversion rates reported earlier.

We provide further insight into HFT market-making behavior by examining how their quoting activities relate to their inventory levels and institutional order flow. If an HFT accumulates an extreme positive inventory position, for example, then we would expect them to increasingly shift to higher limit-order submission activity on the ask side of the limit-order book to unwind their position. In addition, if an HFT learns from institutional net order flow that stock prices are likely to decrease, then we would expect them to shift from submitting bid-side limit orders to mitigate adverse-selection risk and shift toward submitting more ask-side limit orders. Similar to our intuition about HFT inventory changes in the previous test, we expect that current inventory levels and institutional order flow will affect HFT quoting behavior as well.

The key dependent variable used for analyzing HFT quote dynamics is the normalized net order submission activity of the market-making HFT. For each stock i and 15-minute period t, we first calculate the total number of shares in the buy limit orders submitted by the HFT minus the total number of shares in their sell limit orders. We then normalize this net order submission variable by subtracting its trailing mean and dividing the difference by its trailing standard deviation. We denote this variable $Q_{m,j,t}$. Our overall objective is to examine how $Q_{m,j,t}$ is affected by the inventory level of the HFT ($D^k \times I_{m,j,t-1}$), institutional order flow ($L_{j,t-1}$), and the inventory level during an institutional trade execution ($D^k \times I_{m,j,t-1} \times |L_{j,t-1}|$).

The results are reported in Table 7, and we find similar results to our tests of HFT inventory changes. The evidence in Column 1 indicates that HFTs submit more sell orders than buy orders when their inventory is positive, and vice versa. This effect is increasing and nonlinear in the HFT inventory level. In Column 2, we include the signed institutional trade indicator variable $L_{j,t-1}$ and find that HFTs submit more same-direction limit orders than opposite-direction limit orders in the presence of an institutional order. Column 3 indicates that these results are robust to the inclusion of the institutional trade control variables used in the previous test. Finally, the regression model in Column 4 includes the interaction variables $D^k \times I_{m,j,t-1} \times |L_{j,t-1}|$ to test how quote dynamics change during institutional trade executions at different inventory levels. For all inventory levels, we find that HFTs are more aggressive about submitting same-direction limit orders relative to opposite-direction limit orders in the presence of large institutional trades. Overall, these results corroborate our previous evidence showing how HFT inventory dynamics change during large institutional trade executions.

4.2 Abnormal HFT buying and selling activity

The results in the previous subsection indicate that HFT net inventory changes and order submission activity are strongly affected by their current inventory level and institutional order flow. These tests, however, do not paint a complete picture of how HFTs alter their buying and selling activities in the presence of a large institutional trade. While we have shown that net inventory changes by the HFT are positive in the presence of institutional buys, for example, it is still unclear whether these changes are being driven by an increase in buying activity or a decrease in selling activity. It is important to delineate between these two effects. An increase in same-direction trades and orders by the HFT in the presence of a large institutional trade would be more consistent with the "back-running" prediction in Yang and Zhu (2017), suggesting that HFTs are directly competing with these trades. A decrease in opposite-direction trades and orders, on the other hand, would be more consistent with the "phantom liquidity" prediction in Ait-Sahalia and Sağlam (2017a, 2017b), suggesting that, instead, the HFTs are reducing liquidity provision to these large trades. In this section, we examine how abnormal HFT buying and selling activity is affected by institutional order flow.

We break up the HFT net inventory change into its abnormal buy and sell components. Similar to our calculation of the standardized inventory level of the HFT (I), we calculate abnormal HFT buy (sell) volume by taking the sum of the shares bought (sold) by the HFT in a particular stock and 15-minute period, demeaning this sum using the trailing mean from all 15-minute periods in the past month, and then dividing the difference by the trailing standard deviation. We also break up HFT net order submission activity (Q) into abnormal HFT buy limit-order activity and abnormal sell limit-order activity, constructed similarly. This gives us a total of four dependent variables. We examine how these variables are affected by the current inventory level of the HFT and the presence of an institutional buy order ($\mathbf{1}_B$) or institutional sell order ($\mathbf{1}_S$). In these tests, we will also control for characteristics of any large trade that is currently being executed (Agg, $Time$, $TSize$) and the abnormal dollar volume of the stock in that 15-minute period.

Table 8 reports our tests of the determinants of abnormal HFT buying and selling activity. Columns 1 and 2 indicate that buy trades by the HFT significantly decrease and sell trades by the HFT significantly increase when inventory levels are positive, and vice versa when inventory levels are negative. Consistent with our previous evidence, these effects become more pronounced when the inventory levels are extreme, although the magnitude of the effect is a little weaker compared to the previous tests. Importantly, we find that same-direction trade activity significantly increases during an institutional trade execution. We also find a slight increase in opposite-direction trade activity during an institutional trade execution, even after controlling for abnormal stock trading volume in that 15-minute period, although it is significantly lower than the increase in same-direction trade activity (the difference in coefficients and the associated p-values are in the row labeled $\mathbf{1}_B - \mathbf{1}_S$). For the tests examining abnormal buy and sell limit-order activity in Columns 3 and 4, we find similar results, with abnormal same-direction limit-order volume increasing more than abnormal opposite-direction limit-order volume in the presence of a large institutional trade (although the net effect is not statistically significant for abnormal sell limit orders). Overall, this evidence suggests that the same-direction inventory changes by HFTs can be better explained by an increase in same-direction trade and order activity than a decrease in opposite-direction trade and order activity, which is consistent with the back-running prediction in Yang and Zhu (2017).[11]

