High Frequency Trading and Market Volatility:

Is there a Fundamental Association?

P. Joakim Westerholm[♣]

Abstract

The findings in this paper confirm that there is a general, statistical and fundamental negative association between High Frequency Trading [HFT] activity and market volatility. The connection between HFT and volatility is most pronounced during periods of very short intervals, however the association is also statistically significant and negative in data of monthly frequency. Results indicate that technological innovation in market structures through the introduction of Co-location 'Proximity Services' on the Nasdaq-OMX Helsinki [OMXH] accelerated the negative association between HFT and market volatility. The implication of this study is that future regulation must weigh up the role of HFT in dampening intra-day volatility with the systematic risks posed by the sudden evaporation of their order-flow from the market.

Keywords: Volatility, High Frequency Trading

[♣] University of Sydney Business School, H69 Economics and Business Building, NSW 2006, Australia, Ph +612 9351 6454, Email: joakim.westerholm@sydney.edu.au I am indebted to Justin Bercich for exploring the effect of co-location on volatility in his honours thesis and to Euroclear Finland Ltd for providing access to data.

1 Introduction

The purpose of this paper is to investigate if there is any evidence of a common perception that there is an association between High Frequency Trading [HFT] activity and market volatility that needs to be regulated. The evolution and innovation in technology has altered how markets are structured with the imposition of a new breed of market participant over the past decade – Algorithmic traders and High-Frequency traders. Algorithmic traders use automated computer processes to analyze, order and execute trades. HFT is conducted by a sub-group of algorithmic traders that act in a propriety capacity through the use of "extraordinarily high-speed" computer platforms to order and execute trades; and utilize co-location proximity servers with ultra-low latency direct market data feed's (SEC, 2012).

The emergence of HFT as a fundamental driver of trading activity on financial markets is a preeminent issue in the contemporary regulatory discourse. Market participants and regulators are concerned with the significance of computer-driven algorithms and are asking what probability of success human cognitive induced decision-making has when competing against rational algorithmic-driven opponents? Since HFT driven marketable orders would be expected to improve liquidity through the magnitude of trading volume that algorithmic strategies infuse (Hendershott, 2011), low frequency traders and regulatory bodies may be willing to accept HFT participation in modern financial markets provided that it can be shown that HFT is not negatively affecting other dimensions of market quality.

The current regulation of HFT is fragmented in part due to the lack of consensus among the limited but growing academic research on the behavior of high frequency traders. A majority of academic research on the topic provides evidence supportive of the role that HFT play in improving market quality across dimensions of liquidity, price discovery and volatility (see Hendershott, Jones and Menkveld, 2011, and Brogaard, Hendershott and Riordan, 2013). Some studies find contradictory results (for example Zhang (2010))

Hendershott, Jones and Menkveld, (2011) specifically suggest that the credit crisis of 2007/2008 would be an appropriate period to study to see how HFT is related to volatility during periods of extreme volatility.[1] Brogaard, Hendershott and Riordan (2013), find that HFTs overall trade in the direction of reducing transitory pricing errors both on average days and on the most volatile days during a period of relative market turbulence (2008-2009).[2]

In this study I attempt to overcome the frailties of defining HFT behavior by building upon the HFT investor classification framework first implemented by Kirilenko et al. (2011) and using raw trading data where individual investor accounts can be identified. The model is based on the assumption that the trader population has varying investment horizons which can be explicitly identified through their inventory versus turnover levels across the trading day. HFTs are identified by their unique algorithmic trading strategies characterized by extremely high turnover levels and low net inventory positions that oscillate around a mean value close to zero. This paper provides a unique contribution to the current literature through the dynamic implementation of the Kirilenko et al. (2011) classification framework on an equities market and across a prolonged time period. Previous studies that have utilized the model, including Cvitanic (2010), and Kirilenko et at. (2011), have focused on one trading day and futures markets that trade a single security.

