

Menkveld, Albert J.

Working Paper

High Frequency Trading and the New-Market Makers

Tinbergen Institute Discussion Paper, No. 11-076/2/DSF21

Provided in Cooperation with:

Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: Menkveld, Albert J. (2011) : High Frequency Trading and the New-Market Makers, Tinbergen Institute Discussion Paper, No. 11-076/2/DSF21, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:

<https://hdl.handle.net/10419/86769>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Duisenberg school of finance – Tinbergen Institute Discussion Paper

TI 11–076/2/DSF 21

High Frequency Trading and the *New-Market* Makers

Albert J. Menkveld

VU University Amsterdam, Tinbergen Institute, and Duisenberg School of Finance.

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: <http://www.dsf.nl/>

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

High Frequency Trading and The *New-Market* Makers¹

Albert J. Menkveld

first version: December 9, 2010

this version: August 15, 2011

¹Albert J. Menkveld, VU University Amsterdam, FEWEB, De Boelelaan 1105, 1081 HV, Amsterdam, Netherlands, tel +31 20 598 6130, fax +31 20 598 6020, albertjmenkveld@gmail.com and TI Duisenberg School of Finance. I thank Carole Comerton-Forde, Robert Engle, Joel Hasbrouck, Frank Hatheway, Terrence Hendershott, Charles Jones, Sunny Li, Emiliano Pagnotta, Lasse Pedersen, George Sofianos, and Michel van der Wel for comments. I am grateful to both NYSE Euronext and Chi-X for acting as data sponsor and to Bernard Hosman whose exploratory analysis has been invaluable. Menkveld gratefully acknowledges Robert Engle and Boyan Jovanovic for sponsoring NYU visiting positions in 2008-2011, VU University Amsterdam for a VU talent grant, and NWO for a VIDI grant.

High Frequency Trading and The *New-Market* Makers

Abstract

This paper links the recent fragmentation in equity trading to high-frequency traders (HFTs). It shows how the success of a new market, Chi-X, critically depended on the participation of a large HFT who acts as a modern market-maker. The HFT, in turn, benefits from low fees in the entrant market, but also uses the incumbent market Euronext to offload nonzero positions. It trades, on average, 1397 times per stock per day in Dutch index stocks. The gross profit per trade is €0.88 which is the result of a €1.55 profit on the spread net of fees and a €0.68 ‘positioning’ loss. This loss decomposes into a €0.45 profit on positions of less than five seconds, but a loss of €1.13 on longer duration positions. The realized maximum capital committed due to margin requirements is €2.052 million per stock which implies an annualized Sharpe ratio of 9.35.

JEL code: G10

keywords: high-frequency trading, market maker, multiple markets

(for internet appendix click http://dl.dropbox.com/u/1468608/hftatnmm_ia.pdf)

Technology has pushed securities trading into a new era where automated markets compete for order flow and high-frequency traders (HFTs) have become the most active participants. The Securities and Exchange Commission (SEC) recently published a policy document that refers to HFT as “one of the most significant market structure developments in recent years. . . it typically is used to refer to professional traders acting in a proprietary capacity that engage in strategies that generate a large number of trades on a daily basis (SEC (2010, p.45)).” The document further notes that U.S. equity volume has become “dispersed among many automated trading centers that compete for order flow in the same stocks (p.4).” It illustrates the NYSE as an example as it traded 79.1% of its listed stocks in 2005 and only 25.1% in 2009.¹

This paper links these two trends, HFT and increased fragmentation, through an analysis of Chi-X, a market that entered European equity trading in April 2007. It started trading Dutch and German equities and gradually added other countries’ stocks to grow to a 15.7% volume share in 2010, behind the London Stock Exchange (21.3%) but ahead of NYSE-Euronext (15.6%) and Deutsche Boerse (12.6%) (see Federation of European Securities Exchanges, FESE (2010)). Chi-X uses the same system architecture as NASDAQ’s main electronic market.

The paper’s main contribution is to show how one particular type of HFT, modern market making, appears to be a necessary condition for the success of a new venue and vice versa. An analysis of Dutch index stocks shows that Chi-X only generated 1-2% of all trades initially, but it jumped to a double-digit share in about a month’s time after the entry of a new trader ID in both Chi-X and Euronext (see Figure 2). Moreover, the participation rate of this trader closely mirrors Chi-X market share, both in magnitude and through time. Also, the trader’s arrival coincides with a 50% drop in the bid-ask spread which is measured as a generalized inside spread that controls for less depth on entrant market quotes (a precise definition is

¹The Committee of European Securities Regulators (CESR) published a similar discussion document where HFT is a focus area; it solicits research as the HFT chapter’s title is: “The impact of High Frequency Trading (HFT) on market structure is still not fully understood (CESR (2010, p.31)).”

in the figure's caption). This trader is the focus of this study and will be referred to as HFT as it fits the HFT definition: it is lightning fast with a latency (inter-message time) upper bound of 1.67 millisecond, it only engages in proprietary trading, it generates many trades (it is in 14.4% of all trades split almost evenly across both markets), and it starts and ends most trading days with a zero net position. I further note that this trader ID does not seem to be a strictly local player. In my 2007-2008 sample period, it participates in 47.7% of all Chi-X trades (43.7 out of 99.2 million trades where Dutch stock represent about 10% of all Chi-X trades).

A detailed analysis of the identified high-frequency trader shows that it acts as a modern multi-venue market maker. The operation uses capital to produce liquidity. The production side is analyzed by a standard decomposition of trade revenue into a bid-ask spread earned (or paid if the order consumed liquidity) and a 'positioning' revenue based on midquote changes in the life of the nonzero position. The variable costs due to exchange and clearing fees are then subtracted to arrive at a 'gross profit'; that is, the profit does not account for the (unknown) fixed costs of, for example, development of the algorithm, acquisition of hardware, exchange and clearing house membership fees. The 'capital-intensity' is explored by calculating the capital tied up in the operation through margin calls in both markets' clearing houses. The HFT cannot net positions across these clearing houses which is shown to increase capital requirements by a factor of 100.² Combining the daily gross profits with the maximum capital draw-down yields an annualized (gross) Sharpe-ratio of 9.35. Finally, the HFT characterization as a modern market maker is detectable in the overall price process as midquotes are pressured downwards if the HFT is on a long position, upwards if it is on a short position.

It is important to view the surprisingly strong results — high HFT profitability and the correlation between

²See [Duffie and Zhu \(2011\)](#) for an analysis of netting efficiency and counterparty risk for single vs. multiple central clearing counterparties (CCPs).

HFT position and price pressure — in the context of new market entry. A natural way to interpret the results is that the documented profitability is a return for entrepreneurial risk. If there is a large fixed cost of building the technology to connect to both the incumbent and an entrant market and to design the optimal algorithm to run the operation, then one should expect temporary ‘excess returns’ to make it a positive net present value (NPV) project. In addition, there was considerable risk as, at the time, it was not at all clear that Chi-X would be a successful entry given that at least one earlier initiative had failed, i.e., the London Stock Exchange introduced EuroSETS to compete with the NYSE-Euronext system in 2004 (see [Foucault and Menkveld \(2008\)](#)). If the HFT cannot fully diversify such risk, it leads to an additional required return. The market entry context also explains the HFT’s large presence in the market (14.4% of all trades) and its position’s effect on prices. Subsequent entry of other HFT firms should dilute the identified HFT’s relative weight and bring its rents down to a competitive level. The sample is too short to identify any such effect.

The paper’s focus is on one type of HFT that can be characterized as a modern market maker. It is part of a rapidly growing literature that encompasses various types of HFT. [Brogaard \(2010\)](#) studies trading of 26 NASDAQ-labeled HFT firms in the NASDAQ market in 2008-2010. In the aggregate, he concludes that HFT tends to improve market quality. [Kirilenko, Kyle, Samadi, and Tuzun \(2010\)](#) study the behavior of high-frequency traders in the E-mini S&P 500 stock index futures on May 6, the day of the flash crash. [Jovanovic and Menkveld \(2010\)](#) model HFT as middlemen in limit order markets and study their effect on trader welfare. The main idea is that, other than human intermediaries, machines have the ability to process vast amounts of (public) information almost instantaneously. This paper’s main contribution to the literature is its focus on cross-market activity, market structure development, and its detailed analysis of the operation’s production capital, i.e., the capital tied up through margin requirements. This latter part is particularly important in view of the literature on the link between funding and market liquidity (see, e.g., [Gromb and Vayanos \(2002\)](#) and [Brunnermeier and Pedersen \(2009\)](#)).

The paper also fits into a broader literature on algorithmic or automated trading. [Foucault and Menkveld \(2008\)](#) study smart routers that investors use to benefit from liquidity supply in multiple markets. [Hendershott, Jones, and Menkveld \(2011\)](#) show that algorithmic trading (AT) causally improves liquidity and makes quotes more informative. [Chaboud, Chiquoine, Hjalmarsson, and Vega \(2009\)](#) relate AT to volatility and find little relation. [Hendershott and Riordan \(2009\)](#) find that both AT demanding liquidity and AT supplying liquidity makes prices more efficient. [Hasbrouck and Saar \(2010\)](#) study low-latency trading or “market activity in the millisecond environment” in NASDAQ’s electronic limit order book ‘Single Book’ in 2007 and 2008 and find that increased low-latency trading is associated with improved market quality.

The remainder of the paper is structured as follows. Section 1 reviews the institutional background and presents the data. Section 2 discusses the methodology. Section 3 presents the empirical results. Section 4 presents the conclusions.

1 Institutional background

The classic market-making literature views the bid-ask spread, a substantial part of investors’ transaction cost, as a compensation for the cost a market maker incurs. It is comprised of essentially three components: (i) order-handling cost (e.g., the fee an exchange charges to process an order), (ii) the cost of being adversely selected on a bid or ask quote, and (iii) the premium risk-averse market makers require for price risk on nonzero positions (see, e.g., [Madhavan \(2000\)](#) for a literature review). Market makers prefer to operate in a system where these costs are low (e.g., low fees or fast access, see discussion below). In human-intermediated markets, however, it is hard for new venues to compete as the high search cost for humans creates a participation externality: traders prefer to be where the other traders are (see [Pagano \(1989\)](#)). A trader risks missing a trade opportunity when checking off-exchange prices (by phone, for example).

Technology has dramatically changed the nature of competition among venues. Participation externalities are severely reduced when markets change from humans on floors to machines on electronic markets where search costs are significantly reduced.³ This creates more scope for new markets to compete through, for example, lower fees that reduce the order-processing costs. But they can also be competitive on the other two cost components of market-making. A fast matching engine enables a market-making strategy to quickly update quotes on the arrival of public information and thus reduce the risk of being adversely selected. And, a fast response time (referred to as ‘low latency’) makes the venue available to a market-making machine that wants to ‘quickly poll’ for investor interest when trying to offload a costly nonzero inventory position.

