Risk and Return in High-Frequency Trading

Matthew Baron, Jonathan Brogaard, Björn Hagströmer and Andrei Kirilenko[1]

First Draft: August 2012

Current Draft: December 2015

Abstract

This paper investigates the importance of relative latency among high-frequency traders (HFTs) in explaining their cross-sectional variation in performance. While the median HFT realizes an annualized Sharpe Ratio of 2.8 and a four-factor annualized alpha of 14%, the three fastest HFTs each month realize an average Sharpe Ratio of 12.7 and a four-factor alpha of 72%. Consistent with theory suggesting that competition on relative latency leads to a concentrated industry, concentration of HFT revenues and trading volume is high and non-declining over the five-year sample. New entrants are slower, earn lower revenues and are more likely to exit. Latency helps explain the success of HFTs across a variety of strategies, including liquidity provision and cross-market arbitrage.

[1] Contact: Matthew Baron, Johnson Graduate School of Management, Cornell University, e-mail: baron@cornell.edu; Jonathan Brogaard, Foster School of Business, University of Washington, e-mail: brogaard@uw.edu; Björn Hagströmer, Stockholm Business School, email: bjorn.hagstromer@sbs.su.se; and Andrei Kirilenko, Brevan Howard Centre for Financial Analysis, Imperial College Business School, e-mail: a.kirilenko@imperial.ac.uk.

The authors would like to thank Hank Bessembinder, Tarun Chordia, Thierry Foucault, Charles Jones, Terry Hendershott, Andrew Karolyi, Robert Korajczyk, Ananth Madhavan, Katya Malinova, Maureen O’Hara, Neil Pearson, Ryan Riordan, Gideon Saar, Ronnie Sadka, and Wei Xiong for their valuable feedback.

Traditional models of market making argue that competition among market intermediaries should decrease their profits and lead to lower trading costs for other investors (Ho and Stoll, 1983; Weston 2000). Several models of high-frequency trading (HFT) adopt this view (Bongaerts and Van Achter, 2015; Jovanovic and Menkveld, 2015; Aït-Sahalia and Saglam, 2014; Menkveld and Zoican, 2015). Other theories offer a contrasting perspective, that competition on speed makes the HFT industry different and leads to a distinct competitive environment (Budish, Cramton and Shim, 2015; Biais, Foucault and Moinas, 2015; Foucault, Kozhan and Tham, 2015). For example, Budish, Cramton and Shim (2015) theorize that small increases in trading speed lead to discontinuous differences in payoffs, as the fastest HFT responds first to profitable trading opportunities, capturing all the gains. Marginally slower HFTs arrive too late. As a result, small differences in trading speed are associated with large differences in trading revenues.

Consistent with the latter view we present evidence that relative latency, is important in explaining the cross-sectional variation in HFT performance. In particular, small differences in trading speed are associated with large differences in trading revenues. Relative latency appears important to the success of HFTs in both their market-making and cross-market arbitrage strategies. Consistent with Budish, Cramton and Shim (2015), we find that the HFT industry is concentrated among a few firms. In contrast to the traditional view that increased competition over time leads to lower profits, HFT concentration of trading revenues and trading volumes are high and non-declining over the five year sample, despite new HFT entry and a decline in overall HFT latency. We furthermore find that new HFT entrants are slower, earn lower trading revenues, and are more likely to exit, which likely reinforces concentration in the HFT industry.

The industrial organization of HFTs has important implications for market quality. Biais, Foucault and Moinas (2015) and Budish, Cramton and Shim (2015) argue that competition on relative latency can lead to an inefficient and costly arms race. This is because competition on relative latency gives rise to a "positional externality" (Frank, 2005), since a firm that becomes faster lowers the relative latency of its competitors, while the discontinuous difference in payoffs provides strong incentives to become marginally faster than other HFTs through greater technological investment. Foucault, Kozhan and Tham (2015) show that competing on relative speed can increase adverse selection for market makers, resulting in higher bid-ask spreads. Finally, a high concentration of market intermediaries could adversely affect market stability. For example, a dominant HFT that incurs a technical malfunction, suffers large capital losses, or suddenly withdraws from providing liquidity could lead to market fragility.

To test the role of latency for HFTs and study more generally the industrial organization of the HFT market, we use a proprietary transaction-level data set with trader identifiers provided by the Swedish financial supervisory authority, Finansinspektionen. The data contain all trades of Swedish equities from January 2010 to December 2014 from all venues including regulated exchanges, multilateral trading facilities (MTFs) and dark pools. Given the high degree of fragmentation of volume in European equity trading, this is an important feature to get the whole picture of trading. In addition, the five-year length of our data is important in allowing us to trace the "long-term" evolution of the HFT industry, at least relative to the rapid pace of innovation in the industry.

We focus on the 25 largest Swedish stocks by market capitalization, as Hagströmer and Nordén (2013) show that HFT activity is mainly concentrated in these stocks.[1] We classify HFT as those traders who self-describe as HFT through their membership in the *European Principal Traders Association* (*FIA-EPTA*; a lobby organization for principal trading firms formed in June 2011) website and any other firm who, according to its own website, undertakes low-latency proprietary trading.[2] The 16 firms that we identify as HFTs all have international trading operations and none of them are headquartered in Sweden. Thus, it is unlikely that the findings reported in this paper are specific to the Swedish context.

The main measure in this paper is *trading revenues*, captured daily for each HFT firm as the net of purchases and sales, marking end-of-day positions to market.[3] We document that the median HFT firm realizes daily trading revenues of 13,224 Swedish *krona* (SEK; 1 SEK = 0.105 USD as of December 2014) in Swedish equities, generating an annualized Sharpe Ratio of 2.8 and a four-factor annualized alpha of 14%. However, the median HFT firm differs substantially from the most successful HFT firms, whose performance metrics are an order of magnitude

[1] The data availability of the Swedish equity market has made it one of the most analyzed markets in the HFT literature. Hagströmer and Nordén (2013) show that HFTs are highly active in this market, constituting around 30% of the trading volume and more than 80% of the order volume. Other empirical studies on this market are Breckenfelder (2013); Brogaard, Hagströmer, Nordén and Riordan (2015); Hagströmer, Nordén and Zhang (2014); van Kervel and Menkveld (2015); and Menkveld and Zoican (2015).

[2] As a robustness check, we alternatively classified all trading accounts as HFT according to observed trading behavior (e.g., if firm median daily trading volume > X and median end-of-day inventory as a percent of firm trading volume < Y). The alternative specification addresses the possibility that some firms may not advertise themselves as HFTs. Classification based on observed trading behavior produces nearly the exact same list of HFTs as our main approach based on self-reporting.

[3] Since our data set does not convey trading fees or other operational cost, we are unable to directly calculate trading profits. However, in Section III, we analyze regulatory filings of four major HFT firms (Virtu, 2011-2014; Knight Capital Group, 2013-2014; GETCO, 2009-2012; and Flow Traders, 2012-2014), which allow comparison of trading revenues and profits. We do not find any evidence suggesting that higher trading revenues are associated with higher fixed costs and show that HFT revenue variation is a good proxy for variation in HFT profits.

higher. HFT firms exhibit large differences in performance, with trading revenues disproportionally accumulating to a few firms.[4]

We test the connection between HFT latency and trading revenues. Following Weller (2013), our measure of latency is the minimum time from a passive execution to a subsequent aggressive execution by the same firm, in the same stock and at the same trading venue. This measure aims to capture the reaction time involved in a *deliberate* decision to trade (i.e., an aggressive HFT trade) in reaction to a market event (i.e., an HFT's limit order being executed), which the HFT may see as informative. As an example of a strategy our measure may capture, Clark-Joseph (2012) shows that HFTs use the execution of small test orders as a signal to trade on incoming order flow ahead of public order book feeds. Over our five-year sample period we show that the latency of the fastest HFT firm falls substantially, from 62 microseconds (10^{-6} seconds) to 4 microseconds.

We find that relative latency, not nominal latency, drives differences in performance across HFTs. Relative latency measures how fast a HFT firm is relative to other HFTs and is captured by ranking HFTs by their calculated latency measure. Nominal latency measures how fast a HFT is in absolute terms and is captured by the log of a HFT firm's calculated latency measure. We find that being in the top-five, top-three, or even the top-one in terms of latency each month predicts substantially heightened revenues. It is not being fast that allows an HFT to capture trading opportunities, it is being fastest, consistent with Biais, Foucault and Moinas (2015) and Budish, Cramton and Shim (2015). Furthermore, we find that the fastest HFTs capture more trading opportunities but they do not earn higher revenues per trade. Our interpretation of this last finding is that the fastest HFTs are no more accurate on a per-trade

[4] Despite the strong performance of a small number of HFT firms, the average cost of HFT intermediation paid by non-HFTs is small, approximately 0.2 basis points.

basis at processing information than other traders (trade quality), but their latency advantage allows them to capture more trading opportunities (trade quantity).[5]

If the traditional view of market-making competition holds (Ho and Stoll, 1983; Weston 2000), the alpha generated by HFTs and the concentration of revenues should disappear as the industry matures. Budish, Cramton and Shim (2015) furthermore argue that if HFTs compete on relative latency, then increased competition does not drive profit opportunities to zero, since, regardless of how fast the market as a whole becomes, there will always be a relatively fastest firm that can use its speed advantage to adversely select other traders. Additionally, rents will remain concentrated among the fastest HFTs, as slower HFTs will arrive too late to compete.

If small differences in latency matter, then the HFT industry should be characterized by persistence in performance, high concentration and a lack of competition. We find evidence consistent with these predictions. Firms exhibit strong persistence in trading performance. Firm-level trading revenues are high and non-declining over the five-year sample. Similarly, the HFT concentration, in terms of both trading volumes and trading revenues, is high and non-declining. New HFT entrants are slower, earn lower trading revenues and are more likely to exit the market than incumbent HFTs. The characteristics of the HFT industry are thus consistent with relative latency being an important determinant of performance.

Lastly, we investigate the role of relative latency in specific HFT strategies. Some theories view fast traders as using speed and aggressive orders to trade on short-lived information – whether in reaction to news, order flow, or latency arbitrage (e.g., Cartea and Penalva 2012; Foucault, Hombert and Rosu, 2015; Biais, Foucault, and Moinas, 2015; Foucault,

[5] This finding also justifies why we emphasize *Revenues* as the main measure of performance. If strategies are not easily scalable, trade quality measures such as the *Sharpe Ratio* and *Revenues per MSEK Traded* are less relevant for comparing firms (Chen, Hong, Huang, and Kubik, 2004): a firm that has high revenues per trade but that captures few trading opportunities may not be considered a strong performer.

Kazhan, and Tham 2015; and Roşu 2015). Other theories view fast traders as passive liquidity providers who use speed to avoid adverse selection and inventory costs (e.g., Jovanovic and Menkveld, 2015; Aït-Sahalia and Saglam, 2014; Hoffmann, 2014). Menkveld and Zoican (2015) posit that both these types can co-exist in equilibrium.

We examine the role of latency in short-lived information strategies, liquidity provision, as well as cross-market arbitrage. We proxy skill in short-lived information trading by the ability of a market order to predict price changes over the next 10 seconds, and skill in liquidity provision as the ability of a passive order to capture a large realized spread. Relative latency is associated with better performance for both trading activities. To study cross-market arbitrage we examine HFTs' equity trading following changes in index futures prices. In the second after a change in the index futures price, the fastest HFTs are more likely than other HFTs to aggressively trade in individual equities in the direction of the futures price change. The behavior is consistent with pursuing cross-market arbitrage. The fastest HFTs are also less likely to supply liquidity to equities trades in the direction of the futures price change, which is consistent with avoiding adverse selection. We thus find that relative latency helps to explain the outperformance of HFTs pursuing a variety of strategies.

We add to the literature on competition between HFTs and the industrial organization of the HFT industry. Most empirical papers on HFT and algorithmic trading evaluate market quality (e.g., Hendershott, Jones, and Menkveld, 2011; Boehmer, Fong, and Wu, 2015; Riordan and Storkenmaier, 2012; Gai, Yao, and Ye, 2013; Malinova, Park, and Riordan, 2013; Hasbrouck and Saar, 2013). We instead look at HFT from a risk and returns perspective and analyze the incentives and competitive forces that shape the market for HFT. Boehmer, Li, and Saar (2015) study competition between HFT firms within three distinct strategies and show that increased competition is associated with lower volatility and the migration of trading volume to newer venues. Brogaard and Garriott (2015) analyze entry and exit of HFT firms and show that increased HFT competition increases market liquidity. Our paper's contribution is to show that competition on relative speed may naturally lead to a highly concentrated industry.

The paper is organized as follows. Section II describes the data and the empirical methodology; Section III characterizes the performance of the HFT industry; Section IV examines the role of latency in HFT performance; Section V considers the implications of latency for HFT industry organization; Section VI studies how speed may be useful for different HFT trading strategies; and Section VII concludes.

II. Data and Methodology

A. Data

The primary data source is the *Transaction Reporting System* (*TRS*), a proprietary data set provided to us by Finansinspektionen, the Swedish financial supervisory authority. According to the *Markets in Financial Instruments Directive* (*MiFID*), financial institutions in the European Union that are under the supervision of one of the national financial supervisory authorities must report all their transactions with financial instruments to TRS.

The TRS data has two features that make it highly suitable for the analysis of revenues in equity trading. First, the scope of the reporting obligation spans transaction at all trading venues, including regulated exchanges, MTFs, and dark pools. Even with the high degree of fragmentation of volume in European equity trading, the data allow us to observe all trading. Second, TRS contains identifiers (name, business identifier code, and address) for both the trading entity reporting the transaction and its counterparty. If the reporting entity undertakes the transaction as a broker for another financial institution, the identifiers for the client institutions are reported, too. The trader identifiers are necessary to identify HFTs and to analyze revenues in the cross-section of firms. Finally, the TRS data contains standard transaction-level variables such as date, time, venue, price, currency, quantity, and a buy/sell indicator. See Appendix Section A1 for information about the filtering procedures applied to the TRS data.

We limit the sample to the constituents of the leading Swedish equity index, the OMX S30 in order to focus on the most liquid stocks where HFTs primarily operate (Hagströmer and Nordén, 2013). We exclude six stocks that are cross-listed in other currencies, because revenue calculations for such stocks would require transaction data for foreign exchange markets.[6] There is only one index constituent change during the sample period. We include *Kinnevik Investment AB* after its inclusion in the index on July 1, 2014, and we include *Scania AB* up until May 16, 2014, when it ceased trading. The final sample has 25 stocks covering the period January 4, 2010 to December 30, 2014.

We match the TRS transactions to transaction-level data available from the *Thomson Reuters Tick History* (*TRTH*) data base. The TRTH data base is accessed through the *Securities Research Centre of Asia-Pacific* (SIRCA). The purpose of the matching is twofold. First, whereas the TRS data has second-by-second timestamps that may be subject to reporting delays, TRTH has exchange-reported timestamps at a microsecond granularity. Through the matching we can assign accurate microsecond timestamps to the TRS data, which is important for our latency measurement. Second, TRTH also contains order book information recorded on a microsecond frequency synchronized to the transaction data. This enables us to assess the status

[6] The following OMX S30 stocks are excluded due to this constraint: *ABB Ltd, Nokia Corporation, TeliaSonera AB, Nordéa Bank AB, AstraZeneca PLC,* and *LM Ericsson B*.

of the order book just before each TRS transaction, which is necessary to measure for example the effective spread, and to determine whether the trade was initiated by the buyer or the seller (following Lee and Ready, 1993).

B. Trading Venues and Stock Characteristics

All the sample stocks have their primary listing at NASDAQ OMX Stockholm, which is open for continuous electronic limit order book trading from 9 am to 5.25 pm on weekdays. For details about the trading mechanism at NASDAQ OMX Stockholm, see Hagströmer and Nordén (2013). Other important trading venues (by trading volume) in our data are Chi-X, BATS, Turquoise (all based around London) and Burgundy (based in Stockholm). In February 2011, BATS and Chi-X merged at the corporate level, but they maintain separate trading venues throughout our sample period. Burgundy was acquired by Oslo Börs in 2012. All sample stocks are subject to mandatory central counterparty clearing.

Table 1 reports descriptive statistics for the sample stocks. *Market Capitalization* (at closing prices on December 31, 2014) ranges from 13,877 million SEK (henceforth MSEK) for *SSABa*; to 475,595 MSEK for *HMb*, the equivalent of 1.65 to 47.56 billion USD, using the exchange rate of December 31, 2014. In the US equities market, these stocks would be labeled as large or mid cap stocks.

INSERT TABLE 1 ABOUT HERE

Daily Trading Volume refers to trading at NASDAQ OMX Stockholm only and is reported in MSEK. *Daily Turnover* is the Daily Trading Volume divided by Market

Capitalization, expressed in percentage points. *Tick Size* is the average minimum price change. *Quoted Spread* is the average bid-ask spread prevailing just before each trade; and *Effective Spread* is the trade value-weighted average absolute difference between the trade price and the bid-ask midpoint. All spread measures are based on continuous trading at NASDAQ OMX Stockholm, expressed relative to the bid-ask spread, and presented in basis points. The *Tick Size* and the *Quoted Spread* are halved to be comparable to the *Effective Spread*. The *Daily Turnover* across stocks is 0.60% and the *Quoted Spread* and *Effective Spread* vary between 2 and 6 bps. The more liquid stocks in our sample have a turnover and spread similar to the US large-cap stocks studied by Brogaard, Hendershott, and Riordan (2014). The *Tick Size* for many stocks is close to the *Quoted Spread*, indicating that market tightness is frequently bounded by the tick size.

Finally, we report *Volatility*, the average 10-second squared basis point returns, calculated from bid-ask midpoints; and an index for the degree of volume fragmentation. The *Fragmentation Index* is defined as the inverse of a Herfindahl index of trading volumes across the five largest trading venues (BATS, Burgundy, Chi-X, NASDAQ OMX Stockholm, and Turquoise). This procedure implies that fragmentation is measured on a scale from 1 to 5 (the number of trading venues considered).[7] *Volatility* ranges from 3 to 17 squared basis points, and the *Fragmentation Index* varies across stocks between 1.76 and 2.32.