---

[1] Hendershott et.al (2011) write "While we do control for share price levels and volatility in our empirical work, it remains an open question whether AT and algorithmic liquidity supply are equally beneficial in more turbulent or declining markets. Like NASDAQ market makers refusing to answer their phones during the 1987 stock market crash, algorithmic liquidity suppliers may simply turn off their machines when markets spike downward. With access to the right data, the 2007 and 2008 stock markets could prove to be a useful laboratory for such an investigation.

[2] Broogard et.al (2013) report results that "suggest that HFTs use information in the limit order book to demand liquidity and that HFTs often supply liquidity on the thin side of the limit order book. This involves possibly incurring adverse selection costs by supplying liquidity in the direction where less liquidity is available. Such liquidity supply is generally interpreted as beneficial if it reduces transitory volatility."

The association between HFT[3] and volatility is investigated, a) taking advantage of a unique opportunity to obtain data on each transaction of high frequency traders in whole market, b) using the October 2008 introduction of Co-location servers on the NASDAQ OMX Helsinki Stock Exchange [OMXH] as an exogenous trigger of HFT and c) in a period of exceptional changes in volatility during 2008 and 2009.

In summary this paper finds that HFT represents a total of 31.8% of all value traded, hence a fundamental component of trading activity on the OMXH throughout the period. I confirm that there is a general, statistical and fundamental negative association between High Frequency Trading [HFT] activity and market volatility. The associations between volatility and HFT participation and volatility and order imbalance are investigated in VAR regressions models. I find that the connection between HFT and volatility is most pronounced during periods of very short intervals, while the association is also statistically significant and negative in data of monthly frequency. Results also indicate that technological innovation in market structures through the introduction of co-location 'Proximity Services' on the [OMXH] accelerated the negative association between HFT and market volatility. The implications of this study in informing future regulations must weigh up the role of HFTs in dampening intra-day volatility with the systematic risks posed by the sudden evaporation of their order-flow from the market.

2 Institutional Setting

The institutional setting at the Nasdaq OMXH is similar to other Nordic European exchanges where trading is conducted electronically in a central limit order book with no designated liquidity suppliers in any major stock issues and since 2006 trading broker identity is pre-

[3] HFT activity is measured both as a) participation, computing the proportion of the value traded where an identified HFT account is a counterparty to the trade and b) as order imbalance where the difference in HFT buyer initiated trades and HFT seller initiated trades scaled by the total value traded by HFT during a time interval is computed.

trade anonymous, trading opens with an auction at 10 am and closes after a long post trading period that ends at 6.20 pm. The minimum tick-size is 0.01 EURO for most stocks as the exchange has few low price stocks. The Helsinki market has become a significant part of the global portfolio despite its relatively small size. Nasdaq OMXH is home to well known companies in the technology sector, and their presence may also have alerted international investors to the other large companies at the exchange, typically in the industries engineering, forestry and resources. During the period of study, foreign investors held on average 61% of the market capitalization of the exchange, which was equivalent to approximately 235 billion EURO at the end of my data sample. 18% of the largest capitalization company on the exchange Nokia was held by 13f registered US institutional investors during the period and most of the approximately 200 common stocks listed during the period had foreign ownership of more than 1%. Many well known international high frequency trading corporations have announced their participation in OMXH in the period 2007 to 2009. Hence the results I draw from this dataset should have implications for our general understanding of financial markets, particularly in the context of institutional investors who operate globally.

Financial exchanges are today facilitated by publicly-listed for-profit companies, who are required to continuously expand their operations to deliver growth. Exchanges have benefited from HFT investors by introducing market platforms to enable orders to be placed quicker, facilitating higher volumes of trading, liquidity, and ultimately profits to the Exchange. The primary method in which Exchanges profit from HFT investors is through the offering of ‘Proximity Services’ to HFT firms through Co-location servers which directly access Exchange servers. Through the minimization of constraints imposed by latency and inefficient transmission cables, HFT’s are able to process and execute trades almost instantaneously. A second line of revenue is drawn by offering ultra-low latency direct market data feeds to clients.

