Do retail traders suffer from high frequency traders?[*]

Katya Malinova[†] Andreas Park[‡] Ryan Riordan[§]

October 3, 2013

Abstract

Using a change in regulatory fees in Canada in April 2012 that affected algorithmic quoting activities, we analyze the impact of high frequency quoting and trading on market quality, trader behavior, and trading costs and profits. Following the change, algorithmic message traffic, i.e. the number of orders, trades, and order cancellations, dropped by 30% and the bid-ask spread rose by 9%. Using trader-level data, we attribute this change to message-intensive algorithmic traders reducing their activity, and we show that their reduced activity had a negative impact on retail traders' intraday returns, in particular on their returns from limit orders. We further find that institutional traders' intraday returns from market orders increased.

[*]Financial support from the SSHRC (grant number 410101750) is gratefully acknowledged. We thank seminar participants at the 2013 WFA, the 2013 Central Bank Workshop on Microstructure (ECB), Erasmus University Rotterdam, the Free University of Amsterdam, and KU Leuven for comments. Special thanks to Mark Seasholes (WFA discussant), Bernt-Arne Odegaard (ECB discussant), Jonathan Brogaard, Terry Hendershott, Rob McMillan, David Panko, Elvira Sojli, and Mark Van Achter for detailed comments. The TMX Group kindly provided us with databases. The views expressed here are those of the authors and do not necessarily represent the views of the TMX Group.

[†]University of Toronto, katya.malinova@utoronto.ca

[‡]University of Toronto andreas.park@utoronto.ca (corresponding author)

[§]University of Ontario Institute of Technology ryan.riordan@uoit.ca

Although technological innovation has always played a critical role in financial markets, the unprecedented growth of automated, algorithmic trading in equity markets over the past decade has been the source of much controversy. Computer algorithms, capable of making and implementing trading decisions at speeds that are orders of magnitude faster than human reaction times, create, execute, modify, and cancel orders at microsecond speeds. To provide an example, in the U.S. during the Dotcom bull market in 2000, there were on average about 5 million trades and quotes per *day*; in the fall of 2012, at peak times there were up to 5 million trades and quotes per *second*.[1]

The initial growth of algorithmic trading was associated with a decline in trading costs, and it was viewed as a positive development by market participants and academics. For instance, using the introduction of automated quotes on the New York Stock Exchange in 2003, Hendershott, Jones, and Menkveld (2011) documented that an increase in algorithmic trading causally improved liquidity.

Yet market participants now frequently report that quotes change so frequently that the lower bid-ask spreads are only an illusion of liquidity and that quotes evaporate before traders are able to trade against them.[2] Moreover, processing millions of orders, cancellations, and trades is costly and requires that dealers, exchanges, and regulators heavily invest in IT infrastructure.

We analyze the empirical impact of intense quoting activity on market liquidity, on trader behavior, and on trading costs of market participants with different levels of sophistication. As a first step, we want to understand whether the decline in trading costs, documented by Hendershott, Jones, and Menkveld (2011), extends beyond the phase of initially modest use of automation and algorithmic trading. Second, equipped with trader-level data, we aim to understand who benefits from the change in trading costs. In

[1]See Larry Tabb's testimony to U.S. Congress, available at http://www.banking.senate.gov. On August 8, 2011, the U.S. credit rating was downgraded, the number of trades and quotes was almost 2.3 billion; see http://www.nanex.net/aqck2/3528.html.

[2]See The Economist, February 25, 2012: "The fast and the furious".

modern markets, traders use both limit and market orders, consequently, intense competition in quotes and a decline in the bid-ask spreads need not benefit traders who use limit orders. Indeed, Hendershott, Jones, and Menkveld (2011) find that the participation of human market makers declined after the introduction of automated quotes.

We are particularly interested in two groups of traders: unsophisticated retail traders and traders who build large positions and who are, presumably, sophisticated. These latter traders may employ sophisticated technology; in what follows, for brevity, we will refer to this latter group as "institutions."

The catalyst for our analysis is a unique policy change that made a subset of algorithmic trading strategies significantly more expensive. As of April 1, 2012, the Investment Industry Regulatory Organization of Canada (IIROC)[3] began charging a portion of its cost recovery fees based on the number of market messages (e.g., trades and well as order submissions, cancellations, modifications) that a broker-dealer generates.[4] Prior to April 1, 2012, IIROC recovered its fees based on dealers' trades only. Ex ante there was very little information about the amount of the per-message fee, but IIROC announced that around 15% of dealers would see an increase and that the remaining dealers would see a reduction in their fees. Following the introduction of the per-message fee, the total number of messages dropped by around 30%.

The change in the fee structure increased costs for traders who used many messages relative to their trades. Using February 2012, the pre-sample month, we identify traders that have both high message-to-trade ratios and that also use many messages; we refer to these traders as message-intensive. We observe that, subsequent to the change, traders in this subgroup reduce their activities significantly — not just in absolute terms but also

[3]IIROC is a self-regulatory organization that oversees dealers and trading activities and, in particular, performs real-time market surveillance.

[4]IIROC's official language refers to the fee schedule as the "integrated fee model"; see IIROC notice 12-0085; the monthly activity fees are divided into "Message Processing Fees" and "Trade Volume Fees" (where trade volume refers to the number of transactions); see http://www.iiroc.ca/Documents/2012/bf393b26-7bdf-49ff-a1fc-3904d1de3983_en.pdf

relative to the rest of the market.

We document that as message-intensive traders reduce their activities, bid-ask spreads increase significantly. Our regression analysis shows that for every 1% decrease in activity of message-intensive traders (relative to all other market participants), bid-ask spreads increase by .3 basis points. This finding is consistent with a theoretical prediction that market makers require higher compensation on their limit orders, the higher the risk of their order becoming stale (i.e. not reflecting the arrival of new information); see e.g., Copeland and Galai (1983), Foucault (1999) or Bernales (2013). If liquidity providers modify their orders less frequently in response to the per-message fee, the chance of their orders becoming stale may increase.

We then analyze the impact of message-intensive traders on the order submission behavior, trading costs, and intraday returns for the groups of retail and institutional traders. For instance, as high frequency quoting activities are reduced, it is imaginable that the lower competition for liquidity provision increases the probability of execution for limit orders and encourages other traders to trade with limit orders (as opposed to market orders). As limit orders do not pay but earn the bid-ask spread, traders who switch may see lower trading costs. First, we find that there are no changes in the probability that limit orders get executed. Second, institutions trade more with market orders, despite the increased bid-ask spread. Third, total (for market and limit orders) ex-post trading costs increase for institutions and (weakly) for retail traders.

Our data also allows for an analysis of intraday trading returns, computed as the intraday returns per dollar traded from buying and selling a security, with the end-of-day portfolio holdings evaluated at the closing price. A positive return implies that a trader was able to "buy low or sell high" relative to the closing price. Crucially, this measure accounts for adverse price movements that occur after a trade. Retail traders' returns fall post event, suggesting that high-frequency quoting activities and message-intensive strategies benefit retail traders! Decomposing these returns into returns from aggressive market and passive limit orders, we find that the reduction in returns stems from limit order trades — that is too say that retail traders who use limit orders were better off with a higher level of activity of message-intensive traders. Institutional traders' returns, on the other hand, are not affected, on the aggregate, as message-intensive algorithms reduce their activities, although their intraday returns to market orders increase.

Arguably, a subset of algorithmic traders that uses large numbers of messages are high-frequency traders (HFTs), in particular those that engage in market making activities. As Hagströmer and Norden (2013) describe, HFTs are a heterogenous group and they employ a variety of strategies. The term high frequency trading thus cannot describe a single type of strategy. One commonly agreed upon feature is that HFTs act in a proprietary capacity.[5] The group of traders that we identify as message-intensive may include a subset of proprietary algorithms, the group may also include so-called agency algorithms that execute customer orders as agents.

Irrespective of the nature of the algorithm, our classification captures traders that are in large part responsible for the externalities that result from the large number of messages (traders that constitute our message-intensive group account for more than 80% of the messages in our sample). Additionally, traders in this group are on the passive side of 48% of all transactions in our sample period. These traders are thus heavily involved in liquidity provision, an observation that is important when interpreting our results.

Our work contributes to three strands of the literature. First, we contribute to the literature on the behavior of individual types of traders such as retail and institutions. Second, we contribute to the rapidly expanding literature on algorithmic and, to some

[5]For instance, the SEC characterizes as "professional traders acting in a *proprietary capacity* that [among other things], engage in strategies that generate a large number of trades; use of extraordinarily high-speed and sophisticated computer programs for generating, routing, and executing orders; submit numerous orders that are cancelled shortly after submission."

extent, high frequency trading. Third, we contribute to the general literature on equity market quality by proposing to study the redistributive effects.

Barber and Odean (2000) show that active retail traders' portfolios underperform the market. Barber and Odean (2002) show that as investors switch to online brokerages, and trade more, their performance falls. Using a Taiwanese, investor-level dataset, Barber, Lee, Liu, and Odean (2009) find that retail traders lose on their aggressive trades. Complementing this literature, we study the impact of high frequency quoting on retail traders.

Our work also relates to the expanding literature on algorithmic and high frequency trading. Biais and Woolley (2011), Jones (2013), and Chordia, Goyal, Lehmann, and Saar (2013) survey this literature. Our work is closely related to Hendershott, Jones, and Menkveld (2011) who use the introduction of auto-quotes on NYSE in 2003 to show that algorithmic trading improves liquidity and makes quotes more informative. Similarly, Boehmer, Fong, and Wu (2012) use the introduction of co-location services on a number of international exchanges to study the impact of algorithmic trading. We contribute by analyzing trading costs and benefits to both limit and market orders. Subsequent to our study, Lepone and Sacco (2013) confirm our finding on the increase in the bid-ask spread for one of Canada's smaller venues, Chi-X, using a 19-month event window.