C. HFT Identification

[7] If there are *N* trading venues and they all have equal shares of the trading volume, the index takes its maximum value *N*. If all trading volume is concentrated to one venue the index takes its minimum value, which is 1. The index design is similar to the Fidessa Fragmentation Index, more details of which can be found at http://fragmentation.fidessa.com/faq/#faq2

Previous studies classify HFTs according to observed trading behavior (as in, e.g., Kirilenko et al., 2015) or using an exchange-defined classification (e.g., Brogaard, Hendershott and Riordan, 2014). We classify traders as those who self-describe as HFT by including firms that are members of the *FIA-EPTA* or that according to their own website primarily undertake low-latency proprietary trading. The advantage of this approach over a classification based on observed trading behavior is that we can verify that HFTs have the characteristics usually associated with them (e.g., high trading volume, short investment horizon, tight inventory management, see Securities and Exchange Commission, 2010). To include a firm as an HFT, we furthermore require that a firm must trade at least 10 MSEK a day (about 1.05 million USD as of December 2014) for at least 50 trading days, which excludes HFT firms that only trade for a handful of days across the five year sample. The above criteria yield a sample of 16 HFT firms.

Due to confidentiality requirements, we cannot report the financial institutions covered in the proprietary data set. However, to give a sense of the HFTs that trade in Swedish equities, Figure 1 lists HFTs that, according to public records, trade in the sample stocks at NASDAQ OMX Stockholm. These public records are available because NASDAQ OMX Stockholm applied mandatory post-trade transparency for all trades through March 2014, disclosing the counterparties of each trade at an exchange-member level (we access the trader identities through TRTH).[8]

INSERT FIGURE 1 ABOUT HERE

[8] Figure 1 is an incomplete listing of HFTs captured in our proprietary data set because: 1) some HFTs might trade exclusively at other venues than NASDAQ OMX Stockholm, and 2) some HFTs may not be direct members of NASDAQ OMX Stockholm.

Figure 1 shows that several well-known international HFTs, such as Flow Traders, GETCO, IMC and Virtu, are present in the market. All of the HFTs are firms with international trading operations, and none are headquartered in Sweden. As a result, it is unlikely that the findings reported in this paper are specific to the Swedish context.

Figure 1 also reports HFTs trading activity over the sample period. HFTs are ranked by their total trading volume in each month, which is visualized in grayscales ranging from white (low trading volume) to black (high trading volume). While several HFTs consistently trade large volumes throughout the entire sample period, other HFTs enter and exit during the sample period or show a time-varying trading volume ranking position. The dynamic nature of firm entry and exit is important for our analysis of the industrial organization of HFT.

D. HFT Performance Measures

The main performance measure in this paper is *Revenues*, defined as the cumulative cash received from selling shares, minus the cash paid from buying shares, plus the value of any outstanding end-of-day inventory positions, marked to the market price at close. We calculate *Revenues* for each HFT firm, each sample stock, and each trading day. Depending on the application we report *Revenues* for different frequencies of time, for individual HFT firms as well as across all firms in the industry, and for individual stocks or all stocks, but all versions of *Revenues* are aggregates of the same panel of firm-stock-day observations.

We assume zero beginning-of-day inventory positions as a way to overcome potential data errors (e.g., some trades may be missing or misreported in the data). Even minor errors in inventory can accumulate over time, leading to large and persistent (unit root) errors if left uncorrected. Therefore, we zero beginning-of-day inventories so that any potential errors do not affect more than one day. This assumption is relatively innocuous because Table 2 shows that most HFTs usually end the day near a zero position anyway.

In the Appendix Section A2, we compare our main method of calculating trading revenues with three alternative approaches, one of which relaxes the assumption of zero inventory at the start of the trading day (i.e., the trading revenues calculated by cumulating daily net inventory positions over the full sample). We show that the mean and median trading revenues are relatively close across the different methods (though the standard deviation widens, as expected, without the inventory corrections), consistent with the fact that inventory corrections reflect small white noise errors in end-of-day positions.

Other measures of HFT performance reported in this paper and defined in this section include *Trading Volume*, *Revenues per MSEK Traded*, *Returns*, factor model *Alphas* (using one, three, or four factors), and the *Sharpe Ratio*.

Trading Volume is the SEK volume traded by each HFT firm on each trading day. This may be seen as a measure of performance in an environment where HFTs compete to maximize profitable trading opportunities, as described by Rosu (2015). *Revenues per MSEK Traded* is calculated daily as *Revenues* divided by *Trading Volume*. *Revenues* and *Trading Volume* capture performance along the quality-dimension, *Revenues per MSEK Traded* capture performance along the quality-dimension.

Returns are calculated by dividing *Revenues* of each firm by the implied capitalization of the firm. The implied capitalization is calculated for each firm as the maximum position in SEK that an HFT's portfolio takes over the five-year sample. HFTs inventory generally exhibit sharp, well-defined maximum and minimum total portfolio positions. We use the observed maximum position as an approximate of the maximum amount of capital that an HFT would need to execute its specific strategy.[9] However, as moderate assumptions about capitalization are needed to calculate *Returns*, we do not emphasize this measure (or factor model *Alphas*) as our primary measure of HFT performance. *Returns* are calculated at daily frequencies, but throughout the paper it is reported in annualized terms.

Factor model *Alphas* are computed for each HFT over the entire sample using the standard Fama-French model (Fama and French, 1993) and the Carhart (1997) momentum factor. The Fama-French and Carhart daily factors are computed according to the methodology from Fama and French (1993) and Ken French's website, using the full sample of Swedish stocks traded on NASDAQ-OMX Stockholm. Methodological details concerning the construction of these factors and validation exercises can be found in the Appendix Section A3.

Finally, the annualized *Sharpe Ratio* for each HFT firm is calculated using daily observations as $SR_i = \frac{\mu_i - r_f}{\sigma_i} * \sqrt{252}$, where μ_i is the average daily return, r_f is the risk-free rate, and σ_i is the standard deviation of HFT firm *i*'s returns. Under the assumption that the risk-free rate can be neglected as it is nearly zero for much of the sample period, the *Sharpe Ratio* is equivalent to: $\frac{\mu(Revenues)_i}{\sigma(Revenues)_i} * \sqrt{252}$. The equity capitalization is therefore irrelevant for calculating the *Sharpe Ratio*.

Similar to *Revenues per MSEK Traded,* the *Sharpe Ratio* measures performance in the quality dimension. If strategies are not easily scalable (as is suggested by the evidence in Section IV), trade quality measures are less relevant for comparing firms (Chen, Hong, Huang, and Kubik, 2004): a firm that has high revenues per trade but that captures few trading opportunities may not be considered a strong performer. Nonetheless, the *Sharpe Ratio* and *Revenues per*

[9] In Section III we show that HFT returns calculated this way are comparable in magnitude to those from regulatory filings of four major HFT firms (Virtu, 2011-2014; Knight Capital Group, 2013-2014; GETCO, 2009-2012; and Flow Traders, 2012-2014), where one can directly observe book capitalization or net liquid assets available to trade.

MSEK Traded measures provide context as measures of risk-adjusted returns and the quality of trading signals.

E. HFT Latency

Our measure of latency is based on Weller (2013): it aims to capture a firm's quickest deliberate reaction speed to market events *within* each stock and venue. Specifically, we measure *Latency* monthly for each HFT firm as the minimum duration between a passive trade and a subsequent aggressive trade within a given venue and stock. To avoid large outliers we do not consider durations longer than one second.

The latency measure is designed to capture the reaction time involved in a *deliberate* decision to trade (i.e., an aggressive HFT trade) in reaction to a market event (i.e., an HFT's limit order being hit), which the HFT may see as informative. An HFT firm cannot control the timing of the passive order but can only react to it; thus, our latency measure captures reactions to incoming order flow, not how fast an HFT can execute two successive trades.[10]

Measuring latency within a given stock and venue (as opposed to across different stocks or venues) ensures that we are not measuring latency across different settings where the

[10] To ensure that *Latency* is not picking up trades that happen close to each other by chance, we calculate the probability of two successive trades – a passive trade followed by an aggressive trade – occurring by chance within a sub-millisecond interval. We find the probability to be small. Specifically, under the assumption that an HFT's trades within any venue or stock are uniformly distributed across a time period [0,T], the approximate probability that no two passive-to-aggressive trades over a period of T seconds will be within t seconds from each other is given by the formula: $[1-\frac{t}{T}]^{(N_{aggressive}N_{passive})}$, where $N_{aggressive}$ and $N_{passive}$ are the number of aggressive and passive trades, respectively. Applying conservative estimates of the parameters: T = 666,600 trading seconds per month, t = 10 microseconds (approximately the average latency of the top 5 fastest HFTs), and ($N_{aggressive}N_{passive}$) = 24,550 x 46,198 = 1.13 x 10^9 (an upper bound on the aggressive and passive trades corresponding to the 99th percentile of the maximum trading volume in any stock-venue combination for each HFT firm-month combination), the probability that no two passive-to-aggressive trades occur within 10 microseconds of each other by chance is 99.91%. The latency measure statistic is thus almost certainly not due to chance or related to trading volume.

microsecond time-stamp may not be accurately synchronized. Additionally, we can test whether within-market latency also explains success at arbitrage across markets (See Section VI).

Figure 2 plots *Latency* over the sample period 2010-2014. HFTs are grouped by their relative rank of latency per month; the categories are Top 1, Top 2-3, Top 4-5, and all HFTs. Over the sample period latency decreases for HFTs in the top 5: for example, the Top 4-5 HFTs decrease in latency over the sample period from around 1 millisecond to around 10 microseconds, while the Top 1 HFT decreases from around 10 microseconds to one microsecond. Each group in the Top 5 speeds up by around an order of magnitude, while All HFTs (disproportionately picking up the slower HFTs) remains relatively constant at around 1 millisecond over the entire sample period. The fact that HFTs outside the Top 5 do not achieve lower latencies over time is consistent with the finding of Brogaard, Hagströmer, Nordén and Riordan (2015) that not all HFTs choose to be the fastest when given the opportunity to choose a faster colocation technology.

INSERT FIGURE 2 ABOUT HERE

Figure 2 marks various technological upgrades: the introduction of INET in early 2010, a high-capacity trading system capable of handling over 1 million messages per second, and two colocation upgrades at NASDAQ OMX Stockholm in March 2011 and September 2012. *Latency* decreases following the technology upgrades provide suggestive evidence that our latency measure captures reaction time. While it is difficult to assess the impact of the INET upgrade since it comes at the start of the sample, the colocation upgrade of 2011 is followed by a sharp, immediate drop in latency for all HFT subgroups in the Top 5, along with a convergence in relative latency. The 2012 colocation upgrade also is followed by a decline in latency, though a more gradual one, which may reflect that as new colocation options become available it takes time to fully utilize the technology.

III. Characterizing HFT Performance

A. HFT Performance in the Cross-Section

To start, we document the risk and return characteristics of individual HFT firms. In Table 2 we report the cross-sectional distribution of HFT performance, latency, and other trading characteristics. For each variable, we retrieve the time-series average for each HFT firm, and then report the distributional statistics across firms.

INSERT TABLE 2 ABOUT HERE

The median HFT firm realizes an average daily *Revenues* of 13,244 SEK, or 65 SEK *Revenues per MSEK Traded*. It has a daily *Trading Volume* of 176 MSEK, an annualized *Sharpe Ratio* of 2.8, and a four-factor (Fama-French plus Carhart momentum) annualized Alpha of 14%. For comparison, the *Returns* are 16%, suggesting that exposure to well-documented risk factors is not particularly relevant for HFT firms.

We also find considerable variation in the cross-section: HFTs exhibit large differences in performance. The cross-sectional distributions of *Revenues* and *Returns* are skewed towards a few high performers. For example, firms in the top 90th percentile generate average daily *Revenues* of 63,463 SEK per day, compared with 13,244 for the median; a *Sharpe Ratio* of 11.2, compared with 2.8 at the median; *Revenues per MSEK Traded* of 342.59, compared with 65.9 at the median; and a four-factor annualized Alpha of 97%, compared with 16% at the median.

HFTs are diverse in terms of other trading characteristics, too. Beyond performance, we report the distributions of *Aggressiveness Ratio* (the market order volume in SEK divided by *Trading Volume*); *End-of-Day Inventory Ratio* (the end-of-day inventory divided by *Trading Volume*); *Max. Intraday Inventory Ratio* (the maximum intraday portfolio position divided by *Trading Volume*); *Average Trade Size* (in SEK); and *Latency* (in milliseconds). Consistent with the characterization of HFTs in the Securities and Exchange Commission's Concept Release on Equity Market Structure (2010) and with functional-based approaches for HFT classification (Kirilenko et al., 2015), most, though not all, HFTs tend to have low intraday and end-of-day inventories and a small average trade size. HFTs vary in their *Aggressiveness Ratio*, with some nearly all active or passive and others mixed (the average aggressive ratio is 53%). There is substantial variation in *Latency* across HFTs, a finding we take up further in Section IV.

B. Comparison of Trading Revenues to Trading Profits Based on Public Filings

The data do not convey trading fees or other operational costs and so we are unable to directly calculate trading profits. However, regulatory filings of four major HFT firms (Virtu, 2011-2014; Knight Capital Group, 2013-2014; GETCO, 2009-2012; and Flow Traders, 2012-2014) allow comparison of trading revenues and profits. A potential concern in our analysis of HFT performance is that firms with higher trading revenues may have higher fixed costs (i.e., higher trading revenue firms have to pay higher costs to achieve that higher performance), so that trading revenues are not a good proxy for firm profitability. We show that in practice this is not likely the case.

Table 3 reports *trading revenue*, *trading costs*, *trading profit margins*, and *trading returns* calculated from annual reports and IPO prospectuses for four high-frequency trading firms for which public data is available. Trading costs are broken down into several categories (e.g., trading and clearing fees, data costs, financing costs, equipment and technical costs), all expressed as a percent of trading revenues.

INSERT TABLE 3 ABOUT HERE

We make three observations. First, trading profit margins are high, ranging between 27.4% and 59.6% of trading revenue for all four firms. Approximately 50-80% of the HFT costs are trading-related fees: brokerage fees, exchange and clearance fees, financing costs. The fixed costs, including communications and data processing, equipment, administrative and technology costs, make up only 15-30% of the total costs.

Second, as a percentage of trading revenues, the fixed costs do not vary substantially across firms and over time. For example, in 2014, KCG had double the trading revenue of Virtu and five times the trading revenue of Flow Traders, but the total fixed costs as a percentage of trading revenue show no pattern (22.7% for KCG; 17.7% for Virtu; 27.2% for Flow Traders). There is also no clear time trend in fixed costs as a percentage of trading revenues. All else being equal, the stability of the fixed costs suggests that firms with higher trading revenues also have higher profits. As such, HFT revenue variation is likely a close proxy for variation in HFT profits.

Lastly, Table 3 reports *trading returns. Trading returns* are calculated two ways based on different capitalization measures: *trading revenue / (trading assets minus trading liabilities)* and *(trading revenue / book equity)*. From these public filings in which capitalization is directly observable, we find trading returns to range from 59.9% to almost 377%, depending on the firm. This suggests the returns computed in Section III.A are of a reasonable magnitude.

IV. The Role of Speed in Performance

Having documented the performance of HFTs, we now test our main hypothesis about speed and HFT trading revenues. While most theories in which HFTs earn profits posit that fast traders should have an advantage, some theories suggest that traders of different speed can specialize along other dimensions (Weller, 2013; Rosu, 2015). According to these models, a relatively slow market intermediary could compensate by providing deeper liquidity on the book and holding longer positions (i.e. provide greater risk-bearing capacity), thus making similar profits as fast traders in equilibrium. [11]

Motivated by the contrasting theories discussed in the introduction, we test whether latency, and especially relative latency, is associated with increased performance.

We estimate the regression:

$$Performance_{i,t} = \alpha_t + \beta_1 \log(Latency)_{i,t} + \beta_2 \mathbf{1}_{\text{top 1 } i,t} + \beta_3 \mathbf{1}_{\text{top 2-3 } i,t} + \beta_4 \mathbf{1}_{\text{top 4-5 } i,t}$$

[11] The latency measure varies widely in the cross-section of HFT firms. Deliberate differences in latency can be due to differences in trading strategies. For example, cross-market arbitrageurs may focus on minimizing geographical latency by investing in microwave links, whereas market makers may put more emphasis on colocation facilities to quickly access news in the local order flow (Brogaard, Hagströmer, Nordén and Riordan, 2015). Alternatively, slower HFTs may specialize in other dimensions. For example, relatively slow HFTs may provide liquidity on the limit order book and hold longer and larger positions (Weller, 2013; Rosu, 2015). Finally, differences in technological capabilities can persist because technological expertise and trading strategies are closely guarded trade secrets. Frictions in labor markets contracting may prevent the movement of human capital and technical knowledge across firms.

$$+\gamma' controls_{i,t} + \text{ month-FEs} + \epsilon_{i,t} \tag{1}$$

where $Performance_{i,t}$ can be any of HFT performance measures *Revenues*, *Returns*, *Sharpe Ratio*, *Revenues per MSEK Traded*, or *Trading Volume*. All dependent variables are aggregated across stocks to retrieve a firm-month frequency. Specifically, *Revenues* and *Trading Volume* are averaged across trading days; *Returns* and *Revenues per MSEK Traded* are calculated using the firm-month observations of *Revenues* and *Trading Volume*. When used as a dependent variable, the *Trading Volume* is entered in terms of the natural logarithm.

