Staff Working Paper/Document de travail du personnel 2018-8

High-Frequency Trading and Institutional Trading Costs

by Marie Chen and Corey Garriott

www.bank-banque-canada.ca

High-Frequency Trading and Institutional Trading Costs

by

Marie Chen[1] and Corey Garriott[2]

[1] London School of Economics
y.chen145@lse.ac.uk

[2] Financial Markets Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
cgarriott@bankofcanada.ca

Acknowledgements

Abstract

Using data on Canadian bond futures, we examine how high-frequency traders (HFTs) interact with institutions building large positions. In contrast to recent findings, we find HFTs in the data act as small-sized liquidity suppliers, and we reject the hypothesis that they engage in back running, a predatory trading strategy. Using a quasi-experiment in November 2011, in which a number of HFTs started trading the bond future, we run a difference-in-differences event study and find more competition among HFTs improves implementation shortfall, effective spreads, and short-term price impacts for institutional trading in Canadian bond futures.

Bank topics: Financial markets; Market structure and pricing; Financial system regulation and policies
JEL codes: G20, G14, L10

Résumé

Au moyen de données relatives au marché des contrats à terme sur obligations du Canada, nous examinons les interactions entre les opérateurs qui pratiquent la négociation à haute fréquence et les investisseurs institutionnels qui prennent d'importantes positions sur ces contrats. Contrairement aux conclusions d'études récentes, nos résultats indiquent que les opérateurs à haute fréquence jouent le rôle de petits fournisseurs de liquidité, et nous rejetons l'hypothèse selon laquelle ils auraient recours à la stratégie de prédation que constitue le parasitisme (back running). Nous appliquons une approche quasi expérimentale à des observations remontant au mois de novembre 2011, alors qu'un certain nombre d'opérateurs à haute fréquence faisaient leur entrée sur ce marché, et procédons à une étude événementielle fondée sur la méthode des doubles différences. Nos résultats montrent qu'une concurrence accrue de ces opérateurs réduit à la fois le décalage lié à l'exécution des transactions, les écarts de taux effectifs et l'incidence à court terme sur les prix pour les investisseurs institutionnels actifs sur ce marché.

Sujets : Marchés financiers, Structure de marché et fixation des prix, Réglementation et politiques relatives au système financier
Codes JEL : G20, G14, L10

I. Introduction

High-frequency trading (HFT) has expanded its reach, moving beyond its traditional base in equity markets and into markets for fixed income. HFT firms are now the largest participants on US trading platforms for bonds (Fleming et al. 2018) and are active on a variety of fixed-income derivatives on the CME. It is natural that HFT would seek new opportunities, as its trading profits from equity markets have tapered since 2009 (Tabb 2017). However, the fixed-income market is not like the equity market, and while HFT is well-suited for equity, it is not obvious how it can play a substantial role in fixed income. HFT firms are capital-light and hence prefer to trade in small sizes. That works well for trade with the small retail traders common in equity markets (Malinova et al. 2018). But, in fixed income, the retail presence is scant. Rather, most of the participants in fixed income are institutional investors, such as big banks or funds, that move positions sized in the millions of dollars. HFT firms lack the capital to be counterparty to these positions, so it is unclear what role HFT is playing in fixed income.

Since HFT's role is unclear, it is worth examining whether high-frequency traders (HFTs) behave the same way in fixed-income markets as they have been found to behave in equity markets. This can deepen our understanding of HFT's role in general, by showing how it adapts (or stays the same) in different circumstances. The literature on HFT in equity has identified a drawback to HFT, namely that it can use predatory trading to raise costs for institutional investors (Yang and Zhu 2018). Since fixed-income markets are distinguished by an even greater presence of institutional investors, we re-examine recent findings in a fixed-income context to see if the results intensify. Specifically, we take a second look at some of the findings in van Kervel and Menkveld (2019) and also Korajczyk and Murphy (2019), which find evidence pointing to HFT predation on institutions. In contrast to these papers, our data reject the hypothesis of predation. Instead, we find limited evidence of an improvement in institutional trading costs attributable to HFT.

In this paper, we ask whether HFTs in a fixed-income dataset can be found to prey on institutional trading using back running (Yang and Zhu 2018), the mechanism that has motivated recent empirical work. In addition, we also look for corroborating evidence of a change in liquidity metrics attributable to HFT. To study the research question, we sample order-book data on Government of Canada bond futures 2009–2015 from the Canadian derivatives exchange, the Montréal Exchange (MX). We identify HFTs empirically as participants that satisfy three standard criteria taken from the literature (Kirilenko et al. 2017; Weller 2017): a high switching rate, a zero overnight inventory four days out of five, and an average order lifetime of less than one minute. We also identify institutional trading empirically by finding sequences of trades that likely represent the objective of an agency to build a large position. Specifically, we locate "trade strings" by a participant: sequences of trades that are nearby in time, are all in the same direction, and are for a position size of at least CAD \$2 million.

The MX data are advantageous to study the research question. The sample is long and comes from a prominent fixed-income market: The front bond future is the most-traded single security in Canada by value and provides the leading price for price discovery on the Canadian yield curve (Campbell and Hendry 2008). Moreover, the data are advantageous particularly for studying HFTs because HFTs have no other electronic access to markets for Canadian fixed income. There are no other electronic derivatives markets in Canada; there is no electronic access to the underlying bond market in Canada; the CME in the US does not list any fixed-income derivatives on Canadian underlying; and there is no ETF focusing on the Canadian 10-year government sector or even the 5 to 10 sector. So, there is no possibility of cross-market netting or cash-and-carry arbitrage outside of the data sample, meaning the HFT positions we observe are likely to represent their entire exposure to the 10-year Canadian sector.

Using the data, we reject the hypothesis that HFTs prey on institutional investors. Across a cross-section of positions being built by institutions, we find that HFTs hold aggregate positions in the *opposite* direction of the positions—*i.e.*, they are short the future during

2

institutional buying, and they are long the future during selling. This is the opposite sign of the position predicted by theory on back running, in which a back runner strategically builds a position in the *same* direction as its prey. To the contrary, while we do observe HFTs trading large quantities with institutional traders, HFTs hold on aggregate the counterparty position. This is the behaviour of a theoretical (albeit short-term) liquidity provider and not a predator.

Importantly, the rejection of back running holds across three cross-sections of the positions built by institutions: by size, by duration (time to completion), and by long-term price impact. The result is important because, in the theory, the predator is assumed to prey on an investor who is fundamentally informed. So, strictly speaking, the theory only provides a prediction in the case of an informed investor.[1] To test specifically this case, we look at trade strings with large size and long duration, for which the price impact after five days is a statistically significant three basis points. Due to the long-lasting and large price impact (for fixed income), this cross-section of trade strings likely derives from informedness. We reject back running for this group of strings, and we also reject it for a group of strings selected for large *ex post* five-day price impact.

To provide some identification, and to corroborate the result, we exploit a quasi-experiment in late 2011. In November 2011, at least one HFT in the data exited the market and more HFTs entered the market, creating a sudden increase in HFT competition as measured by the Herfindahl-Hirschman index.[2] Using this plausibly exogenous variation in HFT competition, we fit a difference-in-differences (DiD) event study to estimate a treatment effect on three liquidity metrics: the implementation shortfalls of the trade strings; and the effective spreads and per-trade price impacts of the trades composing the trade strings. The control group for the study is the interest-rate future listed on the MX, which is a good control because

---

[1] Older theory on predation does not limit its prediction to the case of informed trading. Brunnermeier and Pedersen (2005), for example, applies to all cases in which an institutional trader has motivation to construct a position in a limited amount of time. The older theory also applies to our case and makes the same prediction about the sign of the predator's position. Still, to ensure the newer theory applies, we focus on informed trading.

[2] To protect the interests of its clients, MX asked that we not disclose the exact numbers of HFTs.

it has a market structure inimical to HFT, thus attracting almost no HFT volume, yet otherwise it shares risk factors and the clientele with bond futures. Similar to other work on HFT competition (Brogaard and Garriott 2019), but in contrast to findings on HFT and institutions, we find HFT competition improved implementation shortfall, effective spreads, and certain per-trade price-impact metrics in the bond future relative to the interest-rate future. This is consistent with HFT acting as an efficient short-term market intermediary and an enforcer of short-term price efficiency, as it has be found to act elsewhere (Brogaard et al. 2014; Menkveld 2013; Hendershott et al. 2011).

Taken as a whole, the results shed light on the role of HFT in fixed income. They are consistent with the interpretation of HFT as a short-term, low-inventory liquidity supplier. We find the HFTs in the data supply liquidity because, as mentioned, they hold the counterparty position to institutional trading in the data. However, this "good news" is quite limited. We also find that the inventory they bear is small in size and held only briefly, as they lay off the inventory almost immediately after acquiring it, almost always maintaining an inventory position of less than $100K—even when trading with trade strings for positions of $100M and more. Thus, HFT's benefit to the market is going to be limited, since it does not take on significant risk. The results from the event study confirm the interpretation, as the HFTs only improve shorter-term price impacts: The price is delayed from going where it will go, but it goes there eventually. Again, this is consistent with the activity of a market intermediary that agrees to take the counterparty position but lays it off quickly to another counterparty, possibly one with long-term risk-bearing capacity, such as a bank.

The results stand in contrast to recent work on HFT and institutions. In our view, this could be due to the setting of fixed income, but it also might be due to a different mix of advantages and disadvantages in our dataset. The obvious disadvantage of our data is the empirical identification of both HFT and institutional trading. The advantage of van Kervel and Menkveld (2019) and Korajczyk and Murphy (2019) is their *ex ante* knowledge of both HFT status and of institutional status. However, as a trade-off, we are able to view a

consolidated market rather than a fragmented market, as the HFTs we identify are unable to offset their exposures in other markets because there are no related markets for related assets, such as the large cross-border market in Korajczyk and Murphy (2019) or the other exchanges in van Kervel and Menkveld (2019). As van Kervel and Menkveld (2019) comment, "An important caveat of this study is that we do not observe the trades by HFTs on these alternative markets." HFTs could be hedging in one market by trading in another market or for another asset. The observed "with-the-flow" trading is quite possibly enabled by a hedge. Indeed, as Menkveld (2013) found, HFTs do more of their activity between markets than within them. Our data do not have the caveat, and we reach an opposite conclusion, suggesting the caveat may be a material impediment.

We contribute to the literature by providing more evidence on HFT and institutions but from a fixed-income market rather than an equity market. In a manuscipt, Tong (2014) found warning signs that HFT presence is correlated with institutional costs. Hirschey (2019) documents that HFTs apppear to have the ability to anticipate short-term order flows and trade ahead of them. Korajczyk and Murphy (2019) and van Kervel and Menkveld (2019) study datasets in which both HFTs and institutional traders are *ex ante* identified and find evidence that HFTs may be back running as described by Yang and Zhu (2018). Last, Garriott and Riordan (2019) study an empirically identified set of long-term investors and find they adversely select other traders, including a set of shorter-term traders.

II.   Data and summary statistics

The data come from Montréal Exchange (MX), Canada's derivatives exchange. The MX's major fixed-income products are the CGB contract, the future on the Government of Canada 10-year bond; and the BAX contracts, a set of futures on the Canadian Banker's Acceptance rate (the rate used for the CDOR index). The contract size for the CGB is par value CAD $100,000; for the BAX, par value CAD $1 million. We refer to the CGB contract as the bond future and to the BAX contracts as the rates futures.

The underlying Government of Canada bonds trade OTC via telephone and chat room and also via three interdealer brokers: Shorcan, Freedom, and Tullet Prebon. There is no direct electronic access to the brokers, and there are no derivatives on specifically Canadian bonds or rates listed on other exchanges, such as the CME, and there are no ETF products covering the 10-year Government of Canada sector (nor a "mid-range" sector such as a 5-10 ETF). For more on Government of Canada securities market structure, see Berger-Soucy et al. (2018).

For every trading day 2009–2014, the data contain one report on every quotation, update, fill and cancel for bond and rates futures. Each report contains fields such as price, quantity, millisecond timestamp, side (buy or sell), initiation (buyer or seller), and information sufficient to distinguish the counterparty as required. We drop early-hour and after-hour activity, defined as activity prior to the 8:20 a.m. market open and after the 4:00 p.m. market close. For bond futures, only one expiry has material volume, so we keep data on only the front contract. For rates futures, there is ample volume through the curve. To choose a good comparison contract, we keep whichever of the first three rates futures by expiry has the highest daily volume.

Figure 1 plots the trading volume of the bond futures and the first three rates futures by expiry. In the figure, bond-future volumes grow through the sample period, whereas volumes for rates futures are level after 2011. The two series begin to diverge after 2012, which is after many HFTs begin trading the bond future.

FIGURE 1 ABOUT HERE

Using the data, we compute daily series on: *Volatility*, the two-week rolling standard deviation of the daily close price log return, annualized; *Time to change*, the number of minutes until a change in the bid or ask price; *Order lifetime*, the number of minutes the average limit order rests; *Percent ticklocked*, the percent of the time the bid-ask spread is equal to the tick size; *Inside depth*, the sum of the quantities of visible limit orders at the bid

and ask prices; *Trades*, the number of trades; and *Trade size*, the size (in millions of dollars) of the trades.

A. *The rates futures are a good control for the bond futures*

Table 1 gives summary statistics on these metrics by year. These statistics are useful for understanding why HFTs are attracted to bond futures but not rates futures.