Hasbrouck and Saar (2011) study low-latency trading, document substantial short horizon activity in NASDAQ's limit order book, and find that low-latency trading and market quality are positively related. Brogaard, Hendershott, and Riordan (2012) show that HMAT aggressive trades permanently add information to prices. Hirschey (2011) shows that aggressive HMAT trades predict subsequent non-HMAT liquidity demand. Kirilenko, Kyle, Samadi, and Tuzun (2011) study HMAT in the E-mini S&P 500 futures market and argue that HMAT may have exacerbated volatility during the May 6th Flash Crash. Menkveld (2011) studies how one HMAT firm enabled a new market to gain market share and how this HMAT firm affected the price discovery process. Ye, Yao, and

Gai (2013) study technological advances in message processing on NASDAQ and finds that a reduction in latency from milliseconds to microseconds led to no improvement in market quality, suggesting that there are diminishing returns from technological improvements. Brogaard, Hagströmer, Norden, and Riordan (2013) find that the introduction of speed differentiation improves both bid-ask spreads and market depth. Our work adds to this literature by identifying how changes in high frequency quoting affect retail and institutional traders.

Jones (2013) describes a number of examples of trading venues that impose a form of message fee when traders exceed certain order-to-trade ratios; Colliard and Hoffmann (2013), Haferkorn and Zimmermann (2013), and Meyer and Wagener (2013) study the 2012 introduction of the French transaction tax. At first sight, a tax on financial transactions (FTT) has a similar flavor as a per-message fee. However, the per-message fee that we study is a new redistribution formula for existing fees and disproportionately affects the few traders that submit the bulk of messages. Additionally, the per-message tax is charged at the broker level and is, to the best of our knowledge, commonly not passed on to traders that do not generate a large numbers of trades or quotes. In contrast, an FTT is paid by all investors.

The rest of the paper is organized as follows. Section I. develops testable implications, both from first principles and from the theoretical literature. Section II. describes the data, the sample, and the details of IIROC's fees. Section III. outlines the trader classification. Section IV. outlines our empirical methodology. Section VI. presents our results on market quality, Section VIII. presents the results on trader level behavior. Section X. discusses the results. Tables and figures are at the end of the paper.

I. Theoretical Background and Testable Predictions

Our empirical strategy exploits the introduction of the fee per message that applies per order, per cancellation of an order, or a trade. Prior to the introduction of the per-message fee, IIROC recovered its surveillance costs only through charging fees for executed trades. After the change, the surveillance costs were recovered through fees on both trades and other messages (namely, submissions and cancellations of limit orders). The per-message fee thus cannot be interpreted as a tax on being fast — it increased the costs for strategies that involved numerous messages per trade, which we would refer to as "message-intensive", but it would have lowered the costs for traders who relied primarily on market orders.

Based on a comment letter to IIROC by Getco, a major electronic market maker worldwide[6], we conjecture that the introduction of a per-message fee increased the costs, in particular, to electronic market making strategies. The notion that market making strategies involve a large number of messages relative to trades is supported theoretically by Baruch and Glosten (2013). In an equilibrium of their model, liquidity providing traders modify their quotes each time they observe the state of a limit order book.

Bid-Ask Spread. In their comment letter, Getco argues that if market makers respond to a fee on message-intensive strategies by updating their quotes less frequently, then market makers would take on "additional risk during the time their quotations are placed on a market". They would thus require a higher risk compensation, and the bid-ask spread would widen (see e.g., Copeland and Galai (1983) or Foucault (1999) for the analysis of this effect). Getco's argument is further supported by Bernales (2013) who presents a model where traders differ in their speed and shows that bid-ask spreads are

[6]See http://docs.iiroc.ca/CommentsReceived.aspx?DocumentID=E5F5A707F5CF494ABB4993A42BFDEF44&LinkID=750&Language=en for a list of all the comments that IIROC received o their proposed fee change.

higher in the absence of traders who are able to update their quotations fast.

Empirical Prediction 1 *After the introduction of the per message fee,*

1. *traders that employ message-intensive strategies reduce their activities;*
2. *quoted bid-ask spread widens.*

Benefits to Liquidity Provision. As we discuss in the introduction, the impact of a wider spread on traders that aim to accumulate positions may depend on the type of orders that they use — they pay the bid-ask spread when trading with market orders, yet they receive it when using limit orders. Furthermore, as high-frequency market makers reduce their activities, competition for liquidity provision may decline, and profits to liquidity providers increase.

To access the benefit to liquidity providers empirically, one must account for the possible arrival of new information subsequent to the posting of a limit order — if the limit order submitter is adversely selected as a consequence, then the price would move against the limit order submitter and the bid-ask spread would overestimate the gain. To measure the gain to the limit order submitter, the bid-ask spread is commonly decomposed into the *price impact*, which is the adverse selection component in that it reflects the price movement against the limit order subsequent to the trade, and the so-called *realized spread*, which can be viewed as a profit to liquidity provision.

Empirical Prediction 2 *As the message-intensive high frequency traders reduced their activities, the realized spread increases.*

Adverse Selection Costs, Trader Behavior, and Intraday Returns. Predicting the impact of message-intensive activities on adverse selection and on trader behavior is challenging. The impact on the adverse selection costs, as measured by the price impact, is uncertain. On the one hand, as the high-frequency market makers reduce their activities, limit orders are less frequently modified in response to the arrival of new information, and they should see prices moving against them subsequent to trade more frequently. On the other hand, as competition for liquidity provision declines, the probability of execution of limit orders may increase.[7] If, as a consequence, informed traders use more limit orders in their strategies, the information content of market orders and their price impact may decline.

Existing theoretical models do not provide us with firm predictions on the impact of message-intensive activities of high frequency traders. Arguably these activities stem from rapid limit order submissions and cancellations; the theoretical literature, on the other hand, focusses on markets where fast, high-frequency traders submit either only market orders (see Martinez and Rosu (2011) and Biais, Foucault, and Moinas (2011)) or both market and limit orders (see Bernales (2013) and Hoffman (2013)).

The net effect of reduced liquidity-providing HMAT activities on traders' costs and benefits from trading with both market and limit orders depends on a number of factors, including the adverse selection costs and the level of trader sophistication. In our analysis, we will split traders into subgroups based on their levels of sophistication and then analyze if there are changes in the price impact, in traders' usage of market vs. limit orders, and in trading costs.

II. IIROC's Fee, Data and Sample Selection

A. *The Canadian Equity Market Structure*

During our sample period, Canada has six trading venues that operate as public limit order books, the Toronto Stock Exchange (TSX), Alpha Exchange, Chi-X, Pure, Omega, and TMX Select and two venues that operate as dark pools, Alpha IntraSpread and

[7] Additionally, as fast traders modify their limit orders less frequently, slower traders face a lower winner's curse and may post more aggressive limit orders (see Cespa and Foucault (2013)).

MatchNow.[8] In July 2012, Alpha, and Alpha IntraSpread became part of the TMX Group (which owned the TSX and TMX Select). Based on IIROC statistics, the market share of the TSX in the first half of 2012 was around 62% of all dollar-volume traded in Canada.

B. IIROC's Per-Message Fee

Investment Industry Regulatory Organization of Canada (IIROC) is a Canadian regulator that oversees investment dealers and trading activities, and it performs real-time market monitoring of all Canadian equity trading marketplaces. IIROC operates on a cost-recovery basis, and it recovers the cost through monthly fees. In a nutshell, IIROC charges trading venues who then pass on the fees to broker-dealers that submit orders and trades to these venues. In turn, brokers may either recover these fees through their commissions or pass them on to their clients on a message-by-message or trade-by-trade basis; it is our understanding that the fees are passed on to message-intensive clients, e.g. high-frequency traders, but only through commissions to retail clients. Before April 1, 2012, IIROC's fees were based on market shares of trading volume; after the change, fees were additionally based on the market share of messages, where a message is a trade, or an order submission, cancellation, or modification. Importantly, the introduction of the per-message fee was not an additional cost but a change in the cost-recovery formula; the total fees levied by IIROC remained constant. Since the per-message fee depends on the overall trading activity, its amount is unknown at the beginning of a month.

According to a research report by CIBC (2013), in 2012 the per-message fee was roughly \$.00022 per message (it fluctuates from month to month). For perspective: a typical HMAT message is a (non-executable) limit order for 100 shares. If executed against a market order, this order receives a maker rebate of, roughly, \$.30 from the

[8]The were some smaller venues that had negligible market share. The TSX Venture is technically a separate exchange that trades only TSXV-listed securities — which we do not include in our sample.

exchange. The per message fee, as known after the first month, is $0.00022 and it is thus much too small to impact the bid-ask spread directly. For traders such as retail who use, on average, about 2 messages per trade, it thus appears to be impractical to charge an additional fee per message (on average, dealers would have to charge additional 0.044 cents per trade).

C. Data

Our analysis is based on proprietary trader-level datasets provided to us by the TMX Group.[9] Data on shares outstanding (based on February 2012), splits, and index composition is from the monthly TSX e-Review publications. Data on the U.S. volatility index VIX is from the CBOE database in WRDS. IIROC's new, per-message fee became effective on April 1, 2012, and monthly charges were levied in early May 2012. Our sample period is March 1, 2012 through April 30, 2012, and we classify traders as message-intensive based on their pre-sample, February 2012 activities.

The TSX data is the output of the central trading engine, and it includes all messages from the (automated) message protocol between the brokers and the exchange. Messages include orders, cancellations and modifications, and trade reports. With the exception of traders' intraday returns, when computing market quality measures, we only include trades that happened in a limit order book during the continuous trading session.[10] Each trade is identified as buyer-initiated or seller-initiated, as the data specifies the active (liquidity demanding) and passive (liquidity supplying) party. Finally, the data contains information on the TSX and the Canadian Best Bid and Offer Prices.

[9] Legal disclaimer: TSX Inc. holds copyright to its data, all rights reserved. It is not to be reproduced or redistributed. TSX Inc. disclaims all representations and warranties with respect to this information, and shall not be liable to any person for any use of this information.

[10] We exclude the first 10 minutes and the last 20 minutes of the day to ensure that our results are not driven by the impact of market opening and closing auctions; the TSX starts disseminating information about the market-on-close auction 20 minutes prior to end of the trading day.

D. Sample Selection

We base our analysis on symbols from the S&P/TSX Composite index, Canada's broadest index and require that the companies remain in the index for the entire sample period. We exclude securities with stock splits, with major acquisitions, with fewer than 10 transactions per day, or that changed cross-listing status during the sample period. We delete Fairfax Financial Holdings (ticker: FFH) because of its high price (>$400; the next highest price is below $90). This leaves us with 248 companies in the final sample.