The independent variables $\mathbf{1}_{\text{top 1 } i,t}$, $\mathbf{1}_{\text{top 2-3 } i,t}$, and $\mathbf{1}_{\text{top 4-5 } i,t}$ are indicators for whether a given firm is ranked in the Top 1, Top 2-3, or Top 4-5 by speed in a given month. Nominal latency is captured using *Latency*, while the $\mathbf{1}_{\text{top 1 } i,t}$, $\mathbf{1}_{\text{top 2-3 } i,t}$, and $\mathbf{1}_{\text{top 4-5 } i,t}$ indicators capture relative speed. Since *Latency* can vary widely across firms, from the microsecond to the second level (see Table 2), the relationship between trading speed and trading revenues is best captured by taking logs. The indicator variables capture the potential non-linear relationship between relative latency and performance: the very fastest firms may do much better than firms that are only moderately faster than their competitors.

The control variables account for characteristics that may affect HFT performance, including measures of their risk-bearing capacity and strategies. These variables include the *End-of-Day Inventory Ratio*, *Max. Intraday Inventory Ratio*, and the *Aggressiveness Ratio*, which are defined in Section III.A and calculated on the monthly level for each HFT. *Max. intraday inventory* is used in the denominator to calculate *Returns* and is thus omitted when *Returns* is the dependent variable. The *Investment Horizon* is the median holding time in seconds across all trades of the firm-month, calculated on a first-in-first-out basis. Continuous independent variables are normalized to be in units of standard deviations. Month fixed-effects absorb time-varying market conditions, including market trading volume and volatility. Following Petersen (2009) and Thompson (2011) standard errors are dually clustered by firm and month to account for correlation across firms and time. Table 4 reports coefficient estimates for various specifications of the model described in Eq. (1).

INSERT TABLE 4 ABOUT HERE

Our first result is that being fast is associated with increased revenue. The first specification sets all slope coefficients except that of nominal latency (β_1) equal to zero, and shows a negative and statistically significant relation between *Revenues* and nominal latency.

The second result is that the effect of relative latency on *Revenues* dominates that of nominal latency. This is seen in the second specification, where the $\mathbf{1}_{\text{top 1}\ i,t}$, $\mathbf{1}_{\text{top 2-3}\ i,t}$, and $\mathbf{1}_{\text{top 4-5}\ i,t}$ indicators of relative latency are included along with the nominal latency variable. The lack of statistical significance and reduced economic magnitude for log(*Latency*) suggest that relative speed matters more than nominal speed. Specifically, the estimates show that being in the top 5 in rank speed predicts increased average daily trading revenues of 9,860 SEK (although β_4 is not statistically significant); with additional revenue of 22,842 SEK for being in the top 3; and an additional revenue of 58,213 SEK for being the top 1. Note that the way the regression is set up implies that the relative latency indicator coefficients are additive. These results hold up in the full specification with all control variables, though the coefficients are marginally smaller.

Several of the control variables are related to trading revenues. For example, a one standard deviation increase in the Max. Intraday Inventory is associated with decreased daily

Revenues of 9,698 SEK, suggesting that HFTs that have tighter inventory management perform better.[12] Similarly, HFTs that are more aggressive earn somewhat higher trading revenues. In Appendix Section A6, we re-estimate Eq. (1) but on the stock-level (i.e., on an HFT-stock-month panel), which allows us to analyze stock-month-level characteristics that may affect HFT performance. We find, for example, that *Revenues* are higher in stocks with greater non-HFT trading volume and *Volatility*, while we find no statistically significant association of HFT performance with the *Fragmentation Index*, *Tick Size* or *Quoted Spread*.

To understand why relative latency is important, we next analyze whether it is primarily related to the trading revenues per trade (quality) or the number of trades (quantity). If HFTs use latency advantages to better obtain and aggregate information in order to predict future price changes, we would expect the fastest HFTs to have the highest revenues per trade. However, according to Table 4, we find the contrary, that there is no association between *Latency* and *Revenues per MSEK Traded*. Instead, we find a strong relationship between trading speed and *Trading Volume*. The results imply that faster HFTs are able to capture a larger trading quantity, but that trading speed is not used to maximize the trading quality. It appears that the fastest HFTs are no more accurate at processing new information than other traders, but their latency advantage allows them to capture the most trading opportunities. This result supports the use of a measure of trade quantity, like *Revenues*, rather than a measure of trade quality, like *Revenues per MSEK Traded*, when evaluating HFT performance.

The results for latency effects on *Returns* is similar to those for *Revenues*. The only difference between *Returns* and *Revenues* is that the former is expressed relative the firm market

[12] In Appendix Section A5, we also show that being fast is associated with better inventory management (lower *End-of-Day Inventory Ratios* and shorter *Investment Horizons*), presumably reflecting the fact that faster traders can more skillfully rebalance their inventory positions intraday without incurring trading losses.

capitalization. The results thus indicate that the HFT firm size does not drive the relation to relative latency. Furthermore, we find that the *Sharpe Ratio* is higher for HFT firms with lower relative latency. This demonstrates that the relation between *Revenues* and relative latency is not driven by the risk of the trading strategies applied.

V. Implications of Speed for the HFT Industry

Section IV shows that relative differences in latency can help explain the variation in HFT performance. If trading speed drives performance, why are HFT firms not competing away the rents, either through new entry, higher trading volumes, or increased investment in speed and trading sophistication? Frictions in labor markets may prevent the movement of human capital and technical knowledge across HFT firms. For instance, technological expertise and trading strategies are closely guarded trade secrets and employees often agree to non-compete and non-disclosure agreements. Also, technology evolves quickly making the optimal arrangement a moving target that may be easier to find with past experience.

If achieving low latency is an informal barrier to competition, then the extant literature on limited competition provides predictions for the HFT industrial organization. Geroski, Gilbert, and Jacquemin (1990) provide an overview of the literature. The prediction for the HFT industry would be persistence in performance, both at the firm-level and industry-wide level. Other predictions include high concentration of HFT revenues and trading volume, along with difficulty of new entry. We examine each of these predictions and find evidence that all apply for the HFT industry.

A. Persistence in Firm-Level Performance

We first test for persistence at the firm-level. Large differences in HFT firm performance could potentially be driven by luck. For instance, in a model of identically skilled HFTs, all engaging in strategies with right-skewed distributions of performance, some will happen to outperform. Persistence shows that something other than luck drives a firm's performance.

There is an extant literature showing that performance for actively-managed mutual funds in period *t* generally does not predict performance in period *t+1* (e.g., Carhart, 1997). However, there is evidence of persistence by some investors (e.g., Jagannathan, Malakhov, and Novikov, 2010, for hedge funds; and Kaplan and Schoar, 2004, for private equity). Nonetheless, the expectation for most types of investors and funds is that ability has little influence on performance. The null hypothesis is thus that HFT firms do not exhibit persistence.

To analyze persistence we regress various measures of performance (*Revenues*, *Revenues per MSEK Traded*, *Returns*, and the *Sharpe Ratio*) on their lagged values, on both the daily and monthly frequency (except for the *Sharpe Ratio*, which is available on the monthly frequency only). Measures of performance are standardized in the cross section. That is, on each day, firm-level performance is centered on the mean and scaled by the standard deviation across firms. The standardization controls for potential time-variation in the mean and variance of returns. We also estimate persistence regressions with rank-order performance, ranking the relative performance of firms as 1, 2, 3, etc. for each measure of performance on each day or month. We estimate the following ordinary least squares (OLS) regression:

$$Performance_{i,t} = \beta\ Performance_{i,t-1} + \epsilon_{i,t}, \tag{2}$$

where $Performance_{i,t}$ is defined as above. Results based on daily observations are reported in Table 5, Panel A, and those using monthly observations are reported in Panel B. Since each measure is standardized in the cross-section, the constant term in the regression equation is mechanically zero. We refer to β as the persistence coefficient, with $\beta = 1$ meaning perfect persistence, and $\beta = 0$ meaning no persistence. Standard errors are dually clustered by firm and day in Panel A and by firm and month in Panel B.

INSERT TABLE 5 ABOUT HERE

We find that HFTs have statistically significant daily persistence coefficients of 0.226 for *Revenues* and 0.433 for *Returns*. On the monthly frequency, we find higher persistence coefficients: 0.604 for *Revenues*, 0.710 for the *Sharpe Ratio*, 0.192 for *Revenues per MSEK Traded*, and 0.428 for *Returns*. Performance is more strongly persistent at the monthly level, which is likely due to the higher daily idiosyncratic risk. The rank order analysis shows similar persistence. Consistent with our earlier argument that *Revenues per MSEK Traded* may be a less relevant performance metric for HFTs, we find lower persistence in this measure. In summary, we find that the performance of HFTs is stable over time, in terms of nominal performance as well as in terms of performance rankings.

B. Persistence in Industry-Wide Performance

Having found evidence of persistence in firm-level performance, we examine overall performance of the HFT industry over our five year sample. Given that the HFT industry is relatively new, we may observe decreasing industry-wide performance if competition is increasing over time. However, in a market with competition on relative speed, performance may not change. Budish, Cramton and Shim (2015), for example, argue that if HFTs compete on relative latency, regardless of how fast the market as a whole becomes, there will always be a relatively fastest firm that can use its speed advantage to adversely slower traders, capturing a stable level of rents regardless of nominal speed differences. Furthermore, if competition on relative speed makes new entry difficult, as we show in Section V.D, difficulty of new entry may keep aggregate HFT performance from declining.

Consistent with these predictions, we find that the HFT industry-wide performance is relatively stable over the five-year sample. Table 6, Panel A, reports average daily statistics aggregated across all HFTs and all stocks and reported in half-year intervals. The statistics include *Total Daily Revenues* (*Revenues* summed across all HFT firms), *Average Daily Revenues* (the *Total Daily Revenues* averaged across HFT firms), *Daily Trading Volume* (the *Trading Volume* summed across all HFT firms and reported in MSEK), *Average Revenues per MSEK Traded* (the ratio of *Total Daily Revenues* and *Daily Trading Volume*), and *Average Daily Returns* (*Returns* averaged across all HFT firms). Time trends for *Total Daily Revenues* and *Average Daily Revenues* are also plotted in Figure 3, Panels A and B.

INSERT TABLE 6 ABOUT HERE

INSERT FIGURE 3 ABOUT HERE

Our results show that *Total Daily Revenues* and *Average Daily Revenues* are relatively stable over the five-year period, while *Daily Trading Volume* trends up and *Average Revenues per MSEK Traded* trends down.[13]

To formally test the movement of the trends, we estimate the following OLS regression on daily observations:

$$Performance_t = \alpha + \beta\ (year - 2010) + \epsilon_t \tag{3}$$

where $Performance_t$ is one of the performance measures described above, and *year* is a continuous variable (e.g., *year* would take on the approximate value of 2014.25 on March 31, 2014). A positive (negative) coefficient on the $(year - 2010)$ variable corresponds to an increasing (decreasing) trend in the HFT industry-wide performance over the period 2010-2014. Newey-West standard errors with 30 day lags are used. The coefficient estimates are presented in Table 6, Panel B.

The tests statistically confirm the aforementioned trends. However, there is statistical evidence that trading revenues are slightly increasing over the sample period: a yearly increase of 20,447 SEK for *Total Daily Revenues*, relative to a baseline of 168,452 SEK in the first half of 2010. This increase is however not statistically significant in terms of *Average Daily Revenues*, which takes the number of HFTs in the industry into account.

We also examine time trends in the cost of HFT intermediation for non-HFTs, which is defined as HFT *Total Daily Revenues* divided by *Non-HFT Trading Volume* (the sum of SEK

[13] Returns slightly trend down, but, given the relatively stability of revenues per firm, this mechanically must mean that average firm capitalization is slightly increasing over the sample, since firm capitalization is a fixed characteristic calculated over the entire sample period.

trading volume that does not involve HFTs), calculated on a daily basis. This measure captures the amount of revenue paid from non-HFTs to HFTs per SEK traded. The cost of HFT intermediation for non-HFTs is relatively small, varying between 0.133 and 0.432 bps, depending on the month, which is about the same order of magnitude as typical exchange fees (see Appendix Section A4). Whereas the exchange fees are direct costs incurred to exchange members (and typically passed on to their clients), the cost of HFT intermediation is extracted indirectly through the trading process. Another way to put the cost of HFT intermediation into perspective is to compare it to the effective spread, which according to Table 1 is between 2 and 6 bps. That is, the cost of spreading the (half) bid-ask spread is more than ten times higher than the cost of HFT intermediation for non-HFTs. As seen in Figure 3, Panel C, and formally tested in the last column of Table 6, Panel B, there is a small upward time trend in the cost of HFT intermediation, starting from around 0.13 bps and increasing by about 0.05 on average per year.

C. HFT Market Concentration over Time

Our third industrial organization measure is the industry concentration. We find that HFT revenues are concentrated among a small number of firms, consistent with limited competition. A high concentration of trading revenues is not surprising given the wide disparity in trading revenues documented in Table 2. Examining market concentration provides two new insights into the HFT industry: first, it quantifies how a few firms dominate the market; second, it displays that concentration is stable over time.

Recall that the trading activity by a subset of HFT firms reported in Figure 1 shows various examples of entry and exit within the sample. In addition, Figure 2 shows that the latency differential across firms is decreasing over time. In light of new entry, exit, and latency convergence, one might expect the concentration of revenues to decrease over time.

In contrast to these expectations, we show that the HFT industry concentration remains high and relatively constant over the five-year sample. We calculate Herfindahl indices, a commonly used measure of concentration of market share or earnings within an industry, for both *Revenues* and *Trading Volume*. The *Herfindahl Industry Concentration by Revenues* is calculated as:

$$Herfindahl_{i,t} = \sum_{i=1}^{N} \left[\frac{Revenues_{i,t}}{HFT\ Revenues_t}\right]^2 \quad (4)$$

where N is the number of firms in month t that earn non-negative trading revenues, $Revenues_{i,t}$ is firm i's total trading revenues in month t, and $HFT\ Revenues_t$ is the trading revenues summed across all HFT firms. *Herfindahl Industry Concentration by Volumes* is calculated using the same formula but considering *Trading Volume* instead of trading *Revenues*. A larger index implies a more concentrated industry: the index is at its minimum $1/N$ when all HFT firms have the same share of the industry and at its maximum of when all activity is concentrated to one firm.

As with the trading performance time trends in Table 6, the Herfindahl indices are first calculated for each trading day, then averaged across trading days of each half-year block in the data set. Standard deviations are reported in parentheses. As alternative measures of industry concentration, the ratio of average revenues or volume of the Top 3 or 5 HFTs to the median HFT are reported for each half-year interval in the sample. The results are reported in Table 7, Panel A, and graphed in Figure 3, Panel D.

INSERT TABLE 7 ABOUT HERE

The Herfindahl industry concentration in terms of volume lies in the range from 0.181 to 0.343. In terms of revenues, the corresponding interval is from 0.289 to 0.401. As a reference, Van Ness, Van Ness, and Warr (2005) find that market makers have an average Herfindahl index (in terms of trading volume) on NASDAQ of 0.14, with a range of 0.037 to 0.439.

We test for a time trend in industry concentration by estimating the following OLS regression:

$$Concentration_t = \alpha + \beta \ (year - 2010) + \epsilon_t \tag{5}$$

where, as in Table 6, *year* is a continuous variable (e.g., *year* would take on the approximate value of 2014.25 on March 31, 2014), and $Concentration_t$ is one of the concentration measures discussed above. The coefficient estimates are reported in Table 7, Panel B.

We find no statistically significant time trend for revenues concentration, either with the Herfindahl index, the ratio of Top 3-to-median, or the ratio of Top 5-to-median). There is a statistically significant decrease in trading volume concentration, but this is not economically significant. The Herfindahl industry concentration by volume decreases by 0.012 per year from its starting value of 0.301 in the first-half of 2010. The Top 3 to Median-HFT concentration measures for volume decreases significantly, but the decrease for the Top 5 is small compared to their 2010 values (the constant term).. This indicates that the dominant group of HFTs in terms of trading volumes is widened from three to five firms over time. Our overall conclusion with respect to industry concentration is however that it is relatively constant over the five years.

D. Entry and Exit

To investigate competition in the HFT industry, we examine the properties of firms that just entered the market. New entrants can potentially introduce competition and drive down both firm-level and industry-wide profits. However, if new entrants cannot simply pay to acquire skill but require experience, then they may earn less and be less likely to survive in the market. We find that new entrants indeed have worse performance than incumbents and more likely to exit the market.

To evaluate the importance of experience, we regress performance measures on indicator variables for the length of time an HFT firm has been active in a given stock, designating less than 1 month, 2 months or 3 months as indicators of new entry.

We estimate the following regression equation:

$$Performance_{i,j,t} = \beta_1 \mathbf{1}_{\text{one-month } i,j,t} + \beta_2 \mathbf{1}_{\text{two-month } i,j,t} + \beta_1 \mathbf{1}_{\text{three-month } i,j,t} + (\text{day x stock})\text{-FEs} + \epsilon_{i,j,t} \tag{6}$$

where $Performance_{i,j,t}$ can be *Revenues, Revenues per MSEK Traded,* or *Returns,* and is defined on a firm-stock-day frequency over the period 2010-2014. We exclude the observations in the first three months of the sample, since new entry then cannot be established. The notation $\mathbf{1}_{one-month\, i,j,t}$ takes the value 1 if firm *i* began trading stock *j* in the last 30 calendar days, otherwise 0; similarly, the $\mathbf{1}_{two-month\, i,j,t}$ dummy corresponds to beginning trading in stock *j* in the last 30-60 days, and the $\mathbf{1}_{three-month\,i,j,t}$ dummy corresponds to the previous 60-90 days.[14] If new entrants are less competitive and perform worse than established firms, the coefficient on the one-month dummy should be negative; and consistent with experience mattering the two- and three-month dummy coefficients should also be negative but closer to zero. The results are reported in Table 8. Standard errors are dually clustered by firm-stock and by month and are reported within parentheses.

INSERT TABLE 8 ABOUT HERE

Looking at *Revenues*, HFTs have statistically significant negative coefficients corresponding to the new entry dummies, with the largest negative value of -3,120 SEK for the 1-month dummy and decreasing (in absolute terms) over time to -2,350 SEK for 2 months and -675 SEK for 3 months. *Returns* are also significantly lower for new entrants in the first and second month. However, *Revenues per MSEK Traded* is statistically insignificant for all three time periods. The fact that the new firms have lower total trading revenues but no statistically difference in *Revenues per MSEK Traded* is consistent with HFTs competing on quantity, not quality, and new firms being less able to compete on capturing quantity.

