Correlated High-Frequency Trading

Ekkehart Boehmer, Dan Li, and Gideon Saar*

This version: February 2016

Abstract

This paper studies correlations between the strategies of high-frequency trading (HFT) firms, which is a manifestation of the extent of competition in which these firms engage when pursuing similar strategies. We use a principal component analysis to show that there are several underlying common strategies and that the competing HFT firms pursuing these strategies generate most HFT activity. We investigate whether competition between HFT firms creates a systematic return factor, but find no supporting evidence for such an influence. However, the short-interval return volatility of most stocks loads negatively on a market-wide measure of correlated HFT strategies. The mitigating impact of HFT competition on stock volatility appears to be driven at least in part by a market-making strategy. The paper ends by documenting a negative relationship between two forms of competition—that between HFT firms and that between trading venues. We investigate a potential driver behind this negative relationship, and show that greater HFT competition within a trading venue helps smaller trading venues become more competitive or viable in terms of posting better prices and narrower spreads.

*Ekkehart Boehmer (eboehmer@smu.edu.sg) is from the Lee Kong Chian School of Business, Singapore Management University; Dan Li (lidan@hku.hk) is from the School of Economics and Finance, University of Hong Kong; and Gideon Saar (gs25@cornell.edu) is from the Johnson Graduate School of Management, Cornell University. We thank Matthew Baron, Evangelos Benos, James Brugler, Bjorn Hagstromer, Adam Joseph-Clark, Boyan Jovanovic, Andrew Karolyi, Albert Menkveld, Sophie Moinas, Dermot Murphy, Maureen O'Hara, Mike Prior, Ioanid Rosu, Mehrdad Samadi, and Giorgio Valente for helpful comments, as well as seminar participants at Cornell University, VU Amsterdam, and the PBC School of Finance at Tsinghua University. Dan Li thanks the Investment Industry Regulatory Organization of Canada (IIROC) for providing her with data used in this study during her work at IIROC as part of a research agreement, and Maureen Jensen for her support and many helpful discussions. All conclusions are those of the authors and do not represent the views of IIROC. Part of this research was conducted while Gideon Saar was visiting the School of Management at Tel Aviv University.

Correlated High-Frequency Trading

1. Introduction

High-frequency trading (henceforth, HFT) constitutes a large portion of stock market activity. HFT firms use computerized algorithms for proprietary trading, and they engage in electronic market making, cross-trading venue price arbitrage, short-term statistical arbitrage, and various other opportunistic strategies.[1] Our goal in this paper is to investigate correlations between HFT firms and the relationship of these correlations with the market environment. The correlated activity of HFT firms represents similarity in their strategies driven by the manner in which they respond to market stimuli and their business models. This similarity has two important dimensions. First, it implies that multiple HFT firms compete for the same profit opportunities, in which case competition could enhance the services they provide to the market and reduce the rent they earn. Second, such similarity could mean that HFT firms would be active or withdraw their activity at the same time or in the same stocks, thereby increasing the fragility of our markets or even creating a risk factor that would affect returns.

The competition dimension is of interest in part because the HFT business is a relatively new phenomenon and not much is known about its industrial organization. The secretive nature of HFT firms and their algorithms does not lend itself to easy characterization of strategies, hence making it difficult to judge the extent of competition in particular areas of their activity. In a free-market system, determining whether HFT firms deploy monopolistic strategies or compete heavily over the same sources of profits is important. Presumably, competition between HFT firms could alleviate the need for regulation because the rent extracted by these firms would decline at the same time that the level of services they provide in terms of liquidity provision and arbitrage would increase. The fragility dimension is of interest because HFT firms have driven many human intermediaries out of the market and have come to dominate market activity. If multiple HFT firms operate in tandem across stocks, their activity has the potential to introduce a systematic component to the trading environment. In other words, idiosyncratic activity in one

[1] High-frequency trading therefore does not include agency algorithms that execute orders on behalf of investors; the term "HFT" is used only to denote proprietary trading operations conducted by standalone firms or the trading desks of larger financial firms.

security can influence HFT activity in multiple securities and hence could create systematic patterns in returns or volatility over and above fundamental influences.

Our investigation of these issues is facilitated by data from the Investment Industry Regulatory Organization of Canada (IIROC), which is the national self-regulatory organization that oversees equity trading venues in Canada. The unique features of these data include the ability to observe the HFT activity of both licensed dealers and HFT firms that use direct market access (DMA) arrangements as well as the ability to track the activity of each HFT firm across all trading venues in Canada. Our sample consists of S&P/TSX60 stocks, and we analyze activity from 30 days that represent bullish, bearish, and neutral trading environments during a period running from June 2010 through March 2011.

We characterize 31 trading firms as "high-frequency traders," and perform a principal component analysis to ascertain if there are underlying common strategies that represent much of the variability in the HFT firms' activity as well as to identify the firms that follow each underlying common strategy. We find extensive competition in at least three underlying common strategies. While there are 17 HFT firms do not appear to pursue one of these common strategies, the 14 firms that follow the three common strategies represent most of HFT activity: 96.21% of the messages that HFT firms send to the market and 78.97% of the volume they trade. Therefore, we find that competing HFT firms generate most of the HFT activity, which suggests that the economic rent earned by these firms for making markets or arbitraging price differences across multiple trading venues is probably not very high.

We analyze the empirical representation of these underlying common strategies (i.e., the principal component scores) by regressing them on various attributes of the market environment. One of the underlying common strategies appears to be a market-making strategy pursued by firms with high market shares that become more active when market making is both needed and more profitable. Another underlying common strategy appears to represent a cross-venue arbitrage strategy that moves liquidity across the market, while the third underlying common strategy could represent a short-horizon directional strategy that is active when volatility is high, the book is thinner, and prices differ across trading venues. Our results demonstrate that there is clear heterogeneity in the strategies that HFT firms pursue, suggesting that it may be more constructive to discuss the impact of HFT on the market environment in the context of specific strategies rather than in the aggregate.

We then investigate a particular concern regarding HFT firms' correlated behavior: whether their strategies form a link across stocks, thereby generating a systematic risk factor and hence increasing the risk of stocks. Our focus is on studying the impact of competition between HFT firms, not the magnitude of HFT activity. Our results do not support the idea that the extent of correlation in HFT strategies creates a common factor in short-interval returns. However, we find that the short-interval volatility of stocks loads negatively on the market-wide measure of competition between HFT firms. This contrasts with the positive relationships that both market-wide volatility and the magnitude of HFT activity have with individual stock volatility, suggesting that this negative relationship constitutes a separate and distinct effect. Our analysis shows that the strongest driver behind this negative relationship is competition between firms that load significantly on the second principal component, which we associate with a market-making strategy, suggesting that the reduction in volatility stems in part from a reduction in transitory price movements (see, for example, Ho and Stoll (1983)).

We end the paper by examining the interrelationship of competition between multiple trading venues and competition between HFT firms. One role for HFT firms in a fragmented market structure is to be the market consolidators that transform the environment into a virtual central electronic limit book (from the perspective of most investors) by moving orders across markets extremely quickly to ensure that prices are the same across trading venues and liquidity exists where it is needed. We show that concentration of trading in the market as a whole is negatively related to competition between HFT firms. This relationship is driven by two specific strategies: cross-venue arbitrage (the first principal component) and market making (the second principal component). We examine one channel through which competition between HFT firms can impact market concentration by investigating how it relates to the viability of trading venues. Specifically, we investigate whether higher correlation in HFT strategies on a specific trading venue increases the percentage of time that the venue features the best prices or the narrowest spreads, which we view as measures of the viability or competitiveness of the venue. We find that smaller trading venues are more viable while the largest trading venue exhibits the opposite picture when HFT firms are more competitive.

The remainder of this paper proceeds as follows. The next section discusses the most relevant literature, while Section 3 provides information about our sample and data, and indicates how we identify HFT firms. Section 4 uses a principal component analysis to gain insights into competition between HFT firms as well as the nature of their strategies. Section 5 investigates the question whether correlated HFT activity exerts a systematic influence on short-interval returns or the volatility of stocks. Section 6 studies how correlated HFT activity interacts with competition between trading venues, and Section 7 presents our conclusions.

2. Review of Relevant Literature

Our paper joins a rapidly expanding body of literature on HFT in financial markets.[2] Among the theoretical contributions are those of Ait-Sahalia and Saglam (2014), Hoffmann (2014), Biais, Foucault, and Moinas (2015), Foucault, Hombert, and Rosu (2015), Han, Khapkp, and Kyle (2015), Jovanovic and Menkveld (2015), and Rosu (2015). In particular, Jarrow and Protter (2012) show that, when HFT firms respond to common signals, their correlated activity affects the market price, increasing market volatility and generating abnormal profits for these firms. While we indeed show that the strategies of HFT firms are correlated, we find that the market-wide measure of cross-sectional correlation between HFT strategies affects the volatility of most stocks negatively, not positively. Budish, Cramton, and Shim (2014) as well as Menkveld and Zoican (2014) model HFT firms that pursue heterogeneous strategies in the market, which we also document empirically.

Many empirical contributions focus on intraday analysis of aggregate HFT behavior (e.g., Carrion (2013), Hasbrouck and Saar (2013), Hirschey (2013), Brogaard, Hendershott, and Riordan (2014), Jarnecic and Snape (2014), and Kirilenko, Kyle, Samadi, and Tuzun (2014)). Several papers use data on trading by individual HFT firms, rather than aggregate behavior, to investigate HFT strategies (Hagstromer and Norden (2013), Baron, Brogaard, and Kirilenko (2014), Benos and Sagade (2015), and Hagstromer, Norden, and Zhang (2014)), although they

[2] For recent surveys on the topic of HFT see Jones (2013) and Goldstein, Kumar, and Graves (2014).

focus predominantly on whether HFT firms are demanding or supplying liquidity. Clark-Joseph (2014) uses index futures data from the CMA to examine exploratory trading by HFT firms.

More closely related to our analysis, Hagstromer and Norden (2013) find that the activity of HFT firms that predominantly supply liquidity mitigates intraday volatility, which complements our finding pertaining to the relationship between HFT inter-firm competition and volatility.[3] Breckenfelder (2013) and Brogaard and Garriott (2014) examine one aspect of competition: the entry and exit of HFT firms. Specifically, Brogaard and Garriott analyze data from one alternative trading system in Canada and show that new entrants take volume away from incumbents even as they increase the overall market share of HFT firms—although the effects decrease with each successive entry. They also find that market liquidity improves after the entry of an HFT firm and deteriorates after an exit (especially when there are only one or two HFT firms trading in a given stock). Breckenfelder uses data from the Stockholm Stock Exchange and finds the opposite result: deterioration of liquidity for entries of HFT firms and improvement for exits. Menkveld (2013) examines the strategy of one HFT firm and makes the case that this particular firm enhances the viability of a new trading venue, which is related to our result that competition between HFT firms is positively related to the viability of smaller trading venues and negatively related to the viability of the dominant trading venue.

Chaboud, Chiquoine, Hjalmarsson, and Vega (2014) investigate algorithmic trading on the foreign-exchange EBS platform. While they observe only the aggregate trading of algorithmic traders, they create a measure of correlated algorithmic trading by comparing the frequency with which algorithmic traders are on both sides of a transaction with a benchmark model under the assumption of independent matching of algorithmic traders and humans. The higher their measure, the fewer triangular arbitrage opportunities are observed in the market (and hence the more efficient are prices), but they find no relationship between their measure and the autocorrelation of returns (or excess volatility). Benos, Brugler, Hjalmarsson, and Zikes (2015) use transactions data from the London Stock Exchange to study how ten HFT firms that are

[3] Brogaard, Riordan, Shkilko, and Sokolov (2014) identify short intervals with large price movements and show that NASDAQ HFT firms in the aggregate supply rather than demand liquidity during these intervals, hence possibly dampening volatility. Hasbrouck and Saar (2013) also find that aggregate HFT activity appears to lower the intraday volatility of NASDAQ stocks.

regulated by the UK Financial Conduct Authority interact with each other. They use vector autoregressions to show that there is a positive dynamic relationship between HFT firms: aggressive buying by one firm tends to follow aggressive buying by another firm (and similarly for aggressive selling), and firms tend to trade in response to the past trading activity of other firms. Benos et al. construct a measure of concurrent directional trading on the part of their ten HFT firms that is meant to capture correlated behavior and show that it has positive contemporaneous and lagged relationships with returns.

Lastly, our paper joins several other papers that use Canadian order-level data to investigate HFT-related issues. In particular, Malinova, Park, and Riordan (2014) study the cost of trading around changes in market structure that affected mainly HFT firms and other algorithmic traders, Comerton-Forde, Malinova, and Park (2015) study the nature of liquidity provision around changes in dark trading regulation, and Korajczyk and Murphy (2015) examine HFT liquidity provision to large institutional trades.

3. Sample and Data

Our data come from the Investment Industry Regulatory Organization of Canada ("IIROC"), which is the national self-regulatory organization that oversees all dealers and equity trading venues in Canada (both exchanges and alternative trading systems). All trading venues are required to provide data feeds to IIROC, which performs both real-time and post-trade market surveillance of trading activities. Traders need to obtain an IIROC membership to directly connect to trading venues in Canada, and IIROC admits only security dealers as members. Other financial firms, such as asset managers, banks, insurance companies, and proprietary trading firms can trade through dealers' brokerage arms or via direct market access (DMA) arrangements provided by dealers. DMA arrangements allow non-dealer trading firms to directly access various trading venues without having to hand over their orders to brokers for execution.

During our sample period (June 2010 through March 2011), Canada has five trading venues organized as electronic limit-order books that trade stocks listed on the Toronto Stock Exchange: Alpha ATS Limited Partnership (ALF), Chi-X Canada ATS (CHX), Omega ATS

(OMG), Pure Trading (PTX), and the Toronto Stock Exchange (TSX).[4] Trading on crossing networks ("dark trading") in Canada during our sample period is limited to essentially one dark pool (MATCH Now) with no more than a 3% market share.[5]

3.1 Sample

Our empirical work is carried out on 30 trading days that are selected to capture variation across market conditions. We rank the daily returns of the S&P/TSX Composite Index from June 2010 through March 2011, and select the 10 worst days (down days), the 10 best days (up days), and the 10 days closest to (and centered on) zero return (flat days). In other words, we take the two extremes in terms of days in which the market went up or down the most as well as the days with the least return movement. This design allows us to examine whether the correlation structure of HFT strategies depends on market conditions (as summarized by the daily return on a broad index).

Our sample stocks consist of 52 constituents of the S&P/TSX 60 index, which accounts for approximately 73% of Canada's total equity market capitalization. Eight stocks are excluded from the 60-stock index, as they were converted from income trusts to corporate structures (five stocks), were delisted (one stock), had their symbols changed (one stock), or were listed for less than one year before the start of our sample period (one stock). For further analysis, we sort the sample by market capitalization and divide it into two subsamples, Large and Small, with equal numbers of stocks in each subsample. Panel A of Table 1 presents summary statistics for the sample stocks used in the study.[6] The mean market capitalization is 19.4 Billion Canadian Dollars (henceforth, CAD), with an average stock price of 39.1 CAD, and average daily volume of 78.2 Million CAD. The panel clearly shows that our sample period encompasses three distinct market conditions: the average daily returns of stocks on down, flat, and up days are −1.72%, −0.08%, and 1.66%, respectively.

[4] Alpha became a stock exchange on April 1st, 2012. In July 2012, Alpha was acquired by the TMX Group, which also owns TSX. During our sample period, however, Alpha and TSX were independent trading venues.
[5] MATCH Now also provides a real-time data feed to IIROC and is included in our data. Liquidnet Canada, another dark pool, executed only a few trades each day during our sample period. As a result, IIROC did not require it to participate in the real-time centralized data feed, instead requiring it to submit trade information manually at the end of the trading day.
[6] Data on stock characteristics are obtained from the Summary Information Database of the Canadian Financial Markets Research Centre (CFMRC).

3.2 Data

The order-level data we obtain from IIROC cover all Canadian trading venues, and contain information about order submissions, cancellations, modifications, and executions with 10-millisecond time stamps. The time stamps from all trading venues are synchronized with the regulator's time stamps and reflect the local time at which the message (which is a general term used for submissions, cancellations, and executions of orders) is processed. The record of each message contains several pieces of information: ticker symbol, order side (buy or sell), trading venue, price, total quantity, non-displayed quantity, broker ID, trader ID, order type (e.g., client orders, inventory trading), various flags (e.g., short sell, market on close, opening trade), and order/trade ID. Trade messages are identified as buyer-initiated or seller-initiated. Events in the order's life, including modification, partial fill, full fill, and cancellation are identified with the same order ID.

One advantage of our data is that the same trader IDs are used on all trading venues in Canada. Most HFT firms in Canada do not operate as licensed dealers but rather gain access through DMA (Direct Market Access) arrangements with one or multiple dealers. We obtain tables that identify the trader IDs of all trading firms that use DMA arrangements. Hence, we are able to accurately detect the activity of each HFT firm on all trading venues irrespective of whether it trades via multiple DMA arrangements with dealers or uses multiple trader IDs.

3.3 Identifying HFT Firms

We use our own procedure to identify HFT firms rather than adopting a classification provided by an exchange. Since we have both trader IDs and a mapping of the trader IDs to the firms, we aggregate all trader IDs that belong to the same firm. We operate at the firm level because there are no rules (to the best of our knowledge) that guide how firms use Trader IDs. One possible concern is that firms assign multiple Trader IDs to the order flow of an algorithm and send orders via DMA arrangements with multiple dealers to make it more difficult for outsiders to ascertain their activity. Routing via multiple dealers could also be driven by the desire to limit dependence on one dealer. We choose to work at the HFT-firm level to make our analysis robust to whatever gaming could be going on in terms of the firm's discretionary assignment of Trader

IDs to its orders.[7] If an HFT firm uses multiple algorithms and actually designates a separate Trader ID to each algorithm, our procedure will lump them together, although the principal component analysis we implement in Section 4 could potentially tease out these separate strategies.