4.3 Large trade predictors and HFT dynamics

In this subsection, we examine the predictors of large institutional trades and the relationship between HFT and predicted institutional trading activity. Based on past literature, we include past price movements, trade imbalances, and changes in the limit-order book in the publicly observable information set used by market-making HFTs to predict large institutional trades. Although other traders can also act on this information, HFTs have an advantage because they can act on this information quickly. Several related studies introduce theoretical models with fast traders, showing that the presence of fast traders can lead to worsened liquidity outcomes for other traders or potential market stability issues (see, e.g., Baldauf and Mollner 2018; Bongaerts and Van Achter 2016). Many large institutional parent orders are executed over time using a series of child orders, and the successive price

[11]In unreported results, we find an increase in opposite-direction order cancellations by HFTs in the presence of institutional trades, suggesting that the phantom liquidity mechanism in Ait-Sahalia and Sağlam (2017a, 2017b) also plays a role in how HFTs interact with these trades.

changes associated with these orders provide a useful signal about subsequent child orders. If HFTs can time their trades with enough precision, then they can trade ahead of the predicted institutional order flow, leading to increased execution costs for the institutional order (Bessembinder et al. 2016). Finally, changes in the shares being bid and offered on the limit-order book can also provide useful information about future prices and large institutional orders (Harris and Panchapagesan 2005), especially in light of the evidence in Table 2 that 43% of institutional share volume is executed using standing limit orders.

As a first step, we examine if publicly observable information is predictive of institutional buy and sell orders using a probit regression framework. The dependent variables of interest are the institutional buy and sell indicator variables ($\mathbf{1}_B$ and $\mathbf{1}_S$). The independent variables include past percentage price changes (r), normalized trade imbalances (y), and normalized limit-order imbalances ($LOIB$). y is defined as the total number of buy orders minus the total number of sell orders, which is then demeaned and divided by its standard deviation using data from the past month. $LOIB$ is defined similarly, except that we use the total share size of all buy limit orders minus the total share size of all sell limit orders. We include four lags of r, y, and $LOIB$ in our probit tests to account for potential informativeness about relatively long-run security values (Foucault, Hombert, and Roşu 2016). We test the following probit regression model for $z \in \{B, S\}$:

$$\Pr(\mathbf{1}_{j,t,z}|\cdot) = \sum_{k=1}^{4} \beta_{k,z} r_{j,t-k} + \lambda_{k,z} y_{j,t-k} + \phi_{k,z} LOIB_{j,t-k} + \delta_j + \varepsilon_{j,t-1}, \tag{7}$$

where t denotes a 15-minute time interval and δ_j denotes stock fixed effects. Standard errors are double clustered by stock and date.

The results from the probit regression tests are reported in the first two columns of Table 9. Column 1 examines the predictability of institutional buy orders. Overall, we find that past returns, trade imbalances, and limit-order imbalances are all positively associated with a higher probability of institutional buying in the upcoming period. The return coefficients from previous lags are also positive, but monotonically decreasing in the lag. Interestingly, we find that normalized trade imbalances are more predictive of institutional buys at higher lags, suggesting that trade imbalances provide information about the overall trajectory of a multiperiod institutional buy order. Finally, we find that limit-order imbalances also provide information about the likelihood of an upcoming institutional buy order, reflecting our earlier observation that 43% of the shares in an average institutional order are executed using standing limit orders. Like past returns, the coefficients on limit-order imbalances are also monotonically decreasing in the lag. The regression reported in Column 2 tests the predictability of institutional sell orders, and we find similar results, in that negative returns, trade imbalances, and limit-order imbalances from previous periods are associated with a higher likelihood of an institutional sell order in the upcoming period.

Having established that past returns, trade imbalances, and order imbalances are predictive of institutional trades, we next move on to examining whether these publicly observable variables are also predictive of HFT inventory changes. We test an OLS regression model of normalized HFT inventory changes ($\Delta I_{m,j,t}$) on these predictive variables. Like in our previous tests of inventory changes, we also include controls for the HFT inventory level ($I_{m,j,t-1} \times D_j$ for $j \in \{1, 2, 3, 4\}$). These results are reported in Column 3 of Table 9. First, we find that the coefficients on the inventory level controls are similar to those reported in our previous inventory change regressions, illustrating the robustness of those results. More importantly, we find that past returns are positively related to future HFT inventory changes. This suggests that HFTs use past returns to predict institutional order flow and compete with those orders. Furthermore, this is consistent with the probit regression results showing that past returns are associated with a higher likelihood of a same-direction institutional order. Interestingly, we find that trade and order imbalances at higher lags are negatively related to HFT inventory changes in the current period. This is consistent with the prediction in Yang and Zhu (2017) that HFT demand is negatively related to the previous period trade imbalance, as HFTs utilize their information advantage relative to this imbalance.

In the final two regressions in Columns 4 and 5 of Table 9, we test how HFT inventory changes depend on current inventory levels and the predicted component of the signed indicator variable for a large institutional trade (L). The predicted component is calculated using a first-stage regression of L on the publicly observable variables from the regressions in Columns 1 to 3. Consistent with these previous tests, we find that the predicted component of L is positively related to inventory changes by the HFT, suggesting that HFTs compete with institutional trades using information that is predictive of these trades. However, one issue is that HFTs might be just chasing momentum or past order flow, and thus not necessarily trading directly in response to large institutional trades. We address this issue by testing the relationship between HFT inventory changes and the predicted component of L for aggressive institutional trades only. We define an institutional trade as "aggressive" if the percentage of its trade volume using marketable limit orders is above the median value for all institutional trades (about 57%). Aggressive institutional trades are more likely to be based on private information because they generally occur with greater urgency. Furthermore, because aggressive trades deplete limit-order book liquidity and move prices, it is more likely that observed price movements and order imbalances are being driven by these trades. The results of this test are reported in Column 5, and we find that HFTs are even more likely to compete with aggressive institutional trades using publicly observable information, with a coefficient on the predicted value of L that is about twice as large as the same coefficient from the regression using all institutional trades in Column 4.