Additionally, we examine whether new entrants are more likely to exit the market by estimating the model described in Eq. (6) with the dummy $Exit_{i,j,t}$ as the dependent variable, which takes the value 1 on day t for firm i if that is the last day firm i for stock j. In this analysis we exclude observations in December 2014, the last month of the analysis, as we are unable to

[14] In accounting for entry and exit, we ignore gaps and just count the overall first and last trading days of a firm in a stock as entry and exit dates. HFT mergers (of which there are two in our sample) are also not counted as entry/exit events.

determine market exit in that month. The model is estimated as a linear probability regression and the results are in Table 8.

We find that new entrant in a stock have a higher probability of exit compared to more established HFTs. The effect is found in the three-month indicator variable. The two-month indicator variable is statistically insignificant at the 5% level, and the one-month indicator variable is slightly negative, opposite of what we expected. The first two months' lack of an effect are likely due to HFTs requiring time to test their ability in the market, even if they are initially unprofitable. The *daily* probability of Exit is increased by a statistically significant and economically large amount of 0.7% in the third month. The increased probability of exit is for *each day*, not the cumulative probability of exit at any future point, making the magnitude substantial.

Why do new entrants perform worse than established firms? Given our previous findings on latency, we examine whether new entrants are slower than the incumbent HFTs. Repeating the analysis in Eq. (6) but replacing the dependent variable with our *Latency* measure, we find statistically significant coefficients on the 1-month and 2-month dummies of 74.67 and 134.1 milliseconds.[15] The positive coefficients on the 1-month and 2-month old HFTs show that new entrants are slower than established firms, which might in part explain their lower trading revenues.

In summary, we find that new entrants perform worse and are more likely to exit than established HFTs. This finding helps explain the continued concentration of revenues among a small subset of established firms: if new entrants tend to exit, then the concentration of revenues continues. A high and steady industry concentration combined with strong firm-level persistence

[15] Our definition of *Latency* does not allow variation across stocks. For the purpose of this regression we construct a firm-stock-day panel, where all observations across stocks for a given firm-day is assigned the same *Latency* value.

of *Revenues* and *Returns* suggests that top performing incumbent HFTs maintain their position in the market. Together, the four measures of industry structure we examine point to limited competition, consistent with relative latency driving performance.

VI. The Role of Latency in Various HFT Strategies

We investigate whether the effect of latency on performance is consistent across various HFT strategies. The analysis of different HFT strategies is motivated by Hagströmer and Nordén (2013), who provide empirical evidence of diverse trading strategies in the HFT industry. Our first investigation distinguishes the role of speed in active trading on short-lived information trading versus passive liquidity provision.

The benefit of low latency in active trading strategies is explored by Foucault, Hombert and Rosu (2015). They show that fast news-traders trade aggressively on news, picking off stale quotes. Furthermore, Biais, Foucault, and Moinas (2015) and Foucault, Kazhan, and Tham (2014) show that fast active traders can benefit from a superior ability to react to cross-market arbitrage opportunities. Chaboud, Chiquoine, Hjalmarsson and Vega (2014) provide empirical evidence of fast traders pursuing cross-market arbitrage. From the perspective of passive traders, Hoffmann (2014) emphasizes that low latency allows liquidity providers to reduce their adverse selection costs, as they can revise stale quotes before they are picked off. Aït-Sahalia and Saglam (2014) add to this that passive traders also benefit in terms of reduced inventory costs. In their model, fast liquidity providers are better able to manage their inventory positions, and can accordingly increase their liquidity supply.

A. Price Impact versus Realized Spread

To capture skill in active trading on short-lived information, we measure active *Price Impact* as the basis point change in the bid-ask spread midpoint from just before to ten seconds after a trade initiated by an HFT firm. To capture skill in passive liquidity provision, we measure passive *Realized Spread* as the basis point difference between the transaction price and the bid-ask spread midpoint ten seconds after a trade where an HFT firm is the liquidity provider. *Realized Spread* captures the benefit of earning a wide bid-ask spread, as well as the ability to avoid supplying liquidity to trades with price impact. Each measure is calculated on a firm-stock-month frequency as the SEK-volume-weighted average across all trades of a given HFT in each stock and each month. We interact both *Realized Spread* and *Price Impact* with a ±1 buy-sell indicator variable, such that a higher coefficient corresponds to better trading performance. As the analysis requires information on the bid-ask spread we limit the measures to trades that can be matched to order book data, both contemporary to the trade and ten seconds later.

We re-estimate Eq. (1) at the firm-stock-month level with *Price Impact* and *Realized Spread* as the dependent variables. As before, the $\mathbf{1}_{\text{top 1 } i,t}$, $\mathbf{1}_{\text{top 2-3 } i,t}$, and $\mathbf{1}_{\text{top 4-5 } i,t}$ are indicators that capture relative latency, whereas log(*Latency*) captures nominal latency. The estimation is run either with or without control variables.

The control variables include both HFT firm characteristics and stock characteristics. The HFT firm characteristics are *End-of-day Inventory*, *Max. Intraday Inventory*, *Investment Horizon*, and *Aggressiveness Ratio*, defined in Section III.A. Since these variables are firm-characteristics, they are assigned the same value across stocks. The stock characteristics are measured on stock-month frequency and are defined in Section II.B: *Volatility*, *Fragmentation Index*, *Tick Size*, and *Quoted Spread*. The *Non-HFT Trading Volume* is measured on the same frequency and is defined as in Section V.B. All continuous independent variables are in units of standard deviations. Standard errors are dually clustered by firm-stock and month. The results are reported in Table 9.

INSERT TABLE 9 ABOUT HERE

As for HFT performance in general, we find that for the individual trading strategies relative latency, not nominal latency, drives performance. The coefficients for the three relative latency dummies are all statistically significant and economically meaningful for both the *Price Impact* and the *Realized Spread*. For example, being in the Top 2-3 by speed increases *Price Impact* and *Realized Spread* by 0.802 and 0.539 bps, respectively, while being in the Top 1 by speed increases *Price Impact* and *Realized Spread* by 1.06 and 0.567 bps, respectively. The coefficient for nominal latency is not statistically significant. These results are robust to inclusion of the control variables. Many control variables, including the *Aggressiveness ratio*, *Non-HFT Trading Volume*, and *Volatility*, have a statistically significant relation to the *Realized Spread*, but not to *Price Impact*. This suggests that skill in passive liquidity provision is more sensitive to market conditions and stock characteristics. *Tick size* and *Quoted spread* also influence the *Realized Spread* and not the *Price Impact*, but this may be mechanical due to the definition of the *Realized Spread*. Both *Price Impact* and *Realized Spread* are significantly related to the *Fragmentation Index.*

We conclude that relative and not nominal latency is important for the performance of both active and passive trading strategies. This is consistent with theoretical models, such as Foucault, Hombert, and Rosu (2015) on active trading; and Hoffmann (2014) and Aït-Sahalia and Saglam (2014) on passive trading. The fact that relative latency is important for passive liquidity strategies is non-trivial, given that the latency measure is constructed by looking at the response time to submit an *aggressive* order.

B. Cross-Market Arbitrage

Lastly, we investigate whether speed is useful for cross-market arbitrage between the futures market and equities. Specifically, we test if faster HFTs are more likely than slower HFTs to *actively* trade in equities in quick response to "news" in the futures market, where "news" is defined to be a price change in the OMXS30 futures above a certain size. We also ask whether faster HFTs are less likely than slower HFTs to be adversely selected in a passive trade in equities markets in response to "news" in the futures market. The investigation is in line with the theoretical setup of active fast trading by Biais, Foucault, and Moinas (2015) and Foucault, Kozhan and Tham (2015).

We estimate the following probit regression, which in essence follows the setup of Hendershott and Riordan (2013), and Brogaard, Hagströmer, Nordén and Riordan (2015):

$$\Pr[Fast\ HFT\ Trades] = \Phi[\beta\ News + \gamma' controls +\ Stock\ FEs]. \tag{8}$$

The unit of observation is a trade. To capture who is trading quickly in response to "news" in the futures market, we look at equities trading in the subsequent 1-second interval following a "news" event in the futures market. We define *Fast HFT* variously as being *Top 1, Top 2-3,* or *Top 4-5* in terms of trading speed within a month; *Slow HFTs* are those not among the top five. The dependent variable is 1 when a *Fast HFT* executes an equities trade in the subsequent 1-second and 0 if a *Slow HFT* does it. Thus, our results can be interpreted as the increased probability of a *Fast HFT* trading in equities relative to a *Slow HFT*, in response to "news" in the futures market.

News is defined to be ±1 (and 0 otherwise) when the return on the OMXS30 futures during a one-second window preceding the stock trade is large, defined as when the return exceeds the top decile among non-zero returns of that month. *News* takes the value +1 if the active party trades in direction of the news, and -1 otherwise if in the opposite direction. This design implies that *News* reflects any event that causes a large price change in the futures index. Note that all sample stocks are constituents of the index underlying the futures contract, making arbitrage activities between the two markets likely (Hasbrouck, 2003).

We also control for the following variables, which may affect the probability of Fast HFTs doing cross-market arbitrage. *Lagged Volatility* is the average second-by-second squared return (multiplied by 1,000) over the previous ten seconds; *Lagged Volume* is the SEK trading volume (divided by 100,000) over the previous ten seconds; *Quoted Spread* is defined as above; and *Depth at BBO* is the average number of shares available at the best bid quote and the best offer quote (divided by 100,000), multiplied by the bid-ask spread midpoint.

Estimates for active trading (i.e., the sample being all trades initiated through the submission of a market order by an HFTs) are reported in Table 10, Panel A. Similarly, the estimates for passive trading (i.e., the sample being all trades where liquidity is provided by an HFT) are presented in Panel B. For computational tractability, regressions are run monthly for all month in 2010-2014. Similar to the Fama and Macbeth (1973) procedure, the monthly coefficients are averaged across months to produce the estimates reported in Table 10. To assess the economic magnitudes, we report marginal effects, which can be interpreted as the increased probability of a *Fast HFT* trading relative to a *Slow HFT* (on the active and passive side, respectively) if the explanatory variable increases by one, conditional on all other explanatory variables being at their unconditional means.

INSERT TABLE 10 ABOUT HERE

We find that the *News* coefficients are positive and statistically significant for active trading (Panel A). Based on the marginal effects, the fastest HFT (*Top 1*) is 0.9% more likely to actively trade in equities subsequent to "news" arrival in the futures market, relative a trader outside of the Top 5. The *Top 2-3* and *Top 4-5* show similar, albeit somewhat weaker, skills. Thus, faster HFTs are more likely to quickly submit market orders in response to changes in the futures index. This is consistent with fast active traders being better positioned to pursue cross-market arbitrage, as modeled by Biais, Foucault, and Moinas (2015) and Foucault, Kozhan and Tham (2015). For passive trading, with the exception of the *Top 1* HFT, the *News* coefficients are negative and statistically significant (Panel B). Thus, faster HFTs are less likely to get caught in a passive equities trades that incur adverse selection costs to the liquidity provider. This is also in line with Foucualt, Kozhan and Tham (2015), who find that the probability of toxic arbitrage is related to the latency of arbitrageurs relative the latency of liquidity providers. The effect is somewhat stronger than for active trading, with the marginal effect indicating that the Top 4-5 fastest HFTs are 1.4% less likely to be adversely selected in the equities markets, conditional on futures market news, relative slower HFTs.

VII. Conclusion

We study the role of latency in the performance of HFT firms. We document a number of statistics consistent with superior investment performance by HFTs. There are large cross-sectional differences in performance in the HFT industry, with trading revenues disproportionally accumulating to a few firms. The fastest firms tend to earn the largest trading revenues. While latency decreases substantially over our five-year sample period, we show that it is relative latency, not nominal latency, that explains the differences in performance across HFTs.

Furthermore, we investigate how the role of relative latency may impact competition by analyzing the HFT industry. If small differences in latency are important, then the HFT industry should be characterized by persistence in performance, high concentration, and difficulty of new entry. We find evidence that is consistent with these predictions. Firm trading performance is persistent, trading revenues are high and non-declining, as is HFT concentration, and new HFT entrants are slower, underperform, and are more likely to exit.

Finally, we examine how speed may be used in specific HFT strategies. We find evidence that relative latency is important for success in trading on short-lived information, liquidity provision, and cross-market arbitrage. By understanding how relative latency is related to trading performance, and also to the structure of the HFT industry, researchers and regulators can better evaluate the role and implications of HFTs in financial markets.

References

Aït-Sahalia, Y. and M. Saglam, 2014, High-frequency traders: Taking advantage of speed, *Working Paper*.

Biais, B., T. Foucault, and S. Moinas, 2015, Equilibrium fast trading, *Journal of Financial Economics*, *116*(2), 292-313.

Boehmer, E., K. Fong, and J. Wu, 2015, International evidence on algorithmic trading, *Working Paper*.

Boehmer, E., D. Li, and G. Saar, 2015, Correlated high-frequency trading, *Working Paper*

Bongaerts, D. and M. Van Achter, 2015, High-frequency trading and market stability, *Working Paper*

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

Brogaard, J. and C. Garriott, 2015, High-frequency trading competition, *Working Paper*

Brogaard, J., B. Hagströmer, L. Norden, and R. Riordan, 2015, Trading fast and slow: Colocation and market quality. Forthcoming in *Review of Financial Studies*.

Brogaard, J., T. Hendershott, and R. Riordan, 2014, High-frequency trading and price discovery, *Review of Financial Studies*, 27 (8), 2267-2306.

Budish, E., P. Cramton, and J. Shim, 2013, The high-frequency trading arms race: Frequent batch auctions as a market design response, *Quarterly Journal of Economics*, 30 (1547), 1621.

Carhart, M., 1997, On persistence in mutual fund performance, *Journal of Finance*, 52 (1), 57-82.

Cartea, A. and J. Penalva, 2012, Where is the value in high frequency trading?, *Quarterly Journal of Finance* 2(3), 1250014.

Chaboud, A., B. Chiquoine, E. Hjalmarsson, and C. Vega, 2014, Rise of the machines: Algorithmic trading in the foreign exchange market. *Journal of Finance*, 69(5), 2045-2084.

Chen, J., H. Hong, M. Huang, and J. Kubik, 2004, Does fund size erode mutual fund performance? The role of liquidity and organization, *American Economic Review*, 94 (5), 1276-1302.

Clark-Joseph, A., 2013, Exploratory trading, *Working Paper*

Fama, E., and French, K., 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics*, 33 (1), 3-56.

Fama, E., and J. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy*, 607-636.

Foucault, T., J. Hombert, and I. Rosu, 2015, News trading and speed. Forthcoming in *Journal of Finance*.

Foucault, T., R. Kozhan, and W. Tham, 2015, Toxic arbitrage. *Working Paper*.

Gai, J., M. Yao, and C. Ye, 2013, The externalities of high frequency trading, *Working Paper*.

Geroski, P., R. Gilbert, and A. Jacquemin, 1990, Barriers to entry and strategic competition, Vol. 41, Taylor & Francis.

Hagströmer, B. and L. Nordén, 2013, The diversity of high-frequency traders, *Journal of Financial Markets*, *16*(4), 741-770.

Hagströmer, B., L. Nordén, and D. Zhang, 2014, How Aggressive Are High-Frequency Traders?, *Financial Review*, *49*(2), 395-419.

Hasbrouck, J., 2003, Intraday price formation in US equity index markets, *Journal of Finance*, 58 (6), 2375-2400.

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

Hendershott, T., C. Jones, and A. Menkveld, 2011, Does algorithmic trading increase liquidity?, *Journal of Finance*, 66 (1), 1-33.

Hendershott, T., and R. Riordan, 2013, Algorithmic trading and the market for liquidity. *Journal of Financial and Quantitative Analysis*, 48, 1001–1024.

Ho, T. and H. Stoll, 1983, The dynamics of dealer markets under competition, *Journal of finance*, *38*(4), 1053-1074.

Hoffmann, P., 2014, A dynamic limit order market with fast and slow traders, *Journal of Financial Economics*, *113*(1), 156-169.

Jagannathan, R., A. Malakhov, and D. Novikov, 2010, Do hot hands exist among hedge fund managers? An empirical evaluation, *Journal of Finance*, 65 (1), 217-255.

Jovanovic, B. and A. Menkveld, 2015, Middlemen in limit-order markets, *Working Paper*.

Kaplan, S. and A. Schoar, 2004. Private equity performance: returns, persistence, and capital flows, *Journal of Finance*, 60 (4), 1791-1823.

Kirilenko, A., A. Kyle, M. Samadi, and T. Tuzun, 2015, The flash crash: The impact of high frequency trading on an electronic market, *Working Paper*.

Lee, C. and M. Ready, 1991, Inferring trade direction from intraday data. *Journal of Finance*, *46*(2), 733-746.

Malinova, K., A. Park, and R. Riordan, 2013, Do retail traders suffer from high frequency traders?, *Working Paper*.

Menkveld, A. and M. Zoican, 2015, Need for speed? Low latency trading and adverse selection, *Working Paper.*

Petersen, M., 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *Review of Financial Studies*, 22 (1), 435-480.

Riordan, R. and A. Storkenmaier, 2012, Latency, liquidity and price discovery, *Journal of Financial Markets*. 15 (4). 416–437.

Rosu, I., 2015, Fast and Slow Informed Trading, *Working Paper*.

Securities and Exchange Commission (U.S.), 2010, Concept Release on Equity Market Structure

Thompson, S., 2011, Simple formulas for standard errors that cluster by both firm and time, *Journal of Financial Economics*, 99 (1), 1-10.