TABLE I ABOUT HERE

Compared to the bond future, the rates-futures market is a tick-locked market with minimal price volatility and enormous depth at the inside quotes. For example, in the first year of the sample, the rates future was tick-locked 90% of the time compared to 29% for the bond future. The rates future had a price volatility of 0.46, an order of magnitude below the volatility of 7.1 of the bond future and two orders of magnitude below the volatility of a typical stock on the TSX (around 40). The same year, the depth at the inside quote was around CAD $2 billion, *three* orders of magnitude above the depth of the bond future ($3.3 million) and *four* orders of magnitude above the depth of a typical TSX stock ($700K). The inside depth at the bid and ask of the rates future could be described as two walls of liquidity that, given the low volatility, hardly budge. Anecdotally, "if you place an order for the BAX on Monday, it might fill on Tuesday." Indeed, the average time to a price change for the rates future was more than 30 minutes, compared to around nine seconds for the bond future, and less than four seconds for the typical stock on the TSX.

As other studies have found, HFTs prefer lower-depth, higher-volatility markets (Cox et al. 2019). It is not profitable for HFTs to trade a contract such as the BAX because they are unable to exit a position quickly except by crossing the spread. (Again, for the rates future, HFTs must wait 30 minutes for a typical queue of limit orders at the same price to clear.) Rather, HFTs make profits from volatile markets with frequent price changes and ample opportunity to quote orders (Baron et al. 2019). Since the rates future is a market with poor structure for HFT, yet a market that otherwise shares interest-rate risks (exposure

7

to the level factor of the yield curve) and otherwise shares the same clientele (big Canadian banks and funds), we argue it is an effective control for liquidity trends in the bond future.

B.   The HFTs are identified empirically

To identify participants as HFTs, we draw from the literature three criteria that their behaviour must satisfy throughout the sample (Kirilenko et al. 2017; Weller 2017): criteria on overnight inventory, on switching, and on order lifetime. We take these criteria from work largely on equity markets. Therefore, we are assuming that HFTs in fixed income are not fundamentally different from those in equity in terms of their willingness to bear inventory, their lack of a preference between buying and selling, and their tendency to adjust their position on the order book at a high frequency.

The overnight inventory criterion requires the participant to carry zero inventory during four days out of five through the sample. This criterion is meant to ensure the participant is reluctant to take risk at term, a defining feature of HFTs. The switching criterion requires the participant to follow buy trades with sells or sell trades with buys at least a fourth of the time. This is meant to ensure the participant does not trade directionally on average, another defining feature of HFTs (versus, *e.g.*, algorithmic order execution). The order lifetime criterion requires the participant to leave its limit orders on the book for an average of one minute or less. This is meant to ensure the participant rapidly adjusts its position on the order book in response to new information, a final defining characteristic of HFT.

We cannot verify the usefulness of our criteria out of sample, so the conclusions based on these criteria are of unknown accuracy for other datasets. Other authors who lack exogenous identification and wish to "follow" our identification strategy should keep a few things in mind: to include caveats such as this; to find some plausible way to confirm the identification such as the one we will present in Figure 6; and, to show the same class of trader can be identified in multiple ways, which shows that the identified class is probably distinct. For example, our set of empirically identified HFTs is robust to the relaxation of any one of

the three criteria. If we were to remove the criterion on either the inventory, switching, or order lifetime, we would still identify the same participants as HFTs, which argues for a substantial distinctness.

Figure 2 plots the volume share of HFTs in the bond and rates futures.

FIGURE 2 ABOUT HERE

As the market-structure statistics in Table I would predict, HFTs have a significant presence in the bond future but lack presence in rates futures. For the bond future, HFTs' volume share increases from around 5% in 2009 to around 30% in 2014. In contrast, HFTs' volume share in rates futures is zero in the beginning and peaks at around 3% in 2012. HFTs' presence in the bond future may explain the continued growth in its volume relative to rates futures in Figure 1.

Table II gives summary statistics on the behaviour of participants split in two groups: HFT and nonHFT. The statistics are averaged over members of the two groups once per year of the sample. The statistics are the previously defined order lifetime, switching rate, overnight inventory, and number of trades, plus two new statistics: *Quotes*, the average daily number of new limit orders inserted on the order book; and *Cancels*, the average daily number of canceled limit orders (orders removed from the order book).

TABLE II ABOUT HERE

The HFT group exhibits markedly different characteristics from the nonHFT group. The order lifetimes of HFTs are two orders of magnitude less than the nonHFTs. For example, in 2012, the average lifetime was 45 seconds compared to 22 minutes. The switching rates are also larger, being twice to three times those of nonHFTs; 27% compared to 9% in 2012. And, the overnight inventories are two orders of magnitude greater than those of nonHFTs; in 2012, only 2 contracts compared to 541. These statistics are the ones used to select the groups, so perhaps it is not surprising that they differ so greatly. They could be said to differ by construction. However, as mentioned, the identification of HFTs is robust to the

9

removal of any one of the three criteria. So, for example, even if only the order lifetime and switching rate criteria were used to distinguish HFTs, the overnight inventories would differ in the exact same way.

Moreover, there are differences in the numbers of quotes, trades, and cancels. The statistic on quotes shows the largest difference: 7647 new orders a day for HFTs compared to 1930 for nonHFTs (in 2012), a difference of four times. Still, these three statistics do not differ as starkly between the groups as they do in equity markets. For example, in equities data, HFTs can trade upward of 10 times as often as the average nonHFT, whereas in this sample the HFTs trade more frequently but not, say, 10 times more frequently. In this it is helpful to remember that the median account in equities data would come from a small investment firm, whereas the median account in fixed-income data is a large institutional trader (since they are practically all large traders), and institutions are sophisticated and active on markets. HFTs in fixed income differ more by their low-risk, frequently modified trading strategy than by the raw number of trades.

C.   The institutional trading is also identified empirically

To study HFTs' activity as big banks or funds build large positions, we identify apparent instances of an agency building a large position empirically. Our strategy is to identify series of trades located nearby in time, that are unidirectional, and that amount to a "large" position. We define a "trade string," a series of at least four consecutive unidirectional (all buy or all sell) trades in which one participant builds a net position at least $2 million in size and with no more than a 20-minute gap between the participant's trades.[3] The gap allows us to distinguish trade strings that derive from potentially different trading intentions or perhaps from different parent orders.[4] Furthermore, we allow for attempts at misdirection by participants, often called "tape-colouring," by allowing trade strings that have 5% of their

---

[3]For example, a participant that purchases $25 million from 9:00 a.m. to 9:10 a.m. and another $25 million from 10:00 a.m. to 10:10 a.m. conducts two separate strings of $25 million and not one of $50 million.

[4]We have tried the analysis for gap sizes of 10 minutes and 30 minutes, and it does not change the results.

volume in violation of unidirectionality.[5] We consider the identification strategy conservative. There are many accounts that mix buy and sell orders from clients who may, individually, be trying to build large positions. Our identification strategy will discard these trades since they will appear to be bidirectional.

To examine cross-sections of the strings, we bin them by three criteria: by their size in volume; by time to completion in duration; and by "informedness" in permanent price impact. *Volume* is the total size of the position constructed by the end of the trade string. The volume bins are $2–$10 million, $10–$25 million, $25–$100 million, and $100+ million. *Duration* is two times the volume-weighted trade times (the measure is multiplied by two because raw duration is a half life). The duration bins are 0–10 min., 10–60 min., and 60+ min. *"Permanent" or long-term price impact* is the change in the midquote price observed some increment of time after the last trade in the trade string. There is one price-impact bin: price impacts of 3+ cents after five days.[6]

There are interesting correlations among the bins. As one would expect, the size and duration are correlated with informedness. In general, the larger the string and the longer it is worked, the greater its long-term (five-day) price impact. Figure 4 shows price impacts observed at many lags after the strings terminate. The price impacts are observed at: string termination, end-of-day, and five 24-hour increments after the string termination. The panels of Figure 4 give the impacts in four panels for four selected cross-sections of trade strings: (a) $2–25M; (b) those of size $25M–$100M and of duration at least 10 minutes; (c) those of size $100+ and of duration at least 60 minutes; and (d) those of long-term price impact of at least three basis points. These were chosen to give examples of uninformed, less-informed, and more-informed strings as well as high-impact strings. For the interested reader, in the Bank of Canada working paper version of this paper, we give the charts for all cross-sections of strings.

---

[5]We have tried the analysis for a 0% violation of unidirectionality, and it does not change the results.

[6]The long-term or five-day price impact is <u>not</u> the per-trade price impacts of individual trades in the trade strings. We also study the per-trade price impacts, but when we write long-term (five-day) price impacts, we mean the price movement after the entire trade string is finished, not after individual trades in the string.

For the \$2–\$25M strings, the immediate price impact is small, almost zero, and quickly vanishes into insignificance by the next day. For the \$25–100M strings lasting more than 10 minutes, the immediate price impact is moderate (one basis point) and stays significant for two days. Even after the impact becomes insignificant, it is only barely insignificant, and the estimated impact remains one basis point even five days out. Arguably, these strings are mildly informed, and are moderately useful for testing Yang and Zhu (2018). Last, for the \$100+ strings worked for longer than an hour, the immediate price impact is three basis points, and it is statistically significant at the maximum interval of five days. This is a large, permanent price impact for fixed income due to the size of the notionals. Arguably, these are among the most informed strings in the sample and are the most useful for testing Yang and Zhu (2018).

III.   HFTs do not back run on aggregate in the data

Before studying whether HFTs back run institutions that are building a large position, we first check to see if HFTs even engage with institutions while they build large positions. This is to check whether HFT is materially interacting with institutions in the spirit of Korajczyk and Murphy (2019). We compute the HFT *counterparty share* of trade strings as the percent of volume in the trade string transacted with any of the HFTs. In Figure 3, we report this counterparty share during four segments of a trade string: during the first quartile of volume, the second, the third, and the fourth. To see whether HFT's share varies by the size of the position an institution is building, we report the counterparty shares for our cross-sections of trade strings by volume: \$2–10M, \$10–25M, \$25–100M and \$100M+.

Institutions transact 26% to 35% of their volume with HFTs depending on the size of the position they are building and the amount of the position they have already built. This

confirms HFTs are interacting with institutions as they build large positions. HFTs are counterparty to 35% of the volume of trade strings \$2–\$10M, 30% of the \$10–25M, 27% of the \$25–100M, and 26% of the \$100M+. Clearly, HFTs trade less with larger strings, so HFTs are in some sense aware of when large flows are passing through the market and attempt to reduce their exposure to these flows. HFTs also reduce their participation during the life of a trade string, as is visible in the figure: HFTs reduce participation for each subsequent quarter of the four volume groupings of trade strings. Still, the largest withdrawal is just seven percentage points (out of 35) for the \$2–\$10 million bin. In contrast to Korajczyk and Murphy (2019), we find a smaller withdrawal overall, and we find the withdrawal is smallest in percentage points for the largest institutional positions.

Next, we move to ask whether HFTs in our data behave consistent with theory on back running. In theory on back running and on predation more generally (Yang and Zhu 2018; Brunnermeier and Pedersen 2005), a predator trader buys while a "victim" participant is buying (or sells when the "victim" is selling). The predator's intention is to build a sizeable position before the victim is finished trading. This impacts the price, but since the victim has a need to trade, the victim has no choice but to pay the impacted price. When the victim is finishing its trading, the predator moves to divest its position. The strategy is profitable because the predator divests at an impacted price. Drawing from the theory, we test whether HFTs are engaging in predation by testing whether they buy during buy trade strings and sell during sell trade strings and whether they reverse this behaviour near the end of a trade string.[7]

To perform the test, we first compute the cumulative order flow of three trading groups: the participant working a string; all HFTs in the market; and all others in the market. The *cumulative order flow* of a group is the cumulative sum of securities bought or sold by the group starting at the timestamp when the trade string began. We observe the cumulative order

---

[7]We do not test whether HFTs begin trading before the trade string as this would be a test for illegal front running, not back running, which presumes the HFT learns of a trade flow contemporaneous to the flow, as in Yang and Zhu (2018) or Brunnermeier and Pedersen (2005).

flows at ten moments during the lifetime of each trade string: at each one-tenth volume increment of the string (so at one-tenth complete by volume, at two-tenths complete by volume, and so on), and then we also observe it thirty minutes after the string finishes. To normalize sell trade strings with buy trade strings, all trades for a sell trade string are reversed in sign. The three order flows aggregate to zero since by definition there is a buyer for every seller.

Figure 5 displays the cumulative order flows in blue (the trade string), red (all HFTs) and green (all others), observed at the stated moments during a trade string, averaged for four selected cross-sections of trade strings: those of size \$2–25M; those of size \$25M–\$100M and of duration at least 10 minutes; those of size \$100+ and of duration at least 60 minutes; and those of long-term price impact of at least three basis points. These are the same as in Figure 4 and were chosen to give examples of uninformed, less-informed, and more-informed strings as well as high-impact strings. For the interested reader, in the Bank of Canada working paper version of this paper, we give the charts for all cross-sections of strings. The result is the same for all of them.