5 Conclusion

High-frequency traders play an important market-making role in modern equities markets, and the continued modernization of financial markets in other asset classes will expand the role of HFTs in these spaces as well. Thus, it is more important than ever to understand and identify the potential benefits and drawbacks of HFT in these markets.

Motivated by the theoretical microstructure literature, we examine how HFTs interact with large institutional trades in modern equities markets and what this ultimately implies for the costs of these trades. Using a regulatory change that increased the cost of sending messages to Canadian exchanges, we show that the price impact of large institutional trades decreased by about 15% following this regulatory change, providing an important link between HFT and institutional trading costs. Spreads also widened following this regulatory change. Together, this evidence suggests that the integrated fee model increased the cross-subsidy from small, uninformed traders to large, informed traders. We also show that HFTs compete with large institutional orders over time, partially because the HFTs are reversing inventory positions accumulated from providing liquidity to past institutional child orders but also because the HFT is competing with the institutional order using information inferred from past child orders. Past returns, trade imbalances, and limit-order book imbalances are all important predictors of large institutional orders, and HFTs appear to use this information to competitively trade with subsequent institutional child orders. Overall, our evidence suggests that HFT is associated with higher execution costs for large and information-based institutional orders and lower costs for small, uninformed orders. Regulatory bodies may wish to incorporate this multifaceted effect of HFT on execution costs when contemplating regulatory changes.

References

Admati, A. R., and P. Pfleiderer. 1991. Sunshine trading and financial market equilibrium. *Review of Financial Studies* 4:443–81.

Ait-Sahalia, Y., and M. Sağlam. 2017a. High frequency market making: Implications for liquidity. Working Paper, Princeton University.

———. 2017b. High frequency market making: Optimal quoting. Working Paper, Princeton University.

Amihud, Y., and H. Mendelson. 1980. Dealership market: Market-making with inventory. *Journal of Financial Economics* 8:31–53.

Anand, A., P. J. Irvine, A. Puckett, and K. Venkataraman. 2012. Performance of institutional trading desks: An analysis of persistence in trading cost. *Review of Financial Studies* 25:557–98.

Anand, A., and K. Venkataraman. 2016. Market conditions, fragility, and the economics of market making. *Journal of Financial Economics* 121:327–49.

Baldauf, M., and J. Mollner. 2018. Fast traders make a quick buck: The role of speed in liquidity provision. Working Paper, Northwestern University.

Bertsimas, D., and A. W. Lo. 1998. Optimal control of execution costs. *Journal of Financial Markets* 1:1–50.

Bessembinder, H., A. Carrion, L. Tuttle, and K. Venkataraman. 2016. Liquidity, resiliency, and market quality around predictable trades: Theory and evidence. *Journal of Financial Economics* 121:142–66.

Bongaerts, D., and M. Van Achter. 2016. High-frequency trading and market stability. Working Paper, Erasmus University.

Breckenfelder, J. 2013. Competition between high-frequency traders and market quality. Working Paper, European Central Bank.

Brogaard, J., T. Hendershott, and R. Riordan. 2014. High frequency trading and price discovery. *Review of Financial Studies* 27:2267–306.

Campbell, J. Y., T. Ramadorai, and A. Schwartz. 2009. Caught on tape: Institutional trading, stock returns, and earnings announcements. *Journal of Financial Economics* 92:66–91.

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

Chan, L. K. C., and J. Lakonishok. 1995. The behavior of stock prices around institutional trades. *Journal of Finance* 50:1147–74.

Choi, J. H., K. Larsen, and D. J. Seppi. Forthcoming. Information and trading targets in a dynamic market equilibrium. *Journal of Financial Economics.*

Chordia, T., R. Roll, and A. Subrahmanyam. 2002. Order imbalance, liquidity, and market returns. *Journal of Financial Economics* 65:111–30.

Collin-Dufresne, P., and V. Fos. 2016. Insider trading, stochastic liquidity, and equilibrium prices. *Econometrica* 84:1441–75.

Comerton-Forde, C., K. Malinova, and A. Park. 2018. Regulating dark trading: Order flow segmentation and market quality. *Journal of Financial Economics.* Advance access published July 4, 2018, 10.1016/j.jfineco.2018.07.002.

Cready, W., A. Kumas, and M. Subasi. 2014. Are trade size-based inferences about traders reliable? Evidence from institutional earnings-related trading. *Journal of Accounting Research* 52:877–909.

Crippen, A. 2014. Buffett, Gates and Munger criticize high-frequency trading. *CNBC*, May 5. www.cnbc.com/2014/05/05/buffett-gates-and-munger-criticize-high-frequency-trading.html

Devani, B., B. Tayal, L. Anderson, D. Zhou, J. Gomez, and G. W. Taylor. 2014. Identifying trading groups—Methodology and results. Discussion Paper, IIROC.

Foucault, T., J. Hombert, and I. Roşu. 2016. News trading and speed. *Journal of Finance* 71:335–82.

Glosten, L. R., and P. R. Milgrom. 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics* 14:71–100.

Hagströmer, B., and L. Norden. 2013. The diversity of high-frequency traders. *Journal of Financial Markets* 16:741–70.

Hansch, O., N. Y. Naik, and S. Viswanathan. 1998. Do inventories matter in dealership markets? Evidence from the London Stock Exchange. *Journal of Finance* 53:1623–56.

Harris, L., and V. Panchapagesan. 2005. The information content of the limit order book: evidence from NYSE specialist trading decisions. *Journal of Financial Markets* 8:25–67.