III. Classification of Traders

In Canada, traders must sent their orders to the exchange through a licensed broker. Brokers commonly organize their trading floors into different "desks" by the type of trader or investor that the desks caters to, for instance, retail, institutions, proprietary clients and so on. Consequently, each electronic message (e.g., an order or a trade) is associated with a unique identifier that belongs, for instance, to licensed individual at a broker's trading desk or to a so-called direct market access (DMA) client (an algorithm that accesses the exchange directly, possibly using co-located facilities).[11] Our data contains these unique identifiers. With the exception of retail traders, for whom we have proprietary information, we classify unique identifiers by their behavior.

Message-Intensive Traders. Using the pre-sample month of February, we base our classification of message-intensive traders on the total number of messages and on the message-to-trade ratios for each unique identifier. A message is defined as any system message that a trader sends to the exchange and that the exchange sends to a trader that relates to an order or trade (including order modifications, order cancellations, and

[11] The Canadian regulator IIROC requires that each direct market access (DMA) client has a unique ID. Consequently, messages from a DMA clients are not mixed with other order flow.

cancellations of immediate-or-cancel or fill-or-kill orders). Our goal is to find unique identifiers that use message-intensive strategies because these will have been negatively affected by the event. For each unique identifier, we compute the number of messages and the number of trades that this market participant submitted across the entire sample of TSX Composite securities plus the 42 most frequently traded ETFs in February 2012.[12] A unique identifier is classified as message-intensive if this identifier is both in the top 5% of message-to-trade ratios and in the top 5% of the total number of messages submitted. We exclude traders that use orders that stay in the order book overnight. We include exchange traded funds in the classification to capture multi-asset and multi-asset-class strategies that are message intensive, such as ETF arbitrage or ETF hedging.[13]

Consistent with our intuition from Section I., the identifiers that we classify as message-intensive are likely involved in liquidity provision. They trade over 77% of their volume with passive limit orders, and they are on the liquidity providing side of around 48% of all transactions in our sample period. While our classification may capture proprietary high frequency algorithms, it may also capture message-intensive agency algorithms that execute trading decisions on behalf of an institutional client. These traders are equally affected by the per-message fee and we are unable to differentiate between the impacts of proprietary and agency high quoting activities. See also Hagströmer and Norden (2013) for a discussion on diversity of high-frequency traders.

With the above caveat in mind, for brevity, in what follows, we frequently use high message algorithmic traders (HMATs) to denote the group of identifiers that we classified as message-intensive.

[12]Specifically, we chose those ETFs that had more than 1,000 trades in February 2012.

[13]We did not include ETFs in the trading cost analysis for a number of reasons. Most importantly, ETFs have designated market makers that maintain tight spreads — and it is possible that ETF providers have a contract with the designated market maker on the maximum size of the spread.

Retail Traders. We obtain information about identifiers that are retail traders from a proprietary dataset. This dataset is based on the trading activity in Alpha IntraSpread, a dark pool in which active orders can only be submitted by retail traders. We obtain all the known retail unique identifiers. While this approach does not classify all identifiers that submit orders on behalf of retail clients, the ones that are classified are indeed retail and combined they are involved in about 10% of the dollar volume. Each identifier is associated with a trading desk at a brokerage, which is typically responsible for retail flow from a large number of the broker's retail clients. It is important to know that for the vast majority of retail clients, the decision of where to send their order rests with the broker; therefore a particular identifier using the Alpha IntraSpread dark pool does not provide any information on the level of sophistication of this identifier's retail clients.

In the Internet Appendix we discuss an alternative classification that we employed; our results are robust. The sets of retail traders and high frequency traders do not overlap.

Institutional Traders. Our goal is to identify traders who handle large order volume and who build or unwind large client positions. For each unique identifier that is linked to a client (non-proprietary) account and that is neither retail nor message-intensive, we compute the per-stock cumulative dollar net position (buy dollar volume minus sell dollar volume) from February 1, 2012 to April 30, 2012. We then classify a unique identifier as an institutional trader for all securities and all days if for at least one stock on one day this identifier has an absolute cumulative net position that exceeds $25M. The $25M bound corresponds to selecting approximately the top 5% of identifiers with regards to their maximum net position.

This classification of institutional traders is aims to capture the traders that accumulate the largest positions with TSX trading. Our results relating to institutional traders should thus be interpreted as relating to traders that trade very large positions (not necessarily in each security). There are caveats to this classification. First, it is possible that a trader, for instance, buys on the TSX and sells on another venue and thus does not actually hold the attributed inventory. Second, it is possible that we capture a smart order router that is programmed to deal with, for instance, all "buy" trades.

Trader Group Summary Statistics. There are 3,516 unique identifiers in February 2012; we classify 94 of these as message-intensive (HMAT), 125 as retail, and 109 as institutions. Figure 3 plots the aggregate number of messages against the aggregate number of trades that each trader generated in February. Message-intensive traders have, by design, large message-to-trade ratios, institutions and retail traders have small order to trade ratios. In February 2012, the average message-intensive identifier submits 250,000 messages per day and is party to (roughly) 5,000 trades. Table III presents some summary statistics for these groups; the presented figures are based on by stock and day computations. The small number of traders that we classify as message-intensive (3.6% of all traders) create most of the messages, on average 82% for the sample period. Furthermore, we classify only around 9% of all unique identifiers, but these are involved in 53% of the dollar-volume (or, per day per stock, 48% of the dollar volume).[14]

IV. Estimation Methodology

The summary statistics in Table I indicate that message-intensive traders reduce their activities substantially both in absolute terms and relative to the rest of the market. The introduction of the fee thus had a substantial direct effect on the behavior of message-intensive traders. We estimate the effect of the reduction in message-intensive activities using two approaches.

[14]Note that volume here is double-counted because we count both the active and the passive side. Thus, for instance, if an HMAT would trade on the passive side in every transaction, then the HMAT share would be 50%.

First, we perform an event study in which we estimate the following relationship:

$$\textit{dependent variable}_{it} = \alpha_1 \, \textit{event}_t + \alpha_2 \, \textit{VIX}_t + \delta_i + \epsilon_{it}; \tag{1}$$

where $\textit{event}_t$ is a dummy that is 0 before April 1 2012, and 1 thereafter; δ_i are firm-level fixed effects; and $\textit{VIX}_t$ is the daily realization of the volatility index VIX.[15] The coefficient of interest is α_1 and it reflects the total effect of the fee change on the dependent variable for the month of April.

Our second estimation approach is to use the fee change event as a binary instrument for message-intensive activities and use it in a two-stage least square instrumental variable regression. We then simultaneously estimate

$$\begin{aligned} \textit{HMAT activity}_{it} &= \beta_1 \, \textit{event}_t + \beta_2 \, \textit{VIX}_t + \delta_i + \epsilon_{it}; \\ \textit{dependent variable}_{it} &= \beta_3 \textit{HMAT activity}_{it} + \beta_4 \, \textit{VIX}_t + \delta_i + \epsilon_{it}, \end{aligned} \tag{2}$$

where our main explanatory variable of interest, $\textit{HMAT activity}_{it}$, is instrumented by its estimated value from the first stage regression. As above, δ_i are firm fixed effects. We use two measures for HMAT activity. The first is the number messages from message-intensive traders as a percent of all messages; using this measure the estimate $\hat{\beta}_3$ describes how a 1% increase in relative HMAT activity affects the dependent variable. The second measure is the logarithm of all messages from the group of message-intensive identifiers. Then the interpretation of $\hat{\beta}_3$ is that it measures how a 1% absolute increase in the level of HMAT activity affects the dependent variable.

Canadian and U.S. volatility are highly correlated. Volatility is known to affect trading variables, and we include the U.S. VIX as a control because it is plausibly exogenous to Canadian market activities, yet captures market-wide volatility. To avoid biases in

[15] The presented regressions include firm fixed effects. In unreported regressions, we also analyzed a specification with a vector C_i of firm-level control variables, such as price and market capitalization. The results were similar.

standard errors stemming from observations that are correlated across time by security or across securities by time or both, we employ standard errors that are double-clustered by both security and date.[16] All regressions include stock fixed effects. To ensure that outliers do not drive our results, we winsorize all dependent variables at the 1% level.

The event study and the instrumental variable regressions relate in that the estimate for the event coefficient, $\hat{\alpha}_1$ from (1) should, on average be the same as the product of the estimates $\hat{\beta}_1 \times \hat{\beta}_3$ from (2). The interpretation of the IV regression is that it establishes a causal relation between HMAT activity and the dependent variable.

V. The Impact of the Fee Per Message on HMAT Activities

Table I shows that the number of HMAT messages falls by roughly 31% from March to April and that the HMAT fraction of all messages falls from roughly 84.4% to 79.5%. Figure‘2 plots the total number of messages in logs, the number of HMAT messages, and the fraction of messages that are created by HMAT identifiers. The level of messages and the percentage pertaining to HMAT identifiers before and after the fee change are significantly lower.[17] HMAT identifiers begin reducing their activities in the last days of March, which can be explained by traders "re-training" their algorithms ahead of the fee change. This early decline implies that we may underestimate the size of the true effects.

Table IV presents the results of the full sample first-stage regression. We include the Kleibergen and Paap (2006) Wald statistic of under-identification, which, in our specification, is $\chi^2(1)$ distributed, and the Kleibergen and Paap (2006) statistic for weak identification (following the Andrews and Stock (2005) critical values), and the Angrist-Pitschke

[16] Cameron, Gelbach, and Miller (2011) and Thompson (2011) developed the double-clustering approach independently at around the same time.

[17] We do not have data on comparable U.S. market activities at the time. However, the market research firm Nanex has a plot of total messages for U.S. markets on its website Nanex.net; see http://www.nanex.net/aqck2/3528.html. While the level of messages is lower in 2012 compared to preceding years, there is no notable decline in messages at the time of our event in April 2012.

F-test. Our results highlight that the event caused a significant decline in HMAT activity in the overall sample and that the event is a valid instrument for our IV approach. The estimated effect of the reduction in the fraction of HMAT messages, in the first column, is 1.6% (and thus lower than the aggregate reduction), the estimated reduction in the level of their activities, in the second column, is 29%. We confirm that after the fee was introduced, HMATs reduced their activities significantly.