Co-location services were first marketed to HFT investors on the investigated NASDAQ OMX Helsinki stock exchange [OMXH] during the early half of 2008. This followed the successful implementation by the NASDAQ OMX of similar services in the US market in 2006-2007. Labelled as 'Proximity Services' the product was offered in response to the demand driven "market needs" of algorithmic trading strategies that required minimal latency times through quicker market access (NASDAQ, 2008). 'Proximity Services' for HFT investors on the OMXH were introduced with 'live' functionality on the 1st October 2008. The service offered HFT investors ultra-low latency access to market serves through their own servers located within the exchange, or co-location.

There is no publicly available information on the take-up rate among HFT investors operating on OMXH, however, an analysis of comparative information available for NASDAQ markets in the US and UK indicate an expectation is would be significant (Hasbrouk and Saar, 2010). Given that HFT investors operate within a highly competitive environment, increasingly low latency of order processing levels has become an imperative to their survival. Evidently, it would be a rational prerequisite for HFT's to acquire such 'Proximity Services' in order to ensure their competitors do not have a significant advantage.

3 Data

The dataset utilized to identify levels of HFT activity incorporates a sample set of all investor level transactions conducted on the NASDAQ OMX Helsinki Stock Exchange [OMXH] over the period of January 2008 to September 2009. This data is refined to include only stocks that are a component of the OMX Helsinki Benchmark GI Index, a market index that includes all large-cap firms. The final dataset includes tick-by-tick data for 38 common stock's that traded across the period and remained a continuous component of the semi-annually reviewed Benchmark index. Information on these stocks is presented in Appendix I. 'Upstairs' trades internalized within brokerage firms, which account for an estimated 6% of the value of daily transactions, are included within the sample (Hasbrouck, 2009).

The data originates from the information provided to the shareholder depository administered by Euroclear Finland Ltd. The dataset has become one of the most trusted sources of investor level data, see for example, Grinblatt and Keloharju (2000, 2001a 2001b), Linnainmaa, Grinblatt and Keloharju (2012) and Linnainmaa and Saar (2012) and Berkman, Koch and Westerholm (2013). The Euroclear information is aligned with tick-by-tick level transaction data provided directly by OMXH and with third-party data from Standard and Poor's Compustat. Exchange-level data for the OMXH index is attained through the Thomson Reuters Tick-by-Tick database. Macroeconomic data on the Finnish economy is extracted from official Government sources. Statistics on macroeconomic factors are quoted directly from the Statistics Finland website, and official economy wide statistics are supplied from the Bank of Finland website.

The final dataset of 38 stock's represents a dominant proportion of trading activity (over 70% of all transactions by volume and value) among the 191 securities that were listed on the OMXH throughout the period of January 2008 to September 2009. This sample of stocks is chosen to control for analytical issues resulting from firm-size affects and liquidity constraints that may skew the HFT activity and Volatility relationship. Furthermore, only stocks with relatively large capitalizations and liquid markets for their stocks are traded by HFT. I contrast this sample of large capitalization stocks to a relevant proxy for market volatility – the OMX Helsinki 25 index which constitutes only large capitalization firms. Finally, the analyzed dataset includes all trades conducted on the OMXH including those conducted during market open times and those that occur after daily trade is halted or through trading dark pools.

The time period for this study has been chosen for its unique characteristics in terms of the introduction of co-location servers on the exchange for the first time and its historically high levels of market volatility. This period includes the Exchanges first steps to differentiate HFT from other investors by offering co-location services that enable quicker access to the main servers. During this period from January 2008 to September 2009 the OMXH25 Index fluctuated from a high of 3021.1 on 2nd January 2008 to a low of 1181.7 on 9th March 2009 after which it took a sharp upturn. This represents the most volatile period on the exchange in recent history and incorporates the primary events that facilitated the credit crisis of 2007 and 2008. This unique period enables the testing of the association between HFT and market volatility across significant events of technology innovation that could reasonably be expected to highlight the correlation between the two variables. It is also the only opportunity to do so, as after September 2009 transactions are reported as net daily transactions per investor account due to the exceptionally high volume making it too ineffective to clear transactions trade by trade.

Table 1 provides summary statistics of the dataset that explains the general trading behavior of participants on the HEX over the period January 2008 to September 2009. The final sample includes 440 trading days and 38 large capitalization individual stock ISIN's. Over $725 billion value of trades were conducted between two counter-parties, through a total of more than 51 million transactions during this time period.