Hirschey, N. 2018. Do high-frequency traders anticipate buying and selling pressure? Working Paper, London Business School.

Ho, T., and H. R. Stoll. 1981. Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics* 9:47–73.

Jovanovic, B., and A. J. Menkveld. 2016. Middlemen in limit-order markets. Working Paper, New York University.

Keim, D. B., and A. Madhavan. 1997. Transactions costs and investment style: An inter-exchange analysis of institutional equity trades. *Journal of Financial Economics* 46:265–92.

Kirilenko, A., A. S. Kyle, M. Samadi, and T. Tuzun. 2017. The Flash Crash: High frequency trading in an electronic market. *Journal of Finance* 72:967–98.

Kyle, A. S. 1985. Continuous auctions and insider trading. *Econometrica* 53:1315–35.

Laming, D. R. J. 1968. *Information theory of choice-reaction times.* Oxford, England: Academic Press.

Lee, C. M. C., and M. Ready. 1991. Inferring trade direction from intraday data. *Journal of Finance* 46:733–46.

Malinova, K., A. Park, and R. Riordan. 2018. Do retail investors suffer from high frequency traders? Working Paper, University of Toronto.

Menkveld, A. J. 2013. High frequency trading and the *new-market* makers. *Journal of Financial Markets* 16:712–40.

Perold, A. 1988. The implementation shortfall: Paper versus reality. *Journal of Portfolio Management* 14:4–9.

Stoll, H. R. 1978. The supply of dealer services in securities markets. *Journal of Finance* 33:1133–51.

Van Kervel, V., and A. Menkveld. Forthcoming. High-frequency trading around large institutional orders. *Journal of Finance.*

Weller, B. M. Forthcoming. Does algorithmic trading deter information acquisition? *Review of Financial Studies.*

Yang, L., and H. Zhu. 2017. Back-running: Seeking and hiding fundamental information in order flows. Working Paper, University of Toronto.

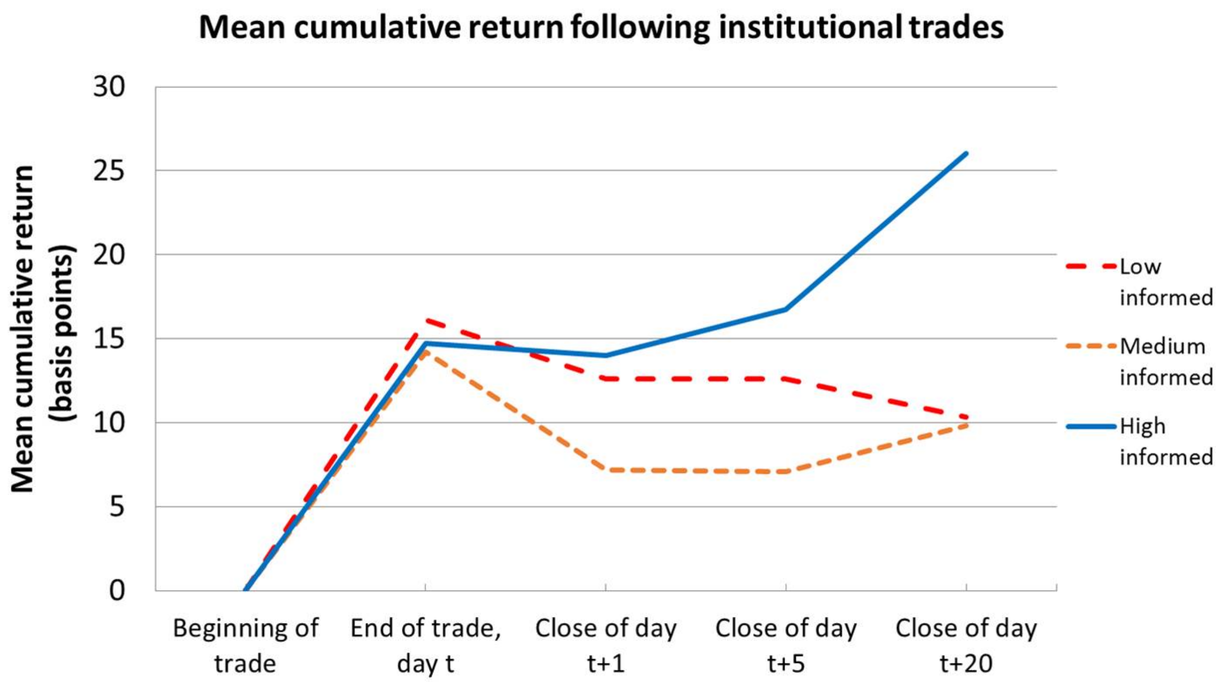

Figure 1. Price evolution around institutional trade by informed type
This graph displays the mean cumulative return relative to the price at the beginning of an institutional trade (p_0). For institutional buy (sell) orders, we calculate the cumulative return at t as (the negative of) $(p_t - p_0)/p_0$. The y-axis is reported in basis points.