The increased ability of venues to compete could explain the proliferation of venues both in the U.S. and in Europe. To the extent that such competition reduces the bid-ask spread and therefore transaction cost, this creates value for investors (on the assumption that they are net liquidity demanders). In addition, new venues might tailor to the needs of different types of investors and produce additional investor utility this way (see, e.g., [Stoll \(2001\)](#)). Such venue heterogeneity however is beyond the scope of this study; the focus is on competition between largely similar systems and the role of a market-making HFT in the viability of new systems. I turn to Europe to analyze the successful entry of a new venue: Chi-X.

Start of Chi-X. The European Union aimed to create a level playing field in investment services when it introduced the *Markets in Financial Instruments Directive* (MIFID) on November 1, 2007. In effect, it enabled the various national exchanges to compete and encouraged new markets to enter.

Instinet pre-empted MIFID when it launched Chi-X on April 16, 2007. Chi-X operates a trading platform which initially only traded Dutch and German index stocks.⁴ By the end of the year, it allowed a consortium

³The search cost has not disappeared entirely; [Hasbrouck and Saar \(2010\)](#) compare the 2-3 milliseconds it takes for participants to respond within a market to the distance between New York and Chicago (both important market centers) which is about 8 milliseconds at the speed of light.

⁴“Chi-X Successfully Begins Full Equity Trading, Clearing and Settlement,” Chi-X press release, April 16, 2007.

of the world's largest brokers to participate in equity through minority stakes.⁵ Before Chi-X, Instinet had operated 'Island' successfully in the U.S. which distinguished itself through subsidization of passive orders (see fee discussion below) and system speed. It is claimed to be one of the fastest platforms in the industry with a system response time, referred to as 'latency', of two milliseconds. At the time of its first anniversary Chi-X claimed it was "up to 10 times faster than the fastest European primary exchange."⁶ In 2005 Instinet sold the U.S. license to NASDAQ. It did keep the international license which led to Chi-X.

In its first 14 months, my sample period, Chi-X traded Belgian, British, Dutch, French, German, and Swiss index stocks. It had captured 4.7% of all trades and was particularly successful in Dutch stocks with a share of 13.6%. In terms of volume, Chi-X overall market share was 3.1% and its Dutch share was 8.4%. Chi-X appears to have used the Dutch index stocks, my sample stocks, to launch what ultimately became a pan-European operation.

Prior to Chi-X entry, Euronext was by far the main trading venue for Dutch stocks. It operated an electronic matching engine similar to Chi-X and competition focused on fees and speed (see discussion below). Dutch stocks also traded as ADRs in the U.S. and in the German Xetra system. They did not yet trade in NASDAQ OMX, Turquoise, or BATS-Europe which entered later on a business model similar to Chi-X: low fees and fast systems.

The broker identified as HFT in this study was a substantial participant in Chi-X. In my sample period, it participated in 43.7 million out of 99.2 million Chi-X trades. It was particularly active in Dutch stocks with participation in 4.9 million out of 8.6 million Chi-X trades.

Matching engine and exchange fee structure. The matching engines of Chi-X and Euronext both run

⁵These brokers were: BNP Paribas, Citadel, Citi, Credit Suisse, Fortis, Getco, Goldman Sachs, Lehman Brothers, Merrill Lynch, Morgan Stanley, Optiver, Société Générale and UBS (op. cit. footnote 6).

⁶"Chi-X Europe Celebrates First Anniversary," Chi-X press release, April 7, 2008.

an electronic limit order book. Investors submit a limit order to the system which summarizes their trade interest. An example of a limit is a buy order for 2000 shares with a price 'limit' of €10. The order is (i) either matched with a standing sell order with a limit price of (weakly) less than €10 in which it executes immediately at the limit sell price or (ii) it is added to the stack of limit buy orders on the buy side of the book. If executed immediately, it is labeled an aggressive order that consumes liquidity. If not, it is a passive order that supplies liquidity. Also, the matching of orders is such that standing orders are ranked by price-time priority. In the example, the standing sell orders with the lowest price get executed first and, among these, the ones that arrived earliest take priority. For a detailed description of generic limit-order markets I refer to [Biais, Hillion, and Spatt \(1995\)](#).

The exchange fees differ significantly across markets both in structure and in level. Euronext charges a fixed fee of €1.20 per trade which for an average size trade (~€25,000) amounts to 0.48 basis point. Highly active brokers benefit from volume discounts which can bring the fixed fee down to €0.60 per trade, which is the fee used for the HFT in subsequent analyses. In addition, Euronext charges a variable fee of 0.05 basis point. The act of submitting an order or cancelling it is not charged (i.e., only executions get charged) unless, on a daily basis, the cancellation-to-trade ratio exceeds 5. In this case, all orders above the threshold get charged a €0.10 fee (~0.04 basis point).

Chi-X conditions on the incoming order's type when charging its fee in what is called a maker-taker model: an aggressive order gets charged 0.30 basis point whereas a passive order receives a rebate of 0.20 basis point in case it leads to an execution. The platform is therefore guaranteed a 0.10 basis point revenue per transaction; the optimality of such fee structure is discussed in [Foucault, Kadan, and Kandel \(2010\)](#). Chi-X does not charge for limit order submissions and cancellations.

Post-trade cost: clearing fee and margin requirement. A trade is not done once two limit orders

have been matched. The actual transfer of the security and the payment is effectuated three days after the transaction; the trade is cleared and settled. This process leads to two types of cost: clearing fees and margin requirements. Chi-X and Euronext also compete at this end of the trading process as they use different clearing houses: EMCF and LCH-Clearnet, respectively. EMCF started as a clearing house at the same time Chi-X entered as a new trading venue: April 16, 2007.

The entry of EMCF triggered a clearing fee war with incumbent clearer LCH-Clearnet. EMCF started by charging €0.30 per trade, 36% less than the LCH-Clearnet fee: €0.47 per trade. Note that these fee levels are substantial as they compare to, for example, the per-trade fee of €1.20 charged by Euronext (see discussion above). On October 1, 2007, six months after EMCF entry, LCH-Clearnet responded by reducing its fee by 34% to €0.31 per trade. At the same time, EMCF reduced its fee to €0.28 per trade. Half a year later, on April 1, 2008, EMCF reduced its fee by 32% to €0.19 per trade; at the same time, LCH-Clearnet reduced its fee by 26% to €0.23. Overall, clearing fees were reduced by 50-60% in the first year after Chi-X entry.

LCH-Clearnet and EMCF are both central counterparty (CCP) clearing houses. They become the counterparty to every trade which removes the risk for each participant that the counterparty to a trade becomes insolvent before the actual transfer is due three days after limit orders were matched. The clearing house itself absorbs such risk and manages it by requiring participants to keep margin accounts with the clearer. The capital held in these accounts is confiscated should the participant become insolvent. The margin requirement is therefore linked to the value owed in as of yet uncleared transactions. LCH-Clearnet uses the SPAN methodology developed by the Chicago Mercantile Exchange which charges for 'specific risk' and 'general market risk'. At the start of the sample, LCH-Clearnet charged 4.8% on the overall net position (marked-to-market) and 3.0% for each stock's position. For example, suppose a broker's yet-to-clear, marked-to-market

position is long €10 million in security XYZ and €20 million short in security ABC at a particular point in the day. In this case, its margin account needs to have at least $€4.8\%|10 - 20| + 3.0\%(|10| + |-20|) = 1.38$ million. On February 9, 2008, LCH-Clearnet increased the specific risk parameter to 6.3% and the general market risk parameter to 4.85%. EMCF on the other hand uses a proprietary system that is opaque to its clearing members. The empirical analysis therefore applies the SPAN methodology to also calculate the margins requirement on the HFT EMCF positions assuming that the schedules are competitive.

2 Data, summary statistics, and approach

Data. The main sample consists of trade and quote data on Dutch index stocks for both Chi-X and Euronext from January 1, 2007 through June 17, 2008. The quote data consist of the best bid and ask price and the depth at these quotes. The trade data contain trade price, trade size, and an anonymized broker ID for both sides of the transaction. The broker ID anonymization was done for each market separately and broker IDs can therefore not be matched across markets — say the first market uses 1,2,3 and the second one uses A,B,C. The Euronext sample also contains a flag that indicates whether the broker's transaction was proprietary (own-account) or agency (for-client). The time stamp is to the second in Euronext and to the millisecond in Chi-X. In the analysis, Chi-X data is aggregated to the second in order to create a fair comparison across markets. A similar sample for Belgian stocks, except for trader ID information, is used for benchmark purposes. The list of all stocks that are analyzed in this study is included as Appendix A. It contains security name, ISIN code, and weight in the local index.

[Figure 1 here]

Summary statistics. Pairing broker IDs systematically across markets (1-A, 1-B, ..., 2-A, 2-B, ...) yields one pair that has all the characteristics of a high-frequency trader.⁷ Figure 1 plots its net position in ‘Unilever’ at various frequencies: minutes, hours, and days. The graphs lead to several observations. First, the series seems to be stationary with mean zero. Second, it seems that the position is zero at the start and at the end of each trading day. Third, the HFT appears to be highly active as positions change frequently and reach up to €300,000 either long or short. These positions appear to last seconds, minutes, or even hours. The Euronext data show that all of its trades were proprietary (the Chi-X data unfortunately do not flag trades as either proprietary or agency). These characteristics match the SEC definition (see page 1) of HFT and this broker ID combination will therefore be referred to as HFT in the remainder of the manuscript.

[insert Figure 2 here]

Figure 2 illustrates how Chi-X entry drove bid-ask spreads down by about 50%, but only after the HFT started to participate in trading. The top graph shows that Chi-X captured a 1-2% share of all Dutch trades in the first few months. It jumped to a double-digit share in August which is exactly the time that the HFT started to trade in both Chi-X and Euronext. The middle graph illustrates that the HFT is a major part of the new market as it is present in 70-80% of Chi-X trades. The bottom graph plots the evolution of the (inside) bid-ask spread of Dutch stocks (index-weighted) benchmarked against the bid-ask spread of Belgian stocks. The latter stocks serve as a useful control sample as they also trade in the Euronext system but get ‘treated’ with Chi-X entry only one year later on April 28, 2008 (which motivates the end date of the graph). Taken together, the plots illustrate that the HFT appears to be the ‘*new-market maker*’. Finally, note that whereas Chi-X and HFT arrival seemed necessary to *reach* lower spread levels, these new equilibrium levels appear viable even in their absence; on December 24 and 31, the HFT was virtually absent in the market, Chi-X

⁷I started with the most active broker ID in Chi-X, paired all Euronext broker IDs, and immediately discovered the HFT this way. It participated in 70-80% of Chi-X trades and therefore became the focus of this study.

share dropped to almost zero, yet spread levels did not bounce back to pre-event levels. A detailed analysis of the change in liquidity supply and trade characteristics is beyond the scope of this study, but is available in [Jovanovic and Menkveld \(2010\)](#).