We use an out-of-sample procedure to identify HFT firms to ensure that the identification is exogenous to the empirical analysis we carry out in the paper. Specifically, our procedure uses data from the 22 trading days in September 2010 to identify the HFT firms, while we carry out the empirical work on the 30 down, flat, and up days as described in Section 3.1.[8] To reduce the number of trading firms that we scrutinize more closely, we rank firms on several criteria and look at those that rank highly on at least one criterion (for example, the number of times per day that the firm's inventory position crosses zero).[9] Our use of multiple criteria is motivated by our desire not to limit our sample to firms that pursue a certain strategy that requires a particular profile (e.g., a high orders-to-trades ratio), but rather to allow for HFT firms that implement a variety of algorithms. We emphasize that some firms subsequently identified as engaged in HFT rank highly on only one criterion.

To help us identify the HFT firms from among the trading firms on our short list, we use two qualitative criteria. First, trading firms that participate in the Toronto Stock Exchange Electronic Liquidity Providers Program are categorized as HFT firms. This is a program that offers fee incentives to firms that use proprietary capital and high-frequency electronic trading algorithms to provide liquidity on the exchange.[10] Second, we search in newspapers and on the web for information about the firms. Some firms have web sites on which they state that they engage in proprietary trading or explicitly state that they pursue high-frequency strategies. Other

[7] For dealers that may have brokerage arms in addition to proprietary trading operations, we exclude orders and trades made in the capacity of an agent, and include in our measurement of their HFT activity only those orders and trades identified in the order type field as proprietary activity.

[8] To ensure that the firms we identify operate during our sample period, which consists of 30 down, flat, or up days (none of which took place during September 2010), we require HFT firms to be active (i.e., to trade) in at least 10 out of the 30 days.

[9] We compute several measures for each trading firm: (1) the orders-to-trades ratio (defined as submissions and cancellations of limit and marketable orders divided by trades), (2) the number of times the firms' intraday inventory positions cross the end-of-day positions (or zero), (3) cross-trading-venue activity in the same time-stamp or a neighboring time-stamp, (4) median time-to-cancellation of non-marketable limit orders, (5) and number of daily trades.

[10] ELPs are either independent proprietary trading firms or proprietary trading desks within large banks or financial firms, and the program requires that they trade passively at least 65% of their volume.

firms are mentioned in newspaper articles as engaging in HFT. We use this information to ensure that we identify HFT firms even if they do not rank highly on some of the quantitative data criteria. Our procedure results in 31 firms identified as HFT firms. Some of these firms have a direct market access (DMA) arrangement with dealers while others are dealers engaged in proprietary trading. Our goal in this identification procedure is to ensure that our sample contains as complete a set of HFT firms that operate on these trading venues as possible.

Panel B of Table 1 provides summary statistics for the 31 HFT firms as well as for categories formed on market share of volume and orders-to-trades ratio. Specifically, MS1 consists of four firms with market share greater than 4%, MS2 consists of six firms with market share between 1% and 4%, and MS3 consists of the rest of the firms. Overall, these HFT firms have 46.4% of the market share in terms of volume. The mean (median) number of times the intraday inventory position of a firm in our sample crosses its end-of-day inventory increases with market share: 4.3 (1.8) for MS3, 13.4 (4.3) for MS2, and 73.0 (70.1) for MS1. Overall, these HFT firms trade often and send many messages to the market. The average number of daily trades of a firm in our sample is 19,445, but it ranges from 102,035 for firms in MS1 to 2,489 for firms in MS3. Similarly, the number of daily messages a firm sends to the market (submissions and cancellations of non-marketable limit orders as well as the number of marketable orders) is 1,063,974, ranging from 5,496,423 for firms in MS1 to 239,157 daily messages on average for firms in MS3.

3.4 Measures of HFT Strategies

Throughout the paper, we use the terms “strategy,” “activity,” and “behavior” of HFT firms somewhat interchangeably. While a strategy is often thought of as a plan of action, we are not privy to the specifics of the algorithms employed by each HFT firm. As empirical researchers who can measure only the outcomes of a strategy (e.g., submissions and cancellations of orders), although it is natural to recognize that the actions we observe are the manifestations of a strategy, and hence treat measures of these actions as representing the strategy. Furthermore, there are multiple ways in which one can operationalize the activity of HFT firms for measurement purposes. For example, one could examine measures that count actions only if they are intentionally initiated by the firm at the instant the action is taken (e.g., the submission or cancellation of an order), or look at trading by an HFT firm and on that basis include executed limit orders for which the timing of execution is not directly under the control of the firm.

The main measure of HFT firm activity that we employ emphasizes actions initiated by the firm. Our measure, MSG, is defined as the sum of three components: the number of submissions of non-marketable limit orders, the number of cancellations of non-marketable limit orders, and the number of marketable order executions. Hence, MSG for an interval (say one second) describes the total number of messages that the HFT firm sends to the market during the interval to initiate a change in its position (either in terms of presence in the limit-order book or to transact immediately).[11] In Sections 5 and 6, we add two additional measures to ensure the robustness of our conclusions. The first measure, TRD, is the number of trades made by an HFT firm in an interval. These trades could occur as the result of submitting marketable orders or the execution of previously submitted limit orders that rested in the book, and therefore TRD measures the change in inventory position of the HFT firm. The second measure, LMT, is comprised of all submissions and cancellations of non-marketable limit orders, and hence describes the actions the firm takes in the interval solely to change its presence in the limit-order book.

Before proceeding to the main analysis, however, we use simple correlations to establish two stylized facts that are directly relevant to the economics of HFT strategies and hence impact the design of our tests in the rest of the paper. First, we look at whether the predominant correlation in HFT strategies involves directional activity or total activity. Second, we examine whether correlations in HFT strategies differ on days in which the market experiences large positive or negative returns.

The first stylized fact is motivated by empirical studies on herding in financial markets (e.g., Wermers (1999), Khandani and Lo (2007, 2011), Choi and Sias (2009), Pedersen (2009),

[11] An AMEND order type is considered as two messages, a cancellation and a resubmission, for the purpose of our measures. Our measures include both non-displayed and displayed orders. A refresh of an iceberg order (when the displayed part is executed and shares from the non-displayed part become displayed) does not lead to a change in price or quantity, hence only the initial order is counted. This standardizes the treatment for the various order types according to their economic functions, and lets us summarize all activity in terms of submissions and cancellations of non-marketable orders and executions of marketable orders.

Brown, Wei, and Wermers (2014)) that recognizes the destabilizing influence that simultaneous actions in one direction (buying or selling) by institutional investors can have on asset prices. If simultaneous actions in one direction are also of critical importance with respect to HFT firms, one should look at how directional HFT activity (buy minus sell orders) rather than total activity (buy plus sell orders) correlates across firms.

Figure 1 compares the magnitudes of the cross-sectional (Panel A) and time-series (Panel B) correlations for total and directional HFT activity. Our cross-sectional correlation measure provides information on whether the strategies of HFT firms are correlated across stocks at a given time. For each 1-second time interval, we compute the correlation coefficient between the activities of pairs of HFT firms across the stocks in the sample, and average the correlations for all pairs of firms in a certain group (e.g., our market share subgroups MS1, MS2, and MS3). The time-series correlation measure provides information on whether the strategies of HFT firms are correlated over time for a given stock. For each stock, we compute the correlation coefficient between the activities of pairs of HFT firms over all time intervals, and average across all pairs of firms in a certain group.[12]

We observe a striking result: the correlations involving total activity are four to ten times the magnitude of those involving directional activity. For example, the 1-second cross-sectional correlation of the total activity measure (MSG) for all HFT firms is 0.200 versus 0.023 for the directional activity measure (NetMSG), and, similarly for the firms with the highest market share, the MSG correlation is 0.360 compared with 0.045 for NetMSG. The picture that emerges is consistent with what one would expect for HFT firms that operate simultaneously on both sides of the market, rather than pursuing very strong directional trading for long periods of time. In other words, it appear as if much of the HFT firms' activity involves either placing buy and

[12] Specifically, the cross-sectional correlation measure is computed as follows. Let Y_{itk} be the measure of HFT firm activity for stock *i*, interval *t*, and firm *k*. If *k* and *l* are two HFT firms, then $Cor_t(Y_k,Y_l)$ is the correlation for a specific interval *t* over the stocks in the sample between the measures of the two firms. There are $0.5*N*(N-1)$ such correlations for a group of *N* HFT firms in each time interval *t*, and our cross-sectional correlation, $Cor(Y_t)$, is computed as the average of these correlations for all pairs of HFT firms. The time-series correlation measure is computed as follows. If *k* and *l* are two HFT firms, then $Cor_i(Y_k,Y_l)$ is the correlation over all intervals in the sample period between the measures of the two firms for stock *i*. There are $0.5*N*(N-1)$ such correlations for a group of *N* HFT firms, and our time-series correlation, $Cor(Y_i)$, is computed as the average of these correlations for all pairs of HFT firms.

sell orders simultaneously or buying and selling very rapidly within the same 1-second interval.[13] Many HFTs presumably design strategies to interact with uninformed order flow. Market microstructure theory suggests that uninformed order flow is non-directional (in contrast to informed order flow), which could potentially explain our finding of low directional correlations of HFT strategies. Such low correlations are somewhat reassuring in terms of market stability because they suggest that directional herding by HFT firms is unlikely to destabilize markets on a regular basis.[14] Given these very low directional correlations, we focus in the rest of the paper on analyzing correlations in total activity.

The second stylized fact we present is motivated by the concern that HFT firms react to adverse market conditions (in terms of declining prices) by changing their strategies and hence bringing on greater fragility. The design of our study is meant to enable us to analyze this issue in greater detail. Specifically, we carry out the empirical work on the ten days with the largest negative index returns from June 2010 through March 2011 (with average daily returns of −1.72%), the ten best days (with average daily returns of 1.66%), and ten days in which the index moved very little (with average daily returns of −0.08%). Figure 2 presents the cross-sectional and time-series correlations of the total HFT activity measure (MSG) separately for these down, flat, and up days, alongside the correlations for the entire 30-day period.

We observe that the cross-sectional correlations for the largest HFT firms (MS1) are almost identical on down, flat, and up days (0.355, 0.361, and 0.362, respectively). Similarly, the time-series correlations exhibit similar patterns on down, flat, and up days (e.g., 0.055, 0.057, and 0.054, respectively, for the set of all HFT firms). We do not observe that these three distinct market environments result in correspondingly varying correlations.[15] This stylized fact, like the previous one concerning lack of correlation in directional trading, is reassuring in terms of

[13] The same conclusions apply when we examine directional versus total activity using other measures of HFT strategies (e.g., TRD or LMT).

[14] It is reasonable to conjecture that this result would hold for a variety of market conditions during normal times. Still, we cannot rule out the possibility that directional trading could become highly correlated if the market experiences extreme stress over a very short timeframe, though we are not aware of such an episode in the Canadian market.

[15] The same conclusion also applies to other measures of HFT strategies (including TRD and LMT). Hasbrouck and Saar (2013) analyze the impact of a low-latency activity measure, which they view as a proxy for the activity of HFT firms, in two periods: one in which the NASDAQ Composite Index went up 4.34% and another in which the index went down 7.99%. Like us, they find that the impact of their proxy on market quality was similar in both periods.

market fragility, although it does not preclude the possibility that the correlations would increase during times of extreme stress such as the Flash Crash in the US. The absence of such extreme episodes in Canadian markets, however, prevents us from examining this possibility empirically. Given these findings, we present in the rest of the paper the analysis for the entire sample period rather than breaking the results down by market conditions.[16]

4. Principal Component Analysis of HFT Strategies

4.1 Correlated Strategies and Competition

The HFT industry is characterized by extreme secrecy. Most HFT firms are private and hence reveal no financial or operating information, and information about the profitability of proprietary trading desks of larger, publicly-listed firms is not disclosed separately to the public. Firms use restrictive covenants in employment contracts as well as litigation to prevent or deter employees from taking code used for trading algorithms when they leave. In general, HFT firms do not reveal information about the operation or even the objectives of their algorithms beyond talking generally about concepts such as "liquidity provision" and "arbitrage."

How can one go about investigating the extent of competition in such an industry? While examining the economic profits of firms would be ideal, it is practically impossible to examine the profitability of most HFT firms or trading desks because the costs associated with hiring and retaining the individuals who develop the algorithms as well as other operating costs are simply unavailable. Researchers can only ascertain firms' net trading revenues (or what is left when a security is bought and then subsequently sold by the firm) from datasets like ours that describe trading activity. Figure 3 shows a histogram of the average daily net trading revenues per stock of the 31 HFT firms in our sample.[17] Some firms have positive trading revenues while others have negative trading revenues. Even if we were to add the liquidity rebates that HFT firms

---

[16] We verified that the results of all the tests we present do not differ materially on down, flat, or up days. We also categorized all 10-minute intervals into three categories based on volatility and three categories based on volume. We computed the correlation measures separately for each category, and looked at whether correlations in HFT strategies differ when we focus on intraday periods (10-minute intervals) distinguished by higher volume or volatility. We could not discern any clear patterns across the categories.

[17] The net trading revenues are computed as follows. We sum the positive cash inflows (how much they get from selling shares) and cash negative outflows (how much they pay for buying shares) for each HFT firm in each stock and on each day. We then assume that shares left at the end of the day are "liquidated" using the end-of-day midquote or closing price, and start every day with zero inventory (see Brogaard, Hendershott, and Riordan (2014)).

obtain from trading venues, which may turn some of the negative revenues into positive ones, we would still be unable to tell whether this trading revenue is sufficient to make economic profits.[18] It is important to stress that revenues without costs cannot be used to make a correct inference about the competitiveness of any industry, including the HFT one. Without the cost structure, we need to rely on other methods to gather evidence on the extent of competition.

The number of competitors is a crucial determinant of competition in an industry. Its importance goes beyond the traditional (static) equilibrium concepts. Since firms must share the collusive profits, a higher number of competitors results in each of them gaining less. As a result, the gain from deviating increases and the long-term benefit of maintaining collusion is reduced. Even in dynamic collusion models, therefore, coordination is more difficult with a larger number of firms (see, e.g., Ivaldi, Jullien, Rey, Seabright, and Tirole (2003)). Because the number of direct competitors in an industry is crucial to the absence of collusion, defining industry boundaries is of paramount importance. Hoberg and Phillips (2016), for example, design a new classification scheme using text-based analysis of firm product descriptions to define industries. We pursue a similar objective: classifying HFT firms that directly compete with each other by identifying firms that follow similar strategies. This enables us to find out exactly how many firms compete in certain "products" (or strategies), and to establish whether the major players are monopolists that pursue markedly distinct strategies or else that multiple firms compete with each other.

We identify competing firms by considering how manifestations of strategies (e.g., the messages HFT firms send to the market, their trades) correlate between HFT firms. The more highly correlated are the strategies of two firms, the more likely it is that they compete in pursuing the same profit opportunities (i.e., respond to the same trading signals) and follow the same business model. In principle, the correlated behavior of firms can also potentially characterize collusive behavior. Specifically, Green and Porter (1984) describe a situation in which collusion results in recurrent episodes of patterns in product prices and firm profits. The assumptions of the model, however, do not fit the HFT industry. For example, the industry in the

[18] Liquidity rebates are payments made by trading venues to HFT firms that take the passive side of trades (i.e., whose limit orders are executed by incoming marketable orders). See Malinova and Park (2014) for a discussion of liquidity rebates in the Canadian market.

Green and Porter model is assumed to be stable over time. However, the HFT industry keeps changing all the time, with low barriers to entry enticing new start-ups to enter while others exit. Second, firms in the model are not able to engage in product differentiation. Coming up with a better algorithm, on the other hand, is the hallmark of the HFT industry, where every player attempts to differentiate itself by having better algorithms that are kept secret. The model assumes that information about the industry and its environment, like the competitors' cost functions, is public. This, of course, is the antithesis of the HFT industry, where almost all firms are private rather than public, and practically all information about a firm's operation is considered proprietary.

The above considerations suggest that correlated behavior in the HFT industry is unlikely to be a manifestation of collusive behavior. As such, we use correlated strategies as a measure of competition and as a tool with which to define the subsets of HFT firms that compete with each other directly in a particular strategy. The higher the correlation, the greater the extent to which HFT firms compete for the same market stimuli (or profit opportunities), and our goal is to investigate this competition and how it interacts with the market environment.[19]

Studies that use datasets of aggregate HFT activity (e.g., Carrion (2013), Brogaard, Hendershott, and Riordan (2014), Jernecic and Snape (2014)) essentially assume that all HFT firms are in the same business. We already know from Hagstromer and Norden (2013) that an HFT firm may pursue a strategy that differs from that of another firm along a particular dimension: percentage of passive trading. We go even further by using a data-driven methodology (principal component analysis) to decompose the correlation matrix of the HFT strategies. This analysis helps us understand HFT strategies in several ways. First, it tells us

[19] Our focus on the number of competitors that pursue correlated strategies is driven by our conclusion that other characteristics that can affect the sustainability of collusion do not apply to the HFT industry (see Ivaldi et al. (2003)). For example, collusion is more difficult to sustain if there are low barriers to entry, both because abnormal profits would entice additional players to enter the market and because the prospect of future entry reduces the effectiveness of retaliation. The constant entry and exit of HFT firms suggest that there are few barriers to operating. Transparency of the industry and its firms facilitates collusion. In classic collusion models, the "price" of the products needs to be observed or readily inferred from market data in order to identify deviations by market participants. For HFTs, where it is difficult to define the product (e.g., liquidity, or lack or arbitrage opportunities) and the trading strategies are complex (often involving hundreds of orders to effect a single trade), it is also difficult to detect a deviation of a firm. Lastly, innovation, the hallmark of the secretive computer algorithms of HFT firms, makes collusion more difficult to sustain because it allows a firm to gain a significant advantage over its rivals, reducing the value of future collusion.

whether there are underlying common strategies that multiple firms follow and that represent much of the variability in the HFT firms' activity. Second, it shows us how close the strategy of each firm is to these underlying common strategies, and therefore helps establish the extent of competition in pursuing each underlying common strategy. Third, we are able to analyze the empirical representation of these underlying common strategies (i.e., the principal component scores) and how they relate to the market environment to gain a better understanding of these strategies.