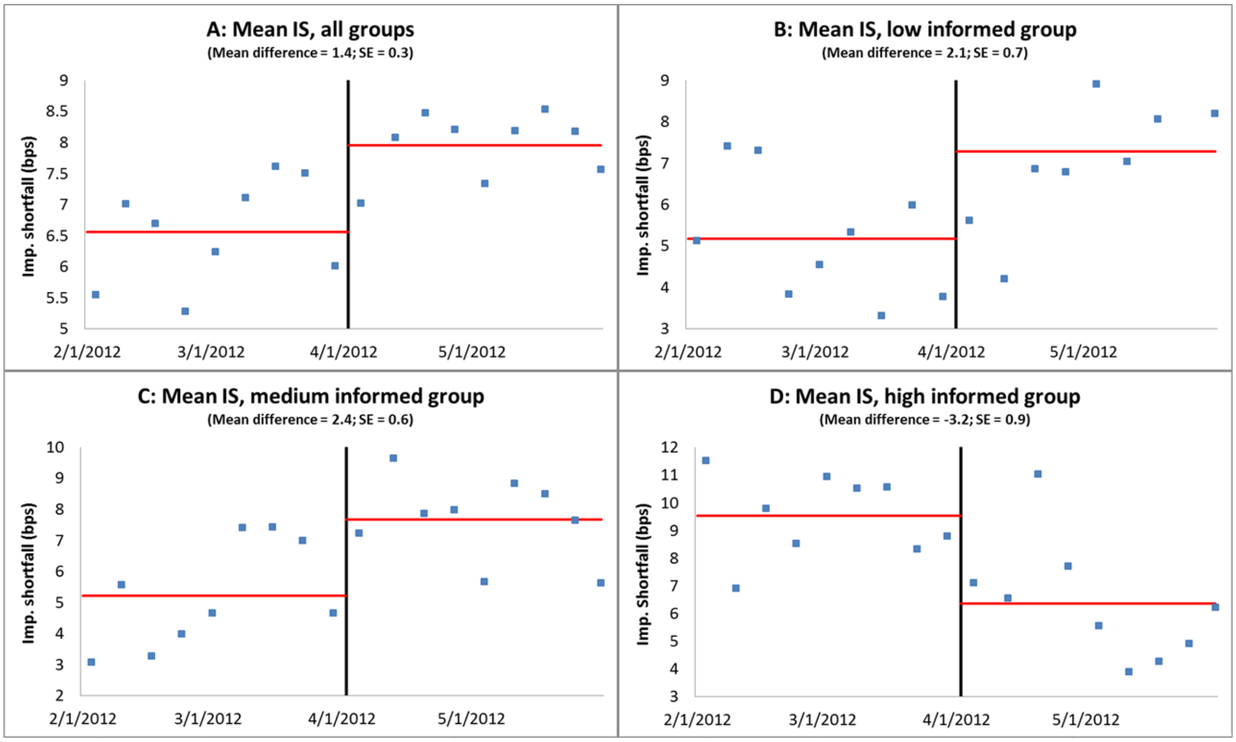

Figure 2. Average implementation shortfall by informed type

This figure displays the weekly mean implementation shortfall (IS) for all traders (panel A) and the three informed groups (panels B to D). The vertical line in each panel indicates the day that the integrated fee model went into effect. The horizontal lines indicate the average IS before and after the fee change. For each panel, we report the mean difference in the horizontal lines and the standard error of the difference in the caption and based on the daily average IS. The y-axis is reported in basis points.

Table 1. Summary statistics for high-frequency trading activity
Percentage of trade volume is the ratio of HFT trading volume to total trading volume (double counted) for each stock-day. *Percentage of orders* is the ratio of total HFT limit-order size to total limit-order size for each stock-day. *Order-to-trade ratio* is the ratio of the number of HFT limit orders to the number of HFT trades. *Aggressiveness* is the percentage of HFT trades that are executed using marketable limit orders, as opposed to passive limit orders. *Trade size* is the number of shares composing a single HFT trade. *Trade value* is the dollar value of the shares composing a single HFT trade. *Inventory ($K)* is the dollar inventory of the highest-volume market-making HFT in a stock and 15-minute period, and Δ*Inventory ($K)* is the change in this variable across 15-minute periods. We also report summary statistics for these variables expressed as a percentage of dollar volume by the HFT in that stock and 15-minute period. For each variable, "PX" represents the Xth percentile of its distribution, and "SD" represents its standard deviation.

	HFT summary statistics (N = 67,787)						
	Mean	Median	P5	P25	P75	P95	SD
Percentage of trade volume (%)	31.6	30.8	11.5	22.0	40.6	53.4	13.1
Percentage of orders (%)	55.4	56.0	21.0	41.3	69.2	85.9	22.9
Order-to-trade ratio	33.1	16.9	5.4	10.5	32.7	119.8	49.5
Aggressiveness (%)	27.8	26.9	7.8	18.2	36.2	50.9	13.3
Trade size (shares)	328	147	111	125	260	1,261	531
Trade value (dollars)	4,354	2,685	459	1,092	5,531	12,133	6,095
Inventory ($K)	3.7	1.3	-105.5	-16.9	23.6	119.4	72.9
Inventory (%)	2.5	0.2	-49.8	-3.3	5.8	63.6	52.1
ΔInventory ($K)	0.0	0.0	-55.6	-7.5	7.5	55.8	48.7
ΔInventory (%)	0.0	0.0	-100.0	-19.4	18.9	100.0	46.9

Table 2. Summary statistics for large institutional trades

Trade size is the dollar value of the institutional trade, measured in millions of dollars. *Number of orders* is the total number of orders submitted by the institution during the execution of the institutional trade. *Number of trades* is the number of smaller trades it takes to execute the large institutional trade. *Aggressiveness* is the percentage of the large institutional trade that is executed using marketable limit orders. *Time to completion* is the number of hours it takes to execute the institutional trade. *Implementation shortfall* is defined in the text and measured in basis points. For each variable, "PX" represents the Xth percentile of its distribution, and "SD" represents its standard deviation.