Van Kervel, V. and A. Menkveld, 2015, High-Frequency Trading around Large Institutional Orders, *Working Paper*

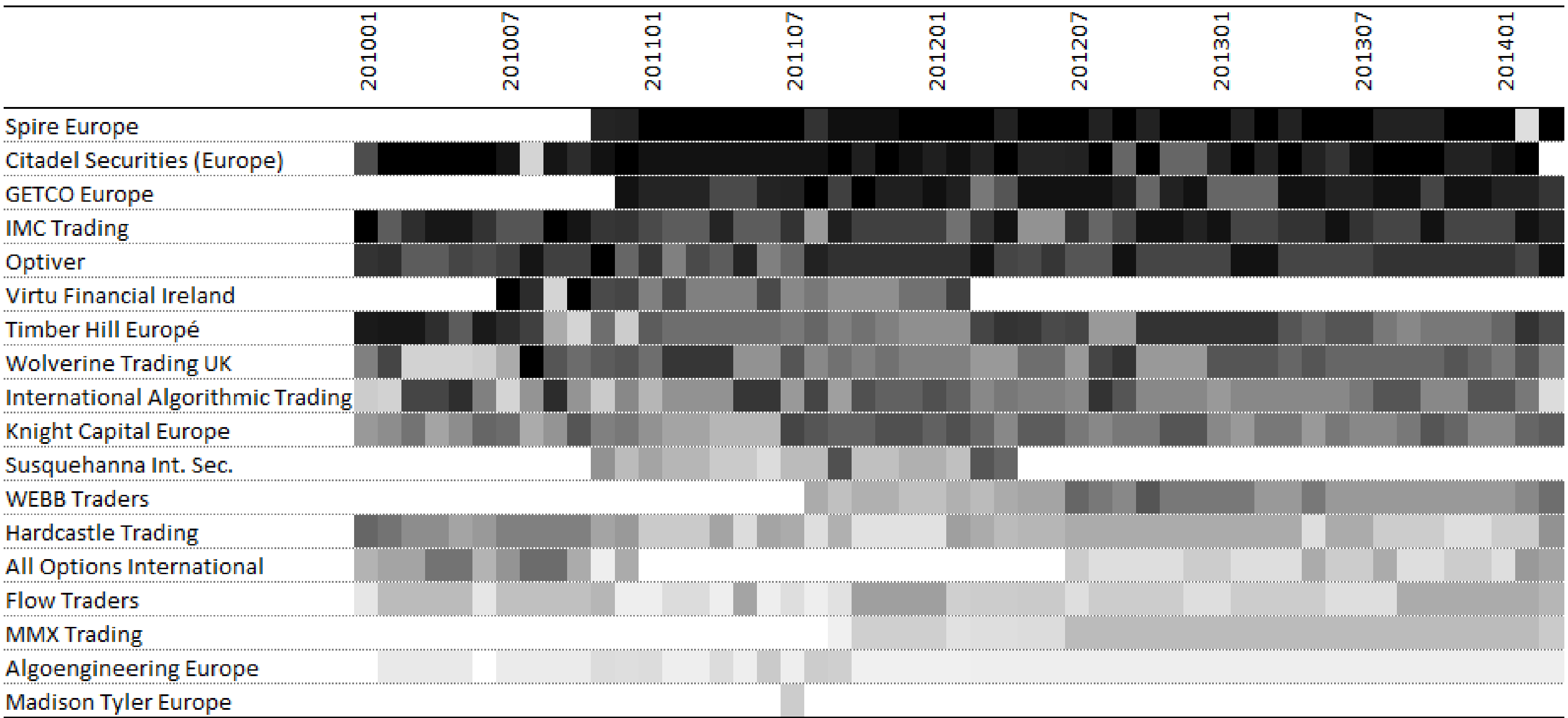
Van Ness, B., R. Van Ness, and R. Warr, 2005, The impact of market-maker concentration on adverse selection costs for NASDAQ stocks, *Journal of Financial Research*, 28 (3), 461-485.

Weller, B., 2013, Intermediation chains and specialization by speed: Evidence from commodity futures markets, *Working Paper*.

Weston, J., 2000, Competition on the Nasdaq and the impact of recent market reforms, *Journal of Finance*, *55*(6), 2565-2598.

Figure 1: HFT activity at NASDAQ OMX Stockholm

This figure shows the monthly ranking of HFT trading volume in the sample stocks at NASDAQ OMX Stockholm. This figure is constructed *entirely from publicly available data and contains no information from our proprietary data set.** The firms included are exchange members at NASDAQ OMX Stockholm that according to public records trade the sample stocks during at least one month in the interval January 2010 to March 2014, and that indicate on their websites that they primarily do low-latency proprietary trading. The ranking of trading volume for each month is visualized on a grayscale, where the HFT firm with the highest volume is black and the HFT firm with the lowest trading volume is white. The set of HFTs presented here is not an exhaustive list of HFTs analyzed in the paper, and the trading volume underlying the ranking excludes trading at other trading venues than NASDAQ OMX Stockholm. The HFT firms are sorted by total trading volume.



* The disclosure of broker IDs in public trade records were made voluntary on March 24, 2014, and the data are thus incomplete after that date.

Figure 2: HFT latency over time

This figure shows how HFT latency develops over time, from January 2010 to December 2014. Latency is measured for each HFT firm and each month as the minimum time recorded from a passive execution to an active execution at the same trading venue in any of the sample stocks. Latency is indicated on the vertical axis on a log scale. *HFT #1* is the latency of the fastest HFT in each month; *HFTs #2-3* is the average latency of the 2nd and 3rd fastest HFTs; *HFTs #4-5* is the average latency of the 4th and 5th fastest HFTs; and *All HFTs* is the average latency across all HFTs for which latency is lower than one second in the given month. The vertical bars indicate microstructure events at NASDAQ OMX Stockholm that are expected to be associated with changes in latency. The sample consists of 25 Swedish stocks.

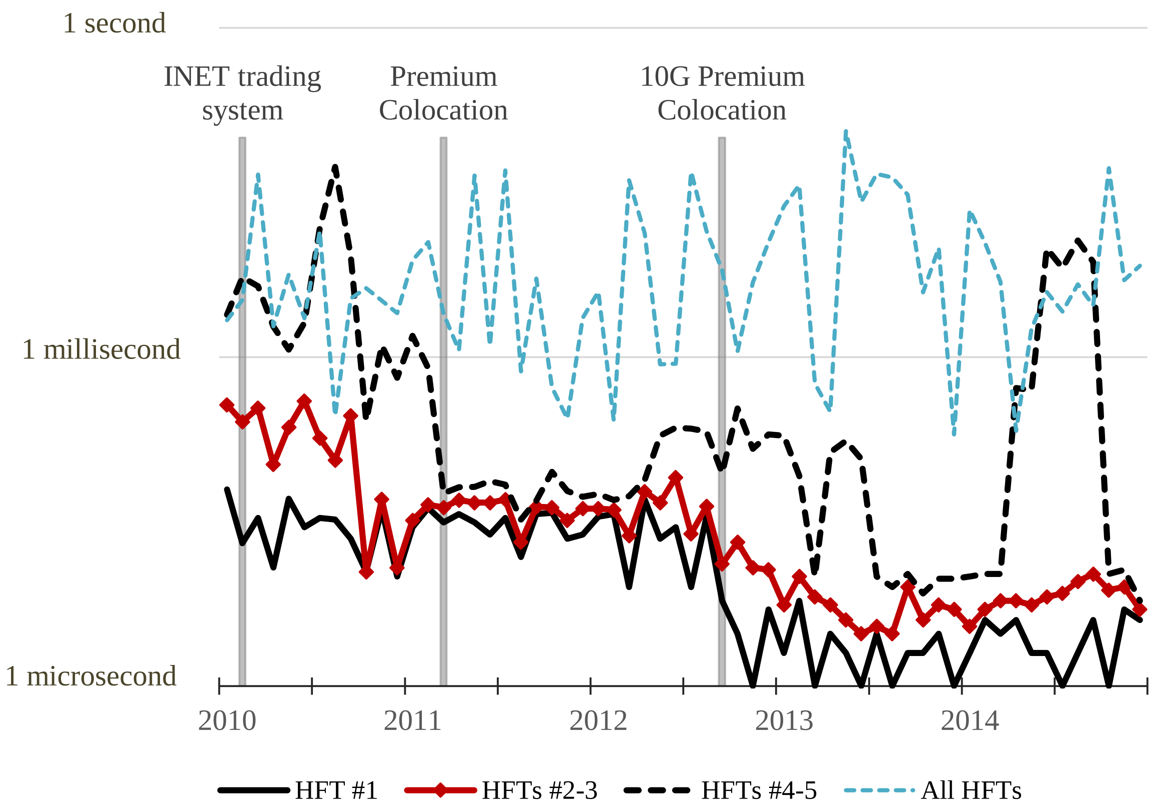
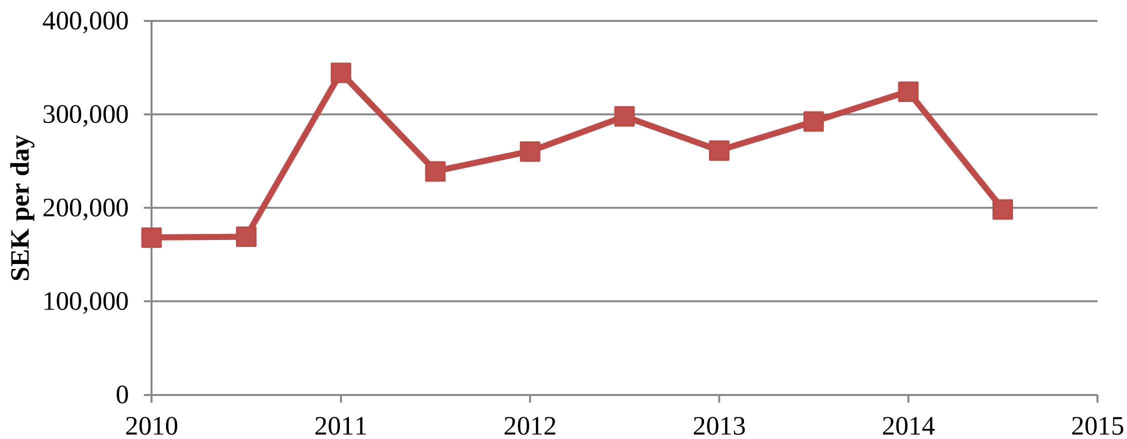


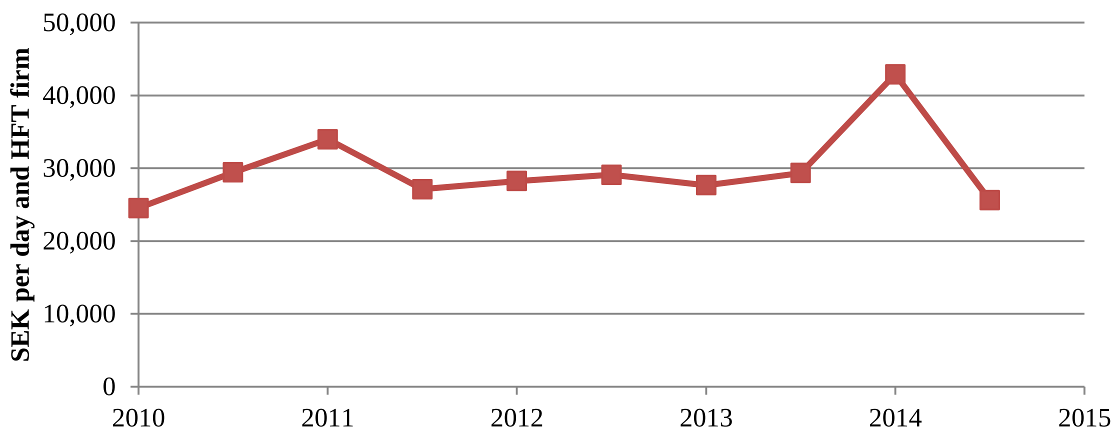
Figure 3: HFT revenues, cost of HFT intermediation, and industry concentration over time

This figure shows various time series trends. Panel A shows the *Total Daily Revenues*, which is the average daily sum of *Revenues* (defined as in Table 2) across all HFT firms. Panel B shows *Average Daily Revenues*, which is the average *Revenues* across all HFT firms active on each given day. Panel C shows the *Cost of HFT Intermediation to Non-HFTs*, calculated as *Total Daily Revenues* divided by the total non-HFT trading volume. Panel D shows two indexes of *HFT Industry Concentration*: Herfindahl indexes calculated on daily *Revenues* or with daily *Trading Volumes*, see Eq. (4). The sample consists of 25 Swedish stocks and 60 months of trading (January 2010 to December 2014). The time series are reported on a biannual frequency.

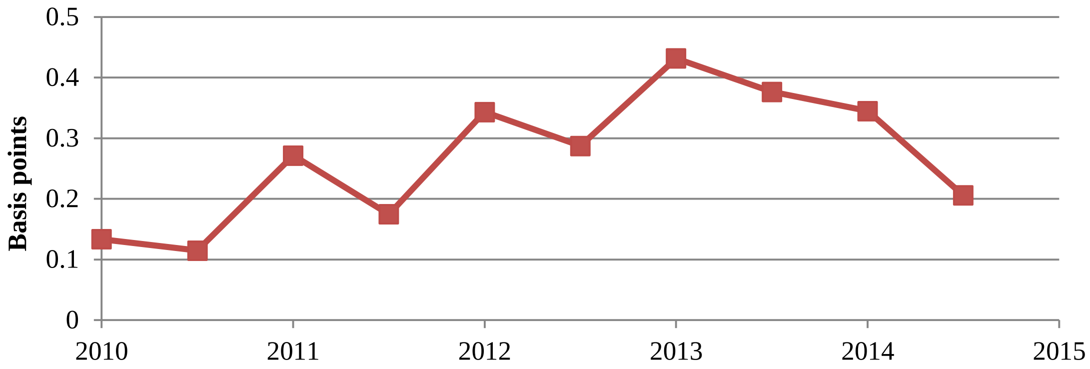
Panel A: Total Daily Revenues



Panel B: Average Daily Revenues



Panel C: Cost of HFT Intermediation to Non-HFTs



Panel D: HFT Industry Concentration

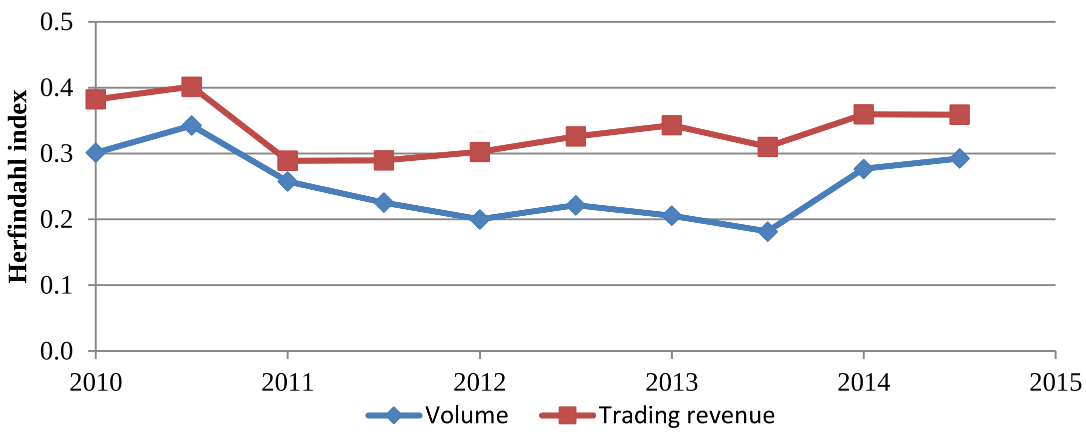


Table 1: Stock characteristics

This table reports summary statistics of the 25 Swedish stocks in the sample. *Market Capitalization* is based on the closing price on December 31, 2014 (expressed in MSEK). All other statistics are calculated as averages across trading days in December 2014. Because SCVb is delisted in May 2014, the metrics for that stock are based on April 2014. *Daily Trading Volume* refers to trading at NASDAQ OMX Stockholm only and is reported in MSEK. *Daily Turnover* is the Daily Trading Volume divided by Market Capitalization, expressed in percentage points. *Tick Size* is the average minimum price change; *Quoted Spread* is the average bid-ask spread prevailing just before each trade; and *Effective Spread* is the trade value-weighted average absolute difference between the trade price and the bid-ask midpoint. All spread measures are based on continuous trading at NASDAQ OMX Stockholm, expressed relative to the bid-ask spread, and presented in basis points. The Tick Size and the Quoted Spread are halved to be comparable to the Effective Spread. *Volatility* is the average 10-second squared returns, calculated from bid-ask midpoints. The *Fragmentation Index* is the inverse of a Herfindahl index of trading volumes across the five largest trading venues (BATS, Burgundy, Chi-X, NASDAQ OMX Stockholm, and Turquoise); a higher value signifies greater fragmentation. The table is sorted by Market Capitalization.