4.2 HFT Firm Loadings on the Principal Components

Principal component analysis is essentially a data reduction technique. In our application, we think about the HFT firms' strategies as the "variables" that we seek to summarize. The input for our analysis is a matrix of data in which there are 31 columns (one for each HFT firm) and the number of rows is equal to the number of stocks times the number of intervals in the sample period. In other words, the stocks are stacked one on top of another and we analyze both the time-series and cross-sectional sources of variation in HFT activity together. The measure of HFT activity that we use to characterize their strategies is MSG, which is comprised of all messages they actively send to the market in an interval (submission of non-marketable limit orders, cancellation of non-marketable limit orders, and marketable limit orders that result in trade executions). We carry out the analysis using 1-second intervals.[20]

The principal component analysis decomposes the correlation matrix (rather than the covariance matrix) of the strategies. In other words, the variables that describe the activity (or strategy) of each HFT firm are standardized to have zero mean and unit variance. By standardizing the variables, we eliminate the possibility that one of them would dominate the procedure because it has much higher scale or range. From an economic perspective, this means that our procedure gives the very active HFTs the same weight as any other HFT (each contributing one unit of variance to the total variance). This helps us focus on the resemblance of strategies to one another even if some firms are larger than others.

[20] We provide the results using only the MSG measure to economize on the presentation in terms of the number of tables and the length of the discussion. We carried out the principal component analysis for other measures as well, and the insights we obtain are not very sensitive to the particular measure used. Similarly, we also used 10-second and 60-second intervals, and the results were very similar to those we present in this section for 1-second intervals.

When conducting the analysis we must first choose how many principal components to retain for subsequent analysis.[21] From an economic perspective, the issue at stake is how many separate underlying strategies are common to a significant number of the HFT firms in our sample. As usual with such data-driven methodologies, that determination is made based on patterns observed in the data. Specifically, we conduct a scree test by plotting the eigenvalues of the principal components and looking for a natural break. Since there are 31 HFT firms, each firm contributes 100/31=3.2% of the variance. However, meaningful principal components would naturally explain more of the variance, and our analysis suggests that the first principal component explains 11.66% of the variance, the second 4.5%, and the third 4.02%, together accounting for over 20% of total variation. We observe that further principal components account for less than 4% of the variance each, and there seems to be a natural break after three components. Hence, we extract three principal components for subsequent analysis.

The second choice when implementing the methodology involves determining the rotation of the principal components. The rotation is meant to help in interpreting the loadings, which are the coefficients of each HFT firm on each of the principal components. We utilize the commonly used varimax orthogonal rotation. The factor loadings using this rotation have a simple interpretation: they are equivalent to bivariate correlations between the HFT firms' strategies and the underlying common strategies represented by the principal components. Our conclusions are robust to using other rotations.[22]

Table 2 presents the loadings from the principal component analysis of the MSG measure. Each principal component can be viewed as representing a certain underlying common strategy that multiple HFT firms follow and which results in the correlated behavior of the firms. In other words, we use the terms "underlying common strategy" and "principal component" interchangeably, although strictly speaking the principal component is a linear combination of

[21] In a principal components analysis, the first component accounts for the largest portion of total variance in the observed HFT strategies and successive principal components account for portions of the variance that were not accounted for by previous principal components.

[22] In particular, we consider whether our conclusions change when using an oblique rotation (promax) that allows the principal components to be correlated. The results we obtain with this rotation are very similar to those with the orthogonal rotation: the loadings are similar in magnitude and the same HFT firms load on the same principal components. Furthermore, the regressions on component scores that we present in Section 4.3 yield similar conclusions irrespective of whether an orthogonal or an oblique rotation is used.

the MSG measures of the 31 HFT firms and therefore represents correlated behavior in terms of their actions. As before, we treat these actions as the outcomes of strategies, and hence the principal component represents a common element in the firms' strategies, which most likely respond to the same market stimuli and represent a similar business model. The larger (i.e., closer to 1) the loading of a particular HFT firm on a principal component is, the greater is the similarity of the HFT firm's strategy to that underlying common strategy. For each principal component, the loadings are sorted from the most positive to the most negative, and the 31 firms are represented by F01 through F31. Looking at the first principal component, for example, we observe that firm F14 has a very large positive loading (0.76) on the first principal component, signifying high correlation with the underlying common strategy represented by this component.

While the choice of a cutoff for the magnitude above which a loading is considered economically significant is somewhat arbitrary, it helps to have some cutoff in mind when considering the results. We consider loadings significant from an economic standpoint if they are greater than 0.35 and mark them with an asterisk in the table. Eight HFT firms have significant loadings on the first principal component, suggesting that it represents a strategy that is common to these eight firms. Similarly, there are six firms with significant loadings on the second principal component, ranging from 0.71 to 0.35, and four HFT firms with significant loadings on the third principal component. Four firms appear to follow more than one strategy: F31 and F28 (with significant loadings on the first and second principal components) and F17 and F20 (with significant loadings on the first and third principal components).

Our principal component analysis therefore shows that there are three underlying common strategies that are followed by multiple firms. On the other hand, 17 HFT firms do not appear to pursue one of these common strategies but rather carry out more unique strategies. How important are the common strategies relative to the unique strategies? The 14 firms that follow the three common strategies represent most of the HFT activity: 96.21% of the messages that HFT firms send to the market and 78.97% of the volume they trade. Therefore, we find that competing (as opposed to monopolist) HFT firms generate most of the HFT activity.

The challenge in interpreting the results of a principal component analysis lies in understanding the economic nature of the underlying common strategies represented by these principal components. It is possible to obtain relevant information by observing whether firms that load more heavily on a certain component share something else in common. In Panel B of Table 1 we categorize firms into groups according to market share or the orders-to-trades ratio. Five out of the six HFT firms that load on the second principal component have high market shares (i.e., they are in MS1 or MS2), suggesting that this could be a market-making strategy. In contrast, the two other principal components are dominated by firms with high orders-to-trades ratios: six out of the eight firms with significant loadings on the first principal component and all four firms with significant loadings on the third principal component are in OT1 or OT2, suggesting that these underlying common strategies require frequent submissions and cancellations of limit orders. To gain greater insight, however, we need to turn to a more structured investigation of these principal components.

4.3 Regressions on Principal Components' Scores

A more direct way of examining the nature of the underlying common strategies represented by the principal components is to investigate another output of the principal component analysis: the component scores. Each principal component is essentially a linear combination of the observed variables (the 31 HFT firms' MSG measures), and the component scores are computed from the observed variables using the estimated loadings provided in Table 2 as weights. There is a separate score for each principle component in each time interval, allowing us to regress them on various market variables and examine the relationships between these component scores and the market environment.

We use a set of 14 variables that describe the market environment. The first two represent the degree of integration of the Canadian market. Specifically, TimeAorB is the percentage of time that the three trading venues with the highest market share post the same bid or ask prices, while TimeSprd is the percentage of time that the three trading venues have the same bid–ask spread. The next five variables represent the state of liquidity in the market as a whole (aggregated across all trading venues): total depth at the Market-Wide Best Bid or Offer (henceforth, MWBBO), total depth up to 10 cents from the MWBBO, depth imbalance at the MWBBO (defined as the absolute value of the difference between the number of shares at the bid and at the ask), depth imbalance up to 10 cents from the MWBBO, and percentage MWBBO spread.[23] The next three variables represent market conditions in the interval: return (computed from the last transaction price in each interval), volatility (computed as the absolute value of return), and the average time between trades in the interval. The last four variables represent information about HFT: the number of trades in which the HFT firms supply liquidity in the interval, the aggregate inventory position of all HFT firms in Canadian dollars (cumulative from the beginning of the day and assuming that all of them start the day with zero inventory), the number of shares traded among the HFT firms in the interval (#HFT), and the number of shares traded between the HFT firms and others (#NoHFT).

We are interested primarily in contemporaneous relationships, which is why the first three columns of Panel A of Table 3 show the results of regressions in which we line up the component scores with observations of the market environment over the same interval. We also provide, in the same table, regressions in which the component scores are regressed on lagged and lead values of the market variables. The idea behind comparing lagged and lead regressions with the contemporaneous relationship is to see whether strategies appear to lead or respond to changes in the market environment. Looking across the columns, it is striking how the signs of the coefficients on almost all variables are the same in the contemporaneous, lagged, and lead regressions. There are some exceptions (e.g., the coefficients on #HFT and #NoHFT on the third principal components), but they seem to involve coefficients that are very small and hence do not represent material influences. The similarity between the contemporaneous, lagged, and lead regressions suggests that HFT strategies respond to certain market or book conditions that surround the intervals. In other words, we do not observe evidence consistent with the idea that HFT strategies create a pronounced change in the market environment such that the conditions in the following interval differ from the conditions in the preceding or contemporaneous intervals.

[23] All measures are time-weighted. Messages are time-stamped to 10 milliseconds. There are orders that are submitted and cancelled (or executed) within the same time stamp. When we consider orders that stay in the book to provide liquidity (e.g., for our measures of depth), we assume that a submitted order stays in the book for 10 milliseconds. The exceptions are the following special order types: immediate or cancel orders (IOC), fill or kill orders (FOK), all or nothing orders (AON), dealer's AG orders (that are generated to fulfil their market making obligations and execute against an incoming order), odd lot orders (OL), and marketable orders, which are assumed to be executed or canceled upon arrival to the market and hence are not added to the limit-order book.

Panel B of Table 3 shows just the coefficients on the contemporaneous regressions together with t-statistics computed from double-clustered standard errors (along both the stocks and intervals dimensions) to focus our attention on the most significant relationships. We noted above that HFT firms that load on the second principal component have high market shares, suggesting that it could represent a market-making strategy. We observe that the coefficients on the market integration variables (TimeAorB and TimeSprd) are both positive, which means that such a strategy is more active when prices and spreads are aligned across the trading venues, and may represent times at which market-making activity is most profitable (i.e., when market-making firms earn the spread rather than lose to changing prices). This underlying common strategy is more active when there is greater depth throughout the book but less depth at the best prices, creating ample opportunities for market making without excessive risk. Finally, the coefficient on HFTliqsup is positive and very large in magnitude (compared to the same coefficient in the regressions of the other two principal components), which means that this strategy is more active when HFT firms supply more liquidity by trading passively, the hallmark of traditional market making. Hence, the regression results support the interpretation that this underlying common strategy represents market making.

The underlying common strategies represented by the other two principal components appear to take advantage of very different trading opportunities. HFT firms that load on the first principal component are more active when prices are more volatile, and there is both greater imbalance and less depth at the best prices. These are times at which some trading venues post better spreads than others (a negative coefficient on TimeSprd) and there appears to be a need to move liquidity across trading venues, suggesting that this underlying common strategy probably represents a cross-venue arbitrage strategy. Firms that load on the third principal component are also more active when prices are more volatile, but also at times at which there is less depth throughout the book and the best prices are not the same across the trading venues (negative coefficient on TimeAorB). This pattern of relationships with the market environment could represent either a price arbitrage strategy or a short-horizon directional strategy that takes advantage of momentum in prices. Such short-horizon directional strategies would be more profitable when there is less depth throughout the book and prices are fragmented across the trading venues.

There are two main takeaways from this exercise. First, there is heterogeneity in the underlying common strategies, as evidenced by the relationships we document between each principal component and variables that represent various aspects of the market environment. Furthermore, there appear to be a significant number of HFT firms that pursue unique strategies, although these represent a small portion of HFT activity in the market. The heterogeneity in HFT strategies is important insofar as the insights generated by thinking about HFT as one "entity" could be rather limited because aggregating HFT activity hides heterogeneity that is important to understanding the manner in which HFT firms interact with markets and ultimately affect them. Second, each underlying common strategy is pursued by multiple HFT firms: eight, six, and four HFT firms load on the first, second, and third principal components, respectively. This represents a high degree of competition in the market because the anonymous nature of trading in electronic limit-order books constitutes a formidable hurdle for collusive behavior. In other words, evidence that each strategy is followed by several firms suggests that there is competitive pressure that drives down possible rents earned by these HFT firms for their activity in the market.[24] While there are also HFT firms with unique strategies in which we do not observe competition, the aggregate economic rents earned by these firms is likely more limited due to their small market share in terms of trading volume.

5. Correlated HFT Activity and Returns

The principal components we document in Section 4 beg the question whether competition among HFT firms could induce a systematic factor that would affect returns. Specifically, high commonality of strategies could mean that at each point in time multiple HFT firms are more active in some stocks and less active in others. It is conceivable that the activity of multiple HFT

[24] The impact of competition between HFT firms on rents could depend on the specifics of the particular strategy. If one thinks of HFT firms as informed traders, the models of Holden and Subrahmanyam (1992) and Foster and Viswanathan (1993) show that it is enough to have two competing informed traders to almost instantaneously eliminate their informational advantage and have their profits vanish in a continuous market. Li (2013) takes another approach. The informed traders in her model are endowed with identical flows of long-lived private information (as opposed to one-shot long-lived private information in the aforementioned two papers). As a result, the aggregate profits of the informed traders decrease in the number of informed traders (specifically, they are inversely proportional to the square root of the number of informed traders), but do not vanish.

firms would impact returns over short intervals. If these firms are more or less active based on the manner in which their algorithms profit from the trading environment as opposed to fundamental factors that should affect stock returns, the extent of correlations in HFT activity could create systematic patterns in returns or volatility over and above fundamental influences.

Here we stress that our focus is not on the question whether the magnitude of HFT activity creates a systematic factor but rather whether the extent of correlations in the activity of multiple HFT firms does so.[25] This is an important distinction: HFT activity undertaken by multiple firms that follow unrelated strategies across stocks is unlikely to generate a systematic influence. On the other hand, HFT activity in which multiple HFT firms engage in strategies that are highly correlated across stocks could potentially turn idiosyncratic occurrences at the microstructure level into systematic patterns in returns or volatility. If such a systematic factor is created in returns, listed companies should have a particular interest in HFT and its impact of their cost of capital. Given the interest in the manner in which HFT affects volatility, the question whether competition between HFT firms systematically influences return volatility in short intervals is also of independent interest.

5.1 Cross-Sectional Correlations Summary Statistics

The measure of competition between HFT firms that we use to study these questions is the cross-sectional correlation measure that was introduced in Section 3.4: for each 1-second time interval, we compute the correlation coefficient between the activities of pairs of HFT firms across the stocks in the sample, and average the correlations for all pairs of firms. Figure 4 shows an intraday plot of the 1-second cross-sectional correlations of the overall activity measure (MSG). We observe a reverse U-shape in which correlations increase rapidly in the first 10 minutes of trading, continue to increase modestly until around noon, and then decline after 3:30 p.m. Over a good portion of the trading day, however, the levels of these correlations seems to be relatively stable.

In Table 4 we present the mean cross-sectional correlation for three variables that describe HFT strategies: MSG, TRD, and LMT. While MSG is a comprehensive measure in the

[25] See Skjeltorp, Sojli, and Tham (2013) for a study of the impact of algorithmic trading on asset prices.

sense that it incorporates every action a firm initiates over the interval, the other two measures represent other aspects of an HFT strategy. Specifically, TRD consists of all trading of an HFT firm over the interval and LMT reflects all actions that a firm undertakes to change its position in the limit order book. Our use of three proxies for HFT strategies is meant to examine whether the results are sensitive to the specific representation of the strategies.[26]

We observe higher correlations for the HFT firms with the largest market shares (MS1). For example, the correlation of the MSG measure for the MS1 subgroup, comprised of the four HFT firms with the largest market share of volume, is 0.359, which is much higher than the correlations for MS2 and MS3 (0.131 and 0.174, respectively, for the 1-second interval). The same pattern can be observed for trades as well as for submissions and cancellations of limit orders. In terms of magnitude, the correlation in trade executions (0.417 for MS1, 0.313 for all HFT firms) is higher than for non-marketable order submission and cancellation (LMT). While in principle such high cross-sectional correlations could be driven by a common factor in returns (e.g., the return on the market), it is interesting to consider whether HFT strategies themselves introduce a systematic component into the trading environment of stocks. We investigate this issue below.

5.2 Regression Analysis

We examine whether the extent of correlations in HFT activity creates systematic patterns in returns or volatility over short intervals by running the following regressions:

$$r_{it} = a_{0i} + a_{1i} Cor\left(Y_t\right) + a_{2i} Y_{mt} + a_{3i} r_{mt} + a_{4i} \left|r_{mt}\right| + a_{5i} Volume_{mt} + error_{it} \tag{1}$$

$$\left|r_{it}\right| = a_{0i} + a_{1i} Cor\left(Y_t\right) + a_{2i} Y_{mt} + a_{3i} r_{mt} + a_{4i} \left|r_{mt}\right| + a_{5i} Volume_{mt} + error_{it} \tag{2}$$

where r_{it} is the return for stock *i* in interval *t*, and $|r_{it}|$ is the absolute value of the interval return (which we use as our measure of interval volatility).[27] Equation 1 originates in a market model for returns, where r_{mt} is the value-weighted return of all stocks in our sample.[28] We are interested

[26] Ex-ante, it is unclear which measure is more important in explaining returns. On the one hand, TRD focuses on trade executions, and as such it could be more tightly related to price changes or returns. On the other hand, most of the activity of HFT firms involves order flow that does not culminate in trades (limit-order submission and cancellation). Including such orders in the representation of the strategy may result in a better description of HFT activity and therefore could be more appropriate for the analysis of how HFT activity impacts returns.
[27] The returns we use are computed from trade prices (using the trade closest to the end of an interval).
[28] The results are similar when we use the equal-weighted rather than the value-weighted return as a proxy for the market portfolio.

in the impact of the cross-sectional correlation, $Cor(Y_t)$, which is a market-wide attribute of the extent of competition (or similarity in strategies) between HFT firms that is computed for each interval in the 30-day period.