In the estimation results of the second stage of the IV regression that we present in the following sections, a *negative* coefficient indicates that the decline in the percentage of HMAT activity led to an *increase* in the respective dependent variable. The coefficients on *event* thus have the opposite sign as the coefficients on % HMAT and log HMAT messages.

VI. The Impact of HMAT on Market Quality

Bid-Ask Spreads. We measure bid-ask spreads by the time weighted quoted half-spread based on the Canadian best bid and offer prices and by the volume-weighted effective half-spread; both are measured in basis points of the prevailing midpoint. For security i the effective half-spread for a trade at time t is defined as

$$espread_{it} = q_{it}(p_{it} - m_{it})/m_{it}, \tag{3}$$

where p_{it} is the transaction price, m_{it} is the midpoint of the quoted spread prevailing at the time of the trade, and q_{it} is an indicator variable, which equals 1 if the trade is buyer-initiated and -1 if the trade is seller-initiated. Our data includes identifiers for the active side (the market order that initiated the trade) and for the passive (the limit order) side of each transaction, precisely signing the trades as buyer- or seller-initiated. From our data we use the prevailing (Canadian) National best quotes at the time of each transaction.

Results. Figure 1 plots the time-weighted quoted spreads alongside the percent of HMAT messages. The figure indicates that as message-intensive traders reduce their activities, the bid-ask spread increases. Panel A in Table V presents our results from estimating (1) and (2). The results support Empirical Prediction 1 and confirm that after the introduction of the fee, the bid-ask spread increases because HMATs reduce their activities. Specifically, the decline in HMAT activity led to an increase in the half-spread by .5 basis points; for every 1% decline in relative HMAT activity, the spread increases by .3 bps, and a 10% total drop in HMAT activity leads to a .17 bps increase in quoted spreads. Similarly, in Panel B in Table V we estimate the effect of the change in HMAT behavior on the effective spread. As with the quoted spread, we observe that the reduction in HMAT behavior led to an increase in the effective spread, of the same magnitude as the change in the quoted spread. The result for the relation to the change in the fraction of HMAT trading is, however, only weakly significant.

Realized Spread. Taken at face value, the increase in the bid-ask spread makes the provision of liquidity more attractive and one would thus predict that, ignoring the per-message fee, benefits to liquidity providers increased subsequent to the introduction of the per message fee. A common measure for these benefits is the realized spread, defined as:

$$rspread_{it} = 2q_{it}(p_{it} - m_{i,t+5\text{ min}})/m_{it}, \tag{4}$$

where $m_{i,t+5\text{ min}}$ is the midpoint 5 minutes after the trade.

Results. Contrary to Empirical Prediction 2, Panel D of Table V shows that the realized spread *decreased* following the reduction in HMAT activity. Consequently, even though the quoted spread increases, liquidity providers receive a smaller portion of the spread.

Price Impact. The decline in the realized spread is driven by a change in the price impact, which reflects the adverse selection component of the spread and is defined as

$$\textit{price impact}_{it} = q_{it}(m_{t+5 \text{ min},i} - m_{it})/m_{it}. \tag{5}$$

The price impact and the effective spread are mechanically related in the sense that the difference of the two is the realized spread, interpreted as the revenue that liquidity providers receive in the transaction. Formally,

$$\textit{espread}_{it} = \textit{priceimpact}_{it} + \textit{rspread}_{it}. \tag{6}$$

Consequently, if the effective spread increases and the realized spread declines, the price impact of orders must have increased. In Panel C in Table V we estimate the effect of the change in HMAT behavior on the price impact. We observe that the reduction of HMAT activities led to an overall increase of the price impact of .8 bps.

The price impact further measures the extent to which a trade moves the midpoint of the bid-ask spread following the trade and it reflects the portion of the transaction costs that is due to the presence of informed liquidity demanders. An increase in the price impact indicates an increase in adverse selection for liquidity providers and typically indicates a worsening in market quality.

Our finding on the increase in the price impact is consistent with the idea that limit orders are more likely to become stale (not reflecting the most recent information), the less frequently they are modified (see, e.g., Bernales (2013) and Hoffman (2013) for the theoretical analysis). The fee per message led to a stark decrease in the message traffic, and in particular, to a stark decline in the limit order cancellations.

We conclude that spreads widen and adverse selection increases as message-intensive traders reduce their activity.

VII. Traders' Order Submission Behavior

We compute four measures to study order submission behavior, by trader groups: first, the fraction of volume that are traded with limit orders; second, the fraction of the submitted volume that are limit orders; third, the fraction of the submitted orders that are limit orders; and fourth, the fraction of limit order volume that is filled. The latter measure can be interpreted as the fill rate or the probability that a limit order executes.

Results. Table VI presents our results on tests of changes in the usage of limit orders. We find that institutions (a) trade more with market orders, submit more market orders relative to limit orders, and that (weakly) their limit orders get filled with lower probability. For retail traders there are no statistically significant differences.

These results highlight that there is heterogeneity in the reaction of traders to changes in HMAT behavior.

VIII. The Impact of HMAT on Intraday Returns

The results thus far indicate that, as message-intensive traders reduced their activities, market order submitters pay a larger spread, limit order submitters receive a smaller portion of the spread, and institutional investors submit more market orders. In this section, we study traders' intraday returns to access who benefits and loses from these changes.

Intraday Returns. Trading costs measured by bid-ask spreads are snapshots and these measures do not fully account for price movements subsequent to a trade. If prices include all information at any point in time, then any price movement subsequent to a trade is the result of new information (or noise). By holding the security, an investor then earns a return on his/her investment. On the other hand, if, for instance, an informed order is split into many small orders and the total information content of the order is only revealed over time, then anyone trading against the split order will lose. Uninformed traders must thus take into account that they may trade at the wrong time, before prices reflect all the available information. Informed investors, on the other hand, must take into account that they may move the price as they accumulate a position.[18]

To account for price movements subsequent to a trade, we compute the intraday returns, by trader group. We measure these returns by computing a trader's profit from buying and selling a security and we value the end-of-day portfolio holdings at the closing price; we then scale this profit measure by the daily dollar volume to obtain the per-dollar return. Formally, the per stock i, per day t profit for a group of traders is

$$intraday\ return_{it} = \left((sell\$vol_{it} - buy\$vol_{it}) + (buy\ vol_{it} - sell\ vol_{it}) \times cprice_{it}, \right) / \$vol_{it} \quad (7)$$

where $sell\$vol_{it}$ and $buy\$vol_{it}$ are the total sell and buy dollar-volumes for trader-group i, $buy\ vol_{it}$ and $sell\ vol_{it}$ are the share-volumes, $\$vol_{it} = sell\$vol_{it} + buy\$vol_{it}$ is the overall dollar volume. The profit from intraday trading is $(sell\$vol_{it} - buy\$vol_{it})$; a positive value means that the trader group "bought low and sold high." The term $(buy\ vol_{it} - sell\ vol_{it})$ is the end-of-day net position (assuming a zero inventory position at the beginning of each day), which we evaluate at the closing price, $cprice_{it}$.

Our approach uses the closing price as the benchmark and we thus implicitly assume that the closing price reflects the total information that was generated during a trading day. We compute three versions of the profit measure: one for all orders, one for all orders where a trader is on the passive, liquidity providing side, and one where the trader is on the active, liquidity taking side.[19]

[18] A common measure used by institutions to compute the costs of an order, in particular one that is split into many small orders, is the "implementation shortfall". This measure is, in essence, the volume weighted price of the order relative to the price that prevailed when the trader started to fill the order. Computing this measure with our data is impossible because we do not know when a trader started and completed filling an order and because our measures are computed for groups of traders. In an untabluated analysis, we employed the preceding day's closing price as the starting price to proxy for the shortfall. We found no significant effects of the fee change event.

[19] Barber, Lee, Liu, and Odean (2009) shows that in their dataset, retail traders lose mostly on their

Results. Panel A in Table VII displays the results from our estimation of the impact of HMAT activities on trader profits. The table shows that retail traders' profits decrease significantly whereas profits for institutional traders are unaffected. Split by active vs. passive trades, we observe that there is evidence that as message-intensive traders reduce their activities, retail traders lose more on their passive limit orders and that institutions gain more on their active marketable orders. The finding suggests that, when limit orders are modified less frequently, it becomes easier for institutions to pick off "stale" orders.

Combined with our earlier results on order submission behavior, we observe that institutions trade more with market orders and they derive higher profits from such orders.

IX. Trader-Level Trading Costs

Small bid-ask spreads are commonly considered to represent high market quality, because when spreads are tight, traders who use market orders face low transaction costs. In today's markets, however, traders use both market orders, for which they pay the spread, and limit orders, for which they receive the spread. Intraday returns that we computed in the previous section explicitly accounted for the usage of both order types. Here we study a related measure of transaction costs.

Namely, we compute the *cum fee total costs* per trader group as the difference of the realized spreads paid and the realized spreads received, weighted by the respective shares of active (market and marketable order) and passive (limit order) volumes and explicitly accounting for the exchange trading fees by adding the taker fees charged by the TSX for active trades to the realized spread and subtracting the maker rebates paid by the TSX for passive trades from the realized spreads..[20] We use the realized spread for active

aggressive orders. The profit measure that we compute is noisy, but we don't find major differences between profits for active and passive trades. As Table III shows, active vs. passive profits for retail traders for the entire sample are, -3.7 bps vs. -3.3 bps.

[20]On the TSX, the trading fees can differ by broker, where high-volume brokers receive the best conditions. We use the lowest taker fee, \$0.0033 per share, for our computations; the highest is \$0.0035.

trades because this measure captures the portion of the spread that compensates the liquidity provider and that would depend on the competitiveness of liquidity provision; the remaining portion is the trader's price impact, which s/he should be paying for.

Ceteris paribus, an increase in realized spreads paid and a decrease in realized spreads received both increase the cum fee total costs. However, if traders are able to trade more with limit orders relative to market orders, they can still reduce their cum fee total costs. This measure thus accounts for changes in trading behavior and the underlying costs.