	Institutional trade statistics (N = 1,173,482)						
	Mean	Median	P5	P25	P75	P95	SD
Trade size ($M)	0.72	0.28	0.11	0.16	0.64	2.53	1.91
Number of orders	234	48	1	11	178	855	2,207
Number of trades	118	50	3	20	124	438	261
Order-to-trade ratio	4.9	1.0	0.1	0.4	1.8	6.4	36.2
Aggressiveness (%)	57.0	61.1	0.0	22.1	96.5	100.0	36.6
Time to completion (hours)	3.0	1.7	0.0	0.1	5.3	6.5	4.0
Implementation shortfall (bps)	7.1	2.5	-97.9	-8.8	23.0	119.3	81.9

Table 3. HFT activity statistics around regulatory fee changes

This table reports cross-sectional summary statistics for the average daily number of trades, orders, and cancellations submitted by all HFTs in each stock during the 3-month periods surrounding the integrated fee model. Cross-sectional statistics are also reported for the percentage changes in HFT trades, orders, and cancellations around the regulatory change. For the mean percentage changes, ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

A. Daily number of HFT trades

	Pre-regulation	Post-regulation	Percentage change
Mean	5,220	4,451	-14.7%***
25th percentile	1,417	1,114	-21.4%
Median	3,089	2,386	-22.8%
75th percentile	6,586	5,587	-15.2%

B. Daily number of HFT orders

	Pre-regulation	Post-regulation	Percentage change
Mean	116,783	91,778	-21.4%***
25th percentile	29,463	21,444	-27.2%
Median	60,590	50,989	-15.8%
75th percentile	161,291	137,355	-14.8%

C. Daily number of HFT cancellations

	Pre-regulation	Post-regulation	Percentage change
Mean	112,611	88,250	-21.6%***
25th percentile	26,929	20,317	-24.6%
Median	57,522	49,182	-14.4%
75th percentile	156,566	125,434	-19.9%

Table 4. Institutional trading costs around regulatory fee changes

This table reports results from OLS regressions that test the effect of the integrated fee model on the spread and price impact for large institutional trades. The dependent variable is the implementation shortfall of large institutional trades. Key dependent variables include Fee and Fee $\times \ln(TSize)$. The regressions in Columns 1 and 2 restrict the sample to the 6-month period surrounding the fee change. Column 3 uses the entire time sample. Columns 4 and 5 use the entire time sample and restrict the minimum institutional trade size to \$500,000 and \$1 million, respectively. All control variables are specified in the main body of the text. Standard errors are double clustered by stock and date. t-statistics are reported in parentheses below the regression coefficients. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	[-3,+3] months		All months	Size > \$500K	Size > \$1M
	(1)	(2)	(3)	(4)	(5)
$\ln(TSize)$	6.742***	8.938***	9.114***	13.141***	15.569***
	(22.07)	(26.6)	(29.67)	(20.69)	(15.72)
Fee	3.002***	3.615***	2.859***	2.057**	2.261**
	(4.25)	(4.50)	(5.26)	(2.46)	(2.20)
$\ln(TSize) \times Fee$	-0.981**	-1.135**	-1.467***	-1.976***	-2.778**
	(-2.16)	(-2.36)	(-4.46)	(-2.73)	(-2.40)
$Mret$	0.245***	0.247***	0.237***	0.282***	0.304***
	(21.21)	(21.53)	(28.18)	(28.01)	(25.42)
$\|Mret\|$	49.126	46.234	67.574	64.475	12.754
	(0.76)	(0.68)	(1.62)	(0.99)	(0.14)
Agg		0.185***	0.166***	0.134***	0.098***
		(22.83)	(36.99)	(17.63)	(9.16)
$Time$		-0.942***	-1.080***	-1.135***	-1.086***
		(-5.76)	(-11.55)	(-8.76)	(-7.26)
$\ln(Dvol)$		-4.059***	-4.405***	-5.483***	-4.262***
		(-6.15)	(-12.12)	(-8.59)	(4.65)
SE clustering	Stock-date	Stock-date	Stock-date	Stock-date	Stock-date
Fixed effects	Stock	Stock	Stock	Stock	Stock
N	279,140	251,584	733,890	263,419	141,739
R-squared	0.061	0.071	0.063	0.077	0.085

Table 5. Institutional trading costs around fee change by informed type
The first three columns of this table report results from OLS regressions that examine the effect of the integrated fee model on the spread and price impact of large institutional trades for each informed group $g \in \{H, M, L\}$. Column 4 reports the results of a pooled regression containing all informed types. The dependent variable in all of the regression tests is the implementation shortfall of large institutional trades. All control variables are specified in the main body of the text. Standard errors are double clustered by stock and date. t-statistics are reported in parentheses below the regression coefficients. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Trader informativeness			
	High	Medium	Low	Pooled
	(1)	(2)	(3)	(4)
$\ln(TSize)$	9.790***	8.797***	8.581***	9.067***
	(14.98)	(19.17)	(10.50)	(24.02)
Fee	0.424	2.129***	4.579***	0.156
	(0.35)	(3.07)	(3.43)	(0.14)
$Fee \times \ln(Tsize)$	-2.693***	-0.889*	-1.069	-2.342***
	(-3.74)	(-1.86)	(-1.26)	(-4.97)
$Fee \times \mathbf{1}_M$				2.138*
				(1.80)
$Fee \times \ln(TSize) \times \mathbf{1}_M$				1.203***
				(3.56)
$Fee \times \mathbf{1}_L$				4.202***
				(2.81)
$Fee \times \ln(TSize) \times \mathbf{1}_L$				1.333***
				(3.54)
$Fee \times (1 + \mathbf{1}_M)$				2.294***
p-value				.001
$Fee \times \ln(TSize) \times (1 + \mathbf{1}_M)$				-1.139***
p-value				.007
$Fee \times (1 + \mathbf{1}_L)$				4.358***
p-value				.000
$Fee \times \ln(TSize) \times (1 + \mathbf{1}_L)$				-1.009**
p-value				.027
Clustering	Stock-date	Stock-date	Stock-date	Stock-date
Controls	Yes	Yes	Yes	Yes
Fixed effects	Stock	Stock	Stock	Stock-Group
N	168,056	327,499	165,766	650,492
R-squared	0.068	0.057	0.061	0.064