Stock Ticker	Market Cap. (MSEK)	Daily Trading Vol. (MSEK)	Daily Turnover (%)	Tick Size (bps)	Quoted Spread (bps)	Effective Spread (bps)	Volatility (sq. bps)	Fragment-ation Index
HMb	475,595	1,358	0.29	1.57	2.06	2.26	2.94	2.11
SHBa	228,731	757	0.33	1.38	2.17	2.36	4.08	2.30
SWEDa	221,307	1,011	0.46	2.58	2.90	3.21	5.69	2.00
SEBa	216,025	771	0.36	2.67	3.19	3.30	5.20	1.93
ATCOa	183,324	839	0.46	2.32	3.04	3.23	5.12	2.29
ASSAb	145,878	529	0.36	1.23	2.36	2.51	3.52	2.28
VOLVb	136,816	1,099	0.80	2.98	3.24	3.34	5.71	1.91
INVEb	129,676	590	0.46	1.79	2.43	2.62	4.12	1.80
SCAb	104,559	632	0.60	2.91	3.42	3.59	5.13	2.13
SAND	95,835	798	0.83	3.27	3.67	3.74	6.55	2.12
SCVb	78,800	825	1.05	2.65	3.66	4.75	6.75	1.91
ATCOb	78,395	262	0.33	2.55	3.88	4.09	5.52	1.85
SKFb	68,879	684	0.99	3.14	3.63	3.85	6.07	2.24
ELUXb	68,807	471	0.68	2.24	3.16	3.29	5.05	2.31
SKAb	67,160	320	0.48	3.07	3.63	3.77	4.89	1.90
ALFA	62,205	463	0.75	3.42	4.06	4.10	6.04	2.05
KINVb	60,097	337	0.56	1.95	3.61	3.98	7.95	2.26
SWMA	49,082	337	0.69	2.04	3.10	3.32	4.38	2.24
TEL2b	40,403	334	0.83	2.64	3.21	3.51	5.97	2.12
GETIb	39,540	221	0.56	2.91	3.81	4.21	4.25	1.99
LUPE	34,964	586	1.68	4.08	5.65	5.68	16.53	1.91
BOL	34,326	538	1.57	4.05	4.87	4.90	7.88	2.06
SECUb	32,861	169	0.51	2.73	3.66	3.88	3.75	2.32
MTGb	15,369	155	1.01	2.05	4.07	4.77	7.46	2.17
SSABa	13,877	370	2.67	1.61	3.82	4.14	8.89	1.76

Table 2: The cross-section of HFT trading performance

This table reports descriptive statistics on HFT trading performance and trading characteristics in the cross-section of HFT firms. The following variables underlying the statistics are first aggregated over all stocks and then averaged across time for each HFT; the resulting cross-sectional distribution across HFT firms is then presented. *Revenues* is the average daily trading revenue for each HFT, calculated as cash received from selling shares, minus the cash paid from buying shares, plus the value of any outstanding positions at the end-of-day, marked to the market price at close; *Trading Volume* is the average daily trading volume for each HFT, measured in MSEK; *Revenues per MSEK Traded* is daily *Revenues* divided by daily *Trading Volume* for each HFT; *Returns* is daily *Revenues* divided by the maximum intraday inventory position over the entire sample (used as a measure of capitalization) and reported in annualized terms; *Sharpe Ratio* is the average monthly ratio for each HFT, reported in annualized terms, of the average daily return divided by the standard deviation of daily returns; the *1-factor Alpha* is the intercept estimated in a regression of daily HFT excess returns on the market return factor; the *3-factor Alpha* is the intercept estimated in a regression daily HFT excess returns on the Fama-French factors; the 4-*factor Alpha* is the intercept estimated in a regression of daily HFT excess returns on the Fama-French and Carhart momentum factors; *Aggressiveness Ratio* is the SEK volume traded using market orders divided by the *Trading Volume*; *End-of-Day Inventory Ratio* is the absolute end-of-day SEK position (netted across stocks) divided by the *Trading Volume*; *Max. Intraday Inventory Ratio* is the maximum absolute intraday SEK inventory position divided by the *Trading Volume*; *Average Trade Size* is the average SEK amount traded in each execution; and *Latency* is the minimum time (in milliseconds) from a passive execution to an active execution in the same stock and same trading venue, measured monthly and averaged across months for each HFT. The sample consists of 25 Swedish stocks and 60 months of trading (January 2010 to December 2014).

	Mean	**Std.Dev.**	**p5**	**p10**	**p25**	**p50**	**p75**	**p90**	**p95**
Revenues (SEK)	21,920	31,378	-13,941	-2,166	-623	13,224	34,131	63,463	111,376
Revenues per MSEK Traded	41.48	219.71	-511.41	-305.94	-12.63	65.90	106.18	342.59	460.25
Returns	0.334	0.455	-0.148	-0.102	0.008	0.156	0.589	0.889	1.535
Sharpe Ratio	5.07	7.06	-1.94	-0.59	0.21	2.80	8.20	11.20	26.72
1-factor Alpha	0.342	0.457	-0.139	-0.036	0.007	0.155	0.592	0.897	1.562
3-factor Alpha	0.348	0.461	-0.143	-0.028	0.012	0.146	0.597	0.940	1.565
4-factor Alpha	0.347	0.461	-0.144	-0.029	0.011	0.145	0.597	0.940	1.559
Trading Volume (MSEK)	333.17	390.82	15.00	22.82	30.97	176.03	521.11	911.70	1307.78
Aggressiveness Ratio	0.53	0.26	0.17	0.17	0.30	0.54	0.73	0.93	0.96
End-of-Day Inventory Ratio	0.12	0.12	0.00	0.01	0.02	0.08	0.20	0.31	0.39
Max intraday Inventory Ratio	0.17	0.15	0.01	0.03	0.05	0.12	0.26	0.37	0.52
Average Trade Size (thousands SEK)	70.61	21.63	37.30	44.89	51.28	68.29	88.13	96.58	115.06
Latency (milliseconds)	83.56	186.99	0.01	0.03	0.10	2.25	39.73	508.11	600.19

(N = 16 firms)

Table 3: Trading revenues, costs, and profits from public filings

This table reports *Trading Revenues*, *Trading Costs*, *Trading Profit Margins*, and *Trading Returns* calculated from annual reports and IPO prospectuses for four high-frequency trading firms for which public data is available. *Trading Costs* are broken down into several categories, all expressed as a percent of trading revenues. *Trading Profit Margin* is *(1-trading costs)*, and *Trading Returns* are calculated two ways based on two capitalization measures: as *Trading Revenues / (Trading Assets Minus Trading Liabilities)* and as *(Trading Revenues / Book Equity)*. All quantities are in million USD, except for the firm Flow Traders, which is in million EUR.

	Virtu				KCG		GETCO				Flow Traders		
	2014	**2013**	**2012**	**2011**	**2014**	**2013**	**2012***	**2011**	**2010**	**2009**	**2014**	**2013**	**2012**
Trading Revenues (in millions)	685.2	623.7	581.5	449.4	1,274.4	903.8	526.6	896.5	865.1	955.2	240.8	200.5	125.1
-- percentage of revenue from proprietary trading	98.5%	98.4%	100%	100%	68.5%	67.0%	89.9%	94.2%			100%	100%	100%
Trading Costs (as percent of Trading Revenues)	60.0%	57.8%	72.6%	62.1%	52.4%	59.0%	62.5%	48.5%	48.6%	40.4%	41.6%	43.7%	47.5%
-- Brokerage, exchange and clearance fees, net	33.7%	31.3%	34.5%	32.9%	23.9%	27.3%	35.3%	32.2%	35.1%	32.1%	{ 15.7%	15.8%	14.8%
-- Communication and data processing	10.0%	10.4%	9.5%	10.3%	11.8%	13.7%	17.2%	9.7%	7.1%	4.5%			
-- Equipment rentals + depreciation and amortization	4.5%	4.0%	15.7%	11.1%	10.4%	11.0%	9.1%	6.2%	6.2%	3.8%	1.8%	1.9%	2.4%
-- Net interest and dividends on securities paid (credit lines, securities borrowing, etc.)	8.6%	7.8%	7.1%	6.0%	5.4%	6.5%	1.0%	0.3%	0.1%	0.0%	12.5%	12.8%	12.3%
-- Other trading costs (administrative & technical costs, overhead, etc.)	3.2%	4.4%	5.8%	1.8%	0.8%	0.5%	0.0%	0.0%	0.0%	0.0%	11.5%	13.3%	18.0%
Trading Profit Margin	40.0%	42.2%	27.4%	37.9%	47.6%	41.0%	37.5%	51.5%	51.4%	59.6%	58.4%	56.3%	52.5%
Trading Revenues / (Trading Assets Minus Trading Liabilities)**	331%	377%	373%		95.5%	60.3%	62.4%				117.9%	118.9%	103.3%
Trading Revenues / (Book Equity)	323%	307%	132%		83.7%	59.9%	80.4%				168.6%	146.0%	123.1%

* Does not include any costs associated with the December 19, 2012 merger agreement with Knight, such as any costs related to the Knight August 1, 2012 incident

** Trading asset include cash and cash equivalents, financial instruments owned, receivables from broker-dealers and clearing organizations, and collateralized agreements. Trading liabilities include short-term borrowings, collateralized financing, financial instruments sold and not yet purchased, payables to broker-dealers and clearing organizations, and other accounts payable.

Table 4: Trading performance and latency

This table analyzes the relationship between trading performance and latency. It reports coefficients estimated from Eq. (1) for five performance measures as dependent variables: *Revenues*, *Returns*, *Sharpe Ratio*, *Trading Volume*, and *Revenues per MSEK Traded* (all defined as in Table 2). *Revenue*, *Returns*, *Revenues per MSEK Traded*, and *Trading Volume* are calculated at a daily level for each HFT aggregated across stocks and then averaged across trading days in each month to get firm-month observations; and the *Sharpe Ratio* is calculated as firm-month observations using the mean and standard deviation of daily observations of *Revenues* aggregated across stocks. We estimate OLS regressions with monthly fixed effects. The independent variables considered are as follows: *Top 1*, *Top 2-3*, and *Top 4-5* are indicator variables for whether a given firm is ranked among the top 1, top 2-3, or top 4-5 firms by *Latency* in a given month (defined as in Table 2). The *End-of-Day Inventory Ratio*, *Max. Intraday Inventory Ratio*, *Investment Horizon* and the *Aggressiveness Ratio* are defined as in Table 2. All continuous independent variables are in units of standard deviations. We omit *Max. Intraday Inventory Ratio* as a control when estimating *Returns* as the dependent variable, because *Max. Intraday Inventory ratio* is used in the denominator to calculate returns. *, ** and *** correspond to statistical significance at the 10%, 5%, and 1%, respectively. Standard errors are dually clustered by firm and month and are reported in the parentheses. The sample consists of 25 Swedish stocks and 60 months of trading (January 2010 to December 2014).

	Revenues			Returns			Sharpe Ratio			Log Trading Volume			Revenues per MSEK Traded		
Log Latency	-15530***	-2216	5569	-.254***	.00495	.0457	-5.2***	-1.51	.997	-1.14***	-.716**	-.232	19.5	136	115
	(5250)	(5666)	(9316)	(.0546)	(.102)	(.11)	(.856)	(1.53)	(1.92)	(.207)	(.28)	(.302)	(47.5)	(102)	(84.4)
Top 1 dummy		58213***	54387***		.794***	.793***		12.8***	12.5***		1.41***	1.29***		318	330
		(15821)	(14140)		(.228)	(.241)		(3.93)	(3.39)		(.381)	(.289)		(194)	(207)
Top 2-3 dummy		22842**	18779**		.605***	.595**		8.32**	7.31**		.995**	.777**		262*	279
		(8861)	(7547)		(.233)	(.233)		(3.68)	(3.01)		(.395)	(.309)		(150)	(171)
Top 4-5 dummy		9860	8256		.341	.336*		2.5	1.89		.348	.226		243*	252*
		(7666)	(6816)		(.214)	(.204)		(2.49)	(1.92)		(.223)	(.206)		(145)	(152)
End-of-Day Inv.			3010			.0657			.219			.053			129
			(2385)			(.077)			(.569)			(.0495)			(145)
Max Intraday Inv.			-9698**			[omitted]			-2.59**			-.715***			26
			(3904)						(1.14)			(.228)			(92.4)
Investment Horizon			-6580			-.0749			-1.7			-.145			8.17
			(9434)			(.0642)			(1.35)			(.218)			(79.8)
Aggressiveness Rat.			8314*			.00336			-.337			.131			4.41
			(4571)			(.0506)			(.704)			(.0954)			(39.3)
Constant	25597***	13537**	15016***	.309***	.0472	.0503	6.3***	2.96*	3.24**	18.7***	18.3***	18.4***	59.8	-69.7	-76
	(7805)	(5867)	(5540)	(.0539)	(.083)	(.0843)	(1.34)	(1.69)	(1.56)	(.243)	(.259)	(.219)	(43.5)	(71)	(76.5)
Month FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.124	0.170	0.231	0.127	0.149	0.157	0.247	0.299	0.351	0.519	0.549	0.675	0.058	0.062	0.076
N	615	615	615	615	615	615	615	615	615	615	615	615	615	615	615

Table 5: Persistence of HFT trading performance

This table analyzes persistence in firms' performance on both a firm-day (Panel A) and a firm-month (Panel B) frequency. The persistence coefficient reported in the table is the β estimated in the regression model given in Eq. (2), where *Performance* can be one of the following dependent variables calculated on a daily basis for each HFT (aggregated across stocks): *Revenues*, *Revenues per MSEK Traded*, *Returns,* and the *Sharpe Ratio* (all defined as in Table 2). For the monthly frequency each of the performance measures defined on a daily frequency are averaged across trading days. The *Sharpe Ratio* is considered only for the monthly regressions and is based on mean and standard deviation of daily observations of *Revenues*. The variables are either in units of standard deviations for each day or month ("standardized"; i.e., in each time period, firm-level performance is centered on the mean and scaled by standard deviation across HFT firms) or on the rank order of the HFTs (from 1 to 16 based on performance). *, ** and *** correspond to p-values lower than 10%, 5%, and 1%, respectively. Standard errors are dually clustered by firm and day (month for Panel B) and are reported in the parentheses. The sample consists of 25 Swedish stocks and 60 months of trading (January 2010 to December 2014).

Panel A: Daily persistence

	Standardized			Rank order		
	Revenues	Revenues per MSEK Traded	Returns	Revenues	Revenues per MSEK Traded	Returns
lag dependent variable	.226***	.0223	.433***	.228***	.0389	.325***
	(.0846)	(.0191)	(.0763)	(.0657)	(.0258)	(.0648)
R-squared	0.056	0.019	0.193	0.140	0.106	0.192
N	8276	8276	8276	8276	8276	8276

Panel B: Monthly persistence

	Standardized				Rank order			
	Revenues	Sharpe Ratio	Revenues per MSEK Traded	Returns	Revenues	Sharpe Ratio	Revenues per MSEK Traded	Returns
lag dependent variable	.604***	.710***	.192***	.428***	.534***	.207**	.168***	.650***
	(.117)	(.0841)	(.0707)	(.129)	(.0641)	(.099)	(.0558)	(.0668)
R-squared	0.372	0.528	0.089	0.224	0.324	0.097	0.084	0.457
N	598	598	598	598	598	598	598	598

Table 6: Long-run trends in HFT industry-wide performance

This table reports long-run trends in various variables related to the HFT industry. *Revenues*, *Returns*, and *Trading Volume* are defined as in Table 2 and are aggregated across stocks. *Total Daily Revenues* (*Average Daily Revenues*) is *Revenues* summed (averaged) across all HFT firms, reported in SEK; *Average Revenues per MSEK Traded* is daily *Revenues per MSEK Traded* aggregated across HFT firms; *Daily Trading Volume* is the daily *Trading Volume* summed across HFT firms; *Average Daily Returns* is the *Returns* aggregated across HFT firms and reported in annualized terms; *Cost of HFT Intermediation for Non-HFTs* is *Revenues* divided by non-HFT trading volume. All measures are calculated on a daily frequency and reported in Panel A as the average across the trading days of each half-year period. Standard errors are given in parentheses. Panel A also reports *T*, the number of trading days in each half-year period. Panel B reports estimates of regressions aimed at identifying time trends, specified for each variable observed at daily frequency. The regression specification is given in Eq. (3). An estimated $\beta > 0$ indicates an increasing trend in the dependent variable. *, ** and *** correspond to p-values lower than 10%, 5%, and 1%, respectively. In Panel B, Newey-West standard errors with 30 day lags are in parentheses. The sample consists of 25 Swedish stocks and 60 months of trading (January 2010 to December 2014).

Panel A: Biannual averages

	Total Daily Revenues (SEK)	Average Daily Revenues (SEK)	Average Revenues per MSEK Traded (SEK)	Daily Trading Volume (MSEK)	Average Daily Returns	Cost of HFT Intermediation for Non-HFTs (bps)	T (days)
2010:1	168,452	24,531	101.24	1617.83	0.25	0.133	124
	(19,744)	(2,795)	(10.69)	(57.13)	(0.0295)	(0.014)	
2010:2	169,154	29,430	109.81	1579.41	0.28	0.115	130
	(26,285)	(4,345)	(17.97)	(65.78)	(0.0442)	(0.015)	
2011:1	344,693	33,976	80.71	4503.11	0.20	0.272	124
	(78,895)	(7,313)	(13.39)	(148.80)	(0.0433)	(0.049)	
2011:2	239,204	27,127	93.20	2530.18	0.27	0.175	130
	(28,191)	(3,133)	(9.58)	(80.20)	(0.0318)	(0.018)	
2012:1	260,454	28,243	65.04	4131.75	0.09	0.343	123
	(33,696)	(3,916)	(7.37)	(87.41)	(0.0108)	(0.054)	
2012:2	297,890	29,102	67.77	4274.82	0.11	0.287	127
	(43,554)	(4,050)	(9.33)	(108.39)	(0.0191)	(0.037)	
2013:1	261,503	27,682	45.11	5634.01	0.14	0.432	122
	(29,656)	(3,024)	(4.510)	(108.63)	(0.0162)	(0.044)	
2013:2	292,436	29,347	58.87	5446.29	0.08	0.376	128
	(35,603)	(3,369)	(5.855)	(128.45)	(0.0094)	(0.039)	
2014:1	324,432	42,910	70.37	4222.76	0.19	0.345	121
	(64,944)	(7,769)	(9.907)	(143.01)	(0.0321)	(0.055)	
2014:2	198,559	25,608	45.04	4856.72	0.12	0.206	129
	(34,619)	(4,695)	(7.777)	(154.05)	(0.0203)	(0.035)	

Panel B: Time trend regressions

	Total Daily Revenues (SEK)	Average Daily Revenues (SEK)	Average Revenues per MSEK Traded (SEK)	Daily Trading Volume (MSEK)	Average Daily Returns	Cost of HFT Intermedia-tion for Non-HFTs (bps)
(year - 2010)	20,447*	1,613	-12.48***	778.8***	-.03332***	.05367***
	(11,299)	(1,379)	(2.922)	(118)	(.009925)	(.01109)
Constant	203,345***	25,677***	106***	1869***	.2585***	.1305***
	(31,414)	(3,617)	(9.54)	(303.7)	(.03125)	(.02512)
R-squared	0.004	0.002	0.024	0.356	0.022	0.030
N	1255	1255	1255	1255	1255	1255

Table 7: Long-run trends in HFT concentration

This table reports long-run trends in various variables measuring the HFT industry concentration in terms of revenues and volumes. The *Herfindahl Concentration by Revenues* or *by Trading Volumes* is calculated as in Eq. (4). As another measure of concentration, the ratio of average *Revenues* or *Trading Volume* of the Top 3 or 5 HFTs to the median HFT is reported. Panel A reports biannual averages and standard deviations (in parentheses), along with *T*, the number of trading days in each time period. Panel B reports estimates of regressions aimed at identifying time trends, specified for each variable observed at daily frequency. The regression specification is given in Eq. (5). An estimated $\beta > 0$ indicates an increasing trend in the dependent variable. *, ** and *** correspond to p-values lower than 10%, 5%, and 1%, respectively. In Panel B, Newey-West standard errors with 30 day lags are in parentheses. The sample consists of 25 Swedish stocks and 60 months of trading (January 2010 to December 2014).