Results. Table VIII displays our results on the cum fee total costs. We observe that net costs for institutions increase significantly, by .4 basis points. A reduction of HMAT activity relative to the market by 1% increases institutional traders' net costs by almost .3 bps. The summary statistics in Table II indicate that institution's realized spreads paid did not change much (-0.1bps; a formal regression analysis, not included in the paper, confirms this). Consequently, the increase in costs for institutions stems primarily from the increased use of market orders. For retail traders, the cum fee total costs weakly increase by about .3 bps, although, in the IV regressions, this effect is not significant.

In summary, the cum fee total total costs for institutional traders increase as message-intensive traders reduce their activities, and there is weak evidence that costs for retail traders increase, too.

X. Discussion and Conclusion

The introduction of the per-message fee in Canada was a unique event that increased the cost for some algorithmic trading strategies, including high frequency market making. Equipped with highly granular data, we use this event to study the impact of high-frequency quoting and trading activities on market quality as well as on trading costs and

For orders that execute against a dark order, the taker fee is \$0.001. We also use the highest possible rebate, \$0.0031 cents. Dark orders that clear against incoming marketable orders receive no rebate.

benefits for different groups of traders.

The event did have a noticeable impact: message-intensive traders, a group that likely includes high-frequency market makers, reduced their messages (trades, orders and cancellations) by almost 30%. The increase in the market-wide bid-ask spread right at the introduction of the fee is starkly visible in the data.

Our results on market-wide changes are generally consistent with findings in the literature, e.g. with Hendershott, Jones, and Menkveld (2011) who observe that increases in algorithmic trading on NYSE improved market quality and reduced adverse selection costs. Our main contribution is in addressing the impact on the costs and benefits of traders who use both market and limit orders. We found it intriguing that net trading costs for retail traders did not go up, even as the spreads increased, highlighting the importance of the order choice. Our findings on intraday returns suggest that high frequency quoting and trading does not affect all groups of traders in the same way but that it may lead to redistribution of gains from trade between retail and institutional traders.[21]

Even though Canada is a smaller market compared to the U.S., studying high frequency trading in Canada is instructive because many of the same high frequency firms are active in Canada (this information is part of the public record). Most of the proposed regulations on HFT include some sort of "tax" on HFT quoting activity, often based on the argument that the high level of HFT quoting activity imposes costs on other market participants because they must process the heavy message traffic. The per-message fee in Canada appears to have strongly affected the "good", liquidity-providing HFTs. The reduction in liquidity providing activities led to the deterioration of market quality, increased adverse selection, and significantly reduced the intraday returns of retail investors.

We emphasize that we cannot capture all the externalities that are associated with HFT activities. First, our classification focusses on message intensive strategies. These

[21] It is important to emphasize that institutions often manage funds on behalf of retail clients and thus a policy change that benefits institutions also benefits their retail clients.

strategies were, arguably, most affected by the per-message fee. There are, however, other HFT strategies that do not require many messages and our analysis has nothing to say about the impact of these strategies. Second, we have no data on the information processing costs of the message traffic, and trading costs and benefits must be weighted against these substantial technology costs. Finally, the introduction of the per-message fee led HFTs to reduce but not to eliminate their activity. If HFTs were to leave the market entirely, the effects may well be very different.

References

Andrews, Donald W. K., and James H. Stock, 2005, *Testing for Weak Instruments in Linear IV Regression* (Cambridge University Press).

Barber, Brad M., Yi-Tsung Lee, Yu-Jane Liu, and Terrance Odean, 2009, Just how much do individual investors lose by trading?, *Review of Financial Studies* 22, 609–632.

Barber, Brad M., and Terrance Odean, 2000, Trading is hazardous to your wealth: The common stock investment performance of individual investors, *The Journal of Finance* 55, 773–806.

———, 2002, Online investors: Do the slow die first?, *Review of Financial Studies* 15, 455–488.

Baruch, Shmuel, and Larry Glosten, 2013, Fleeting orders, working paper Columbia University.

Bernales, Alejandro, 2013, How fast can you trade? high frequency trading in dynamic limit order markets, working paper Banque de France.

Biais, B., T. Foucault, and S. Moinas, 2011, Equilibrium algorithmic trading, Working paper Toulouse University, IDEI.

Biais, Bruno, and Paul Woolley, 2011, High frequency trading, Working paper IDEI.

Boehmer, Ekkehart, Kingsley Fong, and Julie Wu, 2012, International evidence on algorithmic trading, Discussion paper, EDHEC.

Brogaard, Jonathan, Björn Hagströmer, Lars L. Norden, and Ryan Riordan, 2013, Trading fast and slow: Colocation and market quality, working paper University of Washington.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan, 2012, High frequency trading and price discovery, working paper UC Berkeley.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, 2011, Robust inference with multi-way clustering, *Journal of Business Economics and Statistics* forthcoming.

Cespa, Giovanni, and Thierry Foucault, 2013, Illiquidity contagion and liquidity crashes, working paper Cass Business School.

Chordia, Tarun, Amit Goyal, Bruce N. Lehmann, and Gideon Saar, 2013, High-frequency trading, *Journal of Financial Markets* pp. –.

Colliard, Jean-Edouard, and Peter Hoffmann, 2013, Sand in the chips: Evidence on taxing transactions in an electronic market, Working paper European Central Bank http://ssrn.com/abstract=2215788.

Copeland, T.E., and D. Galai, 1983, Information effects on the bid-ask spread, *Journal of Finance* 38, 1457–1469.

Foucault, Thierry, 1999, Order flow composition and trading costs in a dynamic limit order market1, *Journal of Financial Markets* 2, 99–134.

Haferkorn, Martin, and Kai Zimmermann, 2013, Securities transaction tax and market quality – the case of france, Working paper Goethe University Frankfurt SSRN: http://ssrn.com/abstract=2229221.

Hagströmer, Björn, and Lars L. Norden, 2013, The diversity of high-frequency traders, *Journal of Financial Markets*.

Hasbrouck, J., and G. Saar, 2011, Low latency trading, Working paper NYU Stern.

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

Hirschey, Nicolas, 2011, Do high-frequency traders anticipate buying and selling pressure?, Discussion paper, University of Texas at Austin.

Hoffman, Peter, 2013, A dynamic limit order market with fast and slow traders, working paper 1526 European Central Bank.

Jones, Charles, 2013, What do we know about high-frequency trading?, Research Paper No. 13-11 Columbia Business School http://ssrn.com/abstract=2236201.

Kirilenko, A., A. S. Kyle, M. Samadi, and T. Tuzun, 2011, The flash crash: The impact of high frequency trading on an electronic market, Working paper Uniervsity of Maryland.

Kleibergen, Frank, and Richard Paap, 2006, Generalized reduced rank tests using the singular value decomposition, *Journal of Econometrics* 133, 97–126.

Lepone, Andrew, and Alexander Sacco, 2013, The impact of message traffic regulatory restrictions on market quality: Evidence from Chi-X canada, CMCRC working paper University of Sydney.

Martinez, Victor Hugo, and Ionnid Rosu, 2011, High frequency traders, news and volatility, Discussion paper, HEC paris.

Menkveld, A., 2011, High frequency trading and the new-market makers, Working paper VU Amsterdam.

Meyer, Stephan, and Martin Wagener, 2013, Politically motivated taxes in financial markets: The case of the french financial transaction tax, Working paper Karlsruhe Institute of Technology http://ssrn.com/abstract=2211748.

Thompson, Samuel B., 2011, Simple formulas for standard errors that cluster by both firm and time, *Journal of Financial Economics* 99, 1–10.

Ye, Mao, Chen Yao, and Jiading Gai, 2013, The externalities of high frequency trading, Working paper University of Illinois at Urbana-Champaign http://ssrn.com/abstract=2066839.

Table I
Sample Summary Statistics – by Stock

The table reports summary statistics on our sample firms. In total there are 248 firms in our sample. Market capitalization is based on March 1, 2012. The percentage of messages by high messaging algorithmic traders (HMATs) are summed over the entire sample of securities, per day. All other figures are per stock per day averages. The price is the time-weighted mid-point of the national bid-ask spread. Intraday volatility is measured by the average daily 10-minute maximal mid-price fluctuation, scaled by the average midpoint. We also add the overall sample average for the S&P/TSX60 constituents.

	how computed	by	Units	Mean	SD	March	April	Difference	TSX60 mean
quoted spread	time-weighted	stock & day	bps	6.7	8.3	6.4	7.0	0.6	2.4
depth	time-weighted	stock & day	$10,000	4.0	18.7	3.9	4.2	0.4	5.2
effective spread	volume-weighted	stock & day	bps	6.4	8.3	6.2	6.6	0.5	2.3
realized spread	volume-weighted	stock & day	bps	−2.6	5.6	−2.3	−2.9	−0.6	−1.3
5-minute price impact	volume-weighted	stock & day	bps	9.0	10.6	8.5	9.6	1.1	3.5
messages per minute		stock & day		180.2	246.4	206.6	151.3	−55.3	448.8
dollar volume per message		stock & day	$	264.8	462.6	266.9	262.6	−4.3	309.8
dollar volume		stock & day	$ million	17.1	31.6	19.3	14.7	−4.5	49.0
trades per minute		stock & day		5.7	6.8	6.0	5.4	−0.6	13.1
intra-day volatility	10-minute midpoint range	stock & day	bps	28.2	16.5	27.5	28.9	1.4	25.6
price		stock & day	$	24.2	18.5	24.5	24.0	−0.5	38.2
market capitalization		stock	$ billion	6.7	11.6	6.7	6.7	0.0	19.8
trade size		stock & day	$1,000	6.0	8.7	6.2	5.7	−0.5	8.3
total messages		day	million	17.4	4.0	20.0	14.6	−5.3	10.5
% HMAT messages		day	%	82.1	2.9	84.4	79.5	−4.8	84.9

Table II
Sample Summary Statistics – by Trader Group (Part I)

The table reports summary statistics for our by-trader statistics. All figures are per stock per day averages for the respective groups, measured in basis points of the prevailing mid price. Cum-fee total cost is the volume-weighted average of the cum-fee realized spread paid (a trader's effective spread minus the price impact) minus the realized spreads received, after accounting for the exchange's maker-taker fees, as defined by equation (7).