We add three control variables to the regressions in 1.1 and 1.2. First, since our goal is to study the effect of correlation in HFT strategies rather than the magnitude of HFT activity per se, we add as a control the magnitude of HFT activity in the market, Y_{mt}, which is the sum across all HFT firms of the same measure that we use for $Cor(Y_t)$. In other words, in the regressions with MSG (TRD) as the measure of HFT activity, $Cor(Y_t)$ is the cross-sectional correlation of the MSG (TRD) measure in interval t, while Y_{mt} is the sum of MSG (TRD) for all HFT firms and all stocks in the same interval. We also control for $Volume_{mt}$ (the aggregate volume in all stocks) and $|r_{mt}|$ (the absolute value of the market return), as conventional wisdom suggests that the activity of HFT firms is driven by volume and volatility.[29]

Panel A of Table 5 presents the results from running 52 regressions (one for each stock) in which the dependent variable is the stock's 1-second interval return (Equation 1). For each variable, we present the average coefficients and t-statistics (computed using heteroskedasticity-and-autocorrelation-consistent standard errors) from the individual regressions, the number of negative coefficients, the number of negative coefficients that are significant at the 5% level, the number of positive coefficients, the number of significant positive coefficients, and the average R-squared. Panel B of Table 5 presents the regression results for all three interval lengths (not just the 1-second interval), but to economize on the size of the table we show only the $Cor(Y_t)$ coefficients. We add three statistical tests to help judge whether the coefficients of the individual regressions are different from zero. First, we run a seemingly unrelated regression estimation (SURE) of the 52-equation system and provide the p-value from a joint F-test that the $Cor(Y_t)$ coefficients are equal to zero.[30] Second, we provide the nonparametric Sign Test and Wilcoxon test as a robustness check on whether the coefficients from the 52 regressions are equal to zero.

[29] In fact, the decrease in HFT activity in 2011 and 2012 was attributed to a decline in volume and volatility in the market. See, for example "High-speed trading no longer hurtling forward" by Nathaniel Popper, *New York Times*, October 14, 2012, and "High frequency trading loses its luster" by Ivy Schmerken, *Wall Street and Technology*, April 1, 2013.

[30] Given the nature of the data (small N, large T) and our desire to estimate a separate vector of coefficients for each stock, which represent loadings from a multiple-factor model, the natural approach is Seemingly Unrelated Regression Estimation (SURE). We note, however, that because we have identical regressors in all regressions (the

The results do not lend support to the idea that competition between HFT firms creates a systematic factor in returns. In particular, the HFT cross-sectional correlation is statistically significant in only a small fraction of the individual regressions and has approximately equal numbers of positive and negative coefficients. The statistical tests shown in Panel B cannot reject the hypothesis that the coefficients on correlated HFT activity across the 52 stocks are equal to zero. While the R-squared of the individual stock regressions increases with interval length (to 17.33% for the 60-second interval), the picture that emerges on the impact of correlated HFT strategies is the same irrespective of the measure we use (MSG, TRD, or LMT) or the time interval: we find no evidence consistent with the hypothesis that correlated HFT strategies across stocks create a common factor in returns.

Panel A of Table 6 presents the results from running 52 regressions (one for each stock) in which the dependent variable is the stock's 1-second interval return volatility (Equation 2). Here we observe a very different picture: the average loading on the correlation between HFT strategies is negative, and at least 43 of the 52 individual stock regressions have negative coefficients. Panel B of Table 6 shows that the negative effect of competition between HFT strategies on individual-stock short-interval volatility is observed irrespective of whether we use 1-second, 10-second, or 60-second intervals. Furthermore, all statistical tests (an F-test, a Sign test, and a Wilcoxon test) reject the hypothesis that the coefficients on correlated HFT activity across the 52 stocks are equal to zero. The magnitude of the effect is also economically meaningful. An increase in MSG (TRD) correlation from 0 to 0.5 would result in a decrease of 31% (32%) in 1-second volatility.[31] It is instructive to contrast this predominantly negative loading with the predominantly positive loadings on the magnitude of HFT activity and the volatility of the market portfolio (in Panel A). These positive loadings show that when there is more fundamental news in the market as a whole, prices of individual stocks move to a greater extent and HFT activity, be it arbitrage or otherwise, intensifies. The fact that the loadings on the

---

common factors), SURE is essentially the same as regression-by-regression OLS (i.e., it does not provide efficiency gains). We use SURE to provide a joint coefficient test, but add nonparametric tests on the magnitude of the OLS coefficients for robustness.

[31] While an increase in correlation from 0 to 0.5 is a rather large change, even a one-standard-deviation increase in correlation would decrease 1-second interval volatility between 6.3% (MSG) and 16.6% (TRD).

correlation between HFT strategies are predominantly negative suggests that it represents a separate and distinct effect.[32]

Why would greater cross-sectional correlation between HFT firms' strategies decrease rather than increase stock return volatility? Jovanovic and Menkveld (2015) posit that HFT firms trade on "hard information," such as price changes in same-industry stocks or the market index. Greater competition to turn these cross-stock "private" signals into public information implies lower adverse selection costs and hence a lower price impact of trades. As short-interval (e.g. 1-second) volatility is predominantly driven by the price impact of trades (as opposed to the public release of fundamental news, which is a relatively rare occurrence for a stock), it follows that greater cross-sectional correlation between HFT strategies would decrease short-interval volatility. It is also possible that the negative loadings are driven by the impact of correlated HFT strategies on the transitory, rather than the permanent, price impact of trades if they reflect greater competition in market-making activity. If returns and order flows of various stocks are correlated, efficient market making would necessitate algorithms that consider multiple stocks (e.g., Ho and Stoll (1983)). Competition between such market-making algorithms would appear as higher cross-sectional correlations of HFT strategies, and such competition would decrease the transitory price impact of trades, lowering short-interval volatility.

We investigate the channel through which competition between HFT firms lowers the short-interval volatility of stocks by examining competition in each of the underlying common strategies represented by the three principal components from Section 4. We run a regression similar to Equation 2 but replace the single cross-sectional correlation that is computed as the average of paired correlations of all HFT firms with three cross-sectional correlations, each computed as the average of paired correlations only for the firms that load significantly on one of

[32] Out of the 52 stocks, 37 are cross-listed in the U.S. A possible concern is that the models in Equation 1 and Equation 2 are misspecified because they exclude market return, volatility, and volume from the U.S. We therefore used the SPY exchange-traded fund that tracks the S&P 500 index to construct variables at the 1-second frequency for the U.S. market return, volatility, and volume (where we use the SPY volume as a proxy for overall U.S. market volume). The correlations between the U.S. and Canadian variables were modest, which enabled us to add the three U.S. variables to the regressions in Equation 1 and Equation 2. We found that the addition of the U.S. market variables did not materially affect our results. In other words, our measure of correlated HFT strategies did not appear to affect returns (similar to the results we report in Panel B of Table 5), but most stocks had a negative and statistically significant loading on our measure in the volatility regressions (similar to the results we report in Panel B of Table 6).

the principal components.[33] Similarly, rather than having one MSG variable as a control, we have three control variables, each aggregating the messages of the HFT firms that load on one of the principal components.

Table 7 presents the results of the regressions in the same format as Panel A of Table 6. For both the MSG and LMT measures, the second principal component seems to drive a significant portion of the relationship: its correlation is the most negative, the average t-statistic is large, and there are 39 stocks (out of 52) for which this variable has a negative and statistically significant coefficient (compared with 10 positive and significant coefficients). This principal component represents the underlying common strategy that we associate with market making based on the analysis in Section 4. As such, this result suggests that the reduction in volatility is driven by lower transitory volatility due to competition between market makers (Ho and Stoll (1983)). The regressions with the TRD measure show strong results for both the second and third principal components (43 and 41 negative and statistically significant coefficients, respectively), which could suggest that the reduction in volatility is due in part to a short-horizon directional strategy that impounds hard information when there is high volatility and prices are not aligned across the trading venues (which is the strategy we associate with the third principal component).[34]

To summarize, contrary to concerns that correlated HFT activity may increase market fragility, we find that competition between HFT firms that gives rise to correlated HFT activity decreases the short-interval volatility of most individual stocks.

6. Correlated HFT Activity and Market Consolidation

The current market structure for equity trading in the U.S., Canada, and many other countries is characterized by the co-existence of multiple trading venues on which the same stocks can be traded. Such trading fragmentation may create negative externalities in the form of worsened price integrity and higher costs as liquidity is scattered across the trading venues. Against these

[33] There are eight HFT firms with loadings greater than 0.35 on the first principal components, six firms with such loadings on the second principal component, and four firms that load on the third principal component. The three variables that we create are not highly correlated with one another (the highest correlation is 0.3), and hence there is no difficulty in having all three of them in the same regression.
[34] As in Table 6, joint F-tests reject the hypothesis that the coefficients on correlated HFT activity across the 52 stocks for these principal components are equal to zero.

negative externalities, proponents of this structure argue that the competition it induces between trading venues results in lower fees and greater innovation in terms of trading technology and services.[35]

The activity of HFT firms appears very relevant to both sides of the equation: alleviating the negative externalities of fragmentation as well as facilitating competition between trading venues. First, HFT firms can act as market consolidators, transforming the environment into a virtual central electronic limit-order book by moving liquidity from one venue to another and ensuring that prices are the same across the trading venues. Second, a new venue must ensure enough liquidity provision in order to attract trading, and hence the willingness of HFT firms to provide liquidity on new trading venues can be essential in fostering competition between venues.[36] What is less well understood, however, is how competition between HFT firms, which is reflected in the pursuit of similar strategies by multiple firms, affects the competitiveness of trading venues and the concentration of trading in the market.

6.1 Time-Series Correlations Summary Statistics

To study whether competition among HFT firms is related to market consolidation in terms of trading venues we use the time-series correlation measure that was introduced in Section 3.4: for each stock, we compute the correlation coefficient between the activities of pairs of HFT firms over all time intervals, and average across all pairs of firms. The time-series correlation measure provides information on whether the strategies of HFT firms are correlated over time for a given stock. Table 8 presents the mean time-series correlations for the set of all HFT firms as well as for firms in our market share subgroups (MS1, MS2, and MS3).

As with the cross-sectional correlations, the magnitude of the correlation is higher for the HFT firms with the greatest market share. For example, the correlation of the MSG measure for the MS1 subgroup, comprised of the four HFT firms with the largest market share of volume, is 0.249, which is much higher than the correlations for MS2 and MS3 (0.044 and 0.030,

---

[35] See O'Hara and Ye (2011) for a recent study of the consequences of market fragmentation in the U.S. We stress, though, that the focus of much of the recent literature has been on fragmentation caused by crossing networks (see, for example, Buti, Rindi, and Werner (2011)) while the type of fragmentation that exists during our sample period in Canada consists almost exclusively of trading venues that are structured as electronic limit-order books.

[36] Menkveld (2013) notes that a new trading venue in Europe became viable only when a large HFT began trading on it.

respectively, for the 1-second interval). The same pattern can be observed for trades as well as for submissions and cancellations of limit orders. While the correlations we observe when considering the strategies of all HFT firms are rather low (e.g., 0.049 to 0.078 for MSG or 0.036 to 0.067 for TRD, depending on the interval length), the correlations between the activity of firms in MS1, which together constitute 26.3% of the market volume, are much higher and can reach 0.446 for trades or 0.322 for overall message activity in larger stocks (I3).

We also observe that time-series correlations involving larger stocks are higher than those involving smaller stocks, especially for HFT firms in the MS1 subgroup. For example, the correlation between messages for the MS1 group is twice as high for larger stocks as for smaller stocks (0.341 vs. 0.158). Menkveld (2013) shows that the profitability of a single HFT that functions as a market maker is higher in large stocks than in small stocks. Greater profit potential in larger stocks could mean more room for multiple players to follow similar strategies, and hence we observe higher correlations.

6.2 Correlated HFT Strategies and Market Concentration

To examine market concentration, we compute the Herfindahl-Hirschman Index (henceforth HHI) of market share in terms of volume for the five trading venues we investigate.[37] The lower the HHI, the less concentrated the market. The independent variable on which we focus is the time-series correlation measure, $Cor(Y_i)$, which is an attribute of the competition between HFT firms in stock i. We run the following cross-sectional regression on the 52 stocks in the sample:

$$\begin{aligned} HHI_i = a_0 + a_1 Cor\left(Y_i\right) + a_2 Y_i + a_3 Spread_i + a_4 Depth_i + \\ a_5 MktCap_i + a_6 Price_i + a_7 Volatility_i + error_i \end{aligned} \tag{3}$$

Since our goal is to study the effect of correlation in HFT strategies rather than the magnitude of HFT activity per se, we add the magnitude of HFT activity in stock i, Y_i, as a control variable. The next two variables, time-weighted average spread and bid-and-offer (BBO) depth, are meant to control for the liquidity environment of the stock. The last three control variables—market

[37] The HHI is computed as the sum of squared market shares of the trading venues, and with five trading venues it is always between 0.2 (if volume is equally divided among the five trading venues) and 1 (if volume concentrates on one trading venue).

capitalization, price level, and the standard deviation of 30-minute returns over the sample period—are meant to control for heterogeneity in fundamental attributes across stocks.

Panel A of Table 9 provides the results for the three HFT measures: MSG, TRD, and LMT. The results are very strong: the coefficient on *Cor(Y_i)* is negative and highly statistically significant in eight out of the nine regressions. This finding is consistent with the hypothesis that more highly correlated HFT strategies, which we view as a manifestation of competition between HFT firms, are associated with a less concentrated market structure. However, it is important to note that the relationship between these two variables—concentration and the correlation of HFT strategies—is best viewed as being jointly co-determined in equilibrium rather than as, strictly speaking, one's causing the other. While competition between HFT firms could enhance the viability of smaller trading venues, the multiplicity of trading venues could increase the profit opportunities from arbitrage and market making across trading venues and hence lead to intensified competition between HFT firms. Competition between HFT firms and competition between trading venues are therefore tightly linked in a "competition begets competition" manner.

As in Section 5.2, we investigate the channel through which competition between HFT firms interacts with market concentration by examining competition in each of the underlying common strategies represented by the three principal components. We run a regression similar to Equation 3 but replace the cross-sectional correlation that is computed as the average of paired correlations of all HFT firms with a cross-sectional correlation that is computed as the average of paired correlations only for the firms that load significantly on one of the principal components (Cor(Y_i)PC1, Cor(Y_i)PC2, or Cor(Y_i)PC3).[38] Similarly, the magnitude of HFT activity in stock *i*, Y_i, that we use as a control variable is computed only for the firms that load significantly on the principal component we investigate in the regression.

To economize on the size of the table, and because we want to present results for the three strategy measures (MSG, TRD, and LMT) and the three interval lengths (I1, I2, and I3), we

[38] Unlike the case in Section 5.2, the time-series correlation measures for at least two of the three principal components are highly correlated with one another, and hence we cannot have all three of them in the same regression. We therefore replace Cor(Y_i) in Equation 3 with Cor(Y_i)PC1, Cor(Y_i)PC2, and Cor(Y_i)PC3 one at a time.

report in Panel B of Table 9 only the coefficient on the correlation variable from each regression (with heteroskedasticity-consistent *t*-statistics). We observe negative and statistically significant coefficients on $Cor(Y_i)PC1$ and $Cor(Y_i)PC2$ for all interval lengths and strategy measures. However, $Cor(Y_i)PC3$ is not statistically significant in any of the regressions. This means that the negative relationship between HFT competition and market concentration is driven by competition in two specific strategies: cross-venue arbitrage (the first principal component) and market making (the second principal component). This result is very intuitive, and increases our confidence in the interpretation of the underlying common strategies that we identify using the principal component analysis.

6.3 Correlated HFT Strategies and the Competitiveness of Trading Venues

In this section we would like to further our understanding of the result that HFT competition is associated with a less concentrated market by asking whether competition between HFT firms enhances the viability (or competitiveness) of specific trading venues. For each of the five trading venues that are organized as electronic limit-order books, we examine whether higher correlation of HFT strategies on that specific venue increases the percentage of time that it displays the best prices or narrowest spreads.

Panel A of Table 10 presents market share summary statistics for the five trading venues as well as two measures of trading venue competitiveness (or viability): (i) the percentage of time that the trading venue posts either the best bid or the best ask in the market (where the market is defined as the aggregation of all five trading venues), and (ii) the percentage of time that the bid–ask spread on the trading venue is the narrowest spread in the market.[39] We observe that there is a dominant trading venue in Canada with a market share in terms of volume of 69.26%. This largest trading venue displays the best price 92% of the time (averaged across all stocks in our sample), and also has the lowest standard deviation (6.1%) and the highest minimum (71.1%). Other trading venues also frequently display the best prices, although the variability in the cross section is greater. Similarly, the largest trading venue has the narrowest

[39] We note that two or more trading venues could potentially be at the best bid or ask at the same time (or have the same narrowest spread). We also recognize that activity on one trading venue could affect activity on another trading venue as in the multi-market inventory model of Lescourret and Moinas (2015).

spread 76.8% of the time (compared with 54.2%, 36.1%, 26.9%, and 9.4% for the other four trading venues).

Not all HFT firms in our sample are very active on multiple trading venues. We therefore focus on firms undertaking substantial activity on all five trading venues, which we define as sending at least 10,000 messages during our sample period to each of the trading venues. Eight HFT firms satisfy this criterion, and we compute trading-venue-specific time-series correlations that tell us whether the strategies of these HFT firms are correlated over time for a given stock on a particular trading venue. These correlations are similar in nature to our time-series correlations from section 6.1 and 6.2, except they are computed for each trading venue separately using only activity on that trading venue. For each trading venue *v*, we run the following cross-sectional regression:

$$\begin{aligned} C_{iv} = a_0 + a_1 Cor\left(Y_{iv}\right) + a_2 Y_{iv} + a_3 Spread_{iv} + a_4 BBOdepth_{iv} + \\ a_5 MktCap_i + a_6 Price_i + a_7 Volatility_i + error_{iv} \end{aligned} \tag{4}$$

where C_{iv} stands for one of the two competitiveness measures for stock *i* on trading venue *v*, and we use MSG (total messages sent by the HFT firm) as our measure of HFT strategies.[40] As in Section 6.2, including the magnitude of HFT activity on the specific trading venue as a control variable ensures that we pick up the effect of correlation in HFT strategies, not the magnitude of their activity per se. Spreads and depth (now computed from the best prices on a single trading venue) are meant to control for the liquidity of the stock, and the last three variables control for heterogeneity in fundamental attributes across stocks.