FIGURE 5 ABOUT HERE

HFT trading does not satisfy the theoretical prediction. HFT cumulative order flow is always in the opposite direction of the trade string, meaning HFTs, in net, are sellers to buy trade strings and buyers from sell trade strings. They take the counterparty position. This is consistent with a strategy of liquidity supply and not consistent with one of predation as in Yang and Zhu (2018) or in Brunnermeier and Pedersen (2005).

While HFT trading is more consistent with liquidity supply, HFTs do not bear a material quantity of a trade string as a counterparty. They do not take positions of much more than \$100K in notional of the trade string, even when trading with strings of size \$100 million and greater. This is a contrast between the results in this paper and van Kervel and Menkveld (2019). In van Kervel and Menkveld (2019), HFTs take risky, long-term positions that are borne for hours and sometimes for days, even in the placebo sample. One explanation

14

for the difference is that equity-market HFTs or perhaps European HFTs are distinguished from others by being capital-deep and bearing investment inventory, behaving like mutual funds and insurance companies. Another explanation is that, as the authors comment, the observed positions may not represent the true positions of the HFTs, who are offsetting the risk at unobserved alternative, foreign, ETF, or derivative markets. An advantage of the dataset we study in this paper is that we observe a consolidated market.

Importantly, there is no back running observed even for the "more-informed" group of strings, namely the ones sized \$100M+ and worked for durations longer than an hour. Similarly, there is no back running observed for the group of high-impact strings, which are also likely to be informed. In the recent theory (Yang and Zhu 2018), the back-runner trader is assumed to observe the trading of a fundamentally informed investor. Thus the theory only predicts predation to occur when the institutional trader is also an informed trader. We check this case in panels three and four of Figure 5. In panels three and four, the hypothesis is rejected the same way, as it is in every cross-section of the trade strings.

IV.   HFT leads to greater liquidity in a quasi-experiment

As Figure 2 showed, participants that behave like HFTs have been present on MX since at least 2009. This is confirmed by the "Approved Participant" list on the MX website, which lists the names of financial companies approved to directly access MX and clear derivatives. Several of them are well-known proprietary electronic trading firms.[8] Whenever a financial company joins or leaves the MX as an Approved Participant, the MX is required by regulation to issue a public release called a circular that gives the name of the company that joined or left the exchange community. Using these releases and four lists of "well-known" HFT companies given by The Wall Street Journal, Investopedia, Quora, and Medium, we construct a graph of the net entry or exit on MX of financial companies that the popular media identify as "HFT."

---

[8]See the MX Approved Participant list.

As the graph shows, 2011 saw the largest number of net new Authorized Participants that are popularly regarded to be HFTs (at least by sources such as The Wall Street Journal). It is common for members of an industry to move to a new market when profits from other markets taper, and Tabb (2017) finds US equity HFT profits tapering from 2009. 2011 could be a logical year for HFTs to move to nearby foreign and derivatives markets, such as the Canadian, given that it takes some time for HFTs to transfer operations across an international border. Therefore, we search for some event in 2011 to use as a plausibly exogenous shock to the industrial organization.

In Figure 7, we graph a common measure of competition in a market's industrial organization, the Herfindahl-Hirschman index, during the years 2011–2012. The HHI index is computed as the daily sum of the squared market shares of the HFTs, where market shares are the daily volumes of each HFT divided by the total volume over all HFTs.

In the graph, we observe what appears to be a shakeout event in the industrial organization. In November 2011, the industry becomes substantially less competitive than it was before, moving from a pre-November index average of around 0.35 to a sudden maximum of 0.53. Then, the index rapidly falls to 0.30 in January 2012 and continues to reach a new low of 0.19 later in 2012. One logical cause would be the exit of one or more incumbent HFTs and the entry of more entrant HFTs, which is likely given Figure 6. In summary, Figure 7 reveals a likely mechanism that could drive a change in liquidity: increased competition.

We take Figures 6 and 7 as motivation to perform an event study using November 2011 as a plausibly exogenous shock to HFT competition. To provide *prima facie* evidence of an impact on liquidity, we graph the number of empirically identified HFTs against a popular metric used to evaluate institutional trading performance, implementation shortfall (Harris 1998). The measure proxies the all-in price of buying or selling a security compared to a

"market" or "fair" price for the security. The figure graphs the implementation shortfall for the smallest group of trade strings, $2–10M, against the number of HFTs in inverted scale.[9]

FIGURE 8 ABOUT HERE

The inverted scale of the number of HFTs makes clear that there is a high correlation between the two metrics, 0.83. Both the number of HFTs and the implementation shortfall of trade strings sized $2–10M are shocked in November 2011, which is the same month that the Herfindahl-Hirschman index was shocked, and this occurs on the same year that the number of Authorized Participants that are popularly regarded as HFT increased by the largest increment in sample. In the chart, the implementation shortfall improved from a 2011 average of around 0.36 cents to a 2012 average of around 0.07 cents.

In preparation for an event study, Table III gives summary statistics on the trade strings. The statistics are given by year and by the four volume bins: *implementation shortfall*, as defined above; *number*, the raw number of strings in the category; *volume per string*, the par value in millions of the number of contracts purchased or sold in the string; *trades per string*, the number of trades in the string; *market share per string*, the volume of the string in percentage of daily volume (single counted); *average duration*, two times the volume-weighted execution time from start to finish of the string in minutes (multiplied by two since raw duration is a half-life); and *average aggressiveness*, the percentage of the volume in the string executed via marketable orders.

TABLE III ABOUT HERE

To make the event study a difference-in-differences, we locate an asset in the MX data to use as a control. An ideal control would not have HFT activity but would otherwise share risk factors and the clientele with the bond futures. We find a good control in the interest-rate futures listed on the same exchange. They differ from the bond future in having a

---

[9]At MX's request, to guard the business interests of its clients, we remove the y-axis label of the number of HFTs.

market structure inimical to HFT and thus in attracting almost no HFT volume. As Table I showed (see Section 2), the rates-futures market is a tick-locked market with minimal price volatility and enormous depth (billions of dollars) at the inside quotes. It takes an average of 30 minutes for the price to change—anecdotally, "if you place an order for the BAX on Monday, it might fill on Tuesday." It is not profitable for HFTs to trade a contract with such high depth and low volatility: If HFTs must wait 30 minutes for a typical queue of limit orders to clear, they cannot hope to exit a position quickly except by crossing the spread, which is expensive. Since the rates future is a market with poor structure for HFT, yet a market that otherwise shares interest-rate risks (exposure to the level factor of the yield curve) and otherwise shares the same clientele (big Canadian banks and funds), we argue it is an effective control for liquidity trends in the bond future.

Figure 9 shows implementation shortfalls for the bond future and the rates future in the two years around the November 2011 event. The figure shows the implementation shortfalls for four cross-sections of the trade strings for both contracts: those sized $2–10M, those sized $10–25M, those sized $25–100M, and those sized $100M+.

FIGURE 9 ABOUT HERE

Parallel trends appear to hold at least visually for the first three cross-sections and is particularly strong for trade strings sized $2–10M and $25–100M. To confirm this quantitatively, later in the paper we will fit a dynamic DiD model to test for pre-period trends. There is no clear treatment effect for any of the cross-sections other than the strings sized $2–10M. Whatever treatment effect exists is confined to the group of the smallest trade strings.

Indeed, in a difference-in-differences regression, the treatment effect is confined to trade strings sized $2–10M, as well as trade strings with duration 0–10 minutes. Table IV shows the results of a differences-in-differences event study on weekly implementation shortfall observed during the five months before and five months after the November 2011 treatment

date. The model fit is

$$IS_{i,t} = Treatment_{i,t} + BAX_i + Month_t + Controls_{i,t} + \epsilon_{i,t}, \qquad (1)$$

where i indexes the contract (bond or rates futures) and t indexes the week. *Treatment* is a dummy for observations of the bond future in the post-period; *BAX* is a dummy for observations of the rates futures; *Month* is a series of time fixed effects (one for each month). The difference in differences is therefore the variable *Treatment*. The control variables are: *Volume*, the trading volume in contracts; *Volatility*, the 20-day rolling daily return volatility; and *Lag*, the previous week's average level of implementation shortfall.

TABLE IV ABOUT HERE

While the treatment effect for the strings sized \$2–10M and for the strings with duration 0–10 minutes are both significant and economically meaningful—31 basis points and 21.5 basis points relative to the rates future—there is no significant treatment effect for the other cross-sections of trade strings by size or by duration. It would appear HFT improves implementation shortfall for only the smallest, shortest positions built by institutional traders.

A.   Why HFT improves implementation shortfall for only the smallest trade strings

We argue it is intuitive that HFT would improve implementation shortfall for only the smallest strings. HFTs are low-capital, low-risk traders that cannot hold in inventory a substantial portion of large institutional flows. Indeed, Figure 5 shows that the HFTs in sample rarely take a position during a trade string of greater than \$100K.[10] Thus, we see it as natural that HFTs would improve the transaction cost of only the smallest positions that institutions build, since HFTs lack the capacity to act as a consistent counterparty to larger positions.

---

[10]This is not "by construction" since the same HFTs can be identified by excluding the overnight inventory criterion and simply using their switching rate and order lifetime.

To confirm this intuition, we provide two pieces of evidence. First, we show HFTs are improving per-trade executions cost at a shorter term, via effective spreads, but not at a longer term, via per-trade price impacts at the five-minute lag. This helps explain why it is possible to discern an improvement in implementation shortfall for only the smaller positions that institutions build. Smaller positions are built quickly, often under 20 minutes (see Table III), so price impacts at a five-minute lag matter less to the implementation shortfall. However, large positions are built much more slowly and impact the price greatly, so they depend more on longer-term improvements in execution costs. For these positions, price impact or "slippage" is a more important determinant of implementation shortfall than the effective spread.

To show this, we run the difference-in-differences event study on four per-trade liquidity metrics computed for solely the trades in the trade strings (not for all the trades on the market): the per-trade *effective spread*, the signed difference between a trade price and the contemporaneous midquote; and three measures of *per-trade price impact*, the signed difference between the midquote contemporaneous to a trade and the midquote some time interval after the trade. The three measures are pre-trade price impact at five seconds, at 30 seconds, and at five minutes. We compute these metrics only for trades in trade strings to narrow the focus of the metrics to outcomes experienced by financial institutions when negotiating large positions. Table V shows the results.

<div align="center">TABLE V ABOUT HERE</div>

In the table, the quasi-experiment of November 2011 leads to statistical improvements in effective spreads, five-second price impacts, and 30-second price impacts: 13 basis points, 10 basis points and 13 basis points respectively. However, there is no significant improvement in the five-minute price impact, with a t-statistic of only 0.75. The plausibly exogenous variation in HFT competition leads to better effective spreads and shorter-term price impacts, as has been found elsewhere (Brogaard and Garriott 2019), but creates no discernible benefit for longer-term price impacts. Yet longer-term price impact is the liquidity statistic most

relevant to the largest institutional traders, for whom "slippage" or long-term price impact is more important than marginal effective spreads.

The second piece of evidence we give, to confirm the intuition that HFT is too reluctant to take risk to benefit the larger positions built by institutions, is to show the time structure of the treatment effect on the liquidity metrics. We show HFT withdraws its benefit from institutions as they work their orders, and it does so quickly. This means that HFT can be counted on to benefit only the orders that finish quickly, since it withdraws well before an institution finishes its largest orders to trade. Figure 11 shows the time structure of HFT's improvements to implementation shortfall, effective spreads, and per-trade price impact given in volume time.

FIGURE 11 ABOUT HERE

For implementation shortfall, the treatment effect is confined to the first $30 million of the position; and for effective spreads and price impacts, it is confined to the first *$10 million.* The rest of the position is unaffected by the treatment, confirming that HFT reduces its benefits once it runs against its risk and capital limits.

B. *Dynamic difference-in-differences*

Last, to show robustness and to verify parallel trends quantitatively, we fit a dynamic difference-in-differences model to the event data (see, *e.g.*, Callaway and Sant'Anna 2019). Dynamic difference-in-differences, sometimes called staggered difference-in-differences, is a form of the DiD model in which the treatment effect is estimated on multiple increments of the pre- and post-period data. This can be used to estimate the progressive response of a treatment over time and to verify that the data are not trending before the treatment date. If the treatment effect is statistically insignificant in all the pre-period increments, then the model corroborates the assertion of parallel trends. Also, if the treatment is statistically significant in all the post-period data, then the model confirms the treatment effect is associated with the claimed treatment date.

We fit the dynamic DiD model on the same event-study dataset, weekly averages of data observed the five months before and five months after the November 2011 treatment date:

$$metric_{i,t} = CGB_{i,t} \cdot Month_t + Month_t + \epsilon_{i,t}, \qquad (2)$$

where i indexes the contract (bond or rates futures), t indexes the week, $CGB_{i,t}$ is a dummy equalling one for the bond future, and $Month_t$ is a month fixed effect. The treatment effects are therefore $CGB_{i,t} \cdot Month_t$. Unlike classic DiD, there is one treatment effect estimated per month, even in the pre-period. The *metric* variables are the five liquidity metrics tested in this paper: implementation shortfall for strings sized \$2–10M, effective spread, and per-trade price impacts after five seconds, 30 seconds, and five minutes. Since the dynamic DiD is multicollinear with a constant, either a month must be omitted from the analysis or the data must be demeaned. We demean the data so to fit a treatment effect for every month.

Table VI gives the fit of the dynamic difference-in-differences for each liquidity statistic.