	When	Who	Mean	SD	Median	March	April	Difference
effective spread	paid when active	retail	7.1	8.1	4.7	7.0	7.3	0.3
		institutions	6.0	7.6	3.7	5.7	6.2	0.5
		HMAT	5.3	7.1	3.2	5.1	5.6	0.5
	received when passive	retail	6.8	8.4	4.2	6.6	7.0	0.3
		institutions	5.8	7.7	3.5	5.6	6.1	0.5
		HMAT	7.0	8.0	4.7	6.8	7.3	0.4
realized spread	paid when active	retail	3.5	6.5	2.3	3.5	3.5	0.0
		institutions	2.0	5.5	1.1	2.0	1.9	−0.1
		HMAT	2.3	6.4	1.0	2.2	2.4	0.2
	received when passive	retail	−4.5	12.8	−2.8	−4.1	−4.9	−0.8
		institutions	−3.7	8.0	−2.2	−3.3	−4.1	−0.9
		HMAT	−1.8	4.7	−1.0	−1.5	−2.1	−0.6
price impact	caused when active	retail	3.6	6.0	2.1	3.5	3.7	0.3
		institutions	4.0	5.9	2.4	3.7	4.4	0.6
		HMAT	3.0	4.2	1.9	2.9	3.2	0.3
	received when passive	retail	5.6	7.6	3.7	5.3	6.0	0.6
		institutions	4.7	6.2	2.9	4.4	5.1	0.7
		HMAT	4.4	5.0	2.8	4.1	4.7	0.5
cum fee total cost		retail	4.3	7.3	2.9	4.1	4.5	0.4
		institutions	2.9	5.6	1.7	2.7	3.1	0.4
		HMAT	0.9	3.6	0.6	0.8	1.1	0.3

Table III
Sample Summary Statistics – by Trader Group (Part II)

The table reports summary statistics for our by-trader statistics. All figures are per stock per day averages for the respective groups. The percent dollar volume is the share of the dollar volume (of the total dollar volume per day per stock) that is traded by the respective group (volume is double-counted, i.e., a 100 share trade counts for 200 shares because we count both the active and the passive side); % passive volume traded is the fraction of the group's total (active plus passive) volume that a group trades with limit orders; % passive volume submitted is the limit order volume as a fraction of the group's total submitted order volume; % passive orders submitted is the number of limit orders as a fraction of the total number of orders submitted by the group; % passive volume filled is the fraction of the group's total submitted limit order volume that gets executed. Cum-fee total cost is the volume-weighted average of the cum-fee effective and realized spreads paid and received by the group, after accounting for the exchange's maker-taker fees, as defined by equation (7); intraday return is the group's daily profit as defined in equation (8), $profit_{it} = (sell\ \$\ vol_{it} - buy\ \$\ vol_{it}) + (buy\ vol_{it} - sell\ vol_{it}) \times closing\ price_{it}$ normalized by the group's daily dollar volume $sell\ \$\ vol_{it} + buy\ \$\ vol_{it}$; the intraday returns for market (limit) orders are defined similarly, except that only volume and dollar volume traded with market (limit) orders are used in computations (instead of the total volume/dollar volume).

	Who	Units	Mean	SD	Median	March	April	Difference
% dollar volume	retail	bps	10.4	8.5	8.0	10.8	10.0	−0.7
(of the daily total per stock)	institutions	bps	19.1	11.6	17.0	19.4	18.7	−0.7
	HMAT	bps	18.5	8.5	17.9	18.3	18.8	0.6
% passive volume traded	retail	bps	46.3	18.4	47.1	46.4	46.3	−0.1
(of the group's total traded)	institutions	bps	48.9	19.5	49.4	49.7	48.1	−1.6
	HMAT	bps	73.8	13.5	75.8	72.5	75.3	2.8
% passive volume submitted	retail	bps	73.2	14.0	75.3	72.9	73.5	0.5
(of the group's submitted)	institutions	bps	74.9	16.5	78.2	75.3	74.5	−0.8
	HMAT	bps	99.0	0.8	99.2	99.0	99.0	0.0
% passive orders submitted	retail	bps	53.6	18.7	54.2	53.9	53.3	−0.6
(of the group's submitted)	institutions	bps	79.7	15.2	84.0	80.3	79.2	−1.1
	HMAT	bps	98.8	1.1	99.1	98.8	98.8	0.0
% passive volume filled	retail	bps	33.3	20.8	30.3	33.9	32.7	−1.2
	institutions	bps	29.0	14.3	27.9	29.2	28.7	−0.6
	HMAT	bps	3.1	2.7	2.4	2.9	3.4	0.4
intraday return – all trades	retail	bps	−5.1	38.6	−1.5	−3.9	−6.4	−2.4
	institutions	bps	2.9	38.0	0.5	2.4	3.5	1.1
intraday return – market orders	retail	bps	−3.7	47.0	−2.7	−3.2	−4.3	−1.1
	institutions	bps	5.1	51.1	0.9	3.2	7.2	4.0
intraday return – limit orders	retail	bps	−3.3	60.9	2.5	−1.6	−5.2	−3.6
	institutions	bps	−0.8	56.7	0.0	−0.4	−1.2	−0.9

Table IV
Impact of the per-message Fee on HMAT Activity – First Stage

This table presents the results from the first stage regression on the impact of HMAT activity and it thus displays the impact of the IIROC message submission fee change on the percentage of messages generated by HMAT and the log of the total number of HMAT messages. The sample spans March and April 2012; the introduction of per-message fees occurred on April 1st. The explanatory variables are the percentage of total messages that are generated by HMATs and the natural logarithm of the total number of HMAT messages, per stock per day; the variable of interest is the event dummy, IIROC fee$_t$, that is 1 after April 1 and 0 before. Our first stage results are then based on estimating the following equation

$$\%\text{HMAT} = \alpha + \beta_1 \text{IIROC Fee}_t + \beta_2 \text{VIX}_t + \gamma_i + \epsilon_{it}$$

VIX$_t$ is the daily realization of the volatility index VIX, and δ_i are firm fixed effects. We include the F-test, the Kleibergen and Paap (2006) Wald statistic of under-identification, which, in our specification is $\chi^2(1)$ distributed, and the Kleibergen and Paap (2006) statistic for weak identification (following the Andrews and Stock (2005) critical values; for our specification, the 10% maximal IV size critical value is 16.38). Standard errors are double-clustered by firm and date. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	%HMAT	log HMAT messages
IIROC Fee$_t$	-1.61***	-0.29***
	(0.59)	(0.05)
VIX	-0.08	0.02*
	(0.13)	(0.01)
Observations	10,408	10,408
R-squared	0.013	0.071
firms	248	248
F-test	5.6	15.9
p-value F-test	0.0	0.0
under id	5.1	10.9
weak id	7.6	30.0

Table V
Impact of HMAT Activity on Quoted Liquidity

This table presents the results from our event study and from the second stage of our instrumental variable regression on the impact of HMAT on the daily time-weighted quoted bid-ask spread based on the national best prices, the effective spread, realized spread, and price impact. There are three explanatory variables of interest: the "plain" event effect (a dummy that is zero before April 1, 2012 and 1 thereafter), the percentage of total messages generated by HMAT (%HMAT), and the log of the number of HMAT messages. The latter two are estimated in a two-stage least square, and %HMAT and the log of the number of HMAT messages are instrumented with the event dummy, IIROC Fee_t. The sample spans March and April 2012; the introduction of per-message fees occurred on April 1st. The estimated equations are

$$(1)\ L_{it} = \alpha + \beta_1 \text{IIROC Fee}_t + \beta_2 \text{VIX}_t + \delta_i + \epsilon_{it},\quad (2)\ L_{it} = \alpha + \beta_1 \%\text{HMAT} + \beta_2 \text{VIX}_t + \delta_i + \epsilon_{it},\quad (3)\ L_{it} = \alpha + \beta_1 \log(\text{HMAT msg}) + \beta_2 \text{VIX}_t + \delta_i + \epsilon_{it}$$

where L_{it} is either the quoted spread or depth; VIX_t is the daily realization of the volatility index VIX; and δ_i are firm fixed effects. Standard errors are double-clustered by firm and time. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Panel A: Quoted Spread			*Panel B: Effective Spread*			*Panel C: Price Impact*			*Panel D: Realized Spread*		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
IIROC Fee_t	0.49***			0.35***			0.82***			-0.44***		
	(0.14)			(0.13)			(0.19)			(0.13)		
%HMAT		-0.30**			-0.22*			-0.51**			0.28**	
		(0.15)			(0.12)			(0.24)			(0.14)	
log HMAT messages			-1.69***			-1.21**			-2.82***			1.54***
			(0.54)			(0.51)			(0.87)			(0.57)
VIX	0.05***	0.03	0.09***	0.04**	0.03	0.07***	0.14***	0.10	0.21***	-0.09***	-0.07	-0.13***
	(0.02)	(0.05)	(0.02)	(0.02)	(0.04)	(0.02)	(0.04)	(0.10)	(0.05)	(0.03)	(0.06)	(0.04)
Method	OLS	2SLS	2SLS	OLS	2SLS	2SLS	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Obs.	10,408	10,408	10,408	10,408	10,408	10,408	10,408	10,408	10,408	10,408	10,408	10,408

Table VI
HMAT Activity and Retail and Institutions Trading Behavior

This table presents the results from our event study and from the second stage of our instrumental variable regression on the impact of HMAT on for activity variables: the fraction of the group's total volume (in %) that a group trades with limit orders (Panel A); the fraction of submitted order volume (not necessarily traded) that are limit orders (Panel B); the fraction of submitted orders that are limit orders (Panel C); and the fraction of limit order volume that gets executed (Panel D). There are three explanatory variables of interest: the "plain" event effect (a dummy that is zero before April 1, 2012 and 1 thereafter), the percentage of total messages generated by HMAT (%HMAT), and the log of the number of HMAT messages. The latter two are estimated in a two-stage least square, and %HMAT and the log of the number of HMAT messages are instrumented with the event dummy, IIROC Fee$_t$. The sample spans March and April 2012; the introduction of per-message fees occurred on April 1st. The estimated equations are as in Table V. Standard errors are double-clustered by firm and time. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Panel A: % volume traded with limit orders