To present results for the five trading venues and the three interval lengths in a clear fashion, we report in Panel B of Table 10 only the coefficient on $Cor(Y_{iv})$ from each regression (with heteroskedasticity-consistent *t*-statistics). We observe an interesting pattern. For both dependent variables, the coefficients on the smaller trading venues are positive, although not all of them are statistically significant. For the largest trading venue in terms of market share of volume, however, we observe the opposite result: the coefficient is negative and statistically significant in almost all regressions. We observe that the impact of correlated HFT strategies on the viability or competitiveness of a trading venue depends on the nature of that venue: it

[40] Results using other measures of HFT strategies are similar in nature.

benefits smaller venues that introduce competition into the market structure and hence detract from the dominant position of the largest trading venue.[41] This result suggests that at least part of the negative relationship between market concentration and competition between HFT firms that we document is driven by the enhancement of the viability of smaller trading venues (in terms of displaying better prices and smaller spreads) by HFT competition, which increases their market share of trading.

7. Conclusions

Our paper examines correlations between strategies of high-frequency trading firms, which reflect competition between these firms in pursuing similar strategies. There are two motivations for examining these correlations. First, they can teach us about the industrial organization of the HFT space (e.g., how many common strategies there are and how many firms follow them). Second, they may reveal to us whether such competition creates a negative externality in the form of market fragility because they represent the tendency of HFT firms to do the same thing at the same time or with respect to the same stocks.

The first important stylized fact that we establish is that while there is substantial correlation between HFT strategies in terms of total activity (i.e., buy orders plus sell orders), there is little correlation in directional activity (i.e., buy orders minus sell orders). The fact that the correlations in directional activity are low has important implications for the stability of the market. Unlike the literature on institutional investor herding that focuses on how directional trading can negatively affect the market, our study indicates that HFT firms do not seem to have a strongly predominant direction that is common to many firms and that could destabilize the market. The second important stylized fact is that the correlation structure is similar under a variety of market conditions. In other words, the correlations between HFT strategies do not increase on days on which stock prices drop significantly. This is reassuring insofar as services

[41] The disparity in results between the smaller trading venues and the dominant exchange suggests that reverse causality is less likely to be a concern in these regressions. We also feel that the economics of trading on these venues suggests that the percentage of time a trading venue posts the best prices is determined by the strategies of the HFT firms, rather than vice versa, because HFT firms are the dominant players on these trading venues in terms of liquidity provision.

that HFT firms provide, both liquidity provision and the virtual consolidation of a market comprised of multiple trading venues, appear rather robust.

Our principal component analysis delves more deeply into the structure behind correlated HFT activity. Our analysis suggests that there are at least three underlying common strategies that are followed by multiple firms. These underlying common strategies relate differently to market conditions and serve correspondingly different functions. While some HFT firms appear to follow unique strategies that are uncorrelated with those of others, almost half of the firms in our sample (representing most of the HFT firms' activity: 78.97% in terms of volume and 96.21% in terms of messages) compete for the same three types of profit opportunities. If these strategies represent market making and the elimination of arbitrage opportunities, then the rent extracted for providing liquidity and consolidating the fragmented market is probably low, reflecting the highly competitive environment.

While competition may eliminate excess rents, it can also introduce externalities by influencing returns or volatility. Our analysis, however, shows no systematic impact of correlated HFT activity on returns. This is important insofar as HFT does not seem to influence prices in a manner that requires additional compensation from investors. In fact, we find that the short-interval volatility of most individual stocks loads negatively on the market-wide measure of HFT competition. While this could potentially be due to a reduction in either permanent or transitory price movements, our finding that the strongest driver of this relationship is the market-making strategy suggests that it is due at least in part to a reduction in transitory price movements brought about by more efficient cross-stock market making by HFT firms.

The last set of results we provide tie competition between HFT firms to competition between trading venues. We document a strong negative relationship between trading concentration in the market as a whole and competition between HFT firms in two underlying common strategies that we associate with cross-venue arbitrage and market making. Furthermore, we investigate a potential driver behind this negative relationship, and show that greater HFT competition within a trading venue helps smaller trading venues become more competitive or viable in terms of posting better prices and narrower spreads.

Our study was facilitated by the availability of high-quality regulatory data. Regulators focus on judging whether particular algorithms are harming the market (or otherwise implementing illegal strategies) as well as evaluating the overall impact of algorithms on the market to determine whether they necessitate a regulatory response in terms of changing the rules governing interactions between traders. Our study is silent on the first issue, but our insights into correlated HFT activity are important for the second one. Strong competition between HFT firms could in principle benefit markets (e.g., lowering rents earned for services HFT firms provide) but could also come at a cost in terms of the potential for fragility or greater volatility. Our study documents high levels of competition in multiple strategies that generate most of the HFT activity in the market, but at the same time shows that concerns about fragility from such correlated activity could be exaggerated. If anything, we find the opposite result: competition between HFT firms helps investors by lowering the volatility of stocks. We hope that as more high-quality data are made available to academic researchers, additional insights into the impact of HFT will emerge and facilitate a better-informed dialogue concerning this issue among investors and regulators.

References

Ait-Sahalia, Yacine, and Mehmet Saglam, 2014, High frequency traders: Taking advantage of speed, Working paper, Princeton University.

Baron, Matthew, Jonathan Brogaard, and Andrei Kirilenko, 2014, Risk and return in high frequency trading, Working paper, Princeton University.

Benos, Evangelos, James Brugler, Erik Hjalmarsson, and Filip Zikes, 2015, Interactions among high-frequency traders, Working Paper No. 523, Bank of England.

Benos, Evangelos, and Satchit Sagade, 2015, Price discovery and the cross-section of high-frequency trading, Working paper, Bank of England.

Biais, Bruno, Thierry Foucault, and Sophie Moinas, 2015, Equilibrium fast trading, Working paper, University of Toulouse.

Breckenfelder, H.-Johannes, 2013, Competition between high-frequency traders, and market quality, Working paper, Stockholm School of Economics, SHoF & SIFR.

Brogaard, Jonathan A., and Corey Garriott, 2014, High-frequency trading competition, Working paper, University of Washington.

Brogaard, Jonathan A., Terrence Hendershott, and Ryan Riordan, 2014, High frequency trading and price discovery, *Review of Financial Studies* 27, 2267-2306.

Brogaard, Jonathan A., Ryan Riordan, Andriy Shkilko, and Konstantin Sokolov, 2014, High-frequency trading and extreme price movements, Working paper, University of Washington.

Brown, Nerissa C., Kelsey D. Wei, Russ Wermers, 2014, Analyst recommendations, mutual fund herding, and overreaction in stock prices, Working paper, University of Maryland.

Budish, Eric, Peter Cramton, and John Shim, 2014, The high-frequency trading arms race: Frequent batch auctions as a market design response, Working paper, University of Chicago.

Buti, Sabrina, Barbara Rindi, and Ingrid M. Werner, 2011, Dark pool trading strategies, Working paper, University of Toronto.

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

Chaboud, Alain P., Benjamin Chiquoine, Erik Hjalmarsson, and Clara Vega, 2014, Rise of the machines: Algorithmic trading in the foreign exchange market, *Journal of Finance*, LXIX, 5, October, 2045-2084.

Choi, Nicole, and Richard W. Sias, 2009, Institutional industry herding, *Journal of Financial Economics* 94, 469-491.

Clark-Joseph, Adam D., 2014, Exploratory trading, Working paper, University of Illinois.

Comerton-Forde, Carole, Katya Malinova, and Andreas Park, 2015, Regulating dark trading: order flow segmentation and market quality, Working paper, University of Toronto.

Foster, F. Douglas, and S. Viswanathan, 1993, The effect of public information and competition on trading volume and price volatility, Review of Financial Studies 6, 23-56.

Foucault, Thierry, Johan Hombert, and Ioanid Rosu, 2015, News trading and speed, *Journal of Finance*, forthcoming.

Goldstein, Michael A., Pavitra Kumar, and Frank C. Graves, 2014, Computerized and high-frequency trading, forthcoming, *The Financial Review*, 49, 177-202.

Green, Edward J., and Robert H. Porter, Noncooperative collusion under imperfect price information, Econometrica 52, 87-100.

Hagstromer, Bjorn, and Lars Norden, 2013, The diversity of high-frequency traders, *Journal of Financial Markets* 16, 741-770.

Hagstromer, Bjorn, Lars Norden, and Dong Zhang, 2014, How aggressive are high-frequency traders? *The Financial Review*, 49, 395–419.

Han, Jungsuk, Mariana Khapko, and Albert S. Kyle, 2015, Liquidity with high-frequency market making, Working paper, Swedish House of Finance.

Hasbrouck, Joel, and Gideon Saar, 2013, Low-latency trading, *Journal of Financial Markets* 16, 646-679.

Hirschey, Nicholas H., 2013, Do high-frequency traders anticipate buying and selling pressure? Working paper, London Business School.

Ho, Thomas S.Y., and Hans R. Stoll, 1983, The dynamics of dealer markets under competition, *Journal of Finance*, 38, 4, September, 1053-1074.

Hoberg, Gerard, and Gordon Phillips, 2016, Text-based network industries and endogenous product differentiation, forthcoming in the Journal of Political Economy.

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

Holden, Craig W., and Avanidhar Subrahmanyam, 1992, Long-lived private information and imperfect competition, Journal of Finance 47, 247-270.

Ivaldi Marc, Bruno Jullien, Patrick Rey, Paul Seabright, and Jean Tirole, 2003, The economics of tacit collusion, Final Report for DG Competition, European Commission.

Jarnecic, Elvis, and Mark Snape, 2014, The provision of liquidity by high-frequency participants, *The Financial Review*, 49, 371–394.

Jarrow, Robert A., and Philip Protter, 2012, A dysfunctional role of high frequency trading in electronic markets, *International Journal of Theoretical and Applied Finance*, 15, 3, 1-15.

Jones, Charles M., 2013, What do we know about high-frequency trading? Working paper, Columbia Business School.

Jovanovic, Boyan, and Albert J. Menkveld, 2015, Middlemen in Limit Order Markets, Working paper, New York University.

Khandani, Amir E., and Andrew W. Lo, 2007, What happened to the quants in August 2007? *Journal of Investment Management* 5, 29-78.

Khandani, Amir E., and Andrew W. Lo, 2011, What happened to the quants in August 2007? Evidence from factors and transactions data, *Journal of Financial Markets*, 14, 1-46.

Kirilenko, Andrei, Albert S. Kyle, Mehrdad Samadi, and Tugkan Tuzun, 2014, The flash crash: The impact of high frequency trading on an electronic market, Working paper, University of Maryland.

Korajczyk, Robert A., and Dermot Murphy, 2015, High frequency market making to large institutional trades, Working paper, Northwestern University.

Lescourret, Laurence, and Sophie Moinas, 2015, Liquidity supply across multiple trading venues. Working paper. ESSEC Business School.

Li, Su, 2013, Multiple insiders with long lived flow of private information, and high frequency competition, Working paper, U.S. Securities and Exchange Commission.

Malinova, Katya, and Andreas Park, 2015, Subsidizing liquidity: the impact of make/take fees on market quality, Journal of Finance 70, 509-536.

Malinova, Katya, Andreas Park, and Ryan Riordan, 2014, Do retail traders suffer from high frequency traders? Working paper, University of Toronto.

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

Menkveld, Albert J., and Marius A. Zoican, 2014, Need for speed? Exchange latency and liquidity, Working paper, Duisenberg School of Finance - Tinbergen Institute Discussion Paper.

O'Hara, Maureen, and Mao Ye, 2011, Is market fragmentation harming market quality? *Journal of Financial Economics*, 100, 3, 459-474.

Pedersen, Lasse Heje, 2009, When everyone runs for the exit, *International Journal of Central Banking* 5, 177-199.

Rosu, Ioanid, 2015, Fast and slow informed trading, Working paper, HEC.

Skjeltorp, Johannes A., Elvira Sojli, and Wing Wah Tham, 2013, Trading on Algos, Working paper, Norges Bank and Erasmus University.

Wermers, Russ, 1999, Mutual fund herding and the impact on stock prices, *Journal of Finance*, LIV, 2, April, 581-622.

Table 1

Summary Statistics

Our sample consists of 52 stocks from the S&P/TSX60 (S&P). We rank the daily returns of the S&P/TSX Composite index from June 2010 through March 2011, and select the 10 worst days ('Down days'), the 10 best days ('Up days'), and the 10 days closest to and centered on zero return ('Flat days') for a 30-day sample period. Panel A presents summary statistics for the sample stocks: market capitalization, price, standard deviation of 30-minute returns, daily volume (in Canadian dollars), and daily return. Panel B presents summary statistics for the 31 high-frequency trading (HFT) firms that we identify using data from the Investment Industry Regulatory Organization of Canada (IIROC). These 31 firms are further categorized into three subgroups according to market share of volume: MS1 (market share of at least 4%; 4 firms), MS2 (market share of between 1% and 4%; 6 firms), and MS3 (the rest of the HFT firms). Similarly, we categorize the 31 firms according to their orders-to-trades ratios into three subgroups: OT1 (ratio greater than 150; 5 firms), OT2 (ratio between 50 and 150; 6 firms), and OT3 (all other HFT firms). Market share (in terms of volume or trades) is computed by dividing the trading undertaken by each HFT firm by total trading in the market. Trades consist of executions of both marketable orders and non-marketable orders. Sub/Canc is the number of (non-marketable) limit-order submissions and cancellations (i.e., all non-execution messages) that the firm sends to the market, where a modification of an order counts as a cancellation and a resubmission. Messages are all the orders (both marketable and non-marketable) and cancellations that a firm sends to the market. Orders/Trades Ratio is defined for each HFT firm as Messages divided by Trades. The mean orders-to-trades ratio is computed as the average of the orders-to-trades ratios of the individual HFT firms, and therefore due to its non-linear nature and the heterogeneity of the firms is not equal to the cross-sectional mean number of messages divided by the cross-sectional mean number of trades. CrossEndInventory is the number of times per day that an HFT firm's intraday inventory position crosses its end-of-day inventory position. We compute the measures for each HFT firm using all days in our sample period, and then provide in the table cross-firm means and medians for all HFT firms as well as for subgroups by market share or orders-to-trades ratio.

Panel A: Sample Stocks

Days	Stocks		MktCap (Million CAD)	Price (CAD)	StdRet (30min Ret)	CADVolume (1000; daily)	Return (daily)
All	S&P60	Mean	19,412	39.1	0.40%	78,156	-0.04%
		Median	11,401	36.9	0.40%	51,120	-0.04%
	Large	Mean	31,043	39.8	0.38%	115,013	-0.07%
		Median	27,574	37.1	0.38%	116,346	-0.12%
	Small	Mean	7,782	38.5	0.41%	41,299	-0.02%
		Median	7,731	36.7	0.41%	34,018	-0.03%
Down Days	S&P60	Mean	19,203	38.3	0.43%	81,195	-1.72%
		Median	11,299	36.4	0.43%	55,892	-1.56%
	Large	Mean	30,811	39.4	0.43%	123,636	-1.79%
		Median	27,802	36.8	0.43%	111,556	-1.76%
	Small	Mean	7,594	37.1	0.44%	38,755	-1.65%
		Median	7,544	35.8	0.44%	31,962	-1.35%
Flat Days	S&P60	Mean	19,499	39.7	0.34%	70,614	-0.08%
		Median	11,840	37.4	0.34%	50,078	-0.07%
	Large	Mean	31,063	39.8	0.32%	104,056	-0.11%
		Median	27,340	37.3	0.29%	104,396	-0.08%
	Small	Mean	7,935	39.5	0.36%	37,172	-0.04%
		Median	7,929	37.6	0.36%	31,421	-0.06%
Up Days	S&P60	Mean	19,535	39.5	0.39%	82,903	1.66%
		Median	11,574	37.0	0.37%	54,097	1.46%
	Large	Mean	31,254	40.0	0.37%	116,300	1.71%
		Median	28,022	37.3	0.35%	109,959	1.52%
	Small	Mean	7,816	38.9	0.41%	49,507	1.62%
		Median	7,721	36.8	0.41%	35,156	1.31%

Panel B: HFT Firms

HFT Firms		MktShare (Volume)	MktShare (Trades)	Trades (daily)	Sub/Canc (daily)	Messages (daily)	Orders/Trades Ratio	CrossEnd Inventory
All	Mean	1.50%	1.70%	19,445	1,056,838	1,063,974	92.7	14.9
	Median	0.34%	0.44%	5,035	97,146	97,288	40.4	2.4
MS1	Mean	6.58%	8.95%	102,035	5,466,547	5,496,423	49.5	73.0
	Median	7.23%	8.60%	98,048	3,272,863	3,312,989	42.2	70.1
MS2	Mean	2.67%	2.08%	23,730	982,137	995,868	38.3	13.4
	Median	2.79%	2.11%	24,003	147,090	162,729	6.9	4.3
MS3	Mean	0.19%	0.22%	2,489	238,236	239,157	116.5	4.3
	Median	0.05%	0.04%	714	40,129	40,268	71.4	1.8
OT1	Mean	0.11%	0.14%	1,549	767,533	767,949	323.9	1.9
	Median	0.04%	0.04%	425	97,146	97,288	257.9	1.9
OT2	Mean	1.89%	2.96%	33,716	3,132,973	3,140,817	116.3	32.6
	Median	0.37%	0.57%	6,487	273,318	276,226	117.1	6.6
OT3	Mean	1.72%	1.72%	19,637	506,324	514,927	27.8	12.9
	Median	0.42%	0.50%	5,727	26,340	37,771	19.4	2.8

Table 2

Principal Components Analysis: Loadings

This table presents the loadings from a principal component analysis of HFT strategies. The measure of HFT activity that we use to characterize the strategies is MSG, which is comprised of all messages an HFT firm actively sends to the market in an interval (submission of non-marketable limit orders, cancellation of non-marketable limit orders, and marketable limit orders that result in trade executions). We carry out the analysis using 1-second intervals. For the purpose of the analysis, we think of the 31 HFT firms as the "variables," while the observations are all intervals during the 30-day sample period for all sample stocks. The principal component analysis uses the varimax orthogonal rotation, and the first three principal components are retained for further analysis. The loading of an HFT firm on each of the principal components signifies the extent to which the firm's activity corresponds to the underlying common strategy represented by that principal component; it is equivalent to the bivariate correlation between the firm's measure and the principal component, and is therefore between –1 and 1. For each principal component, the loadings are sorted from the most positive to the most negative. We mark with an asterisk all principal components that are greater than or equal to 0.35.