Table 6. Inventory mean reversion

This table displays inventory regression coefficients that represent inventory mean reversion rates as a function of inventory levels. D^k is an indicator variable representing inventory levels within a $(k-1)$ to k standard deviation band around the mean for $k \in \{1,2,3\}$. D^4 is an indicator variable representing inventory levels greater than a 3-standard-deviation band around the mean. L is a signed indicator variable that equals 1 (-1) if an institutional buy (sell) order is being executed in that period. Implied half-life is measured in minutes. Differences in coefficients are reported in the panel below the regression coefficients; for regression Column 4, the differences are calculated as $I_{m,j,t-1} \times (1+|L_{m,j,t-1}|) \times (D^k - D^1)$ for $k \in \{2,3,4\}$. Standard errors are double clustered by stock and date. t-statistics are reported in parentheses below the regression coefficients. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	HFT net inventory change ($\Delta I_{m,j,t}$)			
	Full sample	Inst. buy	Inst. sell	Full sample
	(1)	(2)	(3)	(4)
$I_{m,j,t-1} \times D^1$	-0.276***	-0.323***	-0.329***	-0.216***
	(-126.8)	(-101.4)	(-102.9)	(86.9)
$I_{m,j,t-1} \times D^2$	-0.281***	-0.330***	-0.332***	-0.218***
	(-111.2)	(-105.9)	(-110.5)	(-77.4)
$I_{m,j,t-1} \times D^3$	-0.312***	-0.360***	-0.362***	-0.254***
	(-75.6)	(-67.8)	(-72.0)	(-53.7)
$I_{m,j,t-1} \times D^4$	-0.377***	-0.398***	-0.418***	-0.347***
	(-25.4)	(-22.6)	(-26.4)	(-16.6)
$(I_{m,j,t-1} \times D^1) \times \lvert L_{j,t-1} \rvert$				-0.109***
				(-38.0)
$(I_{m,j,t-1} \times D^2) \times \lvert L_{j,t-1} \rvert$				-0.112***
				(-42.0)
$(I_{m,j,t-1} \times D^3) \times \lvert L_{j,t-1} \rvert$				-0.106***
				(-22.4)
$(I_{m,j,t-1} \times D^4) \times \lvert L_{j,t-1} \rvert$				-0.069***
				(-2.8)
$L_{j,t-1}$				0.020***
				(19.1)
Agg		-0.0010***	-0.0011***	-0.0011***
		(-28.9)	(-31.2)	(-44.8)
$Time$		0.0004***	0.0005***	0.0005***
		(5.0)	(6.7)	(8.9)
$TSize$		-0.0031***	-0.0033***	-0.0032***
		(-8.2)	(-8.4)	(-11.3)

$I_{m,j,t-1} \times (D^2 - D^1)$	-0.005**	-0.007**	-0.003	-0.005**
$I_{m,j,t-1} \times (D^3 - D^1)$	-0.036***	-0.037***	-0.033***	-0.034***
$I_{m,j,t-1} \times (D^4 - D^1)$	-0.101***	-0.076***	-0.089***	-0.090***
Implied half-life (D^1)	32.2	26.7	26.1	26.5
Implied half-life (D^2)	31.5	26.0	25.9	26.0
Implied half-life (D^3)	27.8	23.3	23.1	23.3
Implied half-life (D^4)	22.0	20.5	19.3	19.3
SE clustering	Stock-date	Stock-date	Stock-date	Stock-date
N	1,576,111	414,032	450,795	1,576,111
R-squared	0.147	0.172	0.174	0.155

Table 7. HFT quote behavior

The regression tests in this table examine the relationship between HFT net order submission activity (Q) and HFT inventory levels in the presence of large institutional trades. D^k is an indicator variable representing inventory levels within a $(k-1)$- to k-standard-deviation band around the mean for $k \in \{1,2,3\}$. D^4 is an indicator variable representing inventory levels greater than a 3-standard-deviation band around the mean. L is a signed indicator variable that equals 1 (-1) if an institutional buy (sell) order is being executed in that period. Differences in coefficients are reported in the panel below the regression coefficients; for regression Column 4, the differences are calculated as $I_{m,j,t-1} \times (1+|L_{m,j,t-1}|) \times (D^k - D^1)$ for $k \in \{2,3,4\}$. Standard errors are double clustered by stock and date. t-statistics are reported in parentheses below the regression coefficients. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	HFT net order submission ($Q_{m,j,t}$)			
	(1)	(2)	(3)	(4)
$I_{m,j,t-1} \times D^1$	-0.096***	-0.094***	-0.094***	-0.078***
	(-41.9)	(-41.3)	(-41.0)	(-24.6)
$I_{m,j,t-1} \times D^2$	-0.102***	-0.100***	-0.100***	-0.082***
	(-54.6)	(-54.3)	(-54.2)	(-36.5)
$I_{m,j,t-1} \times D^3$	-0.112***	-0.110***	-0.109***	-0.085***
	(-40.3)	(-39.9)	(-39.9)	(-24.2)
$I_{m,j,t-1} \times D^4$	-0.124***	-0.123***	-0.122***	-0.093***
	(-14.5)	(-14.5)	(-14.5)	(-9.6)
$L_{m,j,t-1}$		0.040***	0.020***	0.019***
		(25.7)	(14.0)	(13.4)
$(I_{m,j,t-1} \times D^1) \times \lvert L_{m,j,t-1}\rvert$				-0.028***
				(-6.4)
$(I_{m,j,t-1} \times D^2) \times \lvert L_{m,j,t-1}\rvert$				-0.031***
				(-10.2)
$(I_{m,j,t-1} \times D^3) \times \lvert L_{m,j,t-1}\rvert$				-0.044***
				(-9.1)
$(I_{m,j,t-1} \times D^4) \times \lvert L_{m,j,t-1}\rvert$				-0.066***
				(-4.5)
$I_{m,j,t-1} \times (D^2 - D^1)$	-0.006**	-0.006**	-0.006**	-0.007**
$I_{m,j,t-1} \times (D^3 - D^1)$	-0.015***	-0.016***	-0.016***	-0.023***
$I_{m,j,t-1} \times (D^4 - D^1)$	-0.027***	-0.028***	-0.029***	-0.052***
SE clustering	Stock-date	Stock-date	Stock-date	Stock-date
Controls	No	No	Yes	Yes
N	1,576,111	1,576,111	1,576,111	1,576,111
R-squared	0.011	0.012	0.013	0.013