4 Methodology

4.1 Definition of HFT activity

HFT's conduct trading operations through a hyper-active algorithmic based strategy, whereby traders buy and sell stocks based on extremely short holding periods with the aim of capturing micro profits. Along with the algorithmic nature of HFT strategies, other characteristics that define their behavior is a tendency to hold low net positions by the end of trading days, and their role as net liquidity providers in equity markets.

A prime limitation in analyzing the HFT is the lack of a universal definition to dichotomize HFT market participants from non-HFT participants. To differentiate and classify trading accounts the Investor classification model developed by Kirilenko et al. (2011) is employed. Investors are defined by the actual transactions they execute and how they operate daily on the exchange. This method is based on an inventory versus turnover analysis as opposed to a traditional trade-based prescription to define Investors. By applying this model to the data it is possible to analyze how different categories of investors operate across the time period. The shortcomings of the model are addressed through a comparison with previous HFT proxy literature.

4.1.1 Investor Classification Model

Kirilenko et al. (2011) successfully applies an inventory versus turnover Investor Classification model to define and describe traders on the US S&P E-Mini Future Contract market in the period surrounding the 2010 ‘Flash Crash’. This method defines Investors as HFT or non-HFT based on their trading behavior, particularly their daily net holding position in the instrument, the level of activity on the market in terms of value traded, and the quantity of transactions executed in which the Investor is party to. The central supposition of the model is that financial exchange markets facilitate a platform for traders with “different holding horizons and trading strategies” to interact. For example, large institutional investors seeking to attain a significant stake in a company will generally accumulate a large buy position over a long period of time. In contrast, other traders will seek to maneuver their trading strategy throughout the day to keep their net position to a minimal value whilst trading a high volume of stock. Other investors may utilize both strategies across different periods.

To apply the Investor Classification model and determine what constitutes HFT activity I process the pure original transaction-level data to identify attributes of particular trading accounts. The initial dataset of over 51 million transactions includes information on all individual trades based on the following data fields – company international security code (ISIN), date, time (to nearest second), executing trader account (by anonymous account ID), counterparty account (by account ID), buyer or seller initiator indicator, price and volume. These fields enable us to manipulate the data to calculate each account's tick-by-tick net holding and total trading positions throughout the day, and determine the transactions that the account initiated.

Standard transaction datasets such as the one used in this paper lack a discernible high frequency trader classification system, which justifies the application of Kirilenko et al. (2011) framework to classify accounts into the following trader categories – Intermediaries, High Frequency Traders, Fundamental Buyers, Fundamental Sellers, Small Traders and Opportunistic Traders. These investors are classified based on the following characteristics:

1) Intermediaries (Int) – are very short horizon investors who buy and sell a large volume of securities, but stay around a relatively low target level of inventory. So, their end of day net position is no more than 5% of the value of daily trading transactions in which they are involved. These investors hold a very small position when the markets close whilst participating in a large volume of intra-day trading. Intuitively, it could be argued that these traders have significantly short-term investment horizons, and generally net out a large majority of their positions by the end of the trading day.

2) High-Frequency Traders (HFT) – can be identified as a subset of Intermediaries and represent the top 7% of trading accounts when ranked by the number of daily transactions in which they are involved. I assess the impact of classifying using different percentages of to accounts by daily transactions, but the 7% threshold isolates a distinct group of trading accounts. Results are not significantly affected by changes in this percentage. Essentially, these accounts are the most active or 'High-Frequency' Intermediaries in the market. This cut-off level has been calculated to designate HFT accounts that are significantly different in the magnitude of trading activity prevalent in Intermediaries. Once an account is designated as HFT it is removed from the Intermediary set.

3) Fundamental Buyers (Fun_Buy) and 4) Fundamental Sellers (Fun_Sell) – are generally institutional investors whose trading accounts mostly buy or sell in one direction during the day. These accounts hold at the end of the trading day a long net portfolio position, in terms of trading value executed, that is greater than 15% of the total values of trades in which they are involved with daily. An increase or decrease of the 15% criteria does not materially affect the composition of included accounts.