TABLE VI ABOUT HERE

Each column shows the set of 10 fit treatment effects, one for each month in the pre- and post-periods for each of the liquidity metrics studied in the paper. During the pre-period (before November 2011), the treatment effects are insignificant in 24 out of 25 cases in the months tested. The exception is a treatment effect for the five-minute price impact (which we argue is weakly affected by HFT if at all).

In the post-period, the treatment effects on implementation shortfall are statistically significant in all five months. However, for the per-trade liquidity metrics, only the effective spread has significant treatment effects in all five post-period months. The five-second price impact is significant in four out of five of the post-period months, but the 30-second price impact is significant in only three, and the five-minute price impact is significant in only two. The strength of the treatment effect for only the most immediate per-trade liquidity metric, the effective spread, is for us further illustration that the impact of HFT is discernible

22

moreso in the shorter-term per-trade liquidity metrics, particularly the immediate metric of the effective spread, and is less discernible in the longer-term metrics, particularly the five-minute price impact.

The fit of a dynamic DiD model is often presented graphically so that the eye can judge whether there is a pre-treatment trend in the data. Figure 12 gives a graphical presentation in which the treatment effect for each month is represented in the graph by a dot with bars for 95% significance. For conciseness, we show only the implementation shortfall, effective spread, and five-second price impact.

FIGURE 12 ABOUT HERE

V. Conclusions

In this paper, we examine recent findings about HFT and institutional trading in the context of a fixed-income market rather than an equity market. Using data from Canadian bond futures, we ask whether empirically identified HFTs use back running, a predatory trading strategy (Yang and Zhu 2018), on empirically identified examples of institutions building large positions. In contrast to recent findings (Korajczyk and Murphy 2019; van Kervel and Menkveld 2019), we reject the hypothesis that HFTs are back running in this dataset. HFTs in Canadain bond futures trade the opposite way a back runner would, which is more consistent with a strategy of liquidity supply or short-term price arbitrage. Moreover, exploiting a quasi-experiment in which many HFTs entered the bond futures market, we find increased HFT competition benefits institutions rather than harms them, at least by measures of implementation shortfall, effective spread and short-term price impact.

The results in our study might differ from recent findings for two reasons. First, it could be the different market structure of fixed income versus that of equity. For example, the bond futures market may be more liquid than markets for individual stocks, meaning that HFTs are simply less able to detect institutional trading (though a greater liquidity would

also raise the reward to successful back running, since a back runner in a liquid market can build a larger position and exit at less cost). Another possibility is that our data have a better mix of advantages and disadvantages for the research question. Although the chief disadvantage is the empirical identification of HFTs and institutions, a compensating advantage is the consolidation of the Canadian bond futures market and a lack of alternative markets or assets that HFTs could use to offset their exposures to bond futures. In other studies, HFTs could be hedging their exposures across markets or assets, which is likely as cross-market trading is a primary activity of HFTs (Menkveld 2013). If so, the empiricist cannot accurately observe the HFTs' true exposure, which is likely to be much less than a partial exposure observed in only one piece of a larger market. With our data, we close off this concern, and we find the opposite result.

From our results, we draw the moral that HFT can play some positive role even in a largely institutional market. Although HFTs are unwilling to take on much risk, their traditional advantages in automation and speed (which decreases adverse selection) seem to enable them here, as elsewhere (Hendershott et al. 2011; Brogaard et al. 2014), to decrease transaction costs at least on the margin. HFT competition appears to be an effective mechanism for enhancing this effect on transaction costs, as it has been found to be elsewhere (Brogaard and Garriott 2019). While we do not find HFT decreases transaction costs for the larger positions that institutions build, this is natural given its low tolerance for inventory and for risk. Thus, while our result is "good news" for HFT, it is also news of a good that ultimately is limited, as HFT does not provide any substitute for the risk-bearing capacity that natural market participants currently obtain from banks.

Figure 1: Trading volume of bond futures and of the front three rates futures

This figure shows average monthly trading volume (in number of contracts) for the CGB (the future on the Government of Canada 10-year bond) and for the front three contracts of the BAX (the future on the Canadian Banker's Acceptance rate) by expiry. The left y-axis is for the CGB, and the right y-axis is for the BAX. *Trading volume* is the total electronic number of contracts traded during normal trading hours.
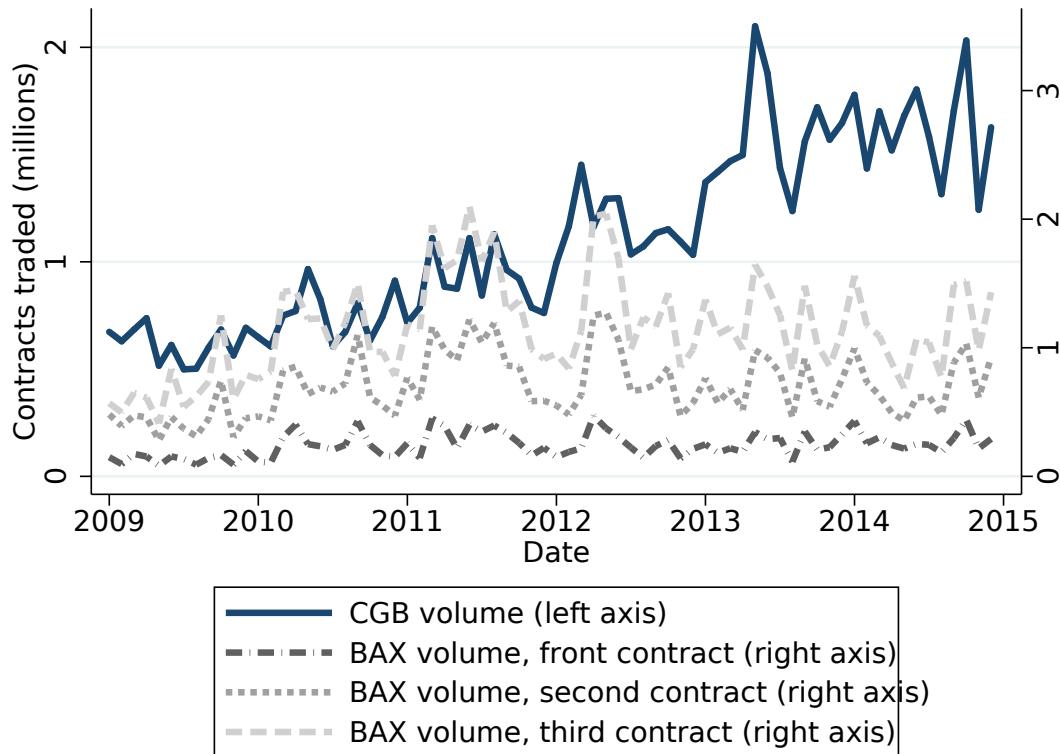
Figure 2: Volume share of high-frequency traders in bond and rates futures

This figure shows average monthly percent HFT volume share for the CGB (the future on the Government of Canada 10-year bond) and for the front three contracts of the BAX (the future on the Canadian Banker's Acceptance rate).
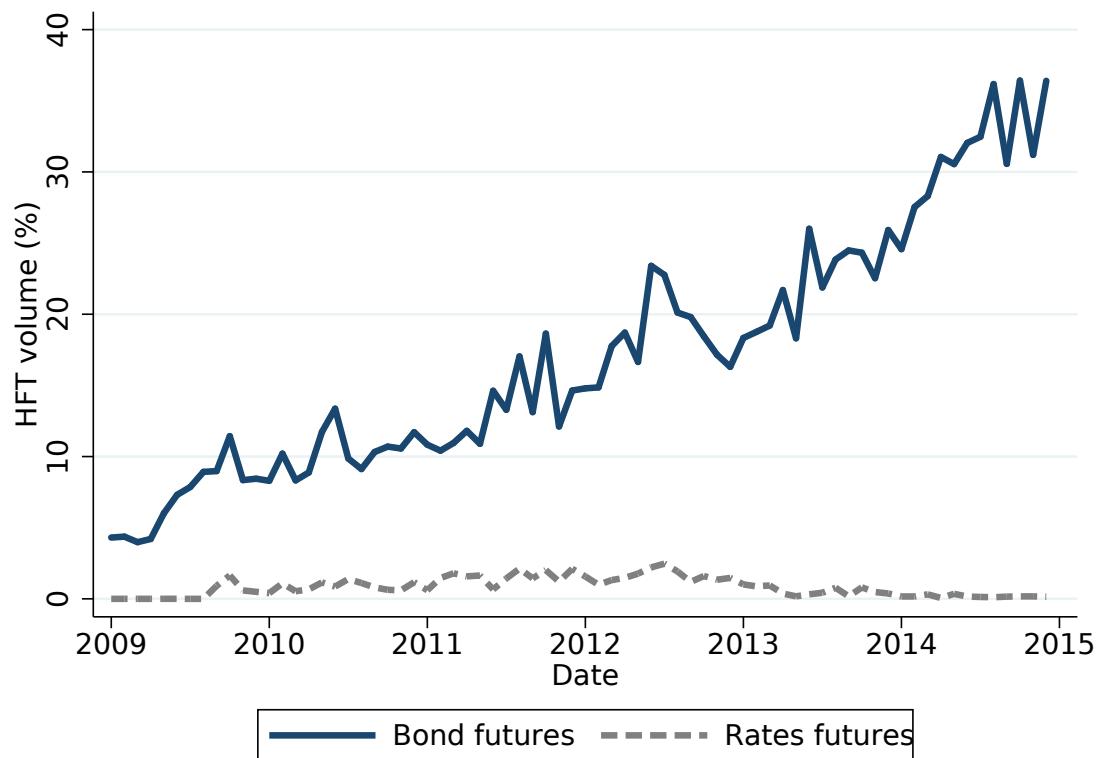
ure 3: HFT counterparty share during four volume quartiles of trade strings, sorted by volume

This figure shows the percent of volume institutions transacted with HFT during a trade string for the CGB contract. A *trade string* is a sequence of trades by a participant, 95% in the same direction (buy or sell), with no more than 20 minutes between any two trades. Each panel is a subsample of trade strings by volume. The graph in each panel divides trade strings into four quartiles at cutoffs of 25%, 50% and 75% of the volume.
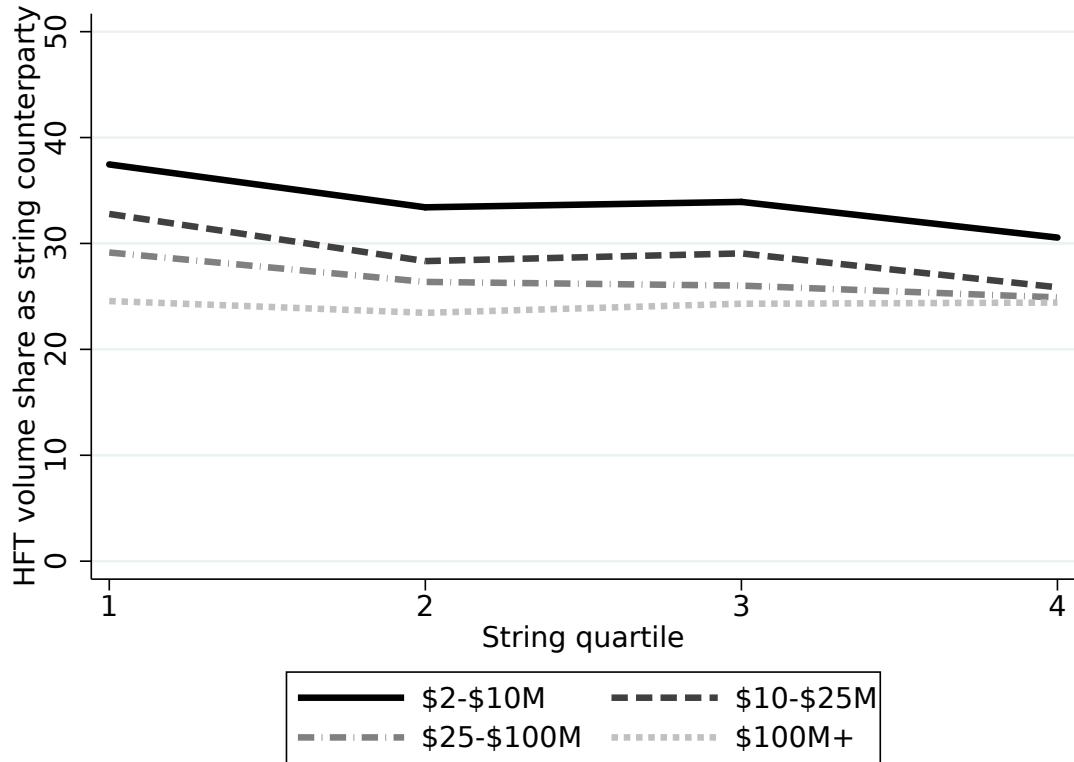
Figure 4: Multi-day price impacts of trade strings

This figure shows the price impact of trade strings in bins by volume, duration or long-term price impact, observed at seven time lags: immediately after the string's finish, at end of day, and at five 24-hour increments after finish. In black is the mean price impact, and in dashed gray is the 95% confidence interval. A *trade string* is a sequence of trades by a participant, 95% in the same direction (buy or sell), with no more than 20 minutes between any two trades. *Volume* is the size of the total net position acquired during the trade string; *duration* is two times the volume-weighted execution time from start to finish of the string in minutes. The first three bins are simply by volume; the last two are by volume and duration to emphasize the importance of duration to the price impact of large-sized strings.