As the HFT appears crucial to both Chi-X take-off and liquidity improvement, the remainder of the analysis is focused on its activity. The analysis is based on its trading for 200 days starting on September 4, 2007 and ending on June 17, 2008. The sample of index stocks is cut into two equal-sized bins of small and large stocks according to their weight in the local index. The remaining analyses are done stock by stock and are summarized in tables that contain index-weighted averages across stocks. The tables also report cross-sectional dispersion by reporting the range (i.e., the lowest and the highest value in a bin) in brackets (cf. [Hasbrouck and Sofianos \(1993\)](#)).

[insert Table 1 here]

Panel A of Table 1 presents statistics to further support the claim that the trader ID combination is a *high-frequency* trader. The HFT trades on average 1397 times per stock per day and is more active in large stocks (1582 times per day) as compared to small stocks (315 times per day). It is fast as its (cross-sectional) average latency (inter-message time) is 1.67 milliseconds.⁸ It is a large market participant as this activity represents a 14.4% participation rate in trades, 15.7% in large stocks and 6.9% in small stocks. The average closing position is -29 shares with a cross-sectional range of -100 shares to +68 shares. This shows that the HFT does not build up towards a long-term position but rather aims at mean-reverting its position quickly. It seems particularly eager to avoid overnight positions as on 69.8% of all trading days it ends the day flat. The

⁸Latency has been proxied by the time between a cancellation and resubmission of a bid or ask quote (e.g., to change the price or quantity on the quote). These cancel-and-resubmit events are identified in the data as a canceled quote followed by a new order in the same direction, of the same size, and within 1000 milliseconds of the canceled order (cf. the definition to construct 'strategic runs' in [Hasbrouck and Saar \(2010\)](#)). Strictly speaking, the latency estimate is an upperbound in case the exchange platform latency is binding. The latency is calculated based on all runs that ultimately led to a trade in January 2008, the median month of the sample.

average standard deviation of its closing position is 1120 shares which is small judged against how actively it trades. Interestingly, the HFT seems to be particularly eager to avoid overnight positions for small stocks as it ends flat 91.0% of the days with an end-of-day position standard deviation of only 120 shares.

Panel B presents some more general trading statistics on the Dutch index stocks to show that these are highly liquid securities. The average stock in the sample trades 15,800 times a day in the incumbent market and 2,200 times a day in the entrant market. Average trade size is €29,000 in the incumbent market and €15,500 in the entrant market. Half the bid-ask spread is, on average, €0.008 which is 1.7 basis points.

Approach. The paper’s main objective is to fully understand the HFT strategy that led to the success of Chi-X and to the lower bid-ask spreads. It takes a micro-economics perspective by separately analyzing the revenue or gross profit⁹ generated by the HFT operation and the capital that is required for it. The profit and capital are then combined to arrive at a standard profitability measure: the (gross) Sharpe ratio.

Inspired by Sofianos (1995), the high-frequency trader’s revenue is decomposed into a spread component and a positioning component. Let n_t^a cumulate HFT aggressive trades (in shares) through time t and let n_t^p cumulate its passive trades. If started off on a zero position, this implies that the HFT net position at time t is $n_t := n_t^a + n_t^p$.

HFT average trading profit over T time units is simply the average net cash flow (assuming it starts and ends at a zero position¹⁰):

$$\bar{\pi}^* = \frac{1}{T} \sum_{t=1}^T -\Delta n_t P_t \quad (1)$$

⁹This revenue or gross profit does not account for the (unknown) fixed costs of, for example, development of the algorithm, acquisition of hardware, exchange and clearing house membership fees.

¹⁰In the data, the HFT closes a day with a zero net position 69.8% of the time (see Table 1). For the days that it does not, it is standard practice to mark-to-market its position at the start and at the end of a trading day. This introduces an additional term in the equation: $(n_T p_T - n_0 p_0)$ where p_t denotes the midquote price. These terms are added in the empirical analysis, but omitted in the main text for expositional reasons.

where $\bar{\pi}^*$ is the average gross profit per time unit (not accounting for trading fees), n_t is the (end-of-time-unit) net inventory position, P_t is the transaction price. Rewriting this sum allows for a natural decomposition of its trading profit into a spread and a ‘positioning’ profit:

$$\bar{\pi}^* = \frac{1}{T} \sum_{t=1}^T n_{t-1} \Delta p_t - |\Delta n_t^a| p_t s_t + |\Delta n_t^p| p_t s_t \quad (2)$$

where p_t is the midquote price (the average of the bid and the ask quote) and s_t is the relative effective half-spread (i.e., $|P_t - p_t|/p_t$).¹¹ In words, equation (2) reads:

profit = realized positioning profit - paid spread aggressive orders + earned spread passive orders

[insert Figure 3 here]

The HFT revenue decomposition is particularly useful to distinguish two contrasting common views on HFT: a friendly view that considers HFT as the new market makers and a hostile view that claims HFT is aggressively picking off other investors’ quotes. These views are illustrated in Figure 3. The top graph plots the simplest inventory cycle for which the HFT generates a profit: it buys one share at a particular price and sells it some time later for a strictly higher price. The subsequent two graphs illustrate two ways (the extremes) by which the HFT might have generated this profit (by providing context to the prices drawn in the top graph). In the middle graph it aggressively picked off on an ask quote that had become stale after the fundamental value jumped. It unloads the position by hitting the bid quote after the bid and ask quotes fully reflect the new fundamental value. The HFT aggressively picked off quotes; it paid the half-spread twice, but made up for it through a (speculative) positioning profit. The situation in the bottom graph is quite the

¹¹These equations hold under the assumption that the HFT engages in either one aggressive trade or one passive trade at each instant of time (each second in my data sample) which is, most likely, not true in the real world. The equation is trivially extended to deal with multiple trades per time unit at the cost of a more burdensome notation.

opposite. The HFT buys the security when an incoming limit sell is matched with the standing HFT bid quote. The fundamental value drops at the time of the incoming order which indicates that the sale might have been information-motivated. The HFT offloads the position when an incoming buy order hits its ask quote. The HFT acts as a market maker; it earned the half-spread twice but suffered a positioning loss as the first incoming market order was information-motivated. The graphs are the simplest way to illustrate how both positioning and order types can be a source of profit or loss to an HFT; they should not be read as a loss on positioning necessarily coincides with a profit on spread or vice versa. The actual decomposition will reveal how important each component is in the overall HFT profit.

In the analysis, the exchange and clearing fees are subtracted from the spread to arrive at a so-called ‘net spread’ paid or earned:

$$\bar{\pi} = \frac{1}{T} \sum_{t=1}^T n_{t-1} \Delta p_t - |\Delta n_t^a| p_t (s_t + \tau^a) + |\Delta n_t^p| p_t (s_t - \tau^p) \quad (3)$$

where $\bar{\pi}$ is the gross profit per time unit (after accounting for trading fees), τ^a denotes the sum of the (proportional) exchange fee and the clearing fee that need to be paid for an aggressive order, τ^p is the equivalent of τ^a for a passive order.

The first term captures HFT ‘positioning’ profit. It cumulates value changes associated with its net position. If it engages in speculative trading this is expected to be positive; on average, it will trade into a long position when the fundamental value is below the midquote and vice versa. The positioning profit might equally be negative if it is adversely selected on its quotes (see, e.g., [Glosten and Milgrom \(1985\)](#)). Or, a negative positioning profit might be willingly incurred to mean-revert out of a non-zero position (see, e.g., [Ho and Stoll \(1981\)](#)). [Hendershott and Menkveld \(2010\)](#) show that in an ongoing market a risk-averse intermediary trades off staying one more time unit on a nonzero position and suffer the price risk versus subsidizing

traffic to get out of the nonzero position. In the process it collects the spread but (willingly) suffers a loss on positions (due to the subsidy it hands out when skewing quotes).

The last two terms in profit equation (3) sum up her result on the ‘net spread’. An aggressive, liquidity-demanding order leads to a negative contribution whereas a passive, liquidity-supplying order might lead to a profit. Aggressive trades pay the effective spread, i.e., the distance between the transaction price and the midquote, pay the clearing fee, and pay the aggressive exchange fee. Passive trades on the other hand earn the effective spread, pay the clearing fee, and either pay an exchange fee (Euronext) or collect a rebate (Chi-X).

Long- vs. short-run positioning profits. To further understand HFT trading profit it would be useful to decompose its positioning profit according to trading horizon. Does the high-frequency trader make money on positions it turns around in a matter of seconds, but perhaps lose money on positions it gets stuck with for hours? If so, how much does each trading horizon’s result contribute to overall positioning revenue? Frequency domain analysis is the most natural way to create such decomposition. [Hasbrouck and Sofianos \(1993\)](#) were the first to propose such decomposition of trading revenue and applied it to NYSE specialist trading (see also, e.g., [Hau \(2001\)](#), [Coughenour and Harris \(2004\)](#)). The approach taken here differs slightly in (i) that it is applied only to positioning profit (after stripping out ‘net spread’ revenues¹²) instead of the full trading profit and (ii) it runs the analysis on the natural clock as opposed to the transaction clock. The positioning profit is decomposed into bins of position durations with boundaries set at 5 seconds, 1 minute, 1 hour, and 1 day. A detailed description of the methodology is in Appendix B.

¹²The net spread part of profit is instantaneous and, if desired, is therefore trivially attributed to the shortest-term bucket.

3 Empirical results

The results are presented in increasing level of detail. First, the HFT's average gross profit, the capital required, and the Sharpe ratio are presented. The gross profit is decomposed into its two main components: positioning profit and net spread. Second, the positioning profit is decomposed according to trade horizon. Third, the net spread result is analyzed by market and is split into its three components: the (gross) spread earned, the exchange fee, and the clearing fee. Finally, the HFT net position data is entered into intraday and interday price changes to understand its impact on price volatility.

3.1 Gross profit versus capital employed

[insert Table 2 here]

Panel A of Table 2 shows that HFT gross profit per trade is €0.88 which is the result of a net spread of €1.55 and a positioning loss of €0.68. It earns an average €1416 per stock per day. The positioning loss is consistent across all stocks as the cross-sectional range is €-1.79 to €-0.07. It is robust evidence against the hostile view of HFT that is based on speculation where the HFT creates an adverse-selection cost for other market participants. In this case, the HFT appears to suffer a consistent positioning loss. The net spread result is consistently positive in the cross-section of stocks: it ranges from €0.25 to €2.15 per trade.

HFT large-stock trades are roughly five times more profitable than its small-stock trades: €0.99 per trade vs. €0.19 per trade, respectively. This differential appears to be entirely driven by a larger net spread earned on large-stock trades as the positioning profit per trade is roughly equal across both size categories. Part of the net spread differential might be attributed to trade size as large stock trades are roughly twice the size of small stock trades (see Table 1). It might also be explained by the larger relative tick size for large stocks

($\sim\text{€}0.01/\text{€}22.97=4.35$ basis point) as compared to small stocks ($\sim\text{€}0.01/\text{€}37.39=2.67$ basis point) (see Table 1).