	Retail traders			Institutional traders		
	(1)	(2)	(3)	(1)	(2)	(3)
IIROC fee$_t$	0.42			-1.80**		
	(0.62)			(0.71)		
%HMAT		-0.26			1.12*	
		(0.42)			(0.58)	
log HMAT messages			-1.44			6.26**
			(2.20)			(2.78)
VIX	-0.32**	-0.34*	-0.28*	0.11	0.20	-0.04
	(0.15)	(0.18)	(0.15)	(0.16)	(0.23)	(0.18)
Method	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Obs.	10,395	10,395	10,395	10,403	10,403	10,403

Panel B: % submitted volume that are limit orders

	Retail traders			Institutional traders		
	(1)	(2)	(3)	(1)	(2)	(3)
IIROC fee$_t$	0.92*			-0.53		
	(0.54)			(0.61)		
%HMAT		-0.57			0.33	
		(0.41)			(0.39)	
log HMAT messages			-3.20*			1.84
			(1.80)			(2.19)
VIX	-0.23*	-0.28*	-0.16	-0.13	-0.11	-0.18
	(0.13)	(0.17)	(0.10)	(0.15)	(0.17)	(0.15)
Method	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Obs.	10,406	10,406	10,406	10,406	10,406	10,406

Panel C:% of submitted orders that are limit orders

	Retail traders			Institutional traders		
	(1)	(2)	(3)	(1)	(2)	(3)
IIROC fee$_t$	-0.51			-1.30**		
	(0.50)			(0.63)		
%HMAT		0.32			0.81	
		(0.33)			(0.51)	
log HMAT messages			1.77			4.50*
			(1.77)			(2.41)
VIX	-0.08	-0.06	-0.13	0.11	0.18	0.00
	(0.11)	(0.12)	(0.10)	(0.14)	(0.21)	(0.15)
Method	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Obs.	10,407	10,407	10,407	10,407	10,407	10,407

Panel D: % of limit order volume that is filled

	Retail traders			Institutional traders		
	(1)	(2)	(3)	(1)	(2)	(3)
IIROC fee$_t$	-1.11			-1.04*		
	(0.78)			(0.56)		
%HMAT		0.69			0.65	
		(0.51)			(0.42)	
log HMAT messages			3.84			3.62*
			(2.48)			(1.95)
VIX	-0.05	0.00	-0.15	0.30*	0.35*	0.21
	(0.28)	(0.29)	(0.23)	(0.18)	(0.20)	(0.19)
Method	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Obs.	10,400	10,400	10,400	10,404	10,404	10,404

Table VII
HMAT Activity and Other Traders' Intraday Returns

This table presents the results from our event study and from the second stage of our instrumental variable regression on the impact of HMAT on the intra-day returns, measured by $\textit{intra-day return}_{it} = (\textit{sell \$ vol}_{it} - \textit{buy \$ vol}_{it}) + (\textit{buy vol}_{it} - \textit{sell vol}_{it}) \times \textit{closing price}_{it}$, scaled by the daily dollar volume. We compute the intraday returns for all trades (Panel A), trades with market orders (using volumes for trades with market orders only) (Panel B), and trades with limit orders (using volumes for trades with limit orders only) (Panel C). There are three explanatory variables of interest: the "plain" event effect (a dummy that is zero before April 1, 2012 and 1 thereafter), the percentage of total messages generated by HMAT (%HMAT), and the log of the number of HMAT messages. The latter two are estimated in a two-stage least square, and %HMAT and the log of the number of HMAT messages are instrumented with the event dummy, IIROC Fee_t. The sample spans March and April 2012; the introduction of per-message fees occurred on April 1st. The estimated equations are as in Table V. Standard errors are double-clustered by firm and time. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

Panel A: intraday return — all trades

	Retail traders			Institutional traders		
	(1)	(2)	(3)	(1)	(2)	(3)
IIROC Fee_t	-3.93**			1.36		
	(1.64)			(1.11)		
%HMAT		2.44**			-0.84	
		(1.17)			(0.78)	
log HMAT messages			13.61**			-4.69
			(5.57)			(3.94)
VIX	0.90*	1.10**	0.57	-0.15	-0.21	-0.03
	(0.47)	(0.53)	(0.41)	(0.28)	(0.35)	(0.25)
Method	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Obs.	10,406	10,406	10,406	10,407	10,407	10,407

Panel B: intraday return — market orders

	Retail traders			Institutional traders		
	(1)	(2)	(3)	(1)	(2)	(3)
IIROC Fee_t	-1.85			5.20***		
	(1.49)			(1.97)		
%HMAT		1.15			-3.19*	
		(0.89)			(1.75)	
log HMAT messages			6.40			-17.98**
			(5.04)			(7.73)
VIX	0.45	0.54	0.29	-0.69	-0.94	-0.26
	(0.44)	(0.45)	(0.36)	(0.48)	(0.73)	(0.50)
Method	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Obs.	10,375	10,375	10,375	10,374	10,374	10,374

Panel C: intraday return — limit orders

	Retail traders			Institutional traders		
	(1)	(2)	(3)	(1)	(2)	(3)
IIROC Fee$_t$	-5.85*			-1.83		
	(3.33)			(1.79)		
%HMAT		3.67			1.13	
		(2.33)			(1.22)	
log HMAT messages			20.36*			6.35
			(11.38)			(6.57)
VIX	1.36	1.66*	0.87	0.59	0.68	0.44
	(0.84)	(0.97)	(0.74)	(0.48)	(0.58)	(0.43)
Method	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Obs.	10,274	10,274	10,274	10,384	10,384	10,384

Table VIII
HMAT Activity and Other Traders' Cum Fee Total Costs

This table presents the results from our event study and from the second stage of our instrumental variable regression on the impact of HMAT on the cum fee total costs of trading, computed as the difference of the realized spread paid and the realized spread received, weighted by the % active and passive volume respectively, and accounting for the exchange's maker-taker fees (see equation (7) in the main text). There are three explanatory variables of interest: the "plain" event effect (a dummy that is zero before April 1, 2012 and 1 thereafter), the percentage of total messages generated by HMAT (%HMAT), and the log of the number of HMAT messages. The latter two are estimated in a two-stage least square, and %HMAT and the log of the number of HMAT messages are instrumented with the event dummy, IIROC Fee_t. The sample spans March and April 2012; the introduction of per-message fees occurred on April 1st. The estimated equations are as in Table V. Standard errors are double-clustered by firm and time. * indicates significance at the 10% level, ** at the 5% level, and *** at the 1% level.

	Retail traders			Institutional traders		
	(1)	(2)	(3)	(1)	(2)	(3)
IIROC Fee_t	0.32*			0.42***		
	(0.19)			(0.11)		
%HMAT		-0.20			-0.26**	
		(0.15)			(0.13)	
log HMAT messages			-1.11			-1.47***
			(0.72)			(0.44)
VIX	0.05	0.04	0.08*	0.00	-0.02	0.04
	(0.05)	(0.06)	(0.04)	(0.03)	(0.05)	(0.03)
Method	OLS	2SLS	2SLS	OLS	2SLS	2SLS
Obs.	10,395	10,395	10,395	10,403	10,403	10,403

Figure 1
Percent HFT of Message Traffic and Spreads

The figure plots the percent of messages that are generated by traders who we classify as HFTs for our sample of TSX Composite securities. The vertical lines mark the event date, April 1, 2012. The solid horizontal lines signify monthly averages.

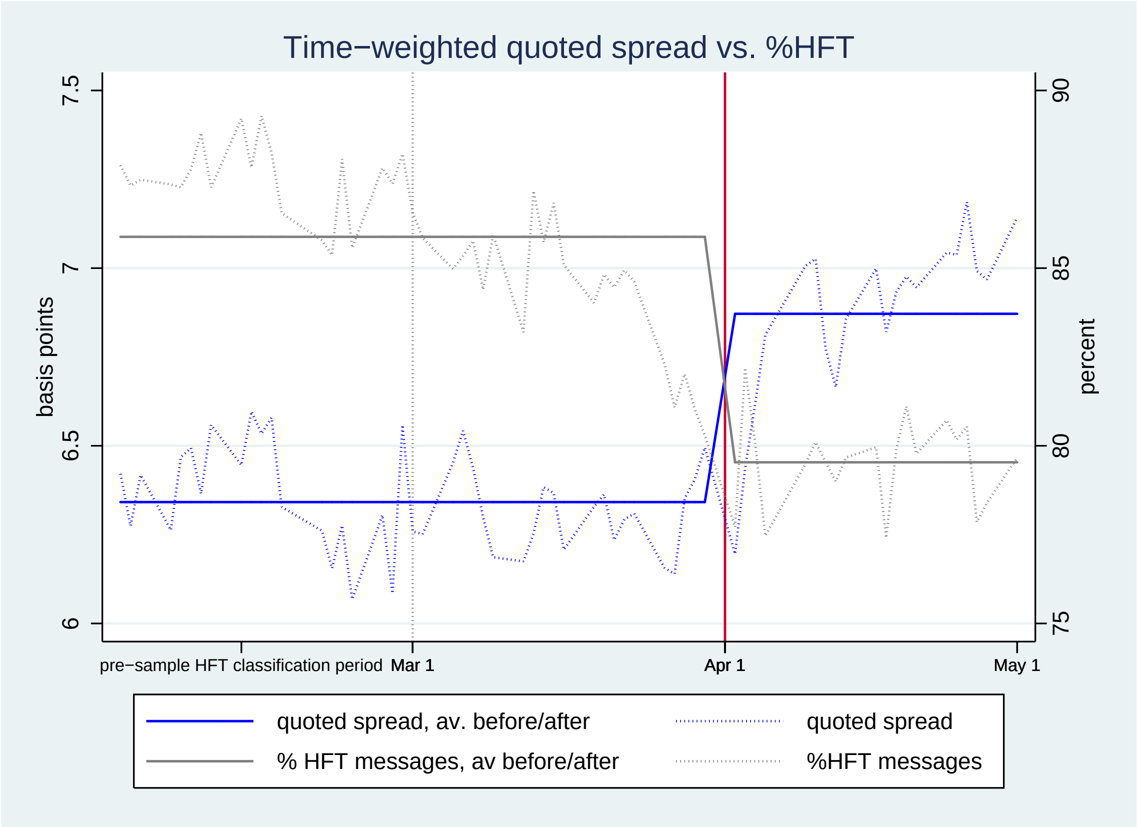


Figure 2
Message traffic on the Toronto Stock Exchange

The figure plots the log of the total number of messages, total HFT and HFT % of submitted on the TSX (not just for our sample). The vertical lines mark the event date, April 01, 2012. The solid horizontal lines signify monthly averages.