(a) For volumes $2–25M

(b) For volumes $25–100M and durations ≥ 10 min.

(c) For volumes ≥ $100M and durations ≥ 60 min.

(d) For price impacts ≥ 3 basis points

28
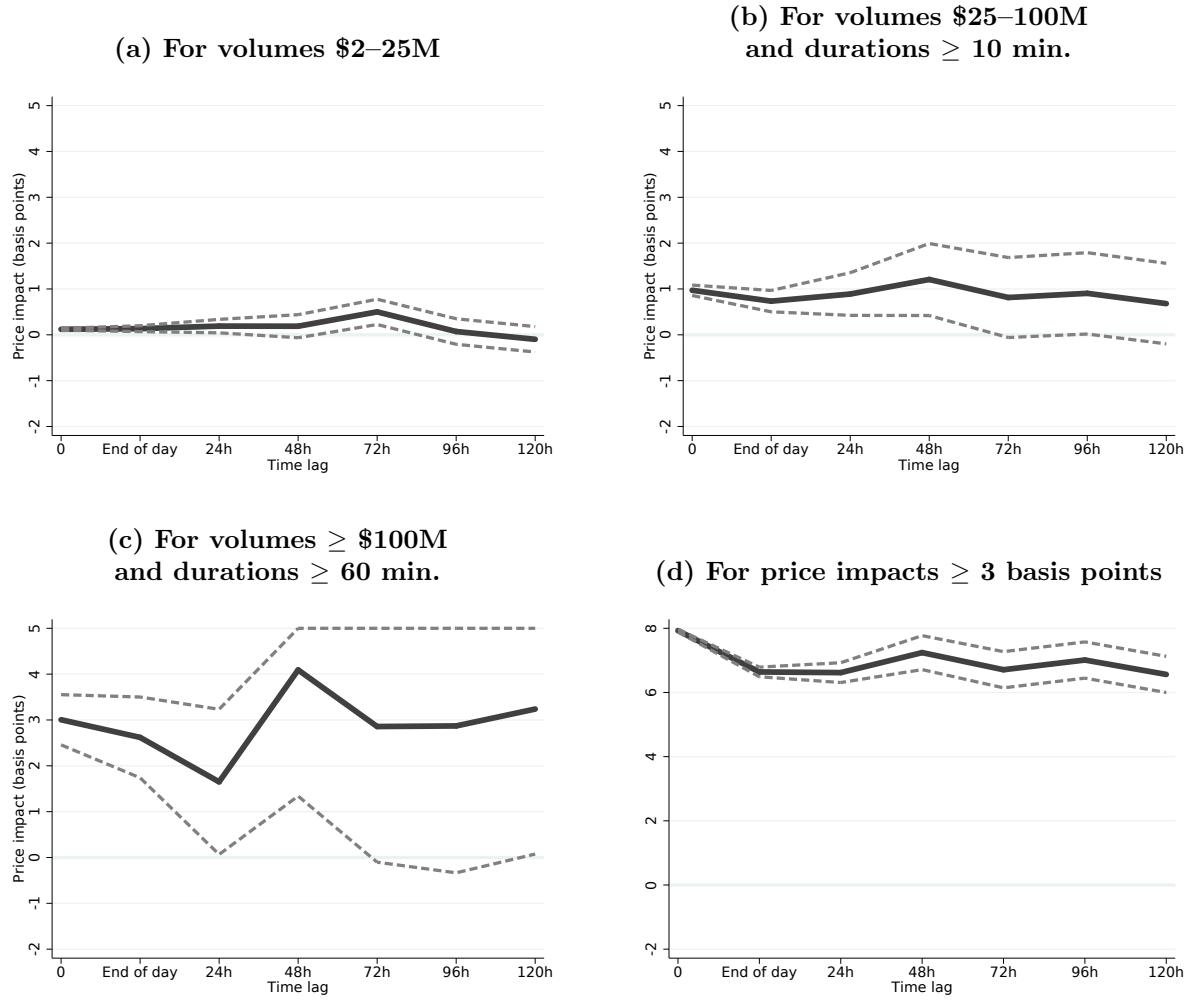
Figure 5: Cumulative order flows during trade strings

This figure shows the cumulative order flows of institutions building a position via a trade string (blue), the net position of HFTs in the market (red), and the net position of all other participants in the market (green), during 11 moments during the life of a trade string: 10 time deciles from string start to finish and at the moment 30 minutes after string end. A *trade string* is a sequence of trades by a participant, 95% in the same direction (buy or sell), with no more than 20 minutes between any two trades. The strings are for the CGB contract, the future on the Government of Canada 10-year bond. Together with Figure **??**, the panels subsample the trade strings in 12 bins by volume and duration. The series sum to zero by definition.

(a) For volumes $2–25M

(b) For volumes $25–100M and durations $\geq$ 10 min.

(c) For volumes $\geq$ $100M and durations $\geq$ 60 min.

(d) For price impacts $\geq$ 3 basis points
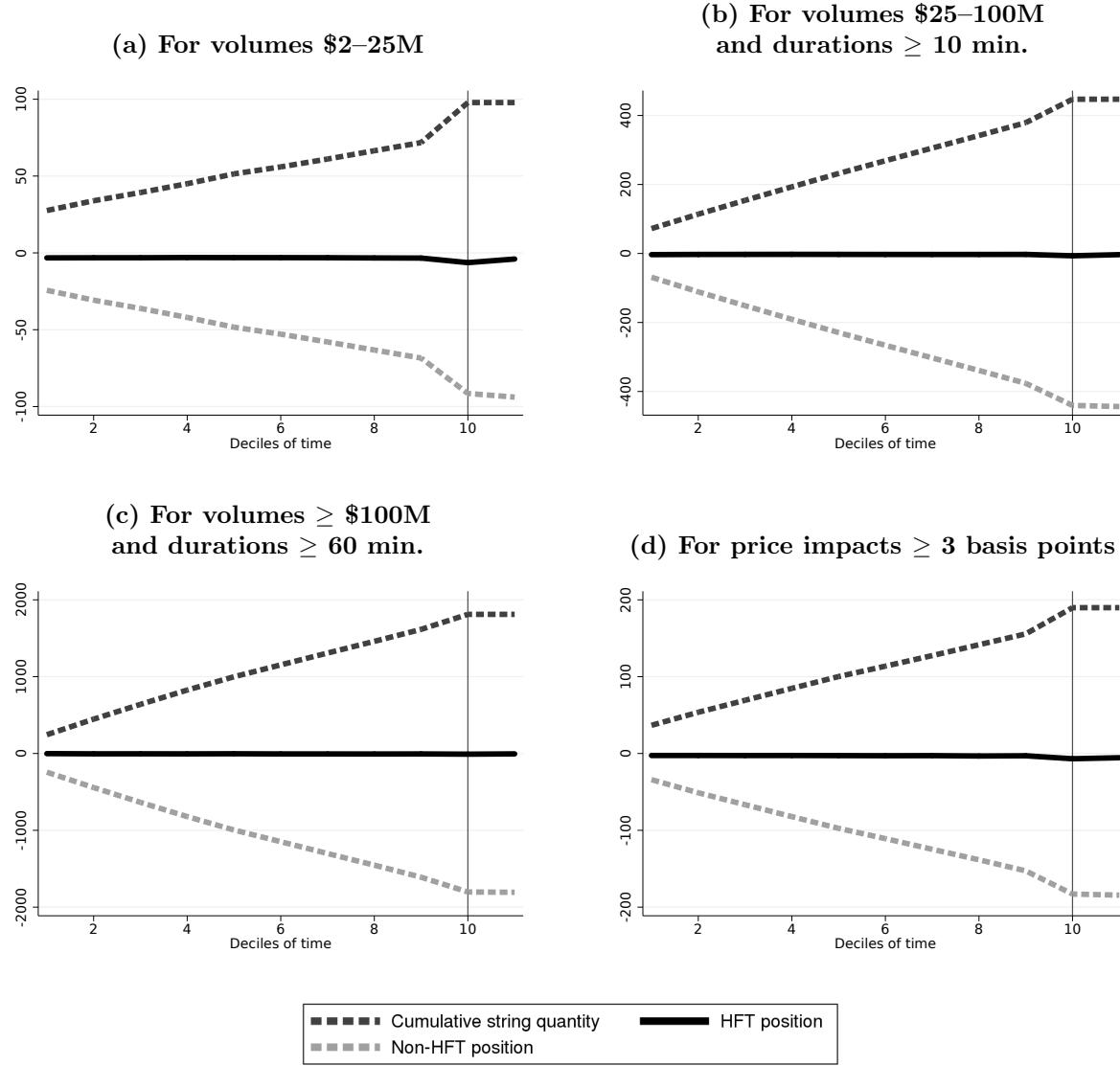
Figure 6: Yearly change in number of Approved Participants popularly regarded as HFTs

In accordance with regulation, the Montréal Exchange (MX) issues a public release called a circular when a new direct-access and clearing member joins or leaves the exchange. The status of being a direct-access and clearing member is called Approved Participant status. The name of the company is listed on the release.

For the purpose of this graph, an MX Approved Participant is considered an "HFT" if its name is mentioned on one of four articles in the popular press listing "well-known" HFT firms. The four articles are from (click for link) The Wall Street Journal, Investopedia, Quora and Medium. The firms listed on these websites are intentionally a superset; most of these firms are not Approved Participants on MX according to the circulars on its website. The websites list the firms named: ATD, Allston Trading LLC, Chicago Trading Company, Chopper Trading, Citadel Securities, DRW Holdings LLC, DV Capital, EWT, Five Rings Capital LLC, Flow Traders, Geneva Trading, GETCO, GSA Capital Partners, GTS, Hudson River Trading, IMC Financial, Jane Street, Jump Trading, KCG Holdings Inc., Knight Capital, Latour Trading, Optiver, Quantlab Financial, RSJ Algorithmic Trading, Spire Europe, Spot Trading, Sun Trading, Susquehanna International Group LLP, Tower Research Capital, Tradebot Systems Inc., Two Sigma Investments LP, Virtu Financial, Wedbush, Wolverine, XR Trading, XTX Markets. This list is intentionally a superset; most of these firms are not Approved Participants on MX according to the circulars on its website.

30
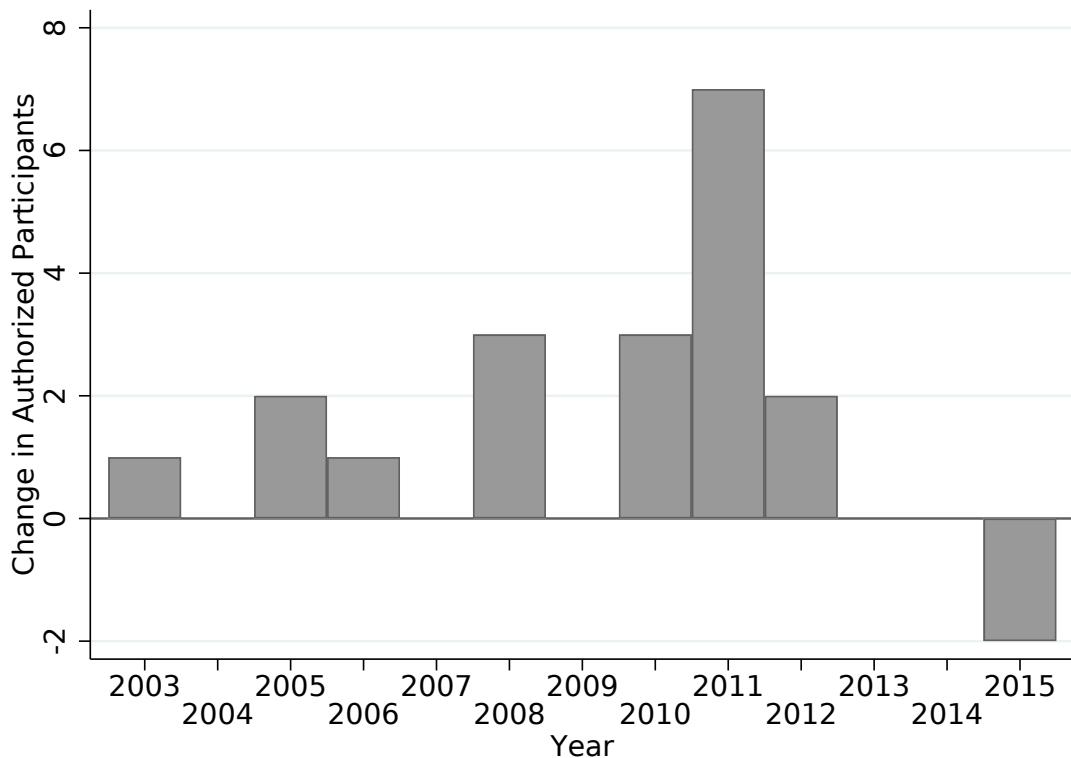
Figure 7: Herfindahl-Hirschman index of HFT volume 2011–2012