Panel B of Table 2 presents statistics on the capital employed in the operation due to margin requirements. On average the margin for specific risk is $\text{€}272,000$ which is an order of magnitude larger than the average margin for general market risk: $\text{€}84,000$.¹³ These margins are five times larger for large stocks as compared to small stocks due to the larger positions the HFT takes in these stocks. The sample period maximum capital employed is a more relevant measure as it is (a lower bound on) the amount of capital the HFT firm needs to make available for the operation. It is roughly five times as high as the average margin required ($\text{€}1,645,000$ for specific market risk and $\text{€}407,000$ for general market risk) which is perhaps surprisingly low given that this maximum is taken over *every second* in the 200 day sample period. This indicates that the HFT is particularly skillful in keeping its position in check.

The general market risk reflects the extent to which the HFT takes on large positions (either long or short) in the entire market which potentially constitutes a systemic risk. If its individual stock positions are perfectly correlated across securities, the HFT takes disproportionate positions in the market. If, on the other hand, they are perfectly negatively correlated, it is a market-neutral operation and less cause for concern. To gauge the size of systemic risk, the actual general market risk (based on the size of the aggregate position) is compared against a zero-correlation benchmark. The values are not very different, $\text{€}84,000$ vs. $\text{€}75,000$ respectively, which indicates that the HFT is not taking on disproportionate market positions. It also shows that the HFT does not actively manage cross-security positions to remain market neutral.¹⁴ In fact, it seems to be running the robots stock by stock.

¹³The general market risk is calculated based on the HFT aggregate position (see Section 1). I choose to allocate this aggregate (across stocks) amount to individual stocks based on a stock's weight in the index.

¹⁴One reason it might not do so is that it could hedge its market exposure in the liquid index futures market.

[insert Figure 4 here]

The high-frequency trader has to keep margin in both clearing systems which is a source of a substantial inefficiency. Figure 4 illustrates the size of the inefficiency by plotting its position in the ‘Unilever’ stock in each market separately and aggregated across markets. The graph shows that the individual market positions are larger in size and mean-revert more slowly than the overall net position. In fact, the process for an individual market position will have an MA(3) character if it is not actively managed. The reason is that trades get settled (and thus disappear from a position) on the third day after a transaction. The aggregate market position shown in the bottom graph (and in Figure 1) clearly is actively managed as mean-reversion to zero is fast and positions are relatively small. Panel B of Table 2 presents the hypothetical margin requirement if the HFT were allowed to net its position across the two clearing houses. The average capital margin is approximately a factor 100 lower for both specific risk and general market risk; the maximum margin is a factor 30-40 lower. Again, comparison with the zero-correlation benchmark shows that the HFT does not take on disproportionate positions in the aggregate market.

Panel C of Table 2 presents Sharpe ratios based on HFT gross profit. The ratio needs the daily excess return which is based on (i) the sample maximum capital employed which is an indication of the standby capital required in the operation, (ii) the daily profit where a nonzero end-of-day position is marked-to-market based on the latest midquote in the day, and (iii) the capital’s required return is set equal to the riskfree rate which is downloaded from Kenneth French’ website (i.e., it is assumed that the HFT can fully diversify its (entrepreneurial) risk). The average daily excess return is 5.86 basis points, its average standard deviation is 9.69 basis points, and the average annualized Sharpe ratio is 9.35.¹⁵ The Sharpe ratio is substantially higher

¹⁵ The annualized Sharpe ratio is calculated as $(\sqrt{250} \text{ days} * \mu(r_t) / \sigma(r_t))$ where $\mu(\cdot)$ and $\sigma(\cdot)$ denote the mean and standard deviation and $r_t := (\pi_t - r_{f_t} * \text{max_cap})$ where π_t is the daily gross profit as in equation (3) with start- and end-of-day positions marked to market, r_{f_t} is the monthly riskfree rate, and max_cap is the full-sample stock-specific maximum capital employed. In the analysis capital is charged the riskfree rate. If, for sake of comparison, it is charged the riskfree rate plus a market risk premium of 6% (effectively assuming a beta of one) the average Sharpe ratio drops to 5.88.

for large stocks as compared to small stocks: 10.77 vs. 1.02 respectively.

3.2 Gross profit: the positioning profit component

[insert Table 3 here]

Table 3 zooms in on the positioning loss reported in Table 2 by decomposing it according to trading horizon. Panel A reports that the overall loss of €-0.68 is composed of a €0.45 profit on ultra-high frequency positions which last less than five seconds, but losses on almost all lower frequency bins. The ultra-high frequency profit is a robust result as the cross-sectional range, €0.15 to €0.59, is entirely in the positive domain. The five seconds to a minute bin shows mixed results with negative and positive values for both small and large stocks. The frequency bins with durations longer than a minute are virtually all negative as ranges are consistently in the negative domain except for large stocks in the lowest duration bin where the range is (-0.13, 0.02).

Panel B shows that the largest loss of €0.67 for durations between one minute and an hour are due to these cycles carrying most weight. As a matter of fact, $7.862/13.833=56.8\%$ of the unconditional variance of HFT net position falls into this duration bin. It also illustrates that the ultra-high frequency bin is disproportionately profitable as only 0.3% of position variance is in this ultra-frequency duration bin.

3.3 Gross profit: the net spread component

[insert Table 4 here]

Table 4 dissects the net spread profit to study its various components. Panel A presents the bottom-line net spread earned (also reported in Table 2). This panel also shows that the HFT is equally active in both

markets: 50.8% of its trades are generated in Chi-X, the remaining 49.2% are generated in Euronext.

Panels A and B of Table 4 present a net spread decomposition for Euronext and Chi-X respectively. Comparing across these two panels leads to a couple of observations. First, in both markets the vast majority of HFT trades are passive: 78.1% in Euronext and 78.0% in Chi-X. The gross spread earned on these passive trades is of similar magnitude: €2.09 in Euronext and €2.52 in Chi-X. This translates into a substantially higher *net* spread in Chi-X relative to Euronext primarily due to the strong fee differential for passive orders: in Chi-X these orders earn a rebate on execution whereas in Euronext they get charged; the average exchange fee is €-0.31 in Chi-X and €0.68 in Euronext. This difference of almost one euro is large relative to the amount earned in the gross spread. Fees turn out to be of first-order importance for the profitability of the HFT. The clearing fee is also substantially lower: €0.17 in Chi-X vs. €0.30 in Euronext. The result is that the net spread result is substantially different across markets: €2.52 in Chi-X vs. €1.11 in Euronext.

In terms of aggressive orders, the most salient difference across markets is that the Euronext gross spread is €-1.26 whereas, surprisingly, the Chi-X gross spread is positive and large: €3.21. The unusual finding of a positive spread for aggressive orders is an artefact of the accounting that takes the midquote in the incumbent market as a reference price. This positive result therefore reveals that the average HFT aggressive order hits a ‘stale’ Chi-X quote if the incumbent midquote has moved past it. For example, the HFT aggressively buys and consumes a Chi-X ask quote of €30.00 if the Euronext midquote is €30.01 which generates a positive gross spread result.

3.4 HFT net position, permanent price change, and price pressure

The results thus far suggest that the HFT is predominantly a market maker: on average, it earns the spread as most of its trades are passive and it suffers losses on its net positions. This section studies whether the HFT

position, given that it is a large intermediary (it participates in 14.4% of all trades, see Table 1), correlates with price change. The microstructure literature suggests it might do so in essentially two ways (see, e.g., [Madhavan \(2000\)](#)): adverse selection and price pressure.

A market maker's quote might be hit by an informed trader in which case the market maker loses money; the intermediary is adversely selected. This implies that the permanent price change (i.e., information) is negatively correlated with the market maker's position change. For example, an information-motivated market sell order causes investors' to rationally infer a negative permanent price change and at the same time makes the market maker's position increase.

The market maker's position also correlates with transitory price changes as its solution to the inventory position control problem is to skew quotes relative to fundamental value, i.e., apply price pressure. A risk-averse market maker who is long relative to its optimal position adjusts its quotes downwards in order to trade out of its position: a lowered ask increases the chance of someone buying and a lowered bid reduces the chance of someone selling. In essence, it subsidizes traffic in order to mean-revert its position where the size of the subsidy is determined by trading off the size of the subsidy against the loss of absorbing price risk on a nonoptimal position (see [Hendershott and Menkveld \(2010\)](#) for a closed-form solution to the stylized control problem).

The interaction of the HFT net position, permanent price change, and price pressure is most naturally captured by estimating a state-space model as proposed in [Hendershott and Menkveld \(2010\)](#). The model is implemented at an intraday frequency, yet recognizes a continuous round-the-clock price process (cf. [Menkveld, Koopman, and Lucas \(2007\)](#)). The idea is simple. The *unobserved* 'fundamental' or 'efficient'

price is characterized by a martingale:¹⁶

$$m_{t,\tau} = m_{t,\tau-1} + \kappa_\tau \tilde{n}_{t,\tau} + \eta_{t,\tau} \quad (\text{where } \tilde{n}_{t,\tau} := n_{t,\tau} - E_{t,\tau-1}(n_{t,\tau})) \quad (4)$$

where (t, τ) indexes time, t runs over days and τ runs over intraday time points (9:00, 9:05, 9:10, ...), m is the efficient price, n is the HFT net position, \tilde{n} is the residual of an AR(2) model applied to n (standard selection criteria indicate that AR(2) is the appropriate model), and η is a normally distributed error term. The efficient price equation uses the residual \tilde{n} rather than n to ensure that m retains its martingale property. It also makes economic sense as a Bayesian update is based on the ‘surprise’ change, leaving out the forecasted change. The five-minute frequency is selected as most of the probability mass in the frequency domain analysis is on minute cycles as opposed to second, hour, or daily cycles (see Panel B of Table 3). For ease of exposition, let the time index $(t, \tau = -1)$ be equal to $(t-1, \tau_{max})$ where τ_{max} is the latest time point in the day. The system therefore runs around the clock and includes overnight price changes. The adverse selection argument predicts that the parameter κ_τ is negative.

A transitory deviation from the efficient price, the *unobserved* ‘pricing error’ (cf. Hasbrouck (2007, p.70)) is modeled as:

$$s_{t,\tau} = \alpha_\tau n_{t,\tau} + \varepsilon_{t,\tau} \quad (5)$$

where s is the (stationary) pricing error and ε is a normally distributed error term that is independent of η .

The $\alpha_\tau n$ part of the pricing error reflects the price pressure exercised by the HFT to revert out of a nonzero position; the argument predicts α_τ to be negative.

¹⁶The intercept is set to zero as the model samples at a five-minute frequency for one year of data. Merton (1980) shows that estimators of second moments (variance, covariances) are helped by frequent sampling, not estimators of first moments (mean). Hasbrouck (2007, p.27) illustrates the trade-off between estimator bias associated with setting the ‘high-frequency intercept’ to zero against the estimator error of setting it equal to the sample mean. He considers it preferable to set it to zero for a one-year sample.