HFT	PC1 Loading	HFT	PC2 Loading	HFT	PC3 Loading
F14	0.76 *	F27	0.71 *	F17	0.51 *
F16	0.67 *	F08	0.52 *	F23	0.48 *
F04	0.57 *	F31	0.41 *	F19	0.47 *
F24	0.54 *	F05	0.40 *	F20	0.43 *
F31	0.48 *	F02	0.39 *	F26	0.33
F28	0.46 *	F28	0.35 *	F12	0.22
F20	0.40 *	F30	0.34	F08	0.20
F17	0.38 *	F06	0.27	F18	0.18
F29	0.33	F12	0.24	F14	0.15
F01	0.32	F26	0.22	F29	0.13
F21	0.26	F10	0.17	F21	0.12
F27	0.26	F01	0.16	F03	0.12
F06	0.26	F20	0.13	F10	0.08
F07	0.23	F04	0.12	F30	0.05
F26	0.23	F24	0.11	F05	0.03
F08	0.22	F09	0.11	F02	0.03
F11	0.15	F23	0.10	F06	0.03
F12	0.08	F11	0.07	F16	0.02
F19	0.07	F18	0.06	F25	0.01
F05	0.04	F14	0.05	F13	0.01
F22	0.03	F03	0.04	F27	0.00
F10	0.01	F15	0.03	F09	0.00
F15	0.01	F25	0.02	F22	-0.01
F09	0.00	F22	0.01	F15	-0.01
F25	0.00	F13	0.01	F04	-0.02
F03	0.00	F29	0.00	F28	-0.07
F13	0.00	F16	-0.03	F31	-0.09
F02	-0.04	F21	-0.07	F24	-0.13
F30	-0.06	F07	-0.12	F07	-0.13
F23	-0.07	F19	-0.12	F11	-0.13
F18	-0.08	F17	-0.16	F01	-0.30

Table 3

Regressions using Principal Component Scores

This table presents regressions of principal component scores on variables that represent the market environment. In the principal component analysis of the MSG measure, the 31 high-frequency trading (HFT) firms are the "variables" while the observations are all 1-second intervals during the 30-day sample period for all sample stocks. The principal component analysis uses the varimax orthogonal rotation, and the first three principal components are retained for further analysis. There are 14 variables that describe the economic environment in each regression. The first two represent the degree of integration of the three trading venues with the highest market share of trading: TimeAorB is the percentage of time that the three trading venues display the same bid or ask prices, while TimeSprd is the percentage of time that the three trading venues have the same bid–ask spread. The next five variables represent the state of aggregate liquidity in the market as a whole (aggregated across all trading venues): total depth at the Market-Wide Best Bid or Offer (henceforth, MWBBO), total depth up to 10 cents from the MWBBO, depth imbalance at the MWBBO (defined as the absolute value of the difference between the number of shares at the bid and at the ask), depth imbalance up to 10 cents from the MWBBO, and percentage MWBBO spreads. The next three variables represent market conditions in the interval: return (computed from the last transaction price in each interval), volatility (computed as the absolute value of return), and the average time between trades in the interval. The last four variables represent information about HFT: the number of trades in which the HFT firms supply liquidity in the interval, the aggregate inventory position of all HFT firms in Canadian dollars (cumulative from the beginning of the day and assuming that all of them start the day with zero inventory), the number of shares traded among the HFT firms in the interval (#HFT), and the number of shares traded between the HFT firms and others (#NoHFT). In Panel A, the first three columns report the coefficients of the contemporaneous regressions in which we line up the component score with observations of the market environment over the same interval. We also provide, side by side, the coefficients from regressions in which the component scores are regressed on lagged and lead values of the economic variables. In Panel B, we report the coefficients from the contemporaneous regressions with t-statistics computed using double-clustered (interval and stock) standard errors.

Panel A: Contemporaneous, Lagged, and Lead Regressions on Component Scores

	Contemporaneous			Lagged			Lead		
	PC1	PC2	PC3	PC1	PC2	PC3	PC1	PC2	PC3
TimeAorB	-0.0330	0.0705	-0.0895	-0.0160	0.0953	-0.0854	-0.0374	0.0729	-0.0892
TimeSprd	-0.0956	0.0386	-0.0108	-0.0911	0.0412	-0.0134	-0.0979	0.0532	-0.0091
MWBBOdep	-1.E-06	-1.E-06	8.E-08	-2.E-06	-2.E-06	1.E-07	-1.E-06	-1.E-06	1.E-07
Dep10	-8.E-08	2.E-07	-8.E-08	-6.E-08	3.E-07	-8.E-08	-5.E-08	3.E-07	-8.E-08
|MWBBOimb|	1.E-06	8.E-07	-5.E-08	2.E-06	2.E-06	-1.E-07	1.E-06	1.E-06	-1.E-07
|Imb10|	-4.E-08	-2.E-07	6.E-08	-5.E-08	-2.E-07	6.E-08	-5.E-08	-2.E-07	6.E-08
%sprd	-4.4624	0.6093	0.9653	-3.9668	0.1172	0.7732	-5.2192	0.1300	1.1132
Ret	16.0262	3.4833	7.8211	4.1540	0.8679	4.4187	3.4556	2.3113	7.2387
|Ret|	285.18	32.34	214.10	110.38	-21.12	159.34	82.47	-3.36	108.35
TBT	-1.1539	-0.7664	-0.2999	-0.6159	-0.4439	-0.2205	-0.5550	-0.3785	-0.1192
HFTliqsup	0.0435	0.2415	0.0248	0.0178	0.0470	0.0010	0.0156	0.0281	0.0027
|Aggpos|	6.E-08	-4.E-08	4.E-08	5.E-08	-2.E-08	2.E-08	3.E-08	-2.E-08	1.E-08
#HFT	-0.0109	-0.1644	-0.0381	-0.0095	-0.0225	-0.0061	-0.0157	-0.0180	-0.0063
#NoHFT	0.0474	-0.1137	-0.0407	-0.0020	-0.0114	-0.0006	-0.0158	-0.0186	-0.0059
Intercept	1.1604	0.5794	0.3486	0.6610	0.3042	0.2705	0.6225	0.2599	0.1757
R^2	13.44%	29.27%	8.64%	2.77%	4.19%	0.61%	1.96%	2.39%	0.42%

Panel B: Contemporaneous Regressions with t-statistics using Double-Clustered Standard Errors

	PC1		PC2		PC3	
	Coef	t-stat	Coef	t-stat	Coef	t-stat
TimeAorB	-0.0330	-0.65	0.0705	3.13	-0.0895	-3.66
TimeSprd	-0.0956	-3.44	0.0386	2.41	-0.0108	-0.99
MWBBOdep	-1.5E-06	-3.65	-1.01E-06	-2.50	7.67E-08	0.55
Dep10	-7.73E-08	-1.81	2.37E-07	4.16	-8.28E-08	-2.77
|MWBBOimb|	1.48E-06	3.61	8.06E-07	2.09	-5.08E-08	-0.39
|Imb10|	-4.05E-08	-0.56	-1.67E-07	-1.84	6.10E-08	1.62
%sprd	-4.4624	-1.90	0.6093	1.03	0.9653	1.48
Ret	16.0262	2.76	3.4833	0.93	7.8211	1.48
|Ret|	285.1800	2.36	32.3437	0.83	214.0975	3.09
TBT	-1.1539	-6.72	-0.7664	-4.70	-0.2999	-3.01
HFTliqsup	0.0435	2.41	0.2415	18.51	0.0248	3.01
|Aggpos|	6.39E-08	1.20	-3.99E-08	-1.18	3.55E-08	1.15
#HFT	-0.0109	-0.56	-0.1644	-10.56	-0.0381	-3.61
#NoHFT	0.0474	3.03	-0.1137	-9.75	-0.0407	-4.09
Intercept	1.1604	4.99	0.5794	3.17	0.3486	2.74
R^2	13.44%		29.27%		8.64%	

Table 4

Cross-Sectional Correlations

This table presents cross-sectional correlations of HFT strategies. We use data from the Investment Industry Regulatory Organization of Canada (IIROC) to identify 31 high-frequency trading (HFT) firms. These 31 firms are further categorized into three subgroups according to market share of volume: MS1 (market share of at least 4%; 4 firms), MS2 (market share of between 1% and 4%; 6 firms), and MS3 (the rest of the HFT firms). The Cross-Sectional Correlation measure indicates whether the strategies of HFT firms are correlated across stocks at a given time. In each time interval, we compute the correlation coefficient between the activities of two HFT firms across the stocks in the sample, and average the correlations for all pairs of firms in a certain group (where ALL consists of the 31 HFT firms). We examine three measures of HFT activity: (i) the number of "messages" (MSG) HFT firms they send to the market, where messages are defined as submissions and cancellations of nonmarketable limit orders as well as executions of marketable limit orders, (ii) trades (TRD), and (iii) submissions/cancellations of nonmarketable limit orders (LMT). The measures representing HFT strategies as well as the correlations are computed separately for three interval lengths: I1 (1-second intervals), I2 (10-second intervals), and I3 (60-second intervals). Our sample consists of 52 stocks from the S&P/TSX60 (S&P). In each panel, we provide the correlation for the sample of S&P/TSX60 as well as two equal subsamples (Large and Small) ranked by market capitalization (with *p*-value for a two-sided test indicating whether Large differs from Small).

Strategy	HFT	**I1**				**I2**				**I3**			
Measure	Group	S&P	Large	Small	*p*-val.	S&P	Large	Small	*p*-val.	S&P	Large	Small	*p*-val.
MSG	MS1	0.359	0.407	0.303	<.001	0.332	0.380	0.271	<.001	0.317	0.348	0.262	<.001
	MS2	0.131	0.164	0.129	<.001	0.120	0.124	0.115	<.001	0.122	0.112	0.120	<.001
	MS3	0.174	0.213	0.243	<.001	0.127	0.145	0.164	<.001	0.108	0.113	0.141	<.001
	ALL	0.198	0.243	0.222	<.001	0.147	0.168	0.160	<.001	0.130	0.135	0.137	<.001
TRD	MS1	0.417	0.460	0.527	<.001	0.436	0.415	0.362	<.001	0.518	0.467	0.454	<.001
	MS2	0.219	0.268	0.346	<.001	0.122	0.124	0.153	<.001	0.155	0.130	0.148	<.001
	MS3	0.189	0.244	0.319	<.001	0.067	0.085	0.113	<.001	0.052	0.059	0.056	0.011
	ALL	0.313	0.386	0.460	<.001	0.152	0.174	0.189	<.001	0.131	0.133	0.127	<.001
LMT	MS1	0.355	0.402	0.301	<.001	0.328	0.374	0.269	<.001	0.313	0.342	0.259	<.001
	MS2	0.127	0.162	0.130	<.001	0.110	0.116	0.108	<.001	0.107	0.102	0.107	<.001
	MS3	0.175	0.214	0.244	<.001	0.129	0.148	0.165	<.001	0.108	0.117	0.142	<.001
	ALL	0.197	0.242	0.222	<.001	0.147	0.168	0.159	<.001	0.127	0.134	0.135	0.113

Table 5

Regressions of Returns on the Correlation of HFT Strategies

This table presents the regressions of individual stock returns on the cross-sectional correlation of high-frequency-trading (HFT) strategies. Our sample consists of 52 stocks from the S&P/TSX60. For each stock, we estimate the following regression over all intervals in the sample period:

$$r_{it} = a_{0i} + a_{1i}Cor\left(Y_t\right) + a_{2i}Y_{mt} + a_{3i}r_{mt} + a_{4i}\left|r_{mt}\right| + a_{5i}Volume_{mt} + error_{it}$$

where r_{it} is the return for stock *i* in interval *t* (computed from the trade price closest to the end of the interval), and r_{mt} is the value-weighted return of all stocks in our sample. We are chiefly interested in the impact of the cross-sectional correlation between HFT firms as measured by Cor(Y_t), which is defined as the average over all pairs of HFT firms of their correlations across stocks (for a particular activity measure). Cor(Y_t) is a market-wide attribute of the extent of competition (or similarity in strategies) between HFT firms in the market, and is computed for each interval in the 30-day period. We use three measures of HFT strategies: MSG (all messages HFT firms send to the market), TRD (all their trades), and LMT (only submissions and cancellations of non-marketable limit orders). We add three control variables to the regressions: the magnitude of HFT activity in the market, Y_{mt} (which is the sum across all HFT firms of the same measure that we use for Cor(Y_t)), $Volume_{mt}$ (the aggregate volume in all stocks) and $|r_{mt}|$ (the absolute value of the market return). For each variable, we present the average coefficient across the 52 stocks, the average t-statistics from the individual regressions (computed using heteroskedasticity-and-autocorrelation-consistent standard errors), the number of negative coefficients, the number of negative coefficients that are significant at the 5% level (from a two-sided test), the number of positive coefficients, the number of significant positive coefficients, and the average R^2. In Panel A we present the coefficients on all variables from the regressions using the 1-second intervals. In Panel B, we present just the coefficient on Cor(Y_t) side by side for the 1-second, 10-second, and 60-second intervals. We add three statistical tests to help judge whether the coefficients of the individual regressions are different from zero. First, we run a seemingly unrelated regression estimation (SURE) of the 52-equation system and provide the p-value from a joint F-test that the Cor(Y_t) coefficients are equal to zero. Second, we provide the nonparametric Sign Test and Wilcoxon test as a robustness check on whether the coefficients from the 52 regressions are equal to zero.

Panel A: Regression of 1-Second Interval Return on the Cross-Sectional Correlation of HFT Strategies

	Variable	Avg. Coef	Avg. t-stat	# Coef < 0	# t-stat < 1.96	# Coef > 0	# t-stat > 1.96	Avg. R-sqrd	# Obs.
MSG	**Intercept**	1.23E-07	0.35	24	4	28	10	4.47%	52
	Cor(Y_t)	-9.63E-08	-0.14	28	3	24	3		
	MSG_{mt}	1.50E-11	0.21	23	0	29	4		
	R_{mt}	8.04E-01	38.38	0	0	52	52		
	$\|R_{mt}\|$	-6.29E-03	-0.67	36	10	16	1		
	$Volume_{mt}$	4.80E-14	0.41	19	0	33	2		
TRD	**Intercept**	1.86E-07	0.65	16	2	36	10	4.51%	52
	Cor(Y_t)	-1.18E-07	-0.27	27	6	25	1		
	TRD_{mt}	-1.49E-10	0.08	24	3	28	4		
	R_{mt}	8.06E-01	38.27	0	0	52	52		
	$\|R_{mt}\|$	-6.78E-03	-0.66	38	7	14	0		
	$Volume_{mt}$	5.10E-14	0.49	19	1	33	2		
LMT	**Intercept**	1.21E-07	0.35	25	4	27	10	4.47%	52
	Cor(Y_t)	-9.16E-08	-0.13	30	3	22	2		
	LMT_{mt}	1.59E-11	0.21	23	0	29	4		
	R_{mt}	8.04E-01	38.38	0	0	52	52		
	$\|R_{mt}\|$	-6.30E-03	-0.67	36	10	16	1		
	$Volume_{mt}$	4.80E-14	0.41	19	0	33	2		

Panel B: Coefficient on Cor(Y_t) from Regressions of Return on the Correlation of HFT Strategies

Variable	**Interval**	Avg. Cor(t) Coef	Avg. Cor(t) t-stat	Joint Test (*p*-value)	# Coef < 0	# t-stat < 1.96	# Coef > 0	# t-stat > 1.96	Sign Test (*p*-value)	Wilcoxon (*p*-value)	Avg. R-sqrd
MSG	**I1**	-9.63E-08	-0.14	0.293	28	3	24	3	0.678	0.584	4.47%
	I2	2.29E-07	-0.14	0.512	30	2	22	1	0.332	0.547	9.44%
	I3	-4.27E-06	-0.11	0.966	27	1	25	0	0.890	0.886	17.33%
TRD	**I1**	-1.18E-07	-0.27	0.296	27	6	25	1	0.890	0.259	4.51%
	I2	-9.43E-07	-0.12	0.498	28	2	24	2	0.678	0.547	9.43%
	I3	-2.01E-06	-0.05	0.491	26	3	26	3	1.000	0.993	17.32%
LMT	**I1**	-9.16E-08	-0.13	0.233	30	3	22	2	0.332	0.461	4.47%
	I2	-7.87E-07	-0.16	0.491	31	2	21	0	0.212	0.439	9.44%
	I3	-2.42E-06	-0.10	0.903	26	1	26	0	1.000	0.929	17.33%

Table 6

Regressions of Interval Volatility on the Correlation of HFT Strategies

This table presents the regressions of individual stock interval return volatility on the cross-sectional correlation of high-frequency trading (HFT) strategies. Our sample consists of 52 stocks from the S&P/TSX60. For each stock, we estimate the following regression over all intervals in the sample period:

$$\left|r_{it}\right| = a_{0i} + a_{1i}Cor\left(Y_t\right) + a_{2i}Y_{mt} + a_{3i}r_{mt} + a_{4i}\left|r_{mt}\right| + a_{5i}Volume_{mt} + error_{it}$$

where $|r_{it}|$ is the absolute value of the return for stock *i* in interval *t* (computed from the trade price closest to the end of the interval), which is our measure of interval return volatility, r_{mt} is the value-weighted return of all stocks in our sample, $|r_{mt}|$ is the absolute value of the market return, and $Volume_{mt}$ is the aggregate volume in all stocks. We are chiefly interested in the impact of the cross-sectional correlation between HFT firms as measured by $Cor(Y_t)$, which is defined as the average over all pairs of HFT firms of their correlation across stocks (for a particular activity measure). $Cor(Y_t)$ is a market-wide attribute of the extent of competition (or similarity in strategies) between HFT firms in the market, and is computed for each interval in the 30-day period. We use three measures of HFT strategies: MSG (all messages HFT firms send to the market), TRD (all their trades), and LMT (only submissions and cancellations of non-marketable limit orders). We also add as a control variable the magnitude of HFT activity in the market, Y_{mt} (which is the sum across all HFT firms of the same measure that we use for $Cor(Y_t)$). For each variable, we present the average coefficient across the 52 stocks, the average t-statistics from the individual regressions (computed using heteroskedasticity-and-autocorrelation-consistent standard errors), the number of negative coefficients, the number of negative coefficients that are significant at the 5% level (from a two-sided test), the number of positive coefficients, the number of significant positive coefficients, and the average R^2. In Panel A we present the coefficients on all variables from the regressions using the 1-second intervals. In Panel B, we present just the coefficient on $Cor(Y_t)$ side by side for the 1-second, 10-second, and 60-second intervals. We add three statistical tests to help judge whether the coefficients of the individual regressions are different from zero. First, we run a seemingly unrelated regression estimation (SURE) of the 52-equation system and provide the p-value from a joint F-test that the $Cor(Y_t)$ coefficients are equal to zero. Second, we provide the nonparametric Sign Test and Wilcoxon test as a robustness check on whether the coefficients from the 52 regressions are equal to zero.