The graph below shows the monthly average level of the Herfindahl-Hirschman Index (HHI) of HFT volumes during, before and after the sample window of the event study. The HHI index is computed as the daily sum of the squared market shares of the HFTs, where market shares are the daily volumes of each HFT divided by the total volume over all HFTs.
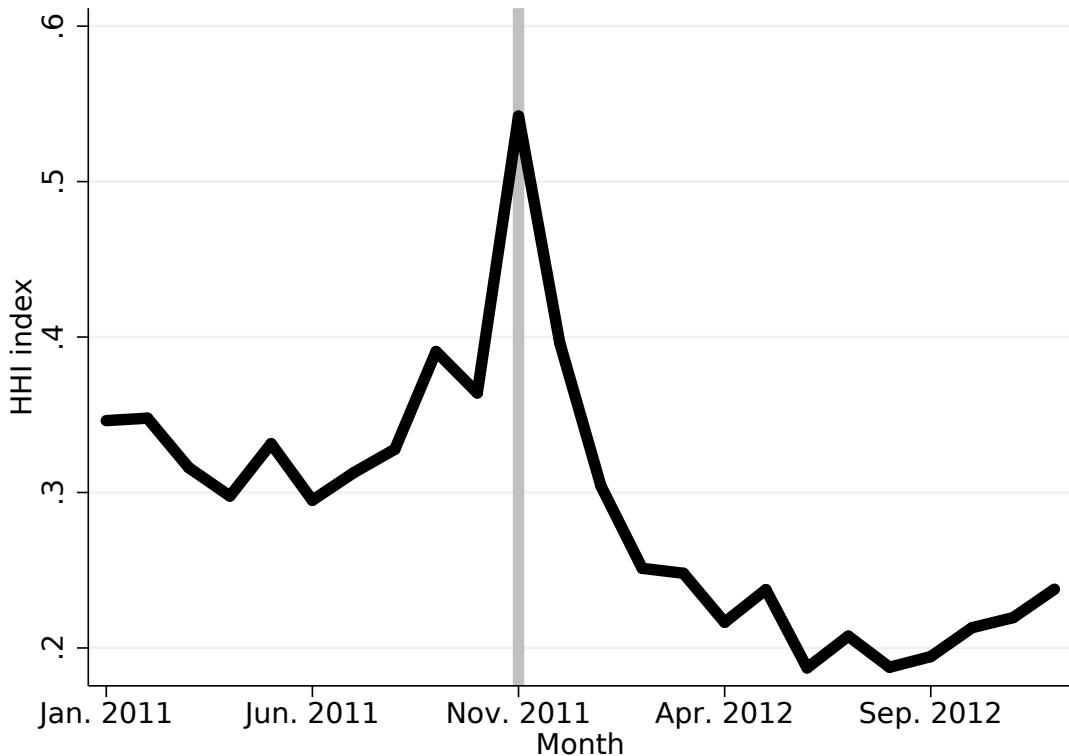
Figure 8: Implementation shorfall of strings sized $2–$10M and number of HFT

This figure plots the average monthly implementation shortfall for trade strings sized $2–$10 million (in solid blue) and the average monthly number of HFTs (in dashed gray) in inverted scale, both for the CGB contract, the future on the Government of Canada 10-year bond, 2009–2014. A *trade string* is a sequence of trades by a participant, 95% in the same direction (buy or sell), with no more than 20 minutes between any two trades. *Implementation shortfall* is the difference between the volume-weighted average price of a string and the midquote contemporaneous to the first trade. At the request of the Montréal Exchange, the number of empirically identified HFTs is not explicitly labeled on the y-axis.
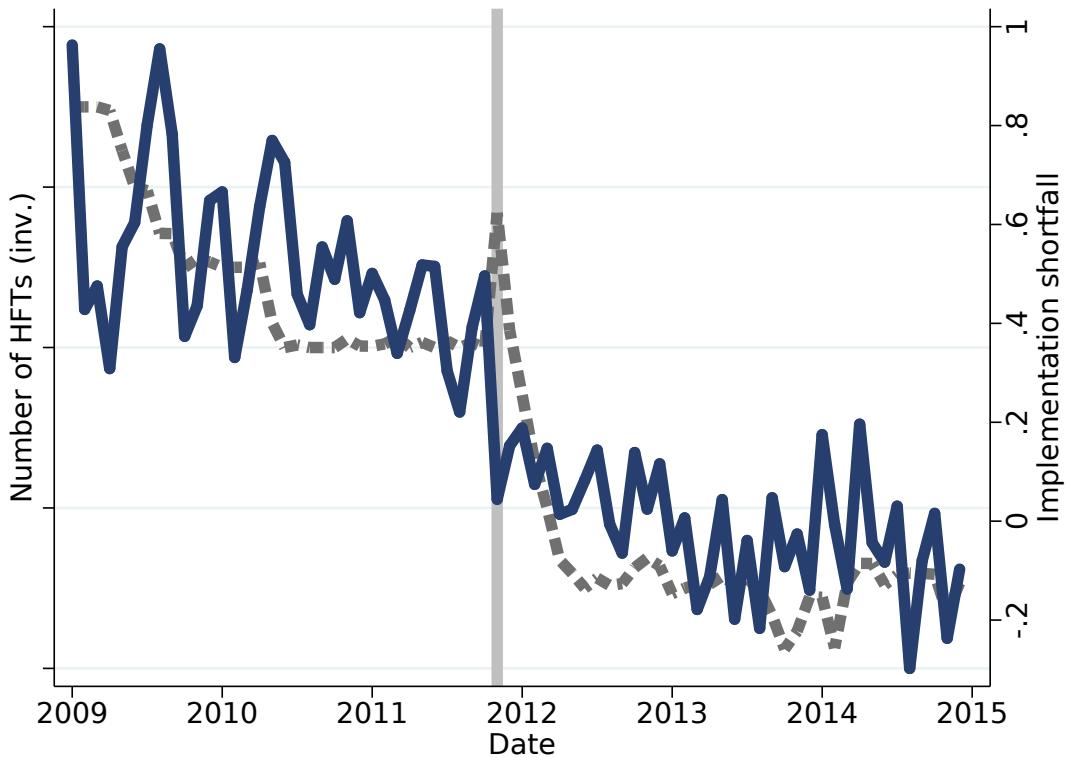
Figure 9: Implementation shortfall of CGB compared with that of BAX, by volume bin

This figure shows the weekly average implementation shortfall of trade strings for the CGB contract (in solid blue), the future on the Government of Canada 10-year bond, compared with that for the BAX contract (in dashed gray), the future on the Canadian Banker's Acceptance rate, 2009–2014. A *trade string* is a sequence of trades by a participant, 95% in the same direction (buy or sell), with no more than 20 minutes between any two trades. *Implementation shortfall* is the difference between the volume-weighted average price of a string and the midquote contemporaneous to the first trade. The vertical black bar is the event month.

(a) $2–$10M

(b) $10M–$25M

(c) $25M–$100M

(d) $100M+
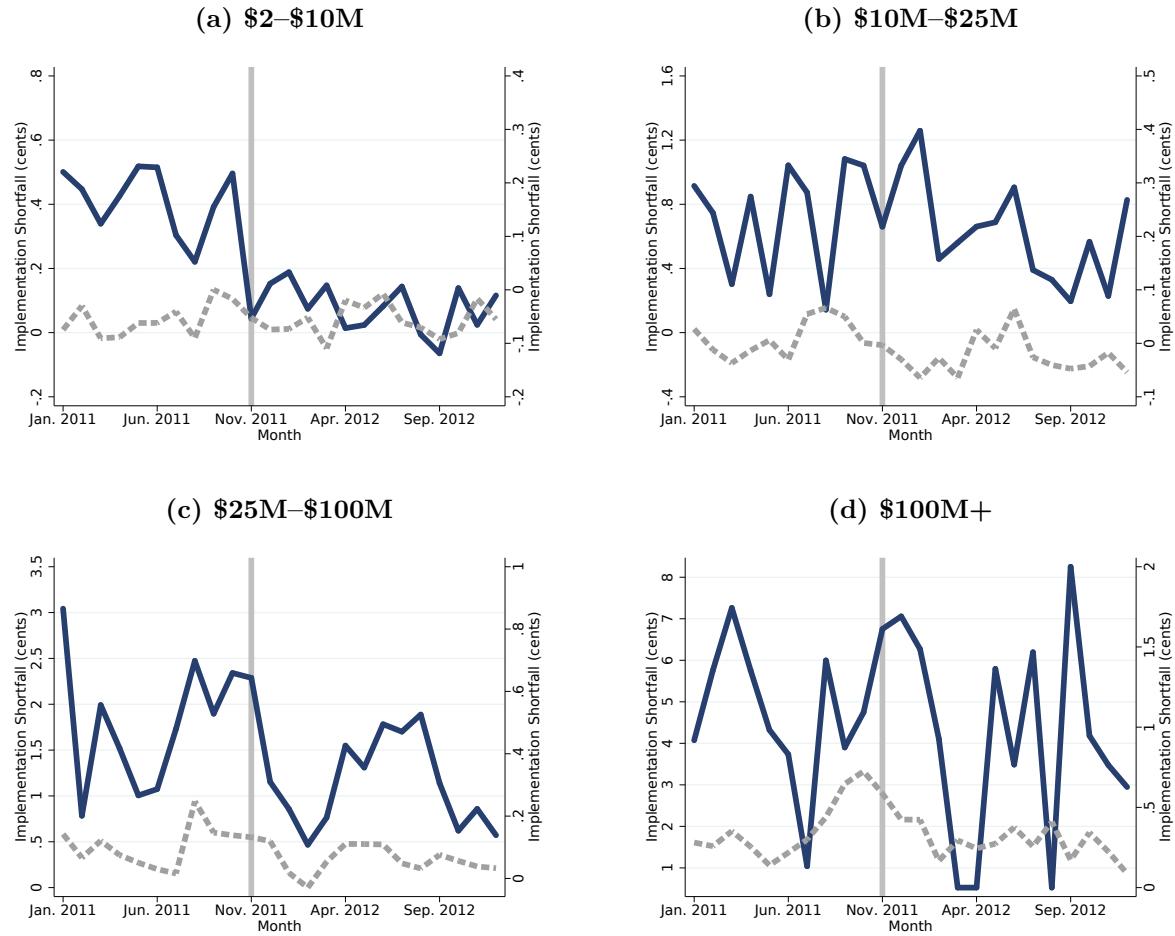
Figure 10: Liquidity metrics for CGB compared to those of BAX, by volume bin

This figure shows liquidity metrics for trades in trade strings for the CGB contract, the future on the Government of Canada 10-year bond, compared with those for the BAX contract, the future on the Canadian Banker's Acceptance rate. A *trade string* is a sequence of trades by a participant, 95% in the same direction (buy or sell), with no more than 20 minutes between any two trades. *Effective spread* is the signed difference between the trade price and contemporaneous midquote; *price impact* is the signed difference between the contemporaneous midquote and the midquote some time interval after the trade.

(a) Effective spread

(b) Five-second price impact

(c) Thirty-second price impact

(d) Five-minute price impact
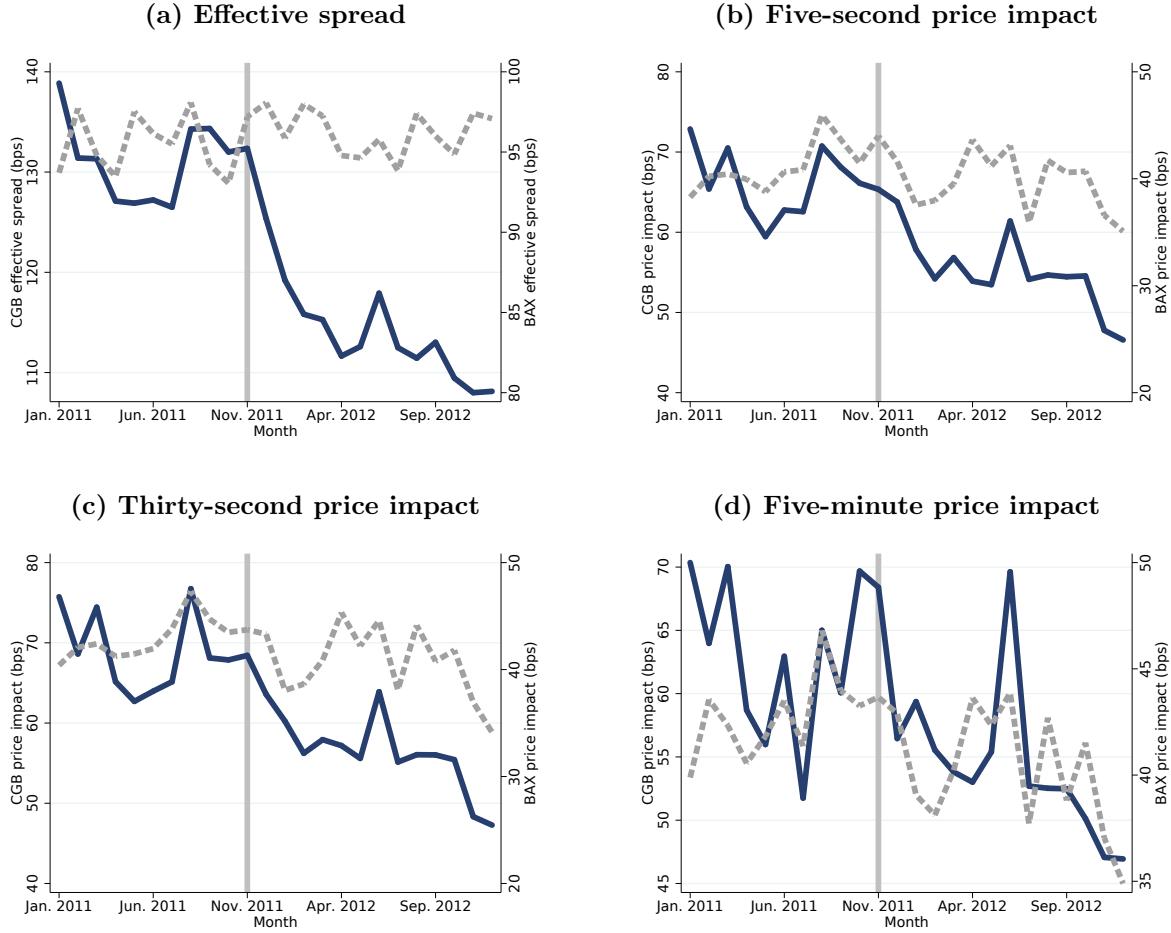
Figure 11: Change in execution costs of CGB trade strings, by volume bin

This figure shows the change in execution-cost metrics computed separately for volume increments of trade strings for the CGB contract, the future on the Government of Canada 10-year bond, after the event. A *trade string* is a sequence of trades by a participant, 95% in the same direction (buy or sell), with no more than 20 minutes between any two trades. The metrics are computed for the first $5 million of the string, the next $5 million, the next $5 million, and so on to $50 million. *Implementation shortfall* is the difference between the volume-weighted average price of a portion of a string and the midquote contemporaneous to the first trade. *Effective spread* is the signed difference between the prices of trades in the portion and the contemporaneous midquotes; *price impact* is the signed difference between the contemporaneous midquote and the midquote some time interval after the trades in the portion.