4) Fundamental Sellers hold at the end of the trading day a short net portfolio position, in terms of trading value executed, that is greater than 15% of the total values of trades in which they are involved with daily. An increase or decrease of the 15% criteria does not materially affect the composition of included accounts.

5) Small Traders (Small) – are involved in transactions that total no more than $10,000 across the trading day.

6) Opportunistic Traders (Opp) – are the trading accounts that remain after the categories 1) to 5) have been classified. These traders may execute algorithmic strategies, however, their

behavior as defined through the volume and value of stock traded is too low to be categorized as an intermediary.

Trading accounts are classified into one of six mutually exclusive categories for each of the 440 individual trading days that the data covers. Hence the possibility that a trader changes strategy is allowed for, which is expected to be less applicable for the more long term investors, but should qualify high frequency traders well as each included account is required to trade with ultra high frequency and low inventory during each specific observations day for which they are included. Descriptive statistics of these Trading categories are produced in Table 2, and represent the behavior of each group across the 21 months analyzed.

4.1.2 Robustness of the investor classification method

While Kirilenko et al's (2011) framework represents a potentially powerful method of classifying trading accounts, there are limitations in applying the model to the Finnish dataset used in this paper. This study applies the model (originally designed for one trading day and one instrument) across all 38 OMXH Benchmark Indexed equity securities and individually for each of the 440 trading days analyzed. Hence, accounts are classified uniquely each trading day with accounts able to shift between categories inter-day depending on how they behave on any given day and the analysis is conducted using observations from those days only when a trader is actively trading according to a HFT strategy.

HFT activity over the sample period from January 2008 to September 2009 is prevalent in 31.76% of average monthly trades by value. This measure oscillates between minimum and maximum levels of 20-45% across the dataset. These results are in-line with expectations and the pervading academic literature. Jarnecic and Snape (2010) find that between 40-64% of trades executed on the London Stock Exchange (LSE) in 2009 were conducted by HFT firms. Furthermore, a report conducted by the CESR (2010) presented an estimate that HFT have a market share of between 25-35% of the activity on the LSE in the first quarter of 2010. Whilst these levels are significantly different those seen in the US, where HFTs currently participate in up to 92% of trades (Ito, 2012) , the fragmented nature of the European market in terms of clearing, settlement and post-trade services may account for these difference. The results attained in this study appear reasonable in comparison to previous research and indicate that the Kirilenko et al's (2011) investor classification model can be applied dynamically across time periods to equities markets that list a large universe of securities.

Analyzing the results in Table 2 it can be noted that HFT investors are net liquidity providers even through this figure is very close to 50%. Finally, the key role that HFT play on the market is exemplified from that fact that whilst they only account for only 0.09% of the unique trading accounts, these traders are involved in over 30% of all transactions by value.

4.2 Definition of Stock Market Volatility

For the purpose of measuring market quality relevant volatility I use realized volatility computed on intra-day intervals of five minutes. Realized volatility has the suitable properties for this research in that it measures short term changes in volatility including both transitory and fundamental volatility[4]. I am interested in the complete short term impact on volatility and hence there I proceed without smoothening the volatility measure using alternatives such as range-based volatility (Parkinson, 1980, and Alizadeh et. al, 2002), GARCH(1,1) adjusted volatility (Bollerslev, 1986), or applying moving average mid-point prices to calculate realized volatility (Andersen et.al, 2003).

Realized volatility is computed in five minute intervals as:

$$VOL_{s,t} = [\ln\left(\frac{S_t}{S_{t-5}}\right)]^2 \qquad (1)$$

[4] Transitory volatility is caused mainly by microstructure effects and frictions while fundamental volatility is caused mainly by changes in the fundamental value of an asset.

Where $VOL_{s,t}$ is the period t realized volatility measure for security s, s_t is the last traded price for security s in the current 5 minute period t, and s_{t-5} is the last traded price for security s in the previous 5 minute period t-5. This measure captures the level of volatility across 5 minute intra-day periods and can be aggregated to daily, weekly and monthly measures by adding up the squared returns for the desired period.