Finally, the *observed* price is modeled as the sum of the efficient price and the pricing error:

$$p_{t,\tau} = m_{t,\tau} + s_{t,\tau} \quad (6)$$

The equations (4), (5), and (6) make up a standard state-space model that is estimated with maximum likelihood using the Kalman filter (see [Durbin and Koopman \(2001\)](#)). The parameterization recognizes potential time-of-day effects by making all parameters depend on τ (including the error terms' variance). To keep the estimation feasible, these parameters are pooled into four intraday time intervals: (open) 9:00-12:00, 12:00-15:00, 15:00-17:30 (close), 17:30-9:00(+1).

[insert Table 5 here]

The estimation results in Table 5 largely support the market-making character of the HFT operation. Panel A reveals that the size of HFT net position increases in the course of the day; its standard deviation is €57,500 in the morning, €71,300 by midday, and €84,400 in the afternoon. The first-order autoregressive coefficient reveals that roughly half of a shock to the HFT position disappears in five minutes which indicates that the five-minute frequency seems appropriate. The AR(1) coefficient is 0.48 in the morning, 0.51 by midday, and 0.40 in the afternoon. The quicker reversals in the afternoon are consistent with the HFT aim to end the day 'flat' (see Table 1); it is too costly to carry an overnight position with its associated price risk. The overnight AR(1), the HFT opening position (after the opening auction) regressed on its position at the previous day close, is 0.02 which is further evidence that the HFT manages its position within the day in such a way so as to avoid a nonzero overnight position.

Panel B reveals that the HFT incurs adverse selection cost most of the day except for when it acquires a position at the market open. The adverse selection parameter, κ_{τ} , is consistently negative for all stocks in

the morning, midday, and afternoon (i.e., the range intervals are strictly in the negative domain). Its intraday pattern is largely monotonic as κ_τ is -0.045, -0.031, and -0.027 basis points per €1000 (surprise) position change, respectively. This might explain why the size of the HFT position gradually rises during the day (cf. Panel A); the HFT increases its activity after the most intense price discovery in the opening hours is over. The HFT is best equipped to intermediate at times when the market is back to ‘normal’ so that it can rely on publicly available ‘hard’ information (e.g., quotes in same-industry stocks, index futures, FX, etc.) to refresh its quotes and thereby minimize adverse selection risk. This argument is developed in detail in [Jovanovic and Menkveld \(2010\)](#). Interestingly, the adverse selection parameter κ_τ is positive for the overnight innovation: 0.027. The high-frequency trader seems to participate in the opening auction when there is an informational opportunity. The size of this ‘favorable selection’ benefit, however, is small relative to the size of adverse selection cost in the remainder of the trading day. This size is measured by the variance of adverse selection ($\kappa_\tau^2 \sigma_\tau^2(\tilde{n})$) which is 0.3 squared basis points per hour for the close-to-open period vs. 37.9, 28.5, and 28.7 squared basis points for the morning, midday, and afternoon sessions, respectively. These sizes in turn are small relative to the size of the permanent price changes in each of these periods: 700, 10,378, 6837, and 7702 squared basis points, respectively.

Panel C shows that the HFT net position generates (transitory) price pressure. The average estimate of the conditional price pressure, α_τ , is negative for all time intervals: -0.059, -0.038, and -0.048 basis points per €1000 position for the morning, midday, and afternoon, respectively. In a stylized market maker model, the size of the conditional price pressure is the optimal control variable that trades off the cost of staying on the position one more time unit (lower pressure) and a higher revenue loss to mean-revert a suboptimal position more quickly (higher pressure) (see [Hendershott and Menkveld \(2010\)](#)). This insight could explain the relatively high pressure in the morning as adverse selection cost is larger, the high pressure in the afternoon as the risk of staying in a nonzero position overnight increases when the market close draws near, and the

low pressure in the midday as both these costs are lower. This negative α_τ result appears robust for all time intervals except for the afternoon session; the range of α_τ estimates is strictly in the negative domain for the morning and the midday, but not for the afternoon. The effect of price pressure is economically significant. Its variance ($\alpha_\tau^2 \sigma_\tau^2(n)$) is 6.7, 4.3, and 94.9 squared basis points for the morning, midday, and afternoon, respectively. To put them into perspective, these numbers are larger than the squared average half spread ($1.7^2=2.89$ squared basis points, see Table 1). Moreover, price pressure variance is more than half of total pricing error variance.

Taken together, these state space results indicate that the HFT does act as a modern market maker and its positions explain most of the transitory pricing error in five-minute midquote returns. The HFT's weight in the market appears large or representative enough to generate visible overall price patterns. The HFT's large effect on pricing error is most likely due to it undercutting the best available rival quote on the side of market that mean-reverts its position; if it is long it will undercut the lowest ask in the market by one tick to increase the probability to get out of the position. The HFT skews the midquote downwards this way. This is, unfortunately, not testable with the data at hand as HFT quotes are not observed (only its trades). The pricing error effect might, however, be entirely driven by its quotes on Chi-X where it receives a rebate if executed. To judge the extent to which this is the case, state space estimation is repeated with the incumbent market midquote as the observed price instead of the generalized inside market midquote used thus far. The results, presented in the internet appendix, appear largely unaffected. This demonstrates that the HFT position effect on prices is not just a 'Chi-X effect'. Finally, five-minute HFT position changes correlate negatively with permanent price changes which is consistent with the positioning losses reported in Table 3; the market maker is adversely selected on its quotes.

4 Conclusion

This paper benefits from proprietary Chi-X and Euronext datasets that contain anonymized broker IDs for trades in Dutch index stocks for a sample period that runs from September 4, 2007 to June 17, 2008. One particular set of broker IDs matched across markets shows the characteristics of an HFT that acts as a market maker in both the entrant market (Chi-X) and the incumbent market (Euronext). In each market, four out of five of its trades are passive, i.e., the HFT was the (liquidity-supplying) limit order in the book that got executed. Its lowest inter-message time is 1.67 millisecond. It trades actively with an average of 1397 trades per stock per day. It makes money on the spread but loses money on its positions. If this positioning loss is decomposed according to duration, one finds that positions that last less than five seconds generate a profit whereas the ones that last longer generally lose money. The HFT is equally active in both markets as roughly half of its trades are on Chi-X and the other half are on Euronext. Overall, it is a significant market participant as it shows up in 14.4% of all trades (aggregated across markets). It is particularly active in Chi-X where it participates in roughly every other trade.

The paper goes on to calculate the capital required for the operation. It studies the two clearing houses associated with each of the markets and retrieves their fee structure and margin requirements. The HFT cannot net its positions across markets and therefore turns out to have to put up a 100 times more capital than what would have been needed if netting were possible. This might in effect explain why the U.S. equity market is most fragmented as netting is possible in the U.S. The maximum (across my sample period) of the capital tied-up due to margining is then taken as a natural measure for the standby capital needed in the operation. This capital measure along with the daily gross profits and the riskfree rate imply an average annualized Sharpe ratio of 9.35.

Round-the-clock price changes are modeled to analyze whether the HFT is visible in the security's market

prices. The market making literature suggests that it might be in two ways: permanent price changes should correlate negatively with HFT position change (the HFT is adversely selected on its quotes) and transitory pricing errors should correlate negatively with the HFT position (the HFT skews quotes to get out of its position). The evidence is supportive. In the trading day, the HFT position generates significant (transitory) price pressure. It is an economically meaningful amount as it is, for example, larger than half the average bid-ask spread. Also, the (surprise) HFT position change correlates negatively with permanent price changes throughout the trading day, but not in the overnight period for which the sign is reversed.

Finally, the paper shows how fees are a substantial part of a high-frequency trader's profit and loss account. It is therefore not surprising that new, low-fee venues have entered the exchange market as they are attractive to these 'modern' market makers. It is shown that such lower fees are, at least partially, passed on to end-users through lower bid-ask spreads. This evidence adds to the regulatory debate on high-frequency traders and highlights that a subset is closely linked to the rapidly evolving market structure that is characterized by the entry of many new and successful trading venues.

Appendix A: List of the sample stocks

This table lists all stocks that have been analyzed in this manuscript. It reports the official ISIN code, the company's name, and the index weight which has been used throughout the study to calculate (weighted) averages.

Dutch index stocks / 'treated' sample			Belgium index stocks / 'untreated' sample		
ISIN code	security name	index weight ^a	ISIN code	security name	index weight ^a
NL0000303600	ing groep	22.3%	BE0003801181	fortis	17.5%
NL0000009470	royal dutch petrol	20.1%	BE0003565737	kbc	15.8%
NL0000009538	kon philips electr	12.1%	BE0003796134	dexia	13.5%
NL0000009355	unilever	11.3%	BE0003793107	interbrew	9.5%
NL0000303709	aegon	7.5%	BE0003470755	solvay	6.5%
NL0000009082	koninklijke kpn	7.4%	BE0003797140	gpe bruxel.lambert	5.7%
NL0000009066	tnt	4.8%	BE0003562700	delhaize group	5.4%
NL0000009132	akzo nobel	4.2%	BE0003810273	belgacom	4.4%
NL0000009165	heineken	3.0%	BE0003739530	ucb	4.1%
NL0000009827	dsm	2.4%	BE0003845626	cnp	2.9%
NL0000395903	wolters kluwer	2.2%	BE0003775898	colruyt	2.3%
NL0000360618	sbm offshore	1.2%	BE0003593044	cofinimmo	2.2%
NL0000379121	randstad	1.1%	BE0003764785	ackermans and van haaren	2.2%
NL0000387058	tomtom	0.6%	BE0003678894	befimmo-sicafi	2.1%
			BE0003826436	telenet	2.1%
			BE0003735496	mobistar	1.5%
			BE0003780948	bekaert	1.1%
			BE0003785020	omega pharma	1.0%

^a: The index weights are based on the true index weights of December 31, 2007. The weights are rescaled to sum up to 100% as only stocks are retained that were a member of the index throughout the sample period. This allows for fair comparisons through time.

Appendix B: Frequency domain decomposition of positioning profit

The first term from equation (3) was labeled the positioning profit, i.e.,

$$\bar{\pi}_{pos} = \frac{1}{T} \sum_{t=1}^T n_{t-1} \Delta p_t \quad (7)$$

where t indexes all seconds in the sample (and T therefore equals 6.12 million¹⁷ (= 200days * 8.5hours * 60mins * 60secs), Δp_t is the log midquote price change, and n_t is HFT net position.

The frequency domain decomposition of this ‘covariance’ term is standard and the discussion below is largely taken from [Hasbrouck and Sofianos \(1993\)](#). A useful reference on frequency domain analysis is [Bloomfield \(2000\)](#). The original series are relabeled x_t for n_{t-1} and y_t for Δp_t for ease of notation.