(a) Implementation shortfall
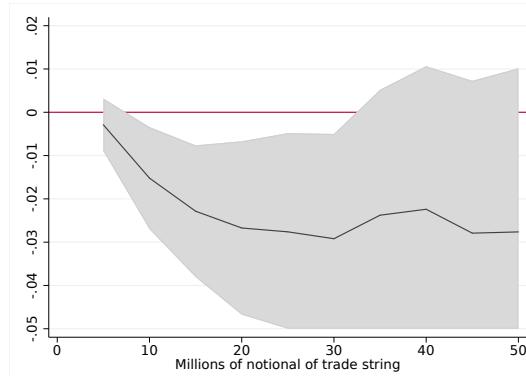
(b) Effective spreads
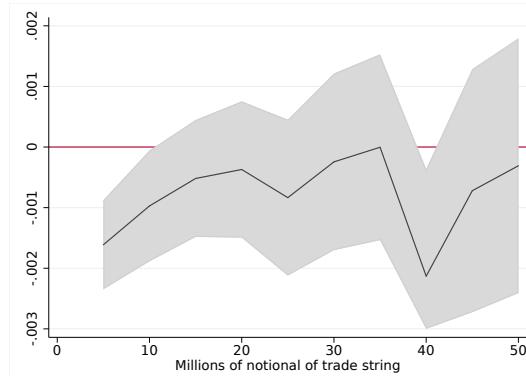
(c) Per-trade price impact
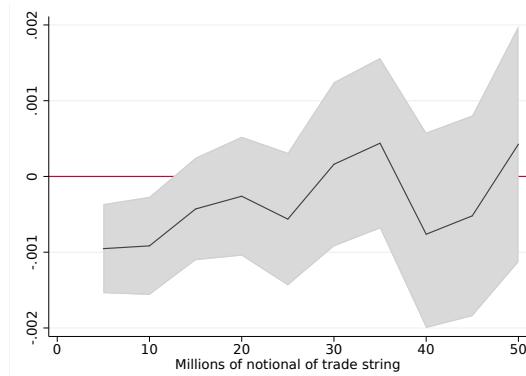
Figure 12: Dynamic difference-in-differences event study on liquidity metrics

This figure portrays the fit treatment effects and 95% confidence interval of the treatment coefficients from a dynamic DiD event study using weekly average data from the CGB (the future on the Government of Canada 10-year bond) and the BAX (the future on the Canadian Banker's Acceptance rate) on various metrics of liquidity. The treatment group is the CGB, and the control group is the BAX. The event window is the five months before and five months after the first month that there was an increase in the number of HFT trading the CGB, November 2011. The metrics are: *Implementation shortfall* (IS), the difference between the volume-weighted average price of a string and the midquote contemporaneous to the first trade; *effective spread*, the signed difference between a trade price and the contemporaneous midquote; and *price impact*, the signed difference between the contemporaneous midquote and the midquote some time interval after the trades in the portion.

(a) Implementation shortfall


(b) Effective spreads

(c) Five-second price impact
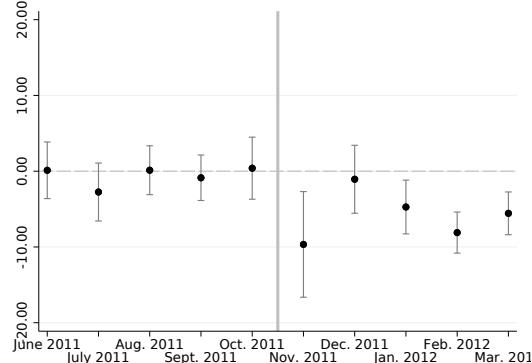
Table I: Average daily statistics for the bond and rates futures, 2009–2014

This table gives yearly averages and standard deviations of the daily average observations of various statistics for the CGB (the future on the Government of Canada 10-year bond) and for the BAX (the future on the Canadian Banker's Acceptance interest rate). *Volatility* is the two-week rolling standard deviation of the daily close price log return, annualized. *Time to change* is the number of minutes until a change in the bid or ask price. *Order lifetime* is the number of minutes the average limit order rests. *Percent ticklocked* is the percent of the time the bid-ask spread is equal to the tick size. *Inside depth* is the sum of the quantities of visible limit orders at the bid and ask prices. *Trades* is the number of trades. *Trade size* is the size (in millions of dollars) of the trades.

Panel A: Statistics for the bond future

	Years					
	2009	2010	2011	2012	2013	2014
Volatility	7.127	5.928	6.376	5.063	5.071	5.252
	(2.758)	(2.045)	(2.060)	(1.667)	(1.962)	(4.049)
Time to change	0.151	0.161	0.141	0.206	0.245	0.250
	(0.039)	(0.053)	(0.038)	(0.079)	(0.101)	(0.085)
Order lifetime	1.53m	2.09m	0.84m	0.68m	0.74m	1.63m
	(0.601)	(1.157)	(0.419)	(0.230)	(0.270)	(0.704)
Percent ticklocked	29.2%	41.2%	49.8%	73.4%	80.4%	89.6%
	(12.51)	(9.127)	(8.361)	(7.103)	(6.893)	(4.016)
Inside depth	$3.25M	$3.24M	$3.12M	$5.01M	$7.55M	$8.38M
	(0.989)	(0.978)	(0.684)	(1.303)	(2.142)	(1.591)
Trades	4853.0	5966.6	8505.4	11944.8	15796.1	17084.7
	(1444)	(1822)	(2187)	(2812)	(4132)	(4397)
Trade size	$0.33M	$0.33M	$0.28M	$0.25M	$0.26M	$0.25M
	(0.068)	(0.068)	(0.046)	(0.041)	(0.033)	(0.039)

Panel B: Statistics for rates futures

	2009	2010	2011	2012	2013	2014
Volatility	0.459	0.762	0.683	0.377	0.319	0.331
	(0.468)	(0.672)	(0.533)	(0.247)	(0.245)	(0.276)
Time to change	30.51	17.34	15.92	49.61	151.6	168.1
	(73.61)	(52.23)	(53.95)	(103.4)	(173.7)	(181.1)
Order lifetime	54.2m	44.5m	45.5m	61.5m	75.3m	95.5m
	(48.02)	(38.02)	(34.60)	(42.54)	(46.85)	(69.32)
Percent ticklocked	90.9%	91.9%	94.4%	97.4%	98.5%	98.8%
	(13.23)	(10.83)	(8.923)	(7.650)	(9.093)	(7.763)
Inside depth	$1944M	$1822M	$2077M	$3990M	$14369M	$17150M
	(2244)	(1880)	(1805)	(2970)	(10260)	(11984)
Trades	238.4	429.9	563.6	445.7	358.1	314.3
	(178.0)	(310.1)	(394.6)	(342.3)	(268.2)	(220.9)
Trade size	$29.1M	$27.8M	$28.4M	$30.5M	$35.2M	$38.9M
	(16.92)	(16.16)	(16.11)	(16.64)	(20.69)	(29.05)

Table II: Summary statistics on participant groups trading the CGB, 2009–2014

This table gives yearly averages and standard deviations of the daily observations of trading statistics for market participants trading the CGB contract, the future for the Government of Canada 10-year bond. The statistics are averaged over the members of the HFT and nonHFT groups. *Order lifetime* is the number of minutes the average limit order rests averaged over members of the participant group. *Switching rate* is the average percent of trades in which a buy follows a sell or a sell follows a buy, averaged over group members. *Overnight inventory* is the end-of-day inventory position in number of contracts, averaged over group members. *Quotes*, *Trades*, and *Cancels* are the daily number of limit-order inserts, trades, and limit-order cancels, averaged over group members.

Panel A: Statistics for HFTs

	Years					
	2009	2010	2011	2012	2013	2014
Order lifetime	0.25m	0.30m	0.49m	0.75m	0.76m	1.01m
	(1.282)	(1.342)	(0.938)	(1.612)	(0.567)	(1.395)
Switching rate	30.0%	28.5%	21.8%	27.3%	27.6%	27.9%
	(9.474)	(10.17)	(8.078)	(11.19)	(10.73)	(12.07)
Overnight inventory	0.362	0.681	1.278	1.545	2.820	5.378
	(1.604)	(3.491)	(3.767)	(3.905)	(7.623)	(12.87)
Quotes	2634.7	4716.1	6707.9	7647.0	8395.0	13859.4
	(4874)	(7941)	(11497)	(10424)	(9750)	(13747)
Trades	648.8	781.2	1266.4	1412.2	1710.3	2670.8
	(354.7)	(555.2)	(837.7)	(1098.4)	(1740.4)	(2820.9)
Cancels	2421.6	4303.7	6047.9	6963.7	7562.6	12663.1
	(4616)	(7583)	(10901)	(9826)	(8811)	(12688)

Panel B: Statistics for nonHFTs

	2009	2010	2011	2012	2013	2014
Order lifetime	18.8m	19.7m	19.5m	21.6m	24.5m	22.6m
	(61.13)	(66.50)	(66.36)	(77.55)	(85.23)	(76.97)
Switching rate	12.9%	12.4%	11.4%	9.2%	8.7%	8.8%
	(13.07)	(11.48)	(12.23)	(10.10)	(8.397)	(8.457)
Overnight inventory	410.6	430.2	468.7	541.6	689.9	682.3
	(580.2)	(575.9)	(649.3)	(751.8)	(948.0)	(947.3)
Quotes	908.6	980.3	2522.8	1928.9	2660.6	2597.4
	(3854)	(4043)	(12023)	(7104)	(9484)	(9495)
Trades	400.7	437.3	598.7	808.5	1060.0	980.2
	(664.1)	(662.2)	(856.4)	(1279.6)	(1794.3)	(1576.5)
Cancels	804.3	866.1	2383.1	1753.2	2399.7	2343.1
	(3634)	(3874)	(11830)	(6898)	(9129)	(9225)

Table III: Summary statistics for bond-future trade strings, by volume bin

This table gives yearly averages of statistics for trade strings on the CGB, the future on the Government of Canada 10-year bond. A *trade string* is a sequence of trades by a participant, 95% in the same direction (buy or sell), with no more than 20 minutes between any two trades. *Implementation shortfall* is the difference between the volume-weighted average price of a string and the midquote contemporaneous to the first trade. *Number* is the count of strings in the category. *Volume per string* is the par value in millions of the position purchased or sold in the string. *Trades per string* is the count of trades in a string. *Market share per string* is the volume per string as a percent of daily volume. *Duration* is two times the volume-weighted execution time from start to finish of the string in minutes (multiplied by two as duration is a half-life). *Aggressiveness* is the percentage string volume executed via marketable orders.