Table 8. HFT buying and selling activity

This table examines how HFT buying and selling activity are affected by their inventory levels and institutional trade activity. The dependent variables in Columns 1 and 2 are normalized HFT buy and sell trade activity, and the dependent variables in Columns 3 and 4 are normalized HFT buy and sell order activity. Standard errors are double clustered by stock and date. t-statistics are reported in parentheses below the regression coefficients. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Buy volume	Sell volume	Buy orders	Sell orders
	(1)	(2)	(3)	(4)
$I_{m,j,t-1} \times D_1$	-0.064***	0.074***	-0.035***	0.059***
	(-16.7)	(18.7)	(-8.0)	(14.3)
$I_{m,j,t-1} \times D_2$	-0.060***	0.081***	-0.035***	0.065***
	(-22.7)	(29.1)	(-11.1)	(20.6)
$I_{m,j,t-1} \times D_3$	-0.064***	0.092***	-0.041***	0.069***
	(-16.0)	(21.7)	(-8.1)	(13.6)
$I_{m,j,t-1} \times D_4$	-0.071***	0.115***	-0.052***	0.070***
	(-8.5)	(11.7)	(-4.7)	(6.3)
$\mathbf{1}_B$	0.061***	0.026**	0.048***	0.027**
	(5.2)	(2.3)	(3.7)	(2.2)
$\mathbf{1}_S$	0.022*	0.056***	0.017	0.036***
	(1.9)	(4.6)	(1.3)	(2.7)
Agg	-0.0007***	0.0003***	0.0003***	-0.0006***
	(-8.5)	(4.0)	(3.0)	(-6.7)
$Time$	0.0008***	0.0005**	-0.0002	0.0009***
	(3.0)	(1.9)	(-0.8)	(3.4)
$TSize$	-0.0011	0.0012	0.0019	-0.0012
	(-0.7)	(0.7)	(1.4)	(-0.9)
$DVol$	0.755***	0.752***	0.658***	0.662***
	(42.5)	(42.6)	(49.5)	(49.7)
$\mathbf{1}_B - \mathbf{1}_S$	0.039***	-0.030***	0.032***	-0.009
p-value	.000	.000	.000	.270
SE clustering	Stock-Date	Stock-Date	Stock-Date	Stock-Date
N	1,576,111	1,576,111	1,576,111	1,576,111
R-squared	0.184	0.184	0.117	0.118

Table 9. Large trade predictors and HFT activity

Columns 1 and 2 represent probit regressions of a large buy and large sell indicator variable, respectively, on publicly observable variables. r is the percentage price change; y is a normalized measure of the net number of shares purchased in that stock; and $LOIB$ is a normalized measure of the net number of shares submitted to the limit-order book. Column 3 is an OLS regression of $\Delta I_{m,j,t}$ on the large trading predictors (r, y, $LOIB$) and HFT inventory levels. Column 4 is an OLS regression of HFT inventory change on HFT inventory level and the predicted value of L. L is predicted using the publicly observable variables in Column 1. Column 5 is similar to Column 4, but uses aggressive institutional trades only for predicting L. Standard errors are double clustered by stock and date. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	Probit regression		OLS regression		
	Inst. buy	Inst. sell	$\Delta I_{m,j,t}$	$\Delta I_{m,j,t}$	Agg. trades
	(1)	(2)	(3)	(4)	(5)
r_{t-1}	0.057***	-0.052***	0.018***		
r_{t-2}	0.047***	-0.041***	0.006***		
r_{t-3}	0.039***	-0.034***	0.003***		
r_{t-4}	0.036***	-0.030***	0.002*		
y_{t-1}	0.004**	-0.001	0.012***		
y_{t-2}	0.006***	-0.004**	-0.001*		
y_{t-3}	0.008***	-0.006***	-0.004***		
y_{t-4}	0.008***	-0.007***	-0.004***		
$LOIB_{t-1}$	0.023***	-0.025***	-0.000		
$LOIB_{t-2}$	0.011***	-0.013***	-0.003***		
$LOIB_{t-3}$	0.007***	-0.008***	-0.002***		
$LOIB_{t-4}$	0.008***	-0.009***	-0.003***		
$I_{m,j,t-1} \times D_1$			-0.282***	-0.277***	-0.279***
$I_{m,j,t-1} \times D_2$			-0.288***	-0.284***	-0.286***
$I_{m,j,t-1} \times D_3$			-0.317***	-0.314***	-0.315***
$I_{m,j,t-1} \times D_4$			-0.379***	-0.378***	-0.379***
Pred. $L_{j,t}$				0.161***	0.338***
Clustering	Stock-date	Stock-date	Stock-date	Stock-date	Stock-date
N	1,490,416	1,490,416	1,490,416	1,490,416	1,490,416
R-squared	0.154	0.157	0.152	0.113	0.118