Panel A: Regressions of 1-Second Interval Volatility on the Cross-Sectional Correlation of HFT Strategies

	Variable	Avg. Coef	Avg. t-stat	# Coef < 0	# t-stat < 1.96	# Coef > 0	# t-stat > 1.96	Avg. R-sqrd	# Obs.
MSG	**Intercept**	8.00E-06	14.41	6	5	46	45	5.47%	52
	$Cor(Y_t)$	-1.72E-05	-12.41	43	39	9	6		
	MSG_{mt}	5.76E-09	26.14	5	5	47	47		
	R_{mt}	-6.29E-03	-0.79	36	10	16	1		
	$\|R_{mt}\|$	8.22E-01	30.07	0	0	52	52		
	$Volume_{mt}$	1.35E-12	2.12	5	3	47	45		
TRD	**Intercept**	1.39E-05	31.02	1	1	51	48	5.81%	52
	$Cor(Y_t)$	-1.78E-05	-33.28	50	49	2	1		
	TRD_{mt}	2.67E-07	23.34	5	4	47	47		
	R_{mt}	-6.01E-03	-0.78	36	10	16	2		
	$\|R_{mt}\|$	7.27E-01	22.38	0	0	52	52		
	$Volume_{mt}$	-3.45E-13	-1.11	43	22	9	3		
LMT	**Intercept**	7.99E-06	14.53	7	5	45	45	5.46%	52
	$Cor(Y_t)$	-1.71E-05	-12.33	43	40	9	6		
	LMT_{mt}	5.74E-09	25.97	5	5	47	47		
	R_{mt}	-6.30E-03	-0.79	36	10	16	1		
	$\|R_{mt}\|$	8.24E-01	30.20	0	0	52	52		
	$Volume_{mt}$	1.38E-12	2.13	5	3	47	46		

Panel B: Coefficient on Cor(Y_t) from Regressions of Interval Volatility on the Correlation of HFT Strategies

Variable	**Interval**	Avg. Cor(t) Coef	Avg. Cor(t) t-stat	Joint Test (*p*-value)	# Coef < 0	# t-stat < 1.96	# Coef > 0	# t-stat > 1.96	Sign Test (*p*-value)	Wilcoxon (*p*-value)	Avg. R-sqrd
MSG	**I1**	-1.72E-05	-12.41	<0.001	43	39	9	6	<0.001	<0.001	5.47%
	I2	-2.43E-04	-8.52	<0.001	46	43	6	5	<0.001	<0.001	9.65%
	I3	-6.51E-04	-3.75	<0.001	45	36	7	4	<0.001	<0.001	14.93%
TRD	**I1**	-1.78E-05	-33.28	<0.001	50	49	2	1	<0.001	<0.001	5.81%
	I2	-1.71E-04	-12.34	<0.001	45	42	7	2	<0.001	<0.001	9.84%
	I3	-2.13E-04	-2.02	<0.001	35	28	17	6	0.018	0.001	14.39%
LMT	**I1**	-1.71E-05	-12.33	<0.001	43	40	9	6	<0.001	<0.001	5.46%
	I2	-2.50E-04	-8.84	<0.001	46	42	6	5	<0.001	<0.001	9.65%
	I3	-7.14E-04	-4.11	<0.001	46	37	6	4	<0.001	<0.001	14.95%

Table 7

Regressions of Interval Volatility on the Correlation of HFT Strategies by Principal Component

This table presents the regressions of individual stock interval return volatility on the cross-sectional correlation of HFT strategies computed separately for the firms that significantly load on the three principal components presented in Table 2. For each stock, we estimate the following regression over all 1-second intervals in the sample period:

$$|r_{it}| = a_{0i} + a_{1i}\text{Cor}(Y_t)\text{PC1} + a_{2i}\text{Cor}(Y_t)\text{PC2} + a_{3i}\text{Cor}(Y_t)\text{PC3} + a_{4i}Y_{mt}\text{PC1} + a_{5i}Y_{mt}\text{PC2} + a_{6i}Y_{mt}\text{PC3} + a_{7i}r_{mt} + a_{8i}|r_{mt}| + a_{9i}\text{Volume}_{mt} + error_{it}$$

where $|r_{it}|$ is the absolute value of the return for stock *i* in interval *t*, which is our measure of interval return volatility, r_{mt} is the value-weighted return of all stocks in our sample, $|r_{mt}|$ is the absolute value of the market return, and Volume_{mt} is the aggregate volume in all stocks. We compute the cross-sectional correlation $\text{Cor}(Y_t)\text{PC1}$ as the average over the correlations of all pairs of HFT firms from among the eight firms that load on the first principal component, and similarly we compute $\text{Cor}(Y_t)\text{PC2}$ ($\text{Cor}(Y_t)\text{PC3}$) for the six (four) firms that load on the second (third) principal component. Each of these cross-sectional correlations is a market-wide attribute of the extent of competition (or similarity in strategies) between HFT firms with significant loadings on a particular principal component. The correlations are computed for three expressions of HFT strategies: MSG (all messages HFT firms send to the market), TRD (all their trades), and LMT (submissions and cancellations of non-marketable limit orders). We also add as a control variable the magnitude of HFT activity of the firms that load on the three principal components: $Y_{mt}\text{PC1}$, $Y_{mt}\text{PC2}$, and $Y_{mt}\text{PC3}$. The table presents for each variable the average coefficient across the 52 stocks, the average t-statistics (using heteroskedasticity-and-autocorrelation-consistent standard errors), the number of negative coefficients, the number of negative coefficients that are significant at the 5% level (from a two-sided test), the number of positive coefficients, the number of significant positive coefficients, and the average R^2.

	Variable	Avg. Coef	Avg. t-stat	# Coef < 0	# t-stat < 1.96	# Coef > 0	# t-stat > 1.96	Avg. R-sqrd	# Obs.
MSG	**Intercept**	9.39E-06	12.03	8	7	44	44	6.13%	52
	Cor(Y_t)PC1	-2.52E-06	-0.75	33	30	19	13		
	Cor(Y_t)PC2	-8.90E-06	-10.70	41	39	11	10		
	Cor(Y_t)PC3	-6.84E-07	-0.76	37	29	15	9		
	MSG_PC1	-1.20E-09	-1.46	34	33	18	14		
	MSG_PC2	7.96E-09	5.88	10	7	42	39		
	MSG_PC3	9.25E-09	7.29	8	7	44	42		
	Rm	-6.66E-03	-0.61	33	8	19	3		
	|Rm|	8.56E-01	28.57	0	0	52	52		
	Volume	2.02E-12	2.89	4	2	48	45		
TRD	**Intercept**	4.37E-05	9.00	1	0	51	50	12.56%	52
	Cor(Y_t)PC1	-1.01E-05	-2.03	42	27	10	3		
	Cor(Y_t)PC2	-1.91E-05	-3.44	48	43	4	0		
	Cor(Y_t)PC3	-2.96E-05	-4.99	48	41	4	3		
	TRD_PC1	-2.35E-08	-0.30	34	15	18	9		
	TRD_PC2	8.13E-07	0.12	25	11	27	12		
	TRD_PC3	4.03E-07	2.05	11	2	41	28		
	Rm	1.51E-03	0.06	23	4	29	4		
	|Rm|	9.87E-01	12.18	0	0	52	52		
	Volume	-2.20E-13	-0.25	31	10	21	11		
LMT	**Intercept**	9.42E-06	12.11	8	7	44	44	6.12%	52
	Cor(Y_t)PC1	-2.54E-06	-0.76	33	31	19	13		
	Cor(Y_t)PC2	-8.65E-06	-10.47	41	39	11	10		
	Cor(Y_t)PC3	-7.39E-07	-0.85	37	29	15	9		
	LMT_PC1	-7.74E-10	-0.95	34	34	18	15		
	LMT_PC2	6.98E-09	5.18	12	7	40	39		
	LMT_PC3	8.96E-09	7.02	7	7	45	42		
	Rm	-6.64E-03	-0.61	33	8	19	3		
	|Rm|	8.58E-01	28.77	0	0	52	52		
	Volume	2.05E-12	2.89	4	2	48	45		

Table 8

Time-Series Correlations

This table presents time-series correlations of HFT strategies. We use data from the Investment Industry Regulatory Organization of Canada (IIROC) to identify 31 high-frequency trading (HFT) firms. These 31 firms are further categorized into three subgroups according to market share of volume: MS1 (market share of at least 4%; 4 firms), MS2 (market share of between 1% and 4%; 6 firms), and MS3 (the rest of the HFT firms). The Time-Series Correlation measure indicates whether the strategies of HFT firms are correlated over time for a given stock. For each stock, we compute the correlation coefficient between HFT activities of any two pairs of HFT firms, and average across all pairs of firms in a certain group (where ALL consists of the 31 HFT firms). We examine three measures of HFT activity: (i) the number of "messages" (MSG) HFT firms they send to the market, where messages are defined as submissions and cancellations of nonmarketable limit orders as well as executions of marketable limit orders, (ii) trades (TRD), and (iii) submissions/cancellations of nonmarketable limit orders (LMT). The measures representing HFT strategies as well as the correlations are computed separately for three interval lengths: I1 (1-second intervals), I2 (10-second intervals), and I3 (60-second intervals). Our sample consists of 52 stocks from the S&P/TSX60 (S&P). In each panel, we provide the correlation for the sample of S&P/TSX60 as well as two equal subsamples (Large and Small) ranked by market capitalization (with *p*-value for a two-sided test indicating whether Large differs from Small).

Strategy	HFT	**I1**				**I2**				**I3**			
Measure	Group	S&P	Large	Small	*p*-val.	S&P	Large	Small	*p*-val.	S&P	Large	Small	*p*-val.
MSG	MS1	0.249	0.341	0.158	<.001	0.238	0.324	0.152	<.001	0.237	0.322	0.153	<.001
	MS2	0.044	0.055	0.034	0.013	0.068	0.077	0.059	0.041	0.089	0.094	0.085	0.342
	MS3	0.030	0.035	0.026	0.007	0.045	0.051	0.038	0.003	0.057	0.065	0.049	0.004
	ALL	0.049	0.061	0.038	<.001	0.064	0.078	0.050	<.001	0.078	0.092	0.063	<.001
TRD	MS1	0.268	0.341	0.195	<.001	0.314	0.400	0.228	<.001	0.351	0.446	0.257	<.001
	MS2	0.056	0.060	0.052	0.232	0.086	0.090	0.081	0.199	0.120	0.126	0.114	0.168
	MS3	0.012	0.012	0.012	0.424	0.019	0.020	0.018	0.276	0.029	0.031	0.027	0.214
	ALL	0.036	0.039	0.032	0.010	0.049	0.055	0.044	<.001	0.067	0.075	0.059	<.001
LMT	MS1	0.245	0.337	0.154	<.001	0.234	0.320	0.149	<.001	0.234	0.318	0.150	<.001
	MS2	0.036	0.047	0.026	0.010	0.055	0.067	0.043	0.015	0.071	0.082	0.061	0.057
	MS3	0.031	0.036	0.026	0.003	0.045	0.053	0.038	0.001	0.058	0.067	0.049	0.001
	ALL	0.048	0.060	0.036	<.001	0.063	0.077	0.048	<.001	0.076	0.092	0.060	<.001

Table 9

Market Concentration and the Correlation of HFT Strategies

This table presents regression results that relate concentration of trading across the trading venues to competition between high-frequency trading (HFT) firms. The five trading venues we investigate are organized as electronic limit-order books and together execute approximately 97.7% of the trading volume during our sample period. To examine market concentration, we compute the Herfindahl-Hirschman Index (henceforth HHI) of market share in terms of volume for the five trading venues for each stock. The HHI is computed as the sum of squared market shares of the trading venues, and the lower the HHI the less concentrated the market. In Panel A, we run the following cross-sectional regression:

$$HHI_i = a_0 + a_1 Cor\left(Y_i\right) + a_2 Y_i + a_3 Spread_i + a_4 Depth_i + a_5 MktCap_i + a_6 Price_i + a_7 Volatility_i + error_i$$

where Cor(Y_i) is the time-series correlation of all HFT strategies that we use to represent competition in the HFT space, and Y_i stands for one of the three measures of HFT strategies we use: MSG (total messages sent), TRD (trades), and LMT (submission of cancellations of non-marketable limit orders). The next two variables, market-wide average spread and bid-and-offer depth, are meant to control for the liquidity environment of the stock. The last three control variables—market capitalization, price level, and the standard deviation of 30-minute returns over the sample period—are meant to control for heterogeneity in fundamental attributes across stocks. We report results side by side for the three time intervals (I1, I2, and I3) with heteroskedasticity-consistent t-statistics. In Panel B, we replace the time-series correlation of all HFT firms with a time-series correlation measure computed only from pairs of HFT firms that load significantly on one of the three principal components (e.g., Cor(Y_i)PC1 for the eight firms that load on the first principal component). We also replace the control variable Y_i with Y_iPC1, which is the magnitude of HFT activity only for the HFT firms that load significantly on that principal component. To economize on the size of the table, and because we want to present results for the three measures of HFT strategies, the three time intervals (I1, I2, and I3), and the three principal components, we report in Panel B only the coefficient on the correlation measure from each of the regressions (with heteroskedasticity-consistent t-statistics).

Panel A: Regressions with Time-Series Correlations of All HFT Firms

	MSG			TRD			LMT		
	I1	I2	I3	I1	I2	I3	I1	I2	I3
Intercept	0.918	0.892	0.771	0.909	0.865	0.642	0.916	0.884	0.756
	(91.37)	(42.12)	(19.74)	(66.01)	(30.27)	(12.64)	(92.94)	(42.98)	(20.38)
Cor(Y_i)	-0.738	-1.135	-1.696	-0.757	-1.234	-0.627	-0.723	-1.093	-1.641
	(-5.28)	(-5.78)	(-5.72)	(-2.53)	(-3.13)	(-1.12)	(-5.29)	(-5.84)	(-5.94)
Y_i	-9.1E-06	-1.3E-04	-7.4E-06	-2.3E-02	-5.0E-02	-6.1E-02	-4.5E-06	-1.3E-04	-3.0E-05
	(-0.32)	(-1.85)	(-0.07)	(-5.64)	(-7.07)	(-5.08)	(-0.18)	(-2.21)	(-0.39)
Spread	0.777	1.134	1.692	0.772	0.968	1.797	0.821	1.194	1.696
	(2.90)	(3.09)	(3.29)	(3.28)	(3.39)	(3.00)	(3.08)	(3.29)	(3.32)
Depth	3.0E-08	-3.1E-08	-1.0E-08	1.5E-07	-6.7E-09	-9.9E-09	5.5E-08	-2.4E-08	-8.4E-09
	(0.53)	(-3.53)	(-5.69)	(2.40)	(-0.64)	(-4.89)	(0.98)	(-2.84)	(-4.76)
MktCap	-7.0E-10	-1.6E-09	-1.8E-09	-2.7E-10	-5.5E-10	-7.1E-10	-6.6E-10	-1.6E-09	-1.6E-09
	(-5.63)	(-7.25)	(-5.35)	(-1.94)	(-2.53)	(-2.04)	(-5.06)	(-6.68)	(-4.48)
Price	2.4E-04	3.8E-04	7.8E-04	1.9E-04	3.2E-04	1.1E-03	2.1E-04	3.4E-04	7.4E-04
	(2.02)	(1.69)	(2.10)	(1.99)	(2.19)	(3.90)	(1.77)	(1.57)	(2.02)
Volatility	-3.170	-8.521	-7.023	0.078	0.635	8.645	-3.228	-8.212	-5.682
	(-1.14)	(-1.64)	(-0.94)	(0.03)	(0.14)	(0.97)	(-1.16)	(-1.59)	(-0.78)

Panel B: Coefficients from Regressions using only HFT Firms with Significant Loadings on a Principal Component

	MSG			TRD			LMT		
	I1	I2	I3	I1	I2	I3	I1	I2	I3
Cor(Y_i)PC1	-0.131	-0.256	-0.440	-0.249	-0.404	-0.702	-0.130	-0.255	-0.432
	(-3.88)	(-4.92)	(-5.79)	(-2.66)	(-3.04)	(-3.80)	(-3.87)	(-4.90)	(-5.73)
Cor(Y_i)PC2	-0.183	-0.264	-0.419	-0.284	-0.393	-0.671	-0.172	-0.264	-0.460
	(-4.24)	(-3.70)	(-3.93)	(-4.29)	(-2.93)	(-3.63)	(-4.33)	(-3.94)	(-4.28)
Cor(Y_i)PC3	0.005	-0.020	-0.044	-0.062	-0.085	-0.031	0.004	-0.019	-0.044
	(0.22)	(-0.52)	(-1.02)	(-0.74)	(-1.12)	(-0.39)	(0.16)	(-0.53)	(-1.05)

Table 10

Trading Venue Competitiveness and the Correlation of HFT Strategies

This table presents summary statistics of measures of competitiveness (or viability) of trading venues and regression results that relate them to competition between high-frequency trading (HFT) firms. The five trading venues we investigate are all organized as electronic limit-order books and together execute 97.7% of the volume during our sample period. We denote the five trading venues in the table by the letters A through E. Panel A provides cross-sectional summary statistics for market share in terms of volume as well as for the two measures of trading venue viability or competitiveness: (i) %TimeBestPrices, defined as the percentage of time that the trading venue posts either the best bid or the best ask in the market (where the market is defined as the aggregation of all five trading venues), and (ii) %TimeSmallSpreads, defined as the percentage of time that the bid–ask spread on the trading venue is the narrowest spread in the market. In Panel B we examine whether correlated activity of HFT firms on a particular trading venue is helpful for the competitive position of the trading venue by manifesting in better prices and spreads. Not all HFT firms in our sample are very active on multiple trading venues. We therefore focus on the firms with substantial activity on all five trading venues, which we define as sending at least 10,000 messages to each of the five trading venues. There are eight HFT firms that satisfy this criterion, and we compute trading-venue-specific time-series correlations that provide information regarding whether the strategies of these HFT firms are correlated over time for a given stock on a particular trading venue. These correlations are similar in nature to the time-series correlation measure from Table 8, except they are computed for each trading venue separately using only activity on that trading venue. For each trading venue *v*, we run the following cross-sectional regression:

$$C_{iv} = a_0 + a_1 Cor\left(Y_{iv}\right) + a_2 Y_{iv} + a_3 Spread_{iv} + a_4 BBOdepth_{iv} + a_5 MktCap_i + a_6 Price_i + a_7 Volatility_i + error_{iv}$$

where C_{iv} is one of the two viability measures, and we use MSG (total messages sent by the HFT firm) as our measure of HFT strategies for computing Cor(Y_{iv}) and Y_{iv}. The next two variables, average spread and bid-and-offer (BBO) depth, are computed separately for each trading venue and are meant to control for the liquidity environment of the stock on that trading venue. The last three control variables—market capitalization, price level, and the standard deviation of 30-minute returns over the sample period—are meant to control for heterogeneity in fundamental attributes across stocks. To economize on the size of the table, and because we want to present results for the five trading venues and the three time intervals (I1, I2, and I3), we report in Panel B only the coefficients on Cor(Y_{iv}) from each regression (with heteroskedasticity-consistent t-statistics).