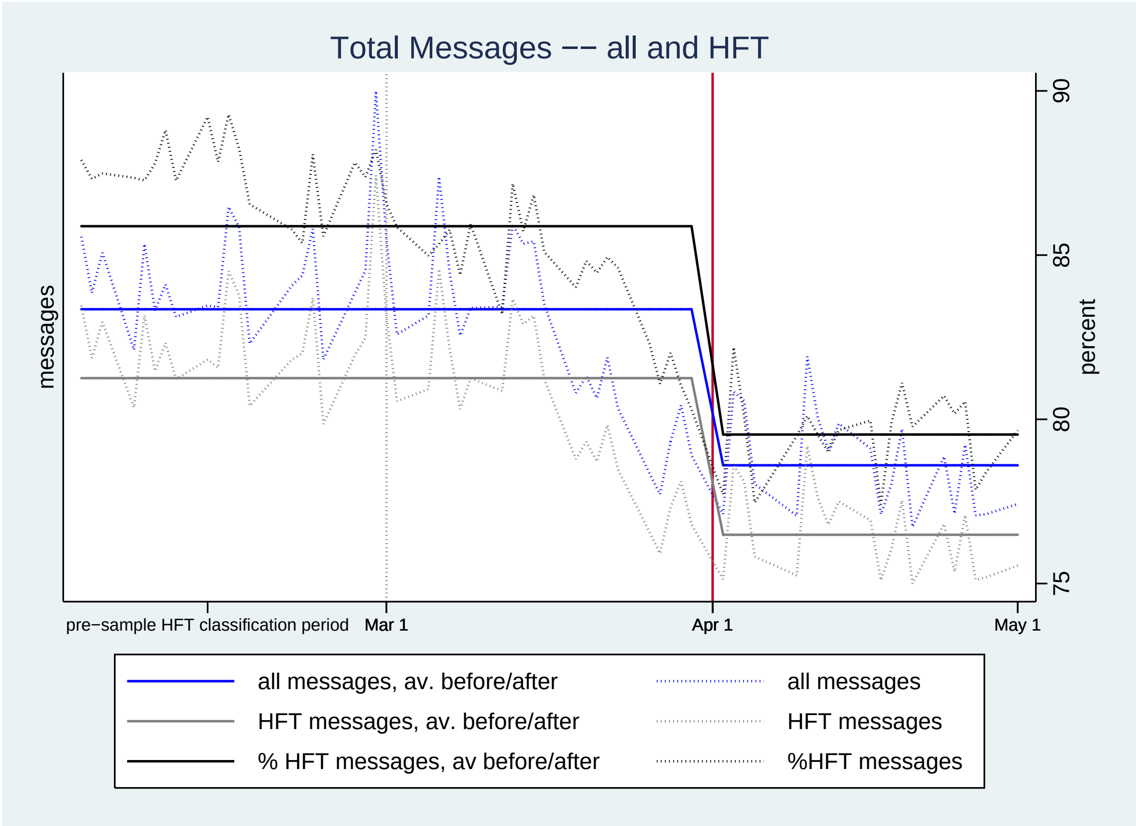


Figure 3
Trader Classification

The figure is a scatter plot of log trades against log total messages; each dot represents the total trades and messages of one specific trader in our data during the pre-sample HFT classification period of February, with grey dots, red squares, blue crosses and black dots indicating, respective, non-HFTs, HFTs, retail traders, and institutional trader. HFTs have, by construction, a large number of traders and messages, and high message-to-trade ratios.

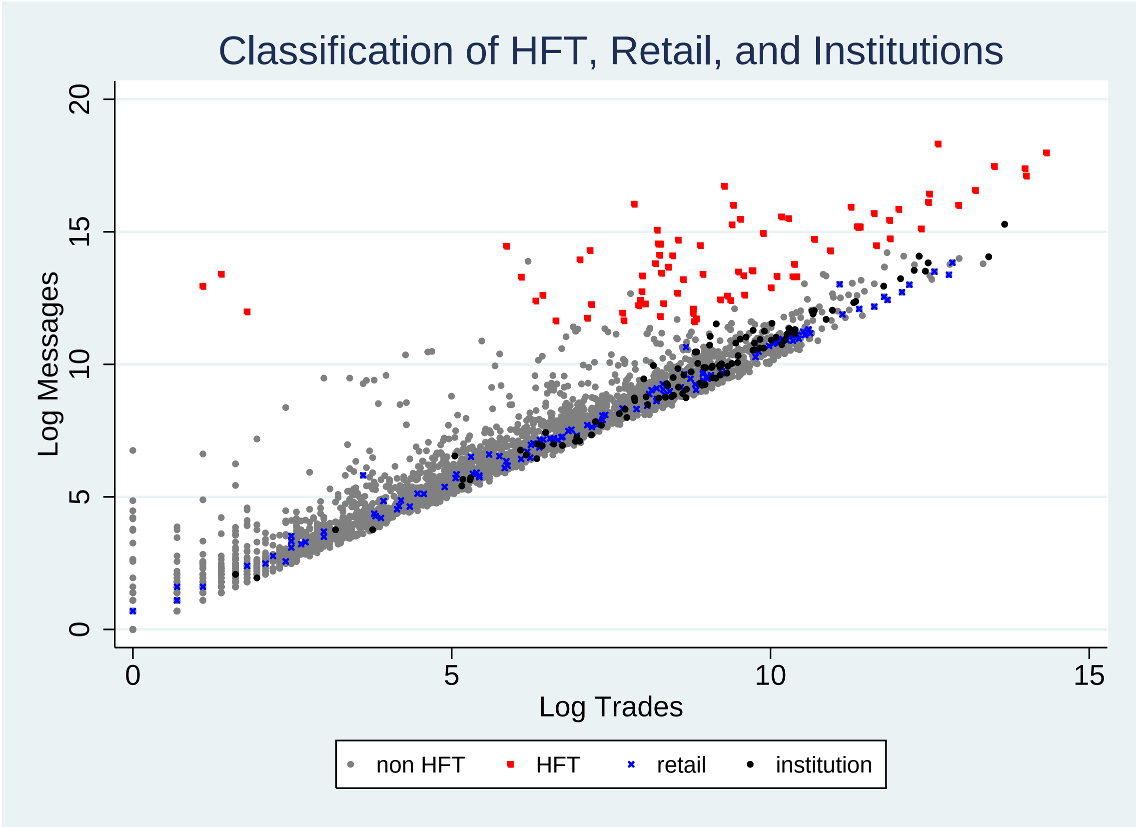
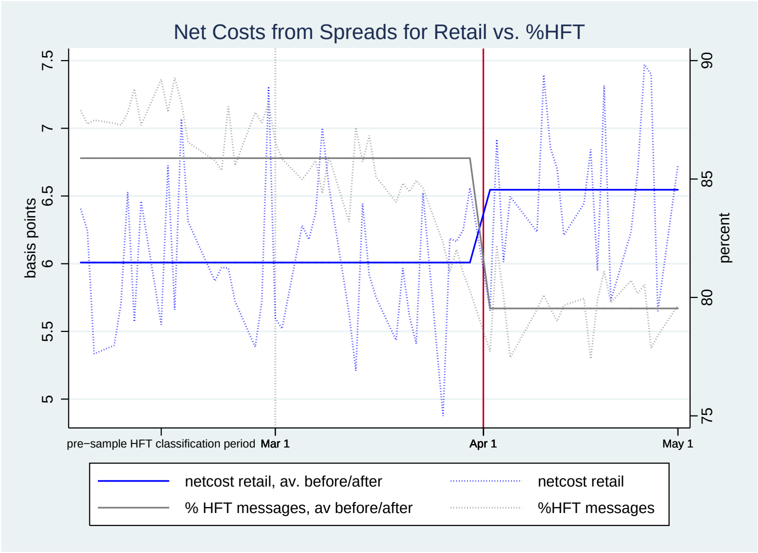


Figure 4
Trading Costs vs. HFT Market Participation

The left panel plots the percentage of messages generated by HFTs and the net costs from bid-ask spreads plus maker/taker fees of retail traders for the full sample. The right panel plots the percentage of messages generated by HFTs and the net costs from bid-ask spreads plus maker/taker fees for institutional traders, for the full sample. The vertical lines mark the event date, April 01, 2012. The solid horizontal lines signify monthly averages.



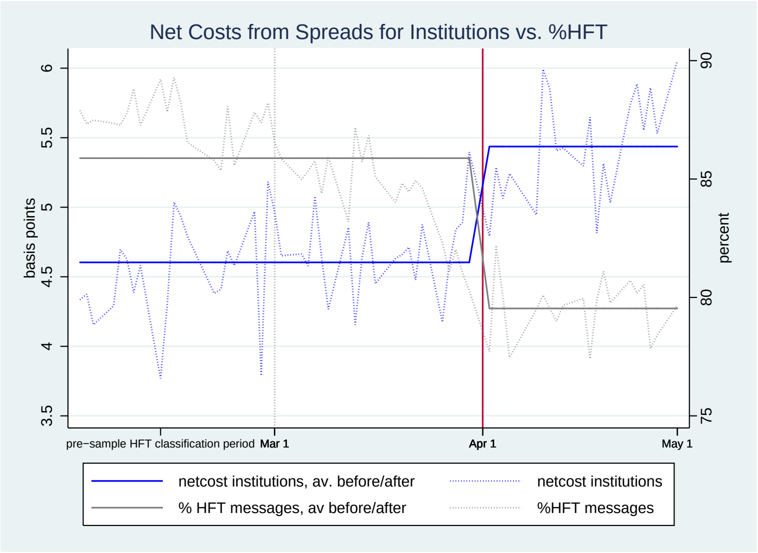


Figure 5
Trading Profits vs. HFT Market Participation

The left panel plots the percentage of messages generated by HFTs and the trading profits/losses from limit order traders for retail traders. The right panel plots the percentage of messages generated by HFTs and the trading profits/losses from market orders for institutional traders. The vertical lines mark the event date, April 01, 2012. The solid horizontal lines signify monthly averages.