4.3 Underlying Variables affecting the HFT to Volatility Relation

To analyze the association between HFT and Market volatility I include a set of control variables. There are three separate sets of control variables identified in the literature to explain market volatility – market cyclical variables, macroeconomic level variables and market related variables.

Market cycle variables may have a significant impact on investor behavior which in turn may correlate with stock market volatility. The two sets of Market cycle variables most causally related to the market are based on economy-wide Price changes and Output levels. Price changes in the Finnish economy are reflected in the official Consumer-Price index which is analyzed to calculate Household price Inflation (InfH) volatility across the period. Another relevant explanatory variable for economy wide price changes is Producer price Inflation (InfP). This measure is calculated from the base Producer price Index that incorporates the evolution and change in commodity prices from the perspective of enterprises. Output volatility levels are most accurately reflected in the official Output Index (OutO) reported by Statistics Finland. This measure smoothens changes in GDP on a monthly adjusted basis. Furthermore, the Industrial Production index (OutP) can be used as a proxy of economy-wide physical output. This index serves as a good indicator for long-term economy production capabilities as non-Industrial short-term variables in the Output function, which tend to be correlated to swings in the economy, are eliminated. Hence, I expect this measure to move in an opposite, but correlated, direction to the stock market index.

Macroeconomic variables that represent the uncertainty in the economy-wide environment can be used as control variables when testing HFT as an explanatory variable for market volatility. Two sets of macroeconomic uncertainty variables are proposed through Inter-Bank Interest rates and domestic currency Exchange rates. Macroeconomic uncertainty in the economy is represented through the volatility of the 3-month Eurepa Interest Rate level (IREA) which are security-backed Euro area inter-bank quoted rates. This is calculated through a process that identifies rates at which the highest rated banks offer loans to each other that are secured by top-grade government securities. Euribor interest rates (IREU) are a similar instrument but instead are un-securitised inter-bank lending rates as quoted by the largest banks in the Euro area. A final macroeconomic control variable is the volatility in the exchange rate of the domestic currency, the Euro, versus its largest trading currency, the US Dollar (ERU), and the internationally weighted instrument, SDR's (ERS).

An important market level control variable expected to be related to volatility is the level of market turnover on the exchange (MKT). This variable is calculated by taking the natural logarithm of the total value traded on the exchange in dollar terms. The conditional volatility for the proposed nine control variables is estimated and then each variable is introduced and evaluated for its contribution when analyzing the HFT and market volatility association.

5 Results and Analysis

Figure 1 shows the share market index development and its volatility during the investigated period 2008 to 2009. It can be seen that the peak in volatility occurs during the latter half of 2008.

Figure 2 depicts the levels of daily market volatility and daily HFT proportion of trading for the period around co-location. HFT activity increases after co-location as expected, while volatility increases as a result of the widespread financial crisis starting to affect European markets about one month later.

5.1 Dynamic association between volatility and HFT

The association between volatility and HFT participation as well as HFT order imbalance is analyzed using a VAR model approach. First I investigate the association between volatility and HFT participation measured as share of value traded during each five minute interval, period t, where a HFT identified account is a counterparty.

$$VOL_{s,t} = \alpha_0 + \sum_{i=1}^{n} \alpha_1 HFT_{s,t-i} + \varepsilon_{s,t} \qquad (2)$$

$$HFT_{s,t} = \beta_0 + \sum_{i=1}^{n} \beta_2 VOL_{s,t-i} + \mu_{s,t} \qquad (3)$$

Where $VOL_{s,t}$ is the realized volatility for stock s in period t as derived from Equation 1. $HFT_{s,t}$ is the fraction of the total value of stock s turnover during period t in which a HFT account is a counterparty. $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$ and $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ are the coefficients tested in the respective Volatility and HFT dependent variable regressions. $\varepsilon_{s,t}$ and $\mu_{s,t}$ are respective error disturbance terms. Second I investigate the association between volatility and HFT participation in a log-log model to determine the magnitude and economic significance of the relationships.