Panel A: Biannual averages

	Herfindahl Industry Conc. by Trading Volumes	Herfindahl Industry Conc. by Revenues	Ratio of Top 3 to Median-HFT Volume	Ratio of Top 3 to Median-HFT Revenue	Ratio of Top 5 to Median-HFT Volume	Ratio of Top 5 to Median-HFT Revenue	T (days)
2010:1	0.301	0.382	3.624	4.540	1.444	2.154	124
	(0.044)	(0.05)	(0.238)	(0.493)	(0.113)	(0.201)	
2010:2	0.343	0.401	2.514	7.002	1.121	1.969	130
	(0.077)	(0.127)	(0.197)	(0.839)	(0.063)	(0.272)	
2011:1	0.257	0.289	3.551	3.371	2.295	2.722	124
	(0.039)	(0.109)	(0.183)	(0.224)	(0.109)	(0.175)	
2011:2	0.226	0.290	4.087	3.174	2.675	2.379	130
	(0.029)	(0.12)	(0.209)	(0.181)	(0.176)	(0.145)	
2012:1	0.200	0.302	1.524	3.075	1.250	2.485	123
	(0.044)	(0.103)	(0.07)	(0.235)	(0.042)	(0.17)	
2012:2	0.221	0.326	2.058	3.306	1.772	2.729	127
	(0.031)	(0.105)	(0.083)	(0.226)	(0.074)	(0.195)	
2013:1	0.206	0.343	1.479	4.671	1.256	2.981	122
	(0.018)	(0.141)	(0.061)	(0.359)	(0.044)	(0.204)	
2013:2	0.181	0.310	1.302	5.378	1.195	3.967	128
	(0.02)	(0.117)	(0.085)	(0.474)	(0.069)	(0.278)	
2014:1	0.277	0.360	2.382	3.573	1.344	2.296	121
	(0.025)	(0.093)	(0.147)	(0.463)	(0.069)	(0.258)	
2014:2	0.293	0.359	1.926	4.302	1.539	2.618	129
	(0.015)	(0.058)	(0.096)	(0.298)	(0.073)	(0.19)	

Panel B: Time trend regressions

	Herfindahl Industry Conc. by Trading Volumes	Herfindahl Industry Conc. by Revenues	Ratio of Top 3 to Median-HFT Volume	Ratio of Top 3 to Median-HFT Revenue	Ratio of Top 5 to Median-HFT Volume	Ratio of Top 5 to Median-HFT Revenue
(year-2010)	-.0117**	-.0026	-.624***	-.0695	-.179*	.0553
	(.0059)	(.0049)	(.189)	(.21)	(.108)	(.192)
Constant	.281***	.343***	4.5***	4.17***	2.27***	2.83***
	(.0164)	(.017)	(.697)	(.665)	(.414)	(.561)
R-squared	0.072	0.001	0.081	0.000	0.017	0.000
N	1255	1255	1255	1255	1255	1255

Table 8: HFT entry and exit analysis

This table analyzes the determinants of HFT entry and exit into stocks. The performance measures *Revenues*, *Revenues per MSEK Traded*, and *Returns* are defined as in Table 3, except that they are not aggregated across stocks. Panel A reports coefficient estimates from Eq. (6), estimated on a panel of firm-stock-day observations. The *one-month dummy* takes the value 1 if firm *i* began trading stock *j* in the last 30 days, otherwise 0; *two*- and *three-month dummy* variables are defined similarly. We exclude the observations in January 2010 as this is when we first observe any firm. In the fourth column, a linear probability regression is estimated using Eq. (6) but with the dummy $Exit_{i,t}$ as the dependent variable, which takes the value 1 on day *t* for firm *i* if that is the last day firm *i* trades. This regression excludes observations in December 2014, the last month of the analysis, since we cannot determine exits. In the fifth column, an OLS is estimated on a firm-stock-month using Eq. (6) but with *Latency* as the dependent variable (since *Latency* is a firm-month variable, it is assigned to be constant across stocks). *, ** and *** correspond to p-values lower than 10%, 5%, and 1%, respectively. Standard errors are dually clustered by firm-stock and month and are reported in the parentheses. The sample consists of 25 Swedish stocks and 60 months of trading (January 2010 to December 2014).

	Revenues (thous. SEK)	Revenues per MSEK Traded	Returns	Daily Probability of Exit (x 10^3)	Latency (in ms): monthly obs.
One-month dummy	-3.12**	95.81	-.0387***	-.789*	74.67**
	(1.335)	(170.3)	(.0132)	(.447)	(31.19)
Two-month dummy	-2.35*	-37.06	-.0203*	.816*	134.1***
	(1.288)	(184.7)	(.0117)	(.454)	(32.65)
Three-month dummy	-.6756	120	-.0178	7.459***	-2.885
	(1.229)	(208.3)	(.0129)	(.518)	(4.466)
Constant	1.563***	67.21***	.0183***	.478***	8.368***
	(.2112)	(2.616)	(.0023)	(.051)	(1.062)
Day x Stock FEs	Yes	Yes	Yes	Yes	(Month x Stock FEs)
R-squared	0.093	0.105	0.078	0.160	0.479
N	217885	217885	217885	217885	10251

Table 9: Price impact, realized spread and latency

This table analyzes revenue per trade by looking at two dependent variables: active *Price Impact* and passive *Realized Spread*. *Price Impact* is the basis point change in spread midpoint from just before to ten seconds after a trade initiated by an HFT firm. *Realized Spread* is the basis point difference between the transaction price and the bid-ask spread midpoint ten seconds after a trade where an HFT firm was the liquidity provider. *Top 1*, *Top 2-3*, and *Top 4-5* are indicator variables for whether a given firm is ranked among the top 1, top 2-3, or top 4-5 firms by speed in a given month. *Latency*, *Average End-of-Day Inventory*, *Maximum Intraday Inventory*, *Investment Horizon*, and the *Aggressive Ratio* are defined as in Table 4. We also control for the following stock-specific variables defined in Table 1: *Non-HFT Trading Volume* (in SEK), *Realized Volatility* (winzorized at the 99.5% quantile), and the *Fragmentation Index*, *Tick Size* and *Quoted Spread*. All continuous variables are in units of standard deviations. *, ** and *** correspond to p-values lower than 10%, 5%, and 1%, respectively. Standard errors are dually clustered by firm-stock and month and are reported in the parentheses. The sample consists of 25 Swedish stocks and 60 months of trading (January 2010 to December 2014).

	Price Impact		Realized Spread	
Log latency	-.333	-.195	-.23**	-.155
	(.219)	(.202)	(.0991)	(.0995)
Top 1 dummy	1.06***	.794**	.567**	.443**
	(.334)	(.329)	(.223)	(.183)
Top 2-3 dummy	.802***	.511	.539**	.391**
	(.306)	(.347)	(.215)	(.167)
Top 4-5 dummy	.55*	.407	.389*	.292
	(.289)	(.289)	(.222)	(.185)
Avg. End-of-Day Inv.		.597**		.0461
		(.248)		(.116)
Max. Intraday Inv.		-.102		-.0509
		(.372)		(.207)
Investment Horizon		-.365		-.159
		(.361)		(.138)
Aggressiveness Ratio		-.13		-.131**
		(.118)		(.0548)
Non-HFT Trading Vol.		-.082		.187***
		(.202)		(.0694)
Volatility		.0839		-.333***
		(.142)		(.0878)
Fragmentation Index		-.334**		-.198***
		(.138)		(.0668)
Tick Size		.197		-.432***
		(.39)		(.141)
Quoted Spread		.625		.722***
		(.506)		(.179)
Constant	3.91***	4.06***	-.0377	.0311
	(.215)	(.186)	(.135)	(.122)
(Month x Stock) Fes	Yes	No	Yes	No
R-squared	0.195	0.028	0.162	0.019
N	10367	10367	10216	10216

Table 10: Cross-market arbitrage in different HFT latency segments

This table presents probit regression estimates corresponding to the probability of initiating a trade (*Active Trading*; Panel A) or supplying liquidity in a trade (*Passive Trading*; Panel B) in the equity markets in response to a change in the futures index price. The dependent variable is 1 when a *Fast HFT* performs the trade and zero if a *Slow HFT* does it. In each month, *Slow HFTs* are those who are not among the top five HFTs in terms of trading speed. *Fast HFTs* are either the top one, top three, or top five HFTs in terms of trading speed. Trades included in the analysis are those performed by HFTs in the sample stocks. *News* is +1 if the active party trades in the direction of the news and -1 if the active party trades in the opposite direction of the news, when the absolute return on the OMXS30 futures during a one-second window preceding the stock trade is "large" (in the top decile for each month among non-zero return), and zero otherwise. We include the following control variables: *Lagged Volatility*, the average second-by-second squared return (multiplied by 1,000) over the previous ten seconds; *Lagged Volume*, the *SEK Trading Volume* (divided by 100,000) over the previous ten seconds; *Quoted Spread*, the difference between the best bid and offer quotes (multiplied by 10,000), divided by the midpoint quote; *Depth at BBO*, the average number of shares available at the best bid quote and the best offer quote (divided by 100,000), multiplied by the midpoint quote. Marginal effects are also reported. Regressions are estimated month-by-month from January 2010 to December 2014; reported coefficients and marginal effects are means across the sixty month. *, ** and *** correspond to p-values lower than 10%, 5%, and 1%, respectively. Standard errors are in parentheses.

Panel A: Active trading

	"Fast" = Top 1 HFT		"Fast" = Top 3 HFTs		"Fast" = Top 5 HFTs	
	Probit (1=Fast HFT)	Marginal effects	Probit (1=Fast HFT)	Marginal effects	Probit (1=Fast HFT)	Marginal effects
Constant	1.397***		2.118***		2.229***	
	(0.26)		(0.18)		(0.17)	
News	0.128***	0.009	0.166***	0.008	0.164***	0.007
	(0.05)		(0.03)		(0.03)	
Lagged Volatility	-0.008***	-0.001	-0.010***	-0.001	-0.008***	-0.001
	(0.00)		(0.00)		(0.00)	
Lagged Volume	-0.004***	0.000	-0.004***	0.000	-0.004***	0.000
	(0.00)		(0.00)		(0.00)	
Quoted Spread	-0.040***	-0.004	-0.033***	-0.001	-0.031***	0.000
	(0.01)		(0.01)		(0.00)	
Depth at BBO	0.105***	0.025	0.057***	0.009	0.043**	0.005
	(0.03)		(0.02)		(0.02)	
Stock FEs	Yes		Yes		Yes	
Average N	109684		222866		277044	
Av. psuedo-R2	0.209		0.176		0.169	

Panel B: Passive trading

	"Fast" = Top 1 HFT		"Fast" = Top 3 HFTs		"Fast" = Top 5 HFTs	
	Probit (1=Fast HFT)	Marginal effects	Probit (1=Fast HFT)	Marginal effects	Probit (1=Fast HFT)	Marginal effects
Constant	0.801***		1.560***		1.678***	
	(0.21)		(0.11)		(0.11)	
News	-0.048	-0.002	-0.123***	-0.015	-0.122***	-0.014
	(0.03)		(0.02)		(0.02)	
Lagged Volatility	0.003	0.000	0.006***	0.001	0.006***	0.001
	(0.00)		(0.00)		(0.00)	
Lagged Volume	-0.001	0.000	0.000	0.000	0.000	0.000
	(0.00)		(0.00)		(0.00)	
Quoted Spread	-0.012**	-0.004	-0.006*	-0.002	-0.005	-0.001
	(0.01)		(0.00)		(0.00)	
Depth at BBO	0.025	0.010	-0.013	-0.001	-0.011	-0.001
	(0.04)		(0.02)		(0.02)	
Stock FEs	Yes		Yes		Yes	
Average N	95268		208771		258409	
Av. psuedo-R^2	0.204		0.169		0.163	

Appendix

A1. TRS data processing

This appendix presents a summary of our data processing. There are on average 8.56 million entries in each month of the TRS data. Post-processing, the number of observations is on average 6.16 million per month.

(a) Time stamp adjustment

The firms reporting to TRS operate in different time zones and the data base does not have a built in functionality to adjust the trade times to a common time zone. To adjust the time stamps we record in which hour the first and last exchange-transaction is executed (at NASDAQ OMX Stockholm or at one of the MTFs) for each firm in each day. All the exchanges open at 8 am (GMT) and close at 4:30 pm (GMT). We adjust the time stamps of firms that do not have their median first and last trades in sync with the opening hours. For example, for a firm with the median first trade hour in a month being 7 am and the median last trade hour being 3 pm, we adjust all transaction time stamps by +1 hour.

(b) Matching to TRTH data

TRS transactions are matched to TRTH transactions using information on stock, trading venue, date, time, price, buy/sell, and quantity. The timestamps of the two databases are allowed to differ by up to one second. Where trader IDs are available in the TRTH data (for NASDAQ OMX Stockholm only), they are added to the matching criterion. TRTH trades are split into two transactions, one for the buyer and one for the seller. If there are several matches to one transaction, the transaction closest in time is considered to be the closest match.

(c) Firms

We analyze trading revenues at the corporation level rather than the branch or division level. Accordingly, we truncate BIC codes (11-letter identifiers of the financial institutions reporting to TRS) to the first four letters that are unique to each corporation. For various reasons, such as mergers and acquisitions, the same corporation may span several (truncated) BIC codes. For example, GETCO acquired Knight Trading in August 2013. We thus treat the (truncated) BIC codes GEEU (GETCO) and NITE (Knight Trading) as separate for the period preceding the merger and as one corporation for August 2013 onwards.

(d) Filtering of trades

Non-proprietary transactions frequently generate more than one entry in TRS. For example, if a broker buys 100 shares on behalf of a client, it may be reported as two transactions in TRS: one transaction where the reporting firm purchases 100 shares at the exchange, and one off-exchange transaction where the reporting firm sells 100 shares to its client.[1] As firms differ in how they report their transactions we need to process the data to make transactions comparable.

For each transaction we seek to retain one representative TRS entry and to attach an entry of the *end investor* associated with that transaction. The *end investor* assignment is done differently depending on the type of trade.

We define *primary transaction* as TRS entries where the counterparty of the trade is a clearing house or the same as the trading venue for the transaction **or** the owner of the trading

[1] In a memorandum on transaction reporting FI provides numerous examples on how different types of trades on behalf of clients may be reported. The memo may be retrieved at http://www.fi.se/upload/90_English/90_Reporting/TRS/memo_transaction_reporting_ver_1_7_2014-03-07.pdf

venue (in the case of dark pools). The definition is motivated by the fact that all exchange transactions must be done through central counterparty (CCP) clearing. All other TRS entries are defined as *secondary transactions*. Of all TRS entries, 81.5% are considered primary transactions.

(i) Primary transaction matched to a secondary transaction of the same firm

To account for several entries reported for the same transaction, we match primary and secondary transactions by *reporting firm*, *stock*, *price*, *quantity*, *date*, and *time*. The time stamps are allowed to differ by no more than one second. The end investor of primary trades matched in this way is set to the client of the secondary trade, if available, and otherwise to the counterparty of the secondary trade. The matched secondary trades are then discarded. Of all primary transactions, 26 % are matched to a secondary transaction in this way.

(ii) Primary transaction matched to a secondary transaction of another firm

To account for the case that the same transaction is reported by both counterparties, we match the reporting firm of primary transactions to the counterparty of secondary transactions. The other matching criteria include *stock*, *price*, *quantity*, *date*, and *time*. As above, the time stamps are allowed to differ by no more than one second. The end investor of primary trades matched in this way is set to the client of the secondary trade, if available, and otherwise to the *reporting firm* of the secondary trade. The matched secondary trades are then discarded. Of all primary transactions, 12 % are matched to a secondary transaction in this way.

(iii) Primary transaction that is not matched to a secondary transaction

Primary transactions that are not matched to a secondary transaction are considered to be on behalf of a client if a client reference is available, and otherwise proprietary. For client (proprietary) trades, the end investor is set equal to the client reference (the reporting firm).

(iv) Secondary transaction that is not matched to a secondary transaction

Secondary transactions that are not matched to a primary transaction, are considered to be on behalf of a client if a client reference is available, and otherwise proprietary. For client (proprietary) trades, the end investor is set equal to the client reference (the reporting firm).

(v) Secondary transactions where the counterparty does not report to TRS

To capture firms that are not obliged to report to TRS, but that still trade our sample stocks, we look for firms that are reported as counterparties but that not show up as reporting firms. For all secondary transactions where such firms appear as counterparties, we create a new entry with the same properties but with opposite direction of trade and with counterparties reversed. This is a way to detect HFT firms that connect to the market through direct market access or sponsored access. Of all secondary transactions, 16 % are subject to this procedure.