	Years					
	2009	2010	2011	2012	2013	2014
Positions of $2–10M						
Implementation shortfall	0.580¢	0.544¢	0.355¢	0.073¢	-0.089¢	-0.037¢
Number	14,977	17,789	20,254	23,687	26,222	26,513
Volume per string	$4.7M	$4.7M	$4.7M	$4.7M	$4.7M	$4.8M
Trades per string	15.5	15.8	18.4	19.8	19.2	20.0
Market share per string	0.32%	0.27%	0.22%	0.17%	0.13%	0.12%
Duration	14.1m	14.9m	16.7m	16.1m	16.9m	21.6m
Agressiveness	40.1%	41.2%	42.3%	44.4%	42.0%	42.1%
Positions of $10–25M						
Implementation shortfall	0.763¢	0.852¢	0.722¢	0.563¢	0.297¢	0.310¢
Number	5,693	6,651	7,814	10,009	12,267	11,730
Volume per string	$14.7M	$14.9M	$14.7M	$14.7M	$15.0M	$14.9M
Trades per string	34.1	37.7	46.2	53.3	50.9	53.6
Market share per string	1.0%	0.83%	0.67%	0.53%	0.40%	0.38%
Duration	24.6m	30.8m	29.4m	27.6m	24.4m	29.5m
Agressiveness	39.5%	43.5%	42.0%	47.0%	45.3%	45.1%
Positions of $25–100M						
Implementation shortfall	1.76¢	1.46¢	1.80¢	1.12¢	1.12¢	1.08¢
Number	2,389	2,767	2,850	4,008	5,241	4,918
Volume per string	$42.0M	$42.7M	$41.9M	$42.5M	$42.5M	$42.7M
Trades per string	98.4	106.6	126.7	157.1	147.6	154.0
Market share per string	2.7%	2.3%	1.8%	1.5%	1.1%	1.1%
Duration	58.8m	62.0m	60.0m	65.5m	53.4m	60.1m
Agressiveness	39.5%	42.8%	40.7%	48.3%	45.8%	45.2%
Positions of $100M+						
Implementation shortfall	6.88¢	4.14¢	4.90¢	3.26¢	2.41¢	2.09¢
Number	212	209	240	321	537	536
Volume per string	$179.1M	$149.5M	$164.0M	$154.2M	$175.2M	$172.5M
Trades per string	653.0	390.8	580.6	523.2	663.4	773.5
Market share per string	10.2%	7.3%	6.7%	5.2%	4.3%	4.1%
Duration	135.7m	121.3m	120.3m	91.2m	118.4m	152.1m
Agressiveness	53.2%	45.9%	46.2%	43.4%	41.0%	48.5%

Table IV: Difference-in-differences event study on implementation shortfall

This table reports regression coefficients from a DiD event study using weekly average data from the CGB (the future on the Government of Canada 10-year bond) and the BAX (the future on the Canadian Banker's Acceptance rate) on implementation shortfall. The treatment group is the CGB, and the control group is the BAX. The event window is the five months before and five months after the first month that there was an increase in the number of HFT trading the CGB, November 2011. *Implementation shortfall* (IS) is the difference between the volume-weighted average price of a string and the midquote contemporaneous to the first trade. The model is,

$$IS_{i,t} = Treatment_{i,t} + BAX_i + Month_t + Controls_{i,t} + \epsilon_{i,t}$$

where i indexes the contract (CGB or BAX) and t indexes the week. *Treatment* is a dummy for observations of the CGB contract in the post-period; *BAX* is a dummy for observations of the BAX contract; *Month* is a series of time fixed effects (one for each month). The difference in differences is therefore the variable *Treatment*. The control variables are: *Volume*, the trading volume in contracts; *Volatility*, the 20-day rolling daily return volatility; and *Lag*, the previous week's average level of implementation shortfall. Columns 1 through 4 report treatment effects for trade strings binned in four groups by volume. Columns 5 through 7 report the treatment effects for trade strings binned in three groups by duration. Coefficients on $Month_t$ are not reported to save space.

	Volume bins				Duration bins		
	$2–10M	$10–25M	$25–100M	$100M+	0-10 min.	10-60 min.	60+ min.
Treatment	-31.00***	-6.64	-70.40	-188.48	-21.53*	-29.35	-85.13
	(-4.89)	(-0.27)	(-1.41)	(-1.63)	(-2.74)	(-1.97)	(-0.86)
BAX	-53.20***	-127.36*	-140.70	-795.06*	-55.25***	-70.72*	-259.06
	(-5.66)	(-2.71)	(-1.87)	(-2.41)	(-7.24)	(-2.26)	(-1.30)
Volume	-86.90	-737.47**	563.15	-332.59	32.70	300.93	-1997.15
	(-0.67)	(-4.36)	(0.43)	(-0.10)	(0.41)	(0.75)	(-1.59)
Volatility	-2.02	-5.78	5.01	-48.29	-0.55	-3.19	-5.11
	(-1.85)	(-1.31)	(0.52)	(-0.85)	(-0.47)	(-0.91)	(-0.22)
Lag	0.01	0.01	-0.15	-0.19	-0.04	-0.23*	-0.17
	(0.15)	(0.08)	(-0.96)	(-1.40)	(-0.16)	(-2.49)	(-1.39)
Constant	52.53***	155.33*	129.86	909.59*	61.41***	58.24	328.49
	(5.10)	(3.21)	(1.19)	(2.46)	(7.32)	(1.48)	(1.48)
N	84	84	84	83	84	84	84
R^2	0.732	0.585	0.420	0.301	0.893	0.413	0.330

t statistics in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table V: Difference-in-differences event study on spreads and price impacts

This table reports regression coefficients from a DiD event study using weekly average data from the CGB (the future on the Government of Canada 10-year bond) and the BAX (the future on the Canadian Banker's Acceptance rate) on effective spread and on various measures of price impact. The treatment group is the CGB, and the control group is the BAX. The event window is the five months before and five months after the first month that there was an increase in the number of HFT trading the CGB, November 2011. *Effective spread* (ES) is the signed difference between the trade price and contemporaneous midquote for trades in the strings, in basis points; *Price impact* (PI) is the signed difference between the contemporaneous midquote and the midquote some time-interval after the trade for trades in the strings, in basis points. Price impacts are computed at five-second, 30-second and five-minute intervals. The model is,

$$metric_{i,t} = Treatment_{i,t} + BAX_i + Month_t + Controls_{i,t} + \epsilon_{i,t}$$

where i indexes the contract (CGB or BAX) and t indexes the week. *Treatment* is a dummy for observations of the CGB contract in the post-period; *BAX* is a control-group fixed effect for observations of the BAX contract; *Month* is a series of time fixed effects (one for each month). The difference in differences is therefore the variable *Treatment*. The control variables are: *Volume*, the trading volume in contracts; *Volatility*, the 20-day rolling daily return volatility; and *Lag*, the previous week's average level of implementation shortfall. Column 1 reports a treatment effect on the effective spread of trades in trade strings. Columns 2 through 4 report the treatment effects on the price impact of trades in trade strings. Coefficients on $Month_t$ are not reported to save space.

	Effective spread	Five-second price impact	30-second price impact	Five-minute price impact
Treatment	-13.24***	-9.52**	-12.99*	-10.84
	(-6.16)	(-3.57)	(-2.52)	(-0.75)
BAX	-28.69***	-18.79**	-25.75**	-26.82
	(-8.27)	(-4.13)	(-4.03)	(-1.91)
Volume	35.28	54.10	41.22	-127.51
	(1.50)	(1.39)	(0.56)	(-0.61)
Volatility	0.11	-0.45	-1.53**	-2.12
	(0.40)	(-0.89)	(-3.90)	(-1.49)
Lag	0.23*	0.15	0.20	0.11
	(2.80)	(1.17)	(1.26)	(1.47)
Constant	100.38***	61.41***	66.86***	73.20***
	(9.20)	(5.90)	(4.87)	(7.79)
N	84	84	84	84
R^2	0.963	0.829	0.748	0.349

t statistics in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table VI: Dynamic difference-in-differences event study on liquidity metrics

This table reports regression coefficients from a dynamic DiD event study using weekly average data from the CGB (the future on the Government of Canada 10-year bond) and the BAX (the future on the Canadian Banker's Acceptance rate) on various metrics of liquidity. The treatment group is the CGB, and the control group is the BAX. The event window is the five months before and five months after the first month that there was an increase in the number of HFT trading the CGB, November 2011. The metrics are: *Implementation shortfall* (IS), the difference between the volume-weighted average price of a string and the midquote contemporaneous to the first trade; *effective spread*, the signed difference between a trade price and the contemporaneous midquote; and *price impact*, the signed difference between the contemporaneous midquote and the midquote some time interval after the trades in the portion. The model is,

$$metric_{i,t} = CGB_{i,t} \cdot Month_t + Month_t + \epsilon_{i,t}$$

where i indexes the contract (CGB or BAX) and t indexes the week. Columns 1 through 4 report treatment effects for trade strings binned in four groups by volume. Columns 5 through 7 report the treatment effects for trade strings binned in three groups by duration. Coefficients on $Month_t$ are not reported to save space (the below are coefficients on $CGB_{i,t} \cdot Month_t$).

Month	Implementation shortfall	Effective spread	Five-second price impact	30-second price impact	Five-minute price impact
Jun. 2011	0.07	-2.43	0.12	-1.10	-0.72
	(0.66)	(-1.03)	(0.06)	(-0.43)	(-0.21)
Jul. 2011	-0.09	-2.69	-2.75	-5.61	-12.16***
	(-1.56)	(-1.01)	(-1.43)	(-1.68)	(-4.44)
Aug. 2011	0.06	2.32	0.13	2.87	1.07
	(0.64)	(0.95)	(0.08)	(0.85)	(0.51)
Sep. 2011	-0.03	5.10	-0.87	-3.12	-3.51
	(-0.58)	(1.76)	(-0.58)	(-1.33)	(-0.99)
Oct. 2011	-0.09	4.74	0.40	-1.17	6.89
	(-0.78)	(1.57)	(0.19)	(-0.48)	(1.05)
Nov. 2011	-0.24**	-6.63***	-9.66**	-4.10	-3.20
	(-3.40)	(-4.88)	(-2.77)	(-1.25)	(-0.68)
Dec. 2011	-0.18*	-6.32**	-1.06	-3.89	-6.09
	(-2.17)	(-2.75)	(-0.47)	(-1.77)	(-1.25)
Jan. 2012	-0.28**	-12.67***	-4.73**	-5.58**	-7.51*
	(-3.42)	(-7.92)	(-2.66)	(-3.13)	(-2.27)
Feb. 2012	-0.36***	-15.66***	-8.11***	-9.69***	-5.59**
	(-7.98)	(-16.22)	(-5.97)	(-4.53)	(-2.85)
Mar. 2012	-0.23***	-17.55***	-5.56***	-8.18***	-5.93
	(-3.59)	(-11.22)	(-3.94)	(-3.88)	(-1.99)
N	84	84	84	84	84
R^2	0.646	0.801	0.754	0.697	0.520

t statistics in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

REFERENCES

Baron, Matthew, Jonathan Brogaard, Björn Hagströmer, and Andrei Kirilenko, 2019, Risk and return in high-frequency trading, *Journal of Financial and Quantitative Analysis* 54, 993–1024.

Berger-Soucy, Léanne, Corey Garriott, André Usche, et al., 2018, Government of canada fixed-income market ecology .

Brogaard, Jonathan, and Corey Garriott, 2019, High-frequency trading competition, *Journal of Financial and Quantitative Analysis* 54, 1469–1497.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan, 2014, High-Frequency Trading and Price Discovery, *The Review of Financial Studies* 27, 2267–2306.

Brunnermeier, Markus K, and Lasse Heje Pedersen, 2005, Predatory trading, *The Journal of Finance* 60, 1825–1863.

Callaway, Brantly, and Pedro HC Sant'Anna, 2019, Difference-in-differences with multiple time periods, *Available at SSRN 3148250* .

Campbell, Bryan, and Scott Hendry, 2008, Price discovery in canadian and us 10-year government bond markets, *Available at SSRN 2695197* .

Cox, Justin, Bonnie Van Ness, and Robert Van Ness, 2019, Increasing the tick: Examining the impact of the tick size change on maker-taker and taker-maker market models, *Financial Review* 54, 417–449.

Fleming, Michael J, Bruce Mizrach, and Giang Nguyen, 2018, The microstructure of a us treasury ecn: The brokertec platform, *Journal of Financial Markets* 40, 2–22.

Garriott, Corey, and Ryan Riordan, 2019, Trading on long-term information, *Available at SSRN 3419065* .

Harris, Lawrence, 1998, Optimal dynamic order submission strategies in some stylized trading problems, *Financial Markets, Institutions & Instruments* 7, 1–76.

Hendershott, Terrence, Charles M Jones, and Albert J Menkveld, 2011, Does algorithmic trading improve liquidity?, *The Journal of Finance* 66, 1–33.

Hirschey, Nicholas, 2019, Do high-frequency traders anticipate buying and selling pressure?, *Available at SSRN 2238516* .

Kirilenko, Andrei, Albert S Kyle, Mehrdad Samadi, and Tugkan Tuzun, 2017, The flash crash: High-frequency trading in an electronic market, *The Journal of Finance* 72, 967–998.

Korajczyk, Robert A., and Dermot Murphy, 2019, High-Frequency Market Making to Large Institutional Trades, *The Review of Financial Studies* 32, 1034–1067.

Malinova, Katya, Andreas Park, and Ryan Riordan, 2018, Do retail investors suffer from high frequency traders?, *Available at SSRN 2183806* .

Menkveld, Albert J, 2013, High frequency trading and the new market makers, *Journal of financial Markets* 16, 712–740.

Tabb, Larry, 2017, Speed: Why it matters and what can be done .

Tong, Lin, 2014, A blessing or a curse? the impact of high frequency trading on institutional investors, in *The Impact of High Frequency Trading on Institutional Investors (October 5, 2015). European Finance Association Annual Meetings*.

van Kervel, Vincent, and Albert J Menkveld, 2019, High-frequency trading around large institutional orders, *The Journal of Finance* 74, 1091–1137.

Weller, Brian M, 2017, Does algorithmic trading reduce information acquisition?, *The Review of Financial Studies* 31, 2184–2226.

Yang, Liyan, and Haoxiang Zhu, 2018, Back-running: Seeking and hiding fundamental information in order flows, *Rotman School of Management Working Paper* .