For an equally spaced time series of length T , the Fourier frequencies are given by $\omega_k = 2\pi k/T$, for $k = 0, 1, \dots, T-1$. The period (length of cycle) corresponding to a frequency $\omega > 0$ is given by $2\pi/\omega$, so the lowest positive frequency in the data corresponds to a component that cycles once over the full sample. The Fourier transform of the data is:

$$J_x(\omega_k) = \frac{1}{T} \sum_{t=1}^T x_t e^{-i\omega_k t}$$

$J_x(\omega_k)$ is the Fourier component of x_t at frequency ω_k . The data may be recovered using the inverse transform:

$$x_t = \sum_{k=0}^{T-1} J_x(\omega_k) e^{i\omega_k t}$$

This expresses x_t as the sum of the components at various frequencies. The usual estimate of the crossproduct between two series, x_t and y_t , is:

$$\hat{M} = \frac{1}{T} \sum_{t=1}^T x_t y_t$$

This estimate is computationally equal to the one formed from the Fourier transforms:

$$\hat{M} = \sum_{k=0}^{T-1} J_x(\omega_k) \overline{J_y(\omega_k)}$$

where the summation is over all Fourier frequencies and the overbar denotes complex conjugation. Written in this fashion, the contributions to the covariance from different frequencies is clearly visible. A particular subset of frequencies, such as the ones corresponding to a certain horizon, may be denoted by K , as subset of $\{k | k = 0, 1, \dots, T-1\}$.

The Fourier transforms were implemented in Python using the library function [matplotlib.mlab.csd](#). The Fourier transforms were smoothed by twice applying the Daniell filter of size 16*256 (see, e.g., [Bloomfield \(2000, p.194\)](#)).

¹⁷Note that the nontrading seconds have been removed for numerical reasons. The procedure still takes overnight positioning profit into account and mechanically adds it to the one second bucket. I believe this profit is a relatively small part of overall profit as the HFT endogenously avoids overnight positions; Table 1 shows that HFT net position is zero at the end of 69.8% of the sample days.

References

- Biais, B., P. Hillion, and C. Spatt (1995). An empirical analysis of the limit order book and the order flow in the Paris Bourse. *Journal of Finance* 50, 1655–1689.
- Bloomfield, P. (2000). *Fourier Analysis of Time Series*. New York: John Wiley.
- Brogaard, J. A. (2010). High-frequency trading and its impact on market quality. Manuscript, Kellogg School of Management.
- Brunnermeier, M. K. and L. H. Pedersen (2009). Market liquidity and funding liquidity. *Review of Financial Studies* 22, 2201–2238.
- CESR (2010). Trends, risks, and vulnerabilities in financial markets. Release No. 10-697.
- Chaboud, A., B. Chiquoine, E. Hjalmarsson, and C. Vega (2009). Rise of the machines: Algorithmic trading in the foreign exchange market. Manuscript, Federal Reserve Board.
- Coughenour, J. F. and L. E. Harris (2004). Specialist profits and the minimum price increment. Manuscript, University of Southern California.
- Duffie, D. and H. Zhu (2011). Does a central clearing counterparty reduce counterparty risk? *Review of Asset Pricing Studies* (forthcoming).
- Durbin, J. and S. J. Koopman (2001). *Time Series Analysis by State Space Models*. Oxford: Oxford University Press.
- FESE (2010). European equity market report - Year 2010. Available at www.fese.be.
- Foucault, T., O. Kadan, and E. Kandel (2010). Liquidity cycles and make/take fees in electronic markets. Manuscript, HEC Paris.
- Foucault, T. and A. J. Menkveld (2008). Competition for order flow and smart order routing systems. *Journal of Finance* 63, 119–158.
- Glosten, L. and P. Milgrom (1985). Bid, ask, and transaction prices in a specialist market with heterogeneously informed agents. *Journal of Financial Economics* 14, 71–100.
- Gromb, D. and D. Vayanos (2002). Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of Financial Economics* 66, 361–407.
- Hasbrouck, J. (2007). *Empirical Market Microstructure*. New York: Oxford University Press.
- Hasbrouck, J. and G. Saar (2010). Low-latency trading. Manuscript, Cornell University.
- Hasbrouck, J. and G. Sofianos (1993). The trades of market makers: An empirical analysis of NYSE specialists. *Journal of Finance* 48, 1565–1593.
- Hau, H. (2001). Location matters: An examination of trading profits. *Journal of Finance* 56(5), 1959–1983.
- Hendershott, T., C. M. Jones, and A. J. Menkveld (2011). Does algorithmic trading improve liquidity? *Journal of Finance* 66, 1–33.
- Hendershott, T. and A. J. Menkveld (2010). Price pressures. Manuscript, VU University Amsterdam.
- Hendershott, T. and R. Riordan (2009). Algorithmic trading and information. Manuscript, University of California, Berkeley.

- Ho, T. and H. R. Stoll (1981). Optimal dealer pricing under transaction cost and return uncertainty. *Journal of Financial Economics* 9, 47–73.
- Jovanovic, B. and A. J. Menkveld (2010). Middlemen in limit-order markets. Manuscript, VU University Amsterdam.
- Kirilenko, A., A. Kyle, M. Samadi, and T. Tuzun (2010). The flash crash: The impact of high frequency trading on an electronic market. Manuscript, U of Maryland.
- Madhavan, A. (2000). Market microstructure: A survey. *Journal of Financial Markets* 3(3), 205–258.
- Menkveld, A. J., S. J. Koopman, and A. Lucas (2007). Modelling round-the-clock price discovery for cross-listed stocks using state space methods. *Journal of Business & Economic Statistics* 25, 213–225.
- Merton, R. (1980). On estimating the expected rate of return on the market. *Journal of Financial Economics* 8, 323–362.
- Pagano, M. (1989). Trading volume and asset liquidity. *The Quarterly Journal of Economics* 104(2), 255–274.
- SEC (2010). Concept release on equity market structure. Release No. 34-61358; File No. S7-02-10.
- Sofianos, G. (1995). Specialist gross trading revenues at the New York Stock Exchange. Manuscript 95-01, New York Stock Exchange.
- Stoll, H. R. (2001). Market fragmentation. *Financial Analysts Journal* 57(4), 16–21.

Table 1: Summary statistics

This table reports summary statistics on trading in Dutch index stocks from September 4, 2007 through June 17, 2008. Panel A presents statistics on the high-frequency trader's transactions; Panel B reports aggregate trading statistics. For each stock size group, it reports the variable weighted average (where index weights are used) and, in brackets, the cross-sectional range, i.e., the group's minimum and maximum value. Latency has been proxied as the minimum time between a cancel-and-resubmit of either a bid or an ask quote. Cancel-and-resubmit events are defined as a canceled quote that is followed by a new order in the same direction, of the same size, and within 1000 milliseconds of the canceled order (cf. the definition to construct 'strategic runs' in Hasbrouck and Saar (2010)). The latency is calculated based on all runs in January 2008, the median month of the sample. Variable units are reported in brackets.

variable	large stocks	small stocks	all stocks
<i>panel A: high-frequency trader trade statistics</i>			
fraction of days with zero closing inventory (%)	66.1 (50.5, 87.5)	91.0 (86.0, 97.0)	69.8 (50.5, 97.0)
daily closing inventory (100-share lots)	-0.35 (-1.00, 0.68)	0.04 (-0.05, 0.17)	-0.29 (-1.00, 0.68)
st. dev. daily closing inventory (100-share lots)	12.9 (2.4, 38.4)	1.2 (0.8, 2.0)	11.2 (0.8, 38.4)
high-frequency trader #trades per day	1582 (344, 2458)	315 (93, 434)	1397 (93, 2458)
high-frequency trader trade participation rate (%)	15.7 (8.6, 19.0)	6.9 (2.1, 9.8)	14.4 (2.1, 19.0)
high-frequency trader latency, min time cancel-resubmit (millisec)	1.17 (1.00, 3.00)	4.56 (1.00, 9.00)	1.67 (1.00, 9.00)
<i>panel B: aggregate trade statistics</i>			
daily number of trades incumbent market, Euronext (1000)	17.0 (7.6, 23.4)	8.5 (5.3, 10.7)	15.8 (5.3, 23.4)
daily number of trades entrant market, Chi-X (1000)	2.5 (0.4, 3.9)	0.5 (0.2, 0.6)	2.2 (0.2, 3.9)
avg trade size incumbent market, Euronext (€1000)	31.4 (14.7, 41.7)	15.1 (10.6, 20.2)	29.0 (10.6, 41.7)
avg trade size entrant market, Chi-X (€1000)	16.8 (7.8, 21.9)	8.1 (6.7, 10.1)	15.5 (6.7, 21.9)
half spread (0.5*(ask-bid)) incumbent market, Euronext (€)	0.007 (0.006, 0.011)	0.016 (0.011, 0.025)	0.008 (0.006, 0.025)
relative half spread incumbent market, Euronext (basis points)	1.6 (1.3, 2.5)	2.2 (1.6, 3.6)	1.7 (1.3, 3.6)
avg transaction price (€)	22.97 (11.01, 27.24)	37.39 (19.41, 52.72)	25.08 (11.01, 52.72)

Table 2: Gross profit, capital employed, and Sharpe ratio

This table reports statistics on the high-frequency trader's gross profit (Panel A), the capital tied up in the operation due to margin requirements (Panel B), and, using both the profit and the capital results along with the riskfree rate, the implied Sharpe ratio (Panel C). For each stock size group, it reports the variable weighted average (where index weights are used) and, in brackets, the cross-sectional range, i.e., the group's minimum and maximum value.

variable	large stocks	small stocks	all stocks
<i>panel A: gross profit</i>			
gross profit per day (€)	1649 (-50, 2751)	55 (-47, 125)	1416 (-50, 2751)
gross profit per trade (€)	0.99 (-0.15, 1.62)	0.19 (-0.18, 0.78)	0.88 (-0.18, 1.62)
positioning profit per trade (€)	-0.69 (-0.90, -0.30)	-0.61 (-1.79, -0.07)	-0.68 (-1.79, -0.07)
net spread per trade (€)	1.68 (0.76, 2.15)	0.80 (0.25, 1.64)	1.55 (0.25, 2.15)
<i>panel B: capital employed</i>			
actual capital employed			
avg margin specific risk (€1000)	308 (62, 441)	62 (13, 83)	272 (13, 441)
avg margin gen market risk (€1000)	95 (18, 139)	18 (4, 25)	84 (4, 139)
avg margin gen market risk, zero corr benchmark (€1000)	85 (17, 122)	17 (4, 23)	75 (4, 122)
max margin specific risk (€1000)	1837 (485, 3206)	522 (90, 1000)	1645 (90, 3206)
max margin gen market risk (€1000)	461 (88, 674)	89 (21, 122)	407 (21, 674)
capital employed if netting across markets were allowed			
avg margin specific risk (€1000)	3.04 (0.79, 3.87)	0.57 (0.20, 0.79)	2.68 (0.20, 3.87)
avg margin gen market risk (€1000)	0.99 (0.25, 1.28)	0.18 (0.07, 0.25)	0.87 (0.07, 1.28)
avg margin gen market risk, zero corr benchmark (€1000)	0.91 (0.24, 1.15)	0.17 (0.06, 0.24)	0.80 (0.06, 1.15)
max margin specific risk (€1000)	59.36 (16.14, 83.50)	15.15 (8.18, 19.89)	52.91 (8.18, 83.50)
max margin gen market risk (€1000)	25.62 (6.35, 33.03)	4.71 (1.69, 6.47)	22.57 (1.69, 33.03)
<i>panel C: Sharpe ratio (based on maximum capital employed)</i>			
avg daily net return in excess of riskfree rate (bps)	6.80 (-1.31, 14.11)	0.32 (-2.03, 2.71)	5.86 (-2.03, 14.11)
st dev daily return (bps)	9.96 (6.60, 20.74)	8.14 (3.30, 22.15)	9.69 (3.30, 22.15)
annualized Sharpe ratio	10.77 (-3.14, 14.68)	1.02 (-3.06, 4.60)	9.35 (-3.14, 14.68)