Panel A: Summary Statistics for Market Share and Competitiveness Measures

		A	B	C	D	E
%Market Share of Volume	Mean	14.32%	11.83%	0.40%	1.91%	69.26%
	Std.Dev.	6.68%	3.43%	0.60%	1.29%	7.99%
	Min	4.77%	3.73%	0.01%	0.42%	45.07%
	25^{th} Perc	9.50%	9.78%	0.05%	1.19%	64.29%
	Median	13.53%	12.21%	0.10%	1.59%	70.44%
	75^{th} Perc	16.67%	14.03%	0.48%	2.33%	75.90%
	Max	34.79%	17.57%	2.70%	7.24%	82.76%
	N	52	52	52	52	52
%TimeBestPrices	Mean	65.2%	83.2%	32.7%	66.0%	92.0%
	Std.Dev.	22.6%	13.7%	17.4%	33.0%	6.1%
	Min	15.0%	39.6%	8.1%	0.1%	71.1%
	25^{th} Perc	50.9%	80.6%	20.8%	37.0%	89.8%
	Median	71.9%	87.0%	28.0%	79.6%	92.5%
	75^{th} Perc	84.9%	91.0%	40.7%	96.3%	96.7%
	Max	92.2%	98.0%	82.8%	99.3%	98.9%
	N	52	52	52	52	52
%TimeSmallSpread	Mean	36.1%	54.2%	9.4%	26.9%	76.8%
	Std.Dev.	24.0%	15.8%	14.3%	19.9%	8.0%
	Min	1.8%	13.2%	0.1%	0.0%	50.4%
	25^{th} Perc	13.4%	44.2%	1.9%	6.4%	72.8%
	Median	37.0%	58.1%	3.3%	27.3%	78.5%
	75^{th} Perc	57.1%	65.3%	10.7%	44.9%	82.5%
	Max	83.5%	83.8%	69.3%	67.6%	89.8%
	N	52	52	52	52	52

Panel B: Coefficients on Cor(Y_{iv}) from Regressions of Competitiveness Measures

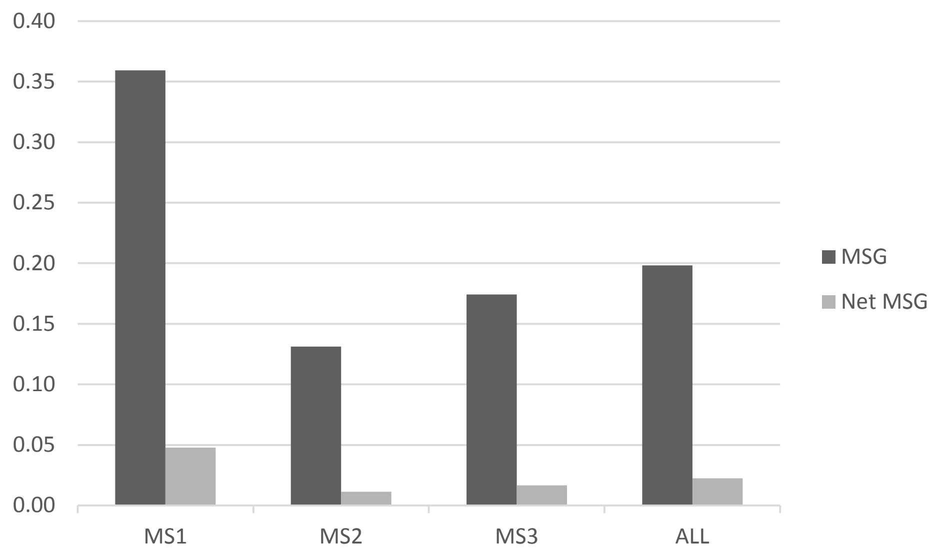
	Interval	Venue A	Venue B	Venue C	Venue D	Venue E
%TimeBestPrices	I1	0.658	0.220	0.970	4.281	-0.340
		(2.15)	(0.92)	(1.27)	(3.20)	(-3.64)
	I2	0.704	0.325	1.474	2.091	-0.199
		(2.85)	(1.46)	(2.17)	(1.11)	(-2.15)
	I3	0.661	0.348	1.427	0.781	-0.096
		(2.70)	(1.76)	(3.31)	(0.53)	(-1.26)
%TimeSmallSpread	I1	1.218	0.891	0.735	2.472	-0.875
		(2.64)	(2.61)	(1.17)	(2.77)	(-4.63)
	I2	1.188	0.891	0.966	1.375	-0.660
		(2.77)	(2.73)	(1.79)	(1.39)	(-3.37)
	I3	1.042	0.794	0.818	0.539	-0.438
		(2.60)	(2.68)	(2.48)	(0.72)	(-2.72)

Figure 1

Correlations of HFT Strategies: Total versus Directional

This figure compares the correlations of HFT strategies for total versus directional measures. We use data from the Investment Industry Regulatory Organization of Canada (IIROC) to identify 31 HFT firms. These firms are further categorized into three subgroups according to market share—MS1 (market share > 4%), MS2 (market share of between 1% and 4%), and MS3 (the rest). We compare the magnitudes of the cross-sectional (Panel A) and time-series (Panel B) correlations for total HFT activity (MSG, defined as buy plus sell orders) and directional HFT activity (NetMSG, defined as buy minus sell orders). The cross-sectional correlation measure indicates whether the strategies of HFT firms are correlated across stocks in a particular time interval. For each 1-second time interval, we compute the correlation coefficient between the activities of two HFT firms across the stocks in the sample, and average the correlations for all pairs of firms. The time-series correlation indicates whether the strategies of HFT firms are correlated over time for a particular stock. For each stock, we compute the correlation coefficient between HFT activities of any two HFT firms, and average these across all pairs of firms.

Panel A: Cross-Sectional Correlations



Panel B: Time-Series Correlations

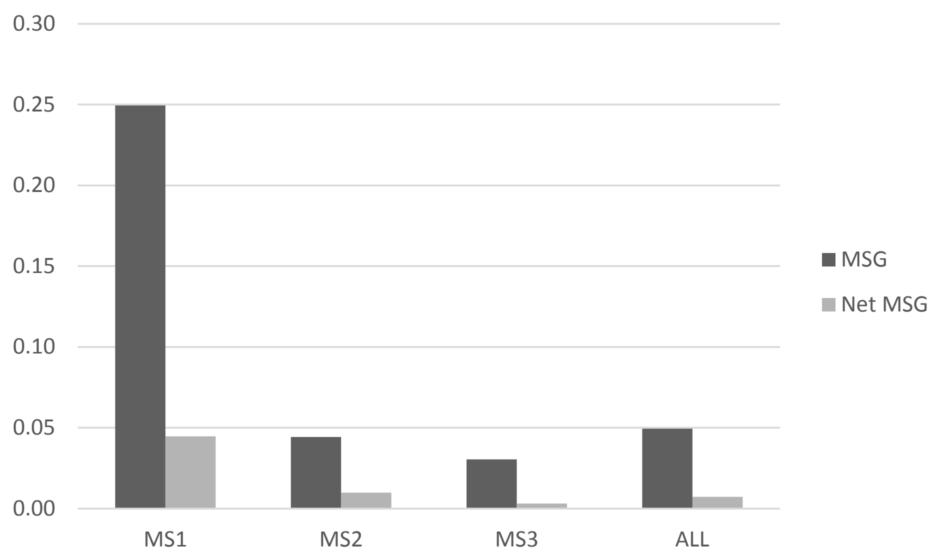
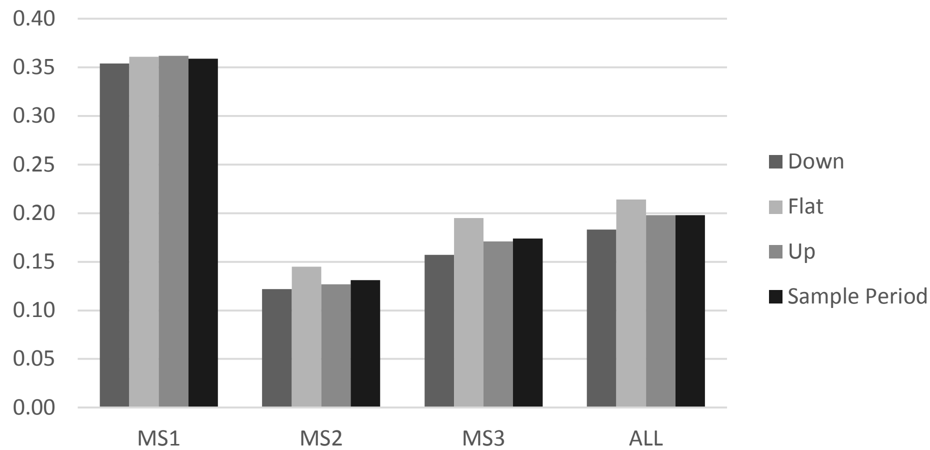


Figure 2

Correlations of HFT Strategies and Market Conditions

This figure compares the correlations of HFT strategies for varying market conditions. We use data from the Investment Industry Regulatory Organization of Canada (IIROC) to identify 31 HFT firms. These firms are further categorized into three subgroups according to market share—MS1 (market share > 4%), MS2 (market share of between 1% and 4%), and MS3 (the rest). We compare the magnitudes of the cross-sectional (Panel A) and time-series (Panel B) correlations for down days, flat days, and up days. Specifically, we rank the daily returns of the S&P/TSX Composite index from June 2010 through March 2011, and select the 10 worst days (Down Days), the 10 best days (Up Days), and the 10 days closest to and centered on zero return (Flat Days), for a total of 30 days (Sample Period). For each period we examine the correlations of HFT strategies in terms of the number of "messages" (MSG) they send to the market, where messages are defined as submissions and cancellations of nonmarketable limit orders as well as execution of marketable limit orders. The cross-sectional correlation measure indicates whether the strategies of HFT firms are correlated across stocks in a particular time interval. In each 1-second time interval, we compute the correlation coefficient between the activities of two HFT firms across the stocks in the sample, and average the correlations for all pairs of firms. The time-series correlation indicates whether the strategies of HFT firms are correlated over time for a particular stock. For each stock, we compute the correlation coefficient between HFT activities of any two HFT firms, and average these across all pairs of firms.

Panel A: Cross-Sectional Correlations



Panel B: Time-Series Correlations

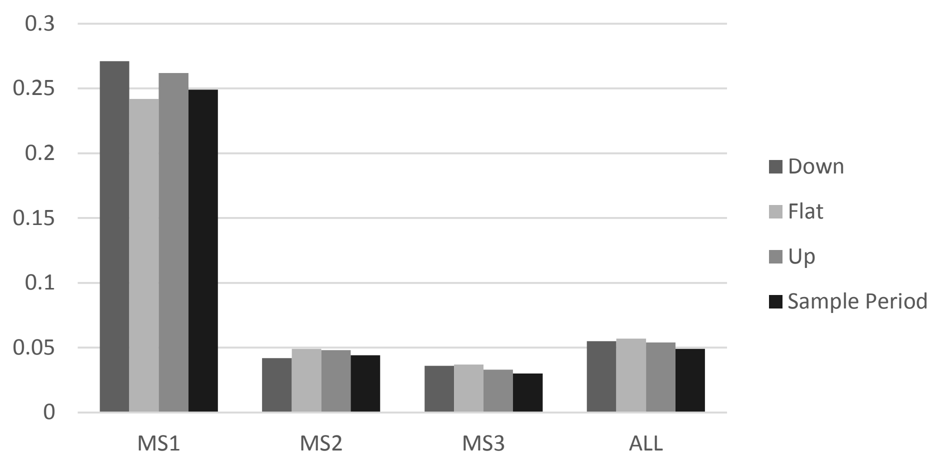


Figure 3
Average Daily Net Trading Revenues of HFT Firms

This figure presents a histogram of the average daily net trading revenues per stock in Canadian dollars of the HFT firms in the sample. We use data from the Investment Industry Regulatory Organization of Canada (IIROC) to identify 31 HFT firms, and compute their average daily net trading revenues per stock as follows. We sum the positive cash inflows (how much they get from selling shares) and negative cash outflows (how much they pay for buying shares) for each HFT firm in each stock and on each day. We then assume that the shares left at the end of the day are "liquidated" using the end-of-day midquote or closing price, and start every day with zero inventory. We compute the average of the stock/day net trading revenues for each of the 31 HFT firms ("X" in the figure) and present them in a histogram.

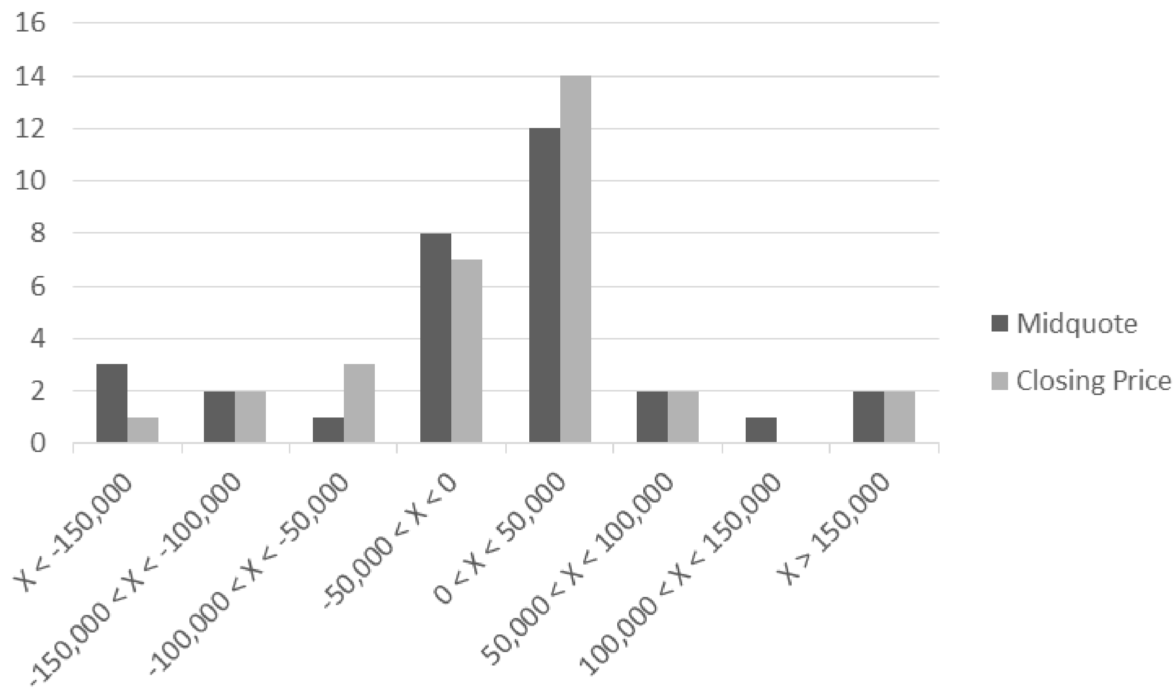


Figure 4

Cross-Sectional Correlation of HFT Strategies over the Trading Day

This figure presents the intraday pattern of cross-sectional correlations of HFT strategies. The measure of HFT activity that we use to characterize the strategies is MSG, which is comprised of all messages an HFT firm actively sends to the market in an interval (submission of non-marketable limit orders, cancellation of non-marketable limit orders, and marketable limit orders that result in trade executions). We carry out the analysis using 1-second intervals. We use data from the Investment Industry Regulatory Organization of Canada (IIROC) to identify 31 HFT firms. In each 1-second time interval, we compute the correlation coefficient between the activities of two HFT firms across the stocks in the sample, and average the correlations for all pairs of firms to obtain the cross-sectional correlation. Our sample consists of 52 stocks from the S&P/TSX60 (S&P). We rank the daily returns of the S&P/TSX Composite index from June 2010 through March 2011, and select the 10 worse days, the 10 best days, and the 10 days closest to and centered on zero return for a 30-day sample period. In the figure, we plot the average over the 30-day sample period of the cross-sectional correlation for each interval during the day.