$$\ln\,(VOL_{s,t}) = \gamma_0 + \sum_{i=1}^{n} \gamma_1 \ln\,(HFT_{m,t-i}) + \varepsilon_{m,t} \qquad (4)$$

$$\ln\left(HFT_{s,t}\right) = \theta_0 + \sum_{i=1}^{n} \theta_2 \ln(VOL_{m,t-i}) + \mu_{m,t} \qquad (5)$$

Where $\ln\left(VOL_{s,t}\right)$ is the natural logarithm of realized volatility for stock s in period t as above and $\ln(HFT_{s,t})$ is the natural logarithm of the fraction of the total value of stock s turnover during period t in which a HFT account is a counterparty to the trade. Thirdly I investigate the full dynamics of the model with lagged dependent and independent variables in the same equation, with the purpose to determine how the association between volatility and HFT participation changes as a result of co-location services, when HFT traders gain a much faster access to the exchanges servers.

$$VOL_{s,t} = \theta_0 + \delta_1 \tau + \sum_{i=1}^{n} \theta_1 HFT_{m,t-i} + \sum_{i=1}^{n} \theta_2 VOL_{m,t-i} + \mu_{m,t} \qquad (6)$$

Where $VOL_{s,t}$ is the realized volatility for stock s in period t as above and $HFT_{s,t}$ is the fraction of the total value of stock s turnover during period t in which a HFT account is a counterparty. τ represents an indicator variable that takes the value 0 up to the end of September 2008 and 1 from October 1, 2008 when co-location services are introduced. Finally I investigate the relation realized volatility to HFT order imbalance[5] to determine if prolonged one sided HFT activity has an impact on contemporaneous and future volatility in the intra-day space. This model is estimated separately for buy initiated and sell initiated HFT trades and in log-log form for all HFT trades:

$$VOL_{s,t} = \alpha_0 + \sum_{i=1}^{n} \alpha_1 HFT\_OI_{s,t-i} + \varepsilon_{s,t} \qquad (7)$$

[5] Order imbalance is computed for each five minute period as value of buy trades initiated by HFT minus value of sell trades initiated by HFT scaled by total value traded by HFT in that period. Order imbalance is scaled by the total traded value so as to eliminate the impact of total trading activity. Actively traded stocks with higher total number of trades per day or a larger daily dollar trading volume are likely to have higher imbalances. The scaling standardizes the imbalance measures (Chordia and Subrahmanyam, 2004).

$$HFT\_OI_{s,t} = \beta_0 + \sum_{i=1}^{n} \beta_2 VOL_{s,t-i} + \mu_{s,t} \tag{8}$$

The results reported in Tables 3, 4 and 5 generally confirms a strong association between volatility and HFT activity in the intra-day time space using five minute frequency. HFT participation is consistently and statistically significantly negatively associated with realized volatility. The association between HFT participation and volatility extends up to 10 lagged 5 minute periods, with the negative association clearest up to 10 minutes prior. The Co-location indicator variable takes a positive sign and is statistically significant indicating that short term realized volatility is higher after the introduction of co-location services.

HTF order imbalance in the two previous periods is positively related to contemporaneous volatility intra-day. Interestingly the results show that HFT order imbalance in previous periods of more than 10 minutes before are negatively related to contemporaneous volatility. There is little relation between previous period volatility and HFT order imbalance. These results show that HFT orders has an increasing impact on realized volatility in future periods with lags up to five 10 minutes while the impact then reverses. This conforms with anecdotal evidence of increased short term volatility around periods of high HFT order imbalance. We consider the overall implications of these findings together with the results from the further analysis in the conclusion section.

5.2 Is there a Fundamental Association between HFT and Volatility?

The previous sections confirm that the two processes of HFT and volatility generally occur during the same or near periods in intra-day space. In order to analyze the fundamental association between HFT and volatility it is necessary to determine whether the association between the two variables is not in fact driven by a third variable. In this section I introduce controls for macroeconomic, cyclical and market factor structures found to impact stock market return volatility in the literature (see e.g. Rigobon and Sack, 2003). I conduct this analysis on monthly data for two reasons, a) we are looking to explain long term fundamental association between HFT and volatility and b) data for most control variables are only available on a monthly frequency.