Table 3: Positioning profit decomposition

This table decomposes the high-frequency trader's positioning profit (reported in Table 2) according to position durations using frequency domain analysis (Panel A). It further decomposes (unconditional) net position variance to establish how much 'mass' is in each duration bracket; this is naturally interpreted as a histogram of durations (Panel B). For each stock size group, it reports the variable weighted average (where index weights are used) and, in brackets, the cross-sectional range, i.e., the group's minimum and maximum value.

variable	large stocks	small stocks	all stocks
<i>panel A: positioning profit decomposition</i>			
ultra-high frequency, period \leq 5sec (€)	0.49 (0.21, 0.59)	0.24 (0.15, 0.39)	0.45 (0.15, 0.59)
high frequency, 5sec<period \leq 1min (€)	-0.30 (-0.41, 0.04)	0.02 (-0.08, 0.20)	-0.25 (-0.41, 0.20)
medium frequency, 1min<period \leq 1hour (€)	-0.67 (-0.99, -0.05)	-0.65 (-1.56, -0.17)	-0.67 (-1.56, -0.05)
low frequency, 1hour<period \leq 1day (€)	-0.18 (-0.40, -0.06)	-0.15 (-0.43, -0.08)	-0.17 (-0.43, -0.06)
ultra-low frequency, 1day<period (€)	-0.03 (-0.13, 0.02)	-0.07 (-0.17, -0.01)	-0.04 (-0.17, 0.02)
total positioning profit per trade (€)	-0.69 (-0.90, -0.30)	-0.61 (-1.79, -0.07)	-0.68 (-1.79, -0.07)
<i>panel B: net position variance decomposition</i>			
ultra-high frequency, period \leq 5sec (€)	0.042 (0.001, 0.060)	0.001 (0.000, 0.001)	0.036 (0.000, 0.060)
high frequency, 5sec<period \leq 1min (€)	0.539 (0.017, 0.761)	0.007 (0.001, 0.011)	0.461 (0.001, 0.761)
medium frequency, 1min<period \leq 1hour (€)	9.182 (0.362, 20.336)	0.145 (0.024, 0.307)	7.862 (0.024, 20.336)
low frequency, 1hour<period \leq 1day (€)	4.503 (0.247, 17.624)	0.085 (0.022, 0.271)	3.858 (0.022, 17.624)
ultra-low frequency, 1day<period (€)	1.885 (0.101, 7.284)	0.045 (0.008, 0.170)	1.616 (0.008, 7.284)
net position variance (mln shares)	16.150 (0.728, 45.866)	0.284 (0.059, 0.760)	13.833 (0.059, 45.866)

Table 4: Net spread decomposition

This table decomposes the high-frequency trader's net spread result (reported in Table 2) along three dimensions: (i) incumbent market (Euronext) or entrant market (Chi-X), (ii) passive or aggressive side of the trade (the passive side of a trade is the standing limit order in the book that is executed against an incoming (marketable) limit order; the latter order is the aggressive side), (iii) (gross) spread or fee. For each stock size group, it reports the variable weighted average (where index weights are used) and, in brackets, the cross-sectional range, i.e., the group's minimum and maximum value.

variable	large stocks	small stocks	all stocks
<i>panel A: high-frequency trader in both markets</i>			
entrant market (Chi-X) trade share (%)	49.8 (43.7, 62.8)	56.5 (51.6, 63.6)	50.8 (43.7, 63.6)
net spread per trade (€)	1.68 (0.76, 2.15)	0.80 (0.25, 1.64)	1.55 (0.25, 2.15)
<i>panel B: high-frequency trader in incumbent market (Euronext)</i>			
#trades per day	770 (216, 1189)	180 (48, 276)	684 (48, 1189)
fraction of passive trades (%)	79.5 (70.5, 82.5)	70.0 (58.7, 81.6)	78.1 (58.7, 82.5)
net spread per trade (€)	0.72 (0.09, 1.27)	-0.07 (-0.44, 1.01)	0.61 (-0.44, 1.27)
net spread per trade, passive orders (€)	1.26 (0.31, 2.03)	0.23 (0.05, 1.50)	1.11 (0.05, 2.03)
gross spread per trade, passive orders (€)	2.25 (1.25, 2.99)	1.17 (0.97, 2.44)	2.09 (0.97, 2.99)
exchange fee per trade, passive orders (€)	-0.68 (-0.71, -0.65)	-0.64 (-0.66, -0.62)	-0.68 (-0.71, -0.62)
clearing fee per trade, passive orders (€)	-0.30 (-0.32, -0.29)	-0.30 (-0.31, -0.29)	-0.30 (-0.32, -0.29)
net spread per trade, aggressive orders (€)	-1.35 (-2.21, -0.80)	-0.75 (-1.12, -0.23)	-1.26 (-2.21, -0.23)
gross spread per trade, aggressive trades (€)	-0.39 (-1.28, 0.17)	0.16 (-0.23, 0.66)	-0.31 (-1.28, 0.66)
exchange fee per trade, aggressive orders (€)	-0.67 (-0.70, -0.63)	-0.63 (-0.64, -0.62)	-0.67 (-0.70, -0.62)
clearing fee per trade, aggressive orders (€)	-0.29 (-0.30, -0.26)	-0.28 (-0.29, -0.27)	-0.29 (-0.30, -0.26)
<i>panel C: high-frequency trader in entrant market (Chi-X)</i>			
#trades per day	812 (128, 1269)	135 (45, 183)	713 (45, 1269)
fraction of passive trades (%)	77.1 (71.4, 81.8)	83.3 (79.0, 90.7)	78.0 (71.4, 90.7)
net spread per trade (€)	2.63 (1.88, 3.17)	1.92 (1.46, 3.05)	2.52 (1.46, 3.17)
net spread per trade, passive orders (€)	2.63 (1.97, 3.15)	1.87 (1.46, 3.14)	2.52 (1.46, 3.15)
gross spread per trade, passive orders (€)	2.46 (1.97, 3.05)	1.90 (1.49, 3.17)	2.38 (1.49, 3.17)
exchange fee per trade, passive orders (€)	0.34 (0.18, 0.45)	0.16 (0.11, 0.21)	0.31 (0.11, 0.45)
clearing fee per trade, passive orders (€)	-0.16 (-0.18, -0.14)	-0.19 (-0.22, -0.17)	-0.17 (-0.22, -0.14)
net spread per trade, aggressive orders (€)	2.61 (1.51, 3.36)	2.21 (1.43, 3.78)	2.55 (1.43, 3.78)
gross spread per trade, aggressive trades (€)	3.30 (1.91, 4.11)	2.65 (1.83, 4.18)	3.21 (1.83, 4.18)
exchange fee per trade, aggressive orders (€)	-0.48 (-0.61, -0.18)	-0.22 (-0.28, -0.18)	-0.45 (-0.61, -0.18)
clearing fee per trade, aggressive orders (€)	-0.21 (-0.22, -0.19)	-0.21 (-0.23, -0.20)	-0.21 (-0.23, -0.19)

Table 5: HFT net position, permanent price change, and price pressure

This table relates inter- and intraday price changes to the HFT net position to analyze whether the HFT can be characterized as a liquidity supplier. A state space model distinguishes transitory and permanent price changes and relates each of them to the HFT net position consistent with the canonical microstructure model of a supplier of liquidity (see [Hendershott and Menkveld \(2010\)](#)):

$$\begin{aligned}
 p_{t,\tau} &= m_{t,\tau} + s_{t,\tau} \\
 m_{t,\tau} &= m_{t,\tau-1} + \kappa_{\tau} \tilde{n}_{t,\tau} + \eta_{t,\tau} \quad (\text{where } \tilde{n}_{t,\tau} := n_{t,\tau} - E_{t,\tau-1}(n_{t,\tau})) \\
 s_{t,\tau} &= \alpha_{\tau} n_{t,\tau} + \varepsilon_{t,\tau}
 \end{aligned}$$

where t indexes days, τ indexes intraday time points (9:00, 9:05, 9:10), ..., p is the inside midquote price (the inside ask is the lowest ask across the entrant and incumbent market, the inside bid is defined similarly), m is a martingale that represents the unobserved ‘efficient’ price, n is the HFT net position, \tilde{n} is the (surprise) innovation in HFT net position where last period’s forecast is obtained based on an AR(2) model for the net position series n , and s represents the unobserved stationary pricing error process. For ease of exposition, let the time index ($t, \tau = -1$) be equal to ($t-1, \tau_{max}$) where τ_{max} is the latest time point in the day. The error terms η and ε are assumed to be mutually independent normally distributed variables with a variance that depends on time of day (τ). In the estimation all time-of-day dependent parameters are pooled into four intraday time intervals: (open) 9:00-12:00, 12:00-15:00, 15:00-17:30 (close), 17:30-9:00(+1). For each stock size group, the table reports the average parameter estimate and in brackets its cross-sectional range, i.e., the group’s minimum and maximum value. Parameter units are in brackets.

variable	time	large stocks	small stocks	all stocks
<i>panel A: net position</i>				
stdev net position, $\sigma(n)$ (€1000)	9:00-12:00	64.6 (22.7, 80.3)	15.4 (7.2, 20.4)	57.5 (7.2, 80.3)
	12:00-15:00	80.3 (25.1, 101.8)	19.0 (7.2, 25.5)	71.3 (7.2, 101.8)
	15:00-17:30	95.4 (24.8, 128.9)	19.9 (7.3, 26.7)	84.4 (7.3, 128.9)
	17:30-9:00 (+1) ^a	0.02 (-0.01, 0.09)	0.04 (-0.03, 0.12)	0.02 (-0.03, 0.12)
ar1 coefficient net position, n	9:00-12:00	0.46 (0.38, 0.68)	0.56 (0.48, 0.73)	0.48 (0.38, 0.73)
	12:00-15:00	0.50 (0.43, 0.69)	0.56 (0.48, 0.72)	0.51 (0.43, 0.72)
	15:00-17:30	0.39 (0.30, 0.64)	0.43 (0.35, 0.62)	0.40 (0.30, 0.64)
	17:30-9:00 (+1) ^a	0.02 (-0.01, 0.09)	0.04 (-0.03, 0.12)	0.02 (-0.03, 0.12)
<i>panel B: permanent price change (Δm)</i>				
cond adverse selection, κ (bp/€1000)	9:00-12:00	-0.036 (-0.073, -0.023)	-0.098 (-0.420, -0.049)	-0.045 (-0.420, -0.023)
	12:00-15:00	-0.024 (-0.053, -0.010)	-0.068 (-0.216, -0.029)	-0.031 (-0.216, -0.010)
	15:00-17:30	-0.019 (-0.042, -0.011)	-0.070 (-0.278, -0.028)	-0.027 (-0.278, -0.011)
	17:30-9:00 (+1)	0.014 (-0.377, 0.101)	0.102 (-0.453, 3.477)	0.027 (-0.453, 3.477)
var adv selection, $\sigma^2(\kappa\tilde{n})$ (bp ² /hr)	9:00-12:00	42.2 (8.8, 61.0)	13.1 (3.7, 54.6)	37.9 (3.7, 61.0)
	12:00-15:00	31.8 (2.5, 68.4)	9.0 (1.1, 23.0)	28.5 (1.1, 68.4)
	15:00-17:30	31.3 (6.2, 70.0)	13.3 (2.6, 74.5)	28.7 (2.6, 74.5)
	17:30-9:00 (+1)	0.3 (0.0, 0.7)	0.7 (0.0, 10.4)	0.3 (0.0, 10.4)