A2. Comparison of HFT revenue calculation methods

We compare four methods of calculating trading revenues. As explained in the main text, adjustments are needed because small data errors in inventory accumulate over time, leading to large and persistent (unit root) data errors if left uncorrected. The four methods are as follows. *No adjustment* is calculated by cumulating daily inventory positions over the full sample; *Method 1 – Benchmark* is the method used throughout the paper that zeros the end-of-day position daily for each HFT (equivalent to assuming that each HFT liquidates any remaining end-of-day position at the daily closing price); *Method*

2 – Intraday Revenues assumes that any remaining end-of-day positions were never purchased in the first place (assuming first-in-last-out inventory accounting); *Method 3 – Intraday Revenues Plus Revenues from Inventory Sold* is similar to *Method 2* but adds back in the revenues from closing end-of-day positions that are in opposite direction of previous day end-of-day inventory (that is, the end-of-day inventory is marked to market only if an offsetting position exists in the previous end-of-day inventory).

Appendix Table 1 reports the firm cross-sectional distribution of HFT trading revenues (as in Table 2) using the four different methods for calculating trading revenues. It shows that inventory adjustments do not alter the main results of the paper in the sense that the mean and median are similar across the different methods (though the standard deviation blows up, as expected, without the inventory adjustments). This is consistent with the fact that inventory corrections reflect small white noise errors in end-of-day positions.

INSERT APPENDIX TABLE 1 ABOUT HERE

A3. Construction of daily Fama-French plus momentum factors for Swedish equities

We construct daily Fama-French and momentum factors for Swedish equities. The data used to construct the factors (daily total excess returns, shares outstanding, and quarterly book values) comes from Compustat Global and covers the period January 2010 to December 2014. We exclude stock-day observations in which the total market capitalization falls below 100 MSEK (about 10.5 million USD as of the exchange rate of December 2014). The four factors (excess market return, small minus large [SML], value minus growth [HML], winner minus loser [WML]) are constructed according to the specifications used to create US factors, as specified on Ken French's website: the value-weighted portfolios consist of top-30%, middle 40%, and bottom-30% of stocks (by market capitalization, book to market, and past-12-

month returns for SML, HML and WML, respective) and are re-sorted every July 1 using data from the previous year's performance.

Appendix Table 2 reports summary statistics corresponding to these traded risk factors for Swedish equities. The statistics include the mean daily log excess return (annualized), its standard error, and the number of observations (i.e., the number of trading days), and are reported for each portfolio. The annualized excess returns on the four portfolios – market excess return, SMB, HML, WML – are 0.160, 0.176, 0.039, and 0.028, which are all positive, as expected.

INSERT APPENDIX TABLE 2 ABOUT HERE

A4. Exchange fees for three exchanges trading Swedish equities

Appendix Table 3 reports exchange fees in 2014 for three stock exchanges (NASDAQ OMX Stockholm, BATS, and Chi-X) trading Swedish equities. Fees range from 0 to 0.325 bps over these selected venues. Exchange fees depend on the side of the trade: "maker" fees are less than "taker" fees (for example, 0.13 vs. 0.325 bps for NASDAQ OMX S30 stocks), and, at Chi-X, makers receive liquidity rebates (negative fees) of about 0.225 bps. For NASDAQ OMX Stockholm, we report the fees for S30 stocks, which are lower than for other stocks; all the stocks in our sample fall into this category.

INSERT APPENDIX TABLE 3 ABOUT HERE

While NASDAQ OMX Stockholm grants preferential prices for liquidity provision under its Liquidity Provider Scheme (LPS), BATS and Chi-X do not (a designated liquidity provider program exists but doesn't have lower fees). Although BATS and Chi-X merged in November 2011, with technology integration complete by April 2012, the trading platforms continue to implement different pricing structures.

A5. Additional analysis of latency and HFT trading characteristics

Appendix Table 4 is an extension of Table 4 and estimates Eq. (1) with various other dependent variables, such as *End-of-day inventory* and *Maximum intraday inventory* (as a percent of daily trading volume), *Investment horizon* (the holding time of the median trade for each month and HFT, in seconds, assuming first-in-last-out accounting), and the *aggressive ratio* (SEK of aggressive trading volume to total trading volume). As before, the variables $\mathbf{1}_{\text{top 1}\ i,t}$, $\mathbf{1}_{\text{top 2-3}\ i,t}$, and $\mathbf{1}_{\text{top 4-5}\ i,t}$ are indicators capturing relative latency, log(*Latency*) captures nominal latency. Continuous independent variables are in units of standard deviations. Control variables include those in Table 4 defined in Section IV, except if a control variable is also the dependent variable, in which case that variable is omitted as a control.

INSERT APPENDIX TABLE 4 ABOUT HERE

The table shows that faster HFTs have lower inventory bounds and shorter investment horizons, presumably reflecting the fact that faster traders can more skillfully rebalance their inventory positions intraday without incurring trading losses. The result for the aggressive ratio is mixed: overall, lower latency HFTs are somewhat less aggressive, though the very fastest HFTs may have slightly higher aggressive ratios (though the coefficients are not statistically significant).

A6. Stock-level analysis of latency and HFT performance

We repeat the analysis in Section IV examining the role of latency but disaggregate performance by stock to test a variety of economically interesting stock-level variables that may influence HFT performance. Specifically, we estimate the following regression equation on a firm-stock-month frequency, with control variables including both HFT firm characteristics and stock characteristics, and month fixed effects,

$$Performance_{i,j,t} = \alpha_t + \beta_1 \log(Latency)_{i,t} + \beta_2 \mathbf{1}_{\text{top 1 } i,t} + \beta_3 \mathbf{1}_{\text{top 2-3 } i,t} + \beta_4 \mathbf{1}_{\text{top 4-5 } i,t}$$
$$+\gamma' controls_{i,j,t} + \text{ month-FEs} + \epsilon_{i,t} \ . \tag{7}$$

As before, the variables $\mathbf{1}_{\text{top 1 } i,t}$, $\mathbf{1}_{\text{top 2-3 } i,t}$, and $\mathbf{1}_{\text{top 4-5 } i,t}$ are indicators capturing relative latency, log(*Latency*) captures nominal latency. Continuous independent variables are in units of standard deviations. Control variables include those in Table 4 defined in Section IV, in addition to stock-specific variables included in Table 8 and defined in Section VI such as: *Non-HFT Trading Volume*, *Realized Volatility*, *Fragmentation*, *Tick Size*, and *Quoted Spread*.

INSERT APPENDIX TABLE 5 ABOUT HERE

Appendix Table 5 reports the results. Across various performance measures (*Revenues*, *Returns*, and the *Sharpe Ratio*), coefficients on $\mathbf{1}_{\text{top 1 } i,t}$ and $\mathbf{1}_{\text{top 2-3 } i,t}$ [but not for log(*Latency*)] are significant, corroborating the results in Table 4 even in the presence of a wide variety of firm-, stock-, and market-level controls. Appendix Table 5 furthermore shows that higher *Revenues* (as well as the *Sharpe Ratio* and the *Revenue per MSEK Traded*) is associated with: greater market fragmentation within a given contract, greater *Tick Size*, higher *Quoted Spread*, higher *Non-HFT Trading Volume*, greater *Volatility*, lower firm-level *Maximum Intraday Inventory*, shorter *Investment Horizons*, and higher *Aggressiveness Ratio*.

Appendix Table 1: Comparison of HFT revenue calculation methods

This table reports the firm cross-sectional distribution of HFT trading revenues (as in Table 2) using four different methods for calculating trading revenues. *No adjustments* is calculated by cumulating daily inventory positions over the full sample; *Method 1 – Benchmark* is the method used throughout the paper that zeros the end-of-day position daily for each HFT (equivalent to assuming that each HFT liquidates any remaining end-of-day position at the daily closing price); *Method 2 – Intraday Revenues* assumes that any remaining end-of-day positions were never purchased in the first place (assuming first-in-last-out inventory accounting); *Method 3 – Intraday Revenues Plus Revenues from Inventory Sold* is similar to *Method 2* but adds back in the revenues from closing end-of-day positions that are in opposite direction of previous day end-of-day inventory (that is, the end-of-day inventory is marked to market only if an offsetting position exists in the previous end-of-day inventory).

	mean	stdev	p5	p10	p25	p50	p75	p90	p95
No adjustments									
Revenues	36,126	409,646	-1,188,071	-136,612	-53,835	41,742	124,947	673,929	751,731
Sharpe Ratio	1.58	4.00	-1.59	-0.72	-0.39	0.80	1.79	2.72	15.93
Revenues per MSEK Traded	161.3	3,297.7	-4,646.9	-4,522.3	-1,803.9	68.2	2,185.2	3,400.2	8,115.5
Method 1 – Benchmark									
Revenues	21,920	31,378	-13,941	-2,166	-623	13,224	34,131	63,463	111,376
Sharpe Ratio	5.07	7.06	-1.94	-0.59	0.21	2.80	8.20	11.20	26.72
Revenues per MSEK Traded	41.5	219.7	-511.4	-305.9	-12.6	65.9	106.2	342.6	460.3
Method 2 - Intraday Revenues									
Revenues	26,361	26,475	-12,449	-10,274	7,624	22,050	52,548	67,059	67,486
Sharpe Ratio	4.88	5.47	-5.31	1.00	1.35	4.12	8.08	10.75	18.87
Revenues per MSEK Traded	205.3	455.3	-292.8	-178.4	52.1	67.3	219.8	486.6	1,757.1
Method 3 - Intraday Revenues Plus Revenues from Inventory Sold									
Revenues	15,669	39,173	-57,219	-27,891	1,685	16,078	25,182	54,296	121,667
Sharpe Ratio	3.17	4.94	-4.45	-2.08	-0.22	1.94	6.62	9.18	14.48
Revenues per MSEK Traded	-90.7	834.6	-2,964.6	-570.1	4.5	62.0	135.7	265.5	1,124.6

Appendix Table 2: Daily Fama-French plus momentum factors for Swedish equities

This table reports summary statistics corresponding to newly created daily Fama-French plus momentum factors for Swedish equities. The mean daily log excess return (annualized), its standard error, and the number of observations (i.e., number of normal trading days) are reported for each of the portfolios. The four factors are constructed according to the specifications used to create US factors, as specified on Ken French's website: the value-weighted portfolios consist of top-30%, middle 40%, and bottom-30% of stocks and are re-sorted every July 1 using data from the previous year's performance. The data (daily total excess returns, shares outstanding, and quarterly book values) come from Compustat Global and covers the period January 2010 to December 2014.

	Mean	S.E.	Daily observations
log market excess returns	0.160	0.083	1255
log large-cap returns	0.152	0.085	1255
log medium-cap returns	0.231	0.070	1255
log small-cap returns	0.347	0.063	1255
log SML returns	0.176	0.068	1255
log growth returns	0.161	0.088	1255
log neutral returns	0.143	0.084	1255
log value returns	0.206	0.090	1255
log HML returns	0.039	0.054	1255
log winner returns	0.171	0.095	1255
log neutral returns	0.155	0.084	1255
log loser returns	0.136	0.088	1255
log WML returns	0.028	0.065	1255

Appendix Table 3: Exchange fees for three exchanges trading Swedish equities

This table reports exchange fees in 2014 for three stock exchanges (NASDAQ OMX Stockholm, BATS, and Chi-X) trading Swedish equities. Exchange fees depend on the side of the trade: "maker" fees are less than "taker" fees, and, at Chi-X, makers receive liquidity rebates (negative fees). NASDAQ OMX Stockholm fees are lower for S30 stocks; all the stocks in this study fall into this category. While NASDAQ OMX Stockholm grants preferential prices for liquidity provision under its Liquidity Provider Scheme (LPS), BATS and Chi-X do not (a designated liquidity provider program exists but it does not have lower fees). Although BATS and Chi-X merged in November 2011, with technology integration complete by April 2012, the trading platforms continue to implement different pricing structures.

	NASDAQ OMX Stockholm for S30 stocks	NASDAQ OMX Stockholm Liquidity Provider Scheme (LPS) for S30 stocks	BATS*	Chi-X*
Maker	0.13 bps	0 bps	0 bps	-0.15 to -0.225 bps**
Taker	0.325 bps	0.5 bps	0.15 bps	0.30 to 0.24 bps

* For non-hidden limit orders

** The exact price within this range depends on volume. The lowest fees are given after total monthly trading volume exceeds 16 billion Euro. Negative values represent liquidity rebates.

Appendix Table 4: Latency and other HFT trading characteristics

This table is an extension of Table 4 and estimates Eq. (1) with various other dependent variables. The table shows that faster HFTs have: lower inventory bounds and shorter investment horizons. The result for the *Aggressiveness Ratio* is mixed: overall, lower latency HFTs are somewhat less aggressive, though the very fastest HFTs may have slightly higher aggressive ratios (though the coefficients are not statistically significant). As in Table 9, Eq. (1) is estimated on a firm-month panel with month FEs and firm-level controls (estimates of control variables omitted to save space). The control variables are the same as Table 9 except if a control variable is also the dependent variable, in which case that variable is omitted as a control. *, ** and *** correspond to p-values lower than 10%, 5%, and 1%, respectively. Standard errors, dually clustered on firm and month, are in parentheses.

	End-of-Day Inventory Ratio		Max. Intraday Inventory Ratio		Investment Horizon		Aggressiveness Ratio	
Log Latency	-.0108	-.0227*	.109***	.0539***	2.73***	1.12***	.052**	.0684***
	(.0141)	(.0119)	(.012)	(.00966)	(.381)	(.298)	(.0225)	(.0235)
Top 1 dummy	-.033	-.0251	-.0188	-.016	-.31	.309	.0823	.0872
	(.0258)	(.0283)	(.0246)	(.0224)	(.758)	(.626)	(.0611)	(.0603)
Top 2-3 dummy	-.0239	-.0149	-.047**	-.0283	-.98	-.0779	.0139	.0153
	(.0223)	(.0259)	(.0219)	(.0202)	(.626)	(.531)	(.0497)	(.0488)
Top 4-5 dummy	-.0183	-.0129	-.0298	-.0193	-.496	.0102	-.0176	-.0148
	(.0228)	(.0259)	(.0218)	(.0196)	(.649)	(.545)	(.0451)	(.0425)
Constant	.0147	.0111	.186***	.18***	4.55***	4.21***	.461***	.459***
	(.0112)	(.0127)	(.0105)	(.00966)	(.308)	(.261)	(.0214)	(.0205)
Firm-level controls	No	Yes	No	Yes	No	Yes	No	Yes
Month FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.110	0.172	0.508	0.689	0.382	0.607	0.061	0.105
N	615	615	615	615	615	615	615	615

Appendix Table 5: Latency and HFT Trading Performance at the stock level

This table analyzes the relationship between trading performance and *Latency*. It is analogous to Table 4 but disaggregated on the stock-level, in order to account for stock-level characteristics. This table reports coefficients estimated from Eq. (1) on a (HFT x stock x month) panel with HFT-month and stock-month controls. The *Top 1/3/5 dummy* is an indicator of whether a given firm is ranked in the top 1/3/5 by speed in a given month, and control variables are described in Table 9. *, ** and *** correspond to p-values lower than 10%, 5%, and 1%, respectively. Standard errors are clustered by firm-stock and month and are reported in the parentheses. The sample consists of 25 Swedish stocks and 60 months of trading (January 2010 to December 2014).

	Trading Revenues		Returns		Sharpe Ratio		Log HFT Trading Volume		Revenues per MSEK Traded	
Log latency	-157	51.5	-.000335	-.00263	.545	1.16	-.535***	-.235**	-8.71	16.6
	(234)	(177)	(.0042)	(.00299)	(1.36)	(.95)	(.106)	(.0926)	(53.4)	(45.4)
Top 1 dummy	2147***	1675***	.0312***	.0227***	6.64**	6.57**	1.53***	.99***	-12.1	74.3
	(590)	(516)	(.0112)	(.00841)	(3.13)	(2.9)	(.229)	(.156)	(99.7)	(106)
Top 2-3 dummy	844*	414	.0271***	.0188**	5.01*	4.76*	1.12***	.565***	-26.1	41.8
	(496)	(443)	(.0104)	(.00826)	(2.77)	(2.63)	(.215)	(.161)	(107)	(112)
Top 4-5 dummy	178	12.4	.0105	.00527	-.845	-1.15	.331**	.0713	-27.5	4.52
	(335)	(260)	(.00731)	(.0056)	(1.56)	(1.57)	(.149)	(.117)	(90.9)	(83.3)
Avg. End-of-Day Inv.		122		.00325		.198		.0148		-31.7
		(159)		(.0028)		(.352)		(.036)		(61.3)
Max. Intraday Inv.		-446**		[omitted]		-.121		-.756***		-10.8
		(209)				(1.01)		(.0767)		(92.1)
Investment Horizon		-278		-.00322		-.591**		-.0643		34.3
		(231)		(.00331)		(.265)		(.0749)		(75.3)
Aggressiveness Ratio		372***		-.00217		-.0029		.158***		-41.3
		(91.9)		(.00172)		(.222)		(.0365)		(25.2)
Non-HFT Volume		812***		.0107***		-1.69		.503***		-71.1
		(157)		(.00314)		(1.83)		(.0469)		(93)
Volatility		84.7		.00923*		1.76		-.0696		73.7
		(173)		(.00539)		(1.56)		(.0525)		(177)
Fragmentation Index		155		.000537		.301		.179***		-24.6
		(108)		(.0021)		(.303)		(.0364)		(29.9)
Tick Size		77.7		.00208		2.18*		.104		123
		(131)		(.0023)		(1.3)		(.066)		(149)
Quoted Spread		104		-.000503		-3.73		-.242***		-172
		(175)		(.00308)		(3.08)		(.0778)		(225)
Constant	781***	970***	.00344	.00748*	.0528	.198	15***	15.2***	76.9	44.3
	(224)	(188)	(.00493)	(.00437)	(.749)	(1.24)	(.113)	(.0707)	(56.5)	(53.7)
(Month x Stock) FEs	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
R-squared	0.121	0.022	0.107	0.013	0.143	0.001	0.499	0.539	0.119	0.001
N	13081	13081	13081	13081	13081	13081	13081	13081	13081	13081