- continued on next page -

- continued from previous page -

variable	time	large stocks	small stocks	all stocks
var perm price change, $\sigma^2(\Delta m)$ (bp ² /hr)	9:00-12:00	10953 (5330, 16029)	7015 (3878, 29296)	10378 (3878, 29296)
	12:00-15:00	7336 (3237, 15730)	3919 (2506, 12627)	6837 (2506, 15730)
	15:00-17:30	8092 (2977, 17613)	5416 (2824, 32240)	7702 (2824, 32240)
	17:30-9:00 (+1)	690 (374, 1070)	764 (436, 3720)	700 (374, 3720)
<i>panel C: price pressure (s)</i>				
cond price pressure, α (bp/€1000)	9:00-12:00	-0.043 (-0.132, -0.017)	-0.148 (-0.540, -0.084)	-0.059 (-0.540, -0.017)
	12:00-15:00	-0.027 (-0.102, -0.007)	-0.101 (-0.431, -0.019)	-0.038 (-0.431, -0.007)
	15:00-17:30	-0.066 (-0.330, 0.108)	0.058 (-0.379, 2.092)	-0.048 (-0.379, 2.092)
var price pressure, $\sigma^2(\alpha n)$ (bp ²)	9:00-12:00	7.1 (1.3, 12.5)	4.5 (1.7, 28.3)	6.7 (1.3, 28.3)
	12:00-15:00	4.5 (0.5, 10.0)	2.9 (0.1, 17.2)	4.3 (0.1, 17.2)
	15:00-17:30	102.0 (0.1, 504.2)	53.6 (0.4, 234.2)	94.9 (0.1, 504.2)
var pricing error, $\sigma^2(s)$ (bp ²)	9:00-12:00	10.1 (1.5, 27.5)	10.6 (1.9, 33.2)	10.2 (1.5, 33.2)
	12:00-15:00	9.5 (0.8, 20.0)	16.2 (6.1, 55.1)	10.5 (0.8, 55.1)
	15:00-17:30	115.4 (15.8, 523.7)	65.8 (1.0, 253.6)	108.2 (1.0, 523.7)

^a: This AR coefficient is based on the HFT opening position (just after the opening auction) relative to its previous day closing position.

Figure 1: Raw data plot high-frequency trader net position by frequency

This figure plots, for the median week in the sample, the high-frequency trader's net position in 'Unilever', the median stock in the large stock group. It plots the series for essentially three frequencies: minutes, hours, and days.



Figure 2: Market share Chi-X, high-frequency trader participation, and bid-ask spread

This figure plots four time series based on trading in Dutch index stocks from January 2, 2007, through April 23, 2008. The top graph depicts the market share of the entrant market Chi-X based on the number of trades; Chi-X started trading Dutch stocks on April 16, 2007. The graph also depicts the high-frequency trader's participation in trades, based on its trading in both the entrant (Chi-X) and in the incumbent market (Euronext). The middle graph replots the high-frequency trader's participation in trades, but now numerator and denominator are entrant market trades only. The bottom graph plots the average bid-ask spread of Dutch stocks relative to the spread of Belgian stocks. The incumbent market is the same for both sets of stocks (Euronext) but the Belgian stocks get delayed 'treatment with Chi-X'; they started trading in Chi-X one year later (April 24, 2008). The bid-ask spread is the inside spread, i.e., the lowest ask across entrant and incumbent market minus the highest bid across these two markets. The spread is calculated daily as a cross-sectional weighted average where weights are equal to stock weights in the local index.

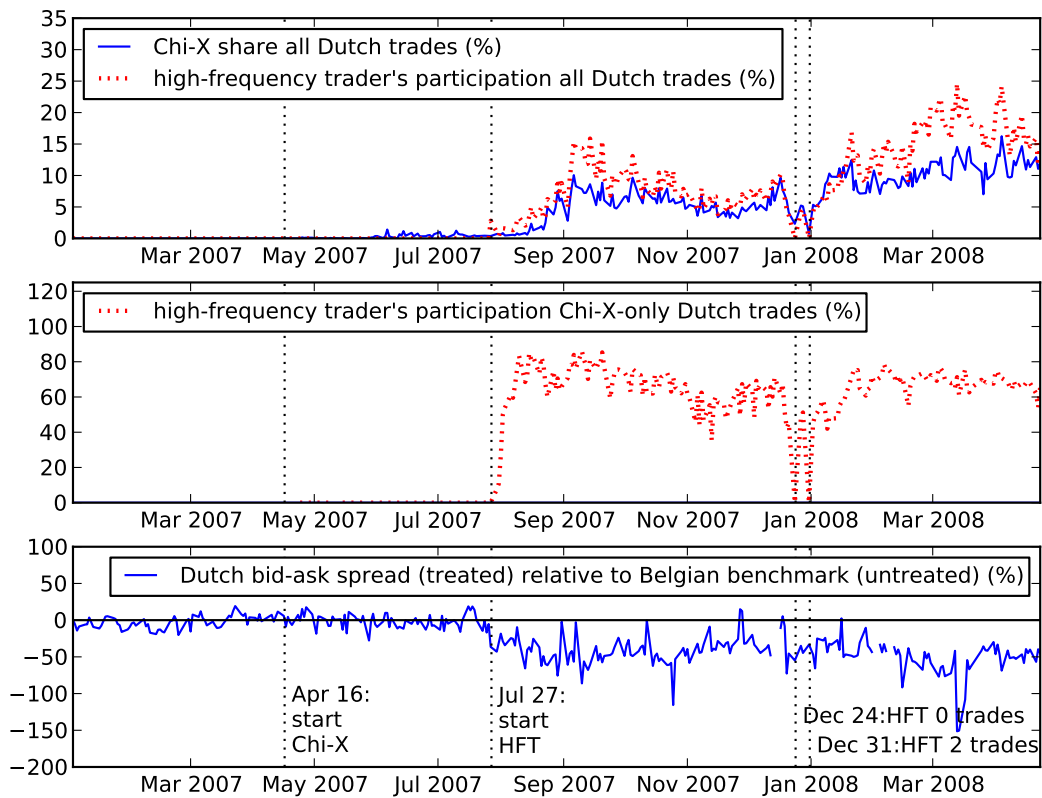


Figure 3: Trade revenue decomposition: spread vs. positioning revenue

This figure illustrates the decomposition of a high-frequency trader's gross profit into a spread and a positioning profit by analyzing the two extremes of how an HFT might generate profit. One extreme is that the HFT aggressively speculates; it consumes liquidity to pursue a fundamental value change that it observes and trades on quickly: spread revenue is negative, positioning revenue is positive. The other extreme is that the HFT passively makes a market; it produces liquidity by submitting bid and ask quotes and suffers a positioning loss when adversely selected by an incoming informed market order (see, e.g., [Glosten and Milgrom \(1985\)](#)). An alternative well-understood source of negative positioning profit is a price concession that a market maker willingly incurs to mean-revert its position (see, e.g., [Ho and Stoll \(1981\)](#)).

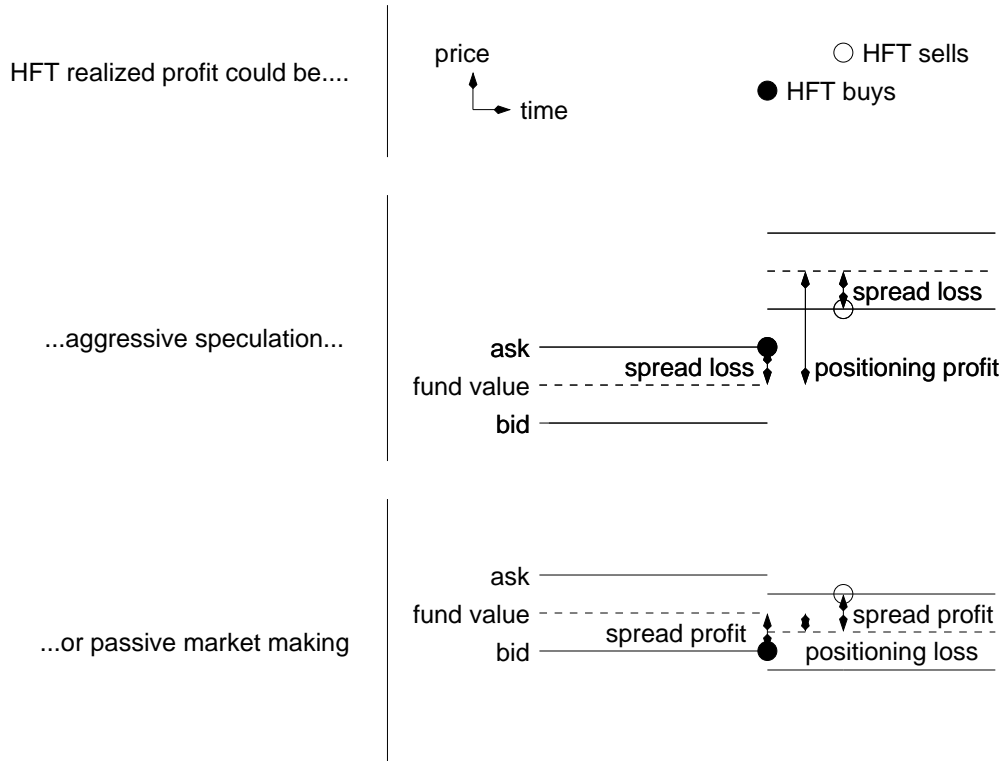


Figure 4: Raw data plot of the high-frequency trader's net position by market

This figure plots, for the median week in the sample, the high-frequency trader's net position in 'Unilever', the median stock in the large size group. It plots the series for Euronext only, Chi-X only, and its aggregated position across markets.