Monthly statistics for the 11 variables used in the regressions are presented in Table 6. As expected the four sets of control variables explicit in both the Macroeconomic and Cyclical factor sections follow similar trajectories through time. The proportion of HFT activity in the market ranges from 18.4% to 42.2% across the 21 months analyzed. This measure indicates that HFT behavior is not fundamentally rooted in the market structure but rather fluctuates through time.

The regression model seeks to test the impact of the independent variable, HFT activity, on the dependent variable, Market volatility, whilst controlling for a third set of variables that may impact on the association between the two. A sequence of four equations are estimated concurrently to ensure that only relevant control variables are included. Given that only 21 monthly observations are available the model utilizes the GMM method and bootstrapping techniques in order to meet the assumptions posed by OLS regression models. Bootstrapping the monthly observations I assume that both the response and covariates are random. In order to account for statistical limitations, I follow the volatility literature and estimate conditional volatility measures at the monthly of frequency by fitting an EWMA model to market volatility and a GARCH (1,1) model to the Market-level, Cyclical and Macroeconomic variables, see Appendix 3.

The estimated equations are:

$$VOL_t = \alpha_1 + \theta_1 InfH_t + \theta_2 InfP_t + \theta_3 OutO_t + \theta_4 OutP_t + \varepsilon_t \tag{9}$$

$$VOL_t = \alpha_1 + \delta_1 ERU_t + \delta_2 ERS_t + \delta_3 IREA_t + \delta_4 IREU_t + \vartheta_n MACRO_t + \epsilon_t \quad (10)$$

$$VOL_t = \alpha_1 + \gamma_1 MKT_t + \rho_n CYCLICAL_t + \vartheta_n MACRO_t + \tau_t \quad (11)$$

$$VOL_t = \alpha_1 + \beta_1 HFT_t + \gamma_1 MKT_t + \rho_n CYCLICAL_t + \vartheta_n MACRO_t + \mu_t \quad (12)$$

Observed values at time *t* for Market Volatility, HFT activity, Market-level, Cyclical and Macroeconomic variables in Table 6 are included in the regression as follows. $MACRO_t$, $CYCLICAL_t$, MKT_t represent a set of macroeconomic, cyclical and market-level inputs from preceding regressions that were estimated to have an impact on Volatility at the 10% significance level. Thus, variables that fit each model are included into the next regression and this process continues into the last regression equation or until they no longer provide explanatory power and are annulled from the model. Regressors at a 10% significance level or higher remain in the model for the second regression. This process is repeated for cyclical, then market-level factors. Control variables that attain a 10% significance level after this third regression are included as a set of control variables in the final regression Equation (9) which tests the association between HFT activity and stock market volatility.

As a result of the control variable selection process, Equation (12) is estimated with the dependent variable conditional market volatility vs. the monthly fraction of HFT, the log of market turnover (MKT), Exchange Rate Volatility Euro vs. USD (ERU), Exchange Rate Volatility Euro vs SDR (ERS) and Interest Rate Volatility in the 3 Month Eurepa interbank rate (IREA), hence the estimated equation becomes:

$$VOL_t = \alpha_1 + \beta_1 HFT_t + \gamma_1 MKT_t + \rho_1 ERU_t + \rho_2 ERS_t + \rho_3 IREA_t + \mu_t \quad (13)$$

The results reported in Table 7 provide strong evidence to assert a fundamental negative association between HFT activity and stock market volatility. The HFT coefficient parameter estimate β_1 is statistically significantly different from at -0.48 (p-value 0.029). This finding indicates that conditional market volatility has been on average 0.48% lower for every 1% increase in HFT activity. An interpretation of this result is that HFT investors who aggressively trade at the bid-ask spread contribute to market quality by lowering the volatility of asset prices. It is evident that periods of high volatility are significantly and fundamentally related to lower HFT activity regardless of whether HFT are supplying or demanding liquidity.

The accuracy and efficiency of the model explaining stock market volatility is improved by adding the HFT activity variable into the regression with an increase in R-squared values from 0.69 to 0.74. There is also clear evidence that the four control variables are statistically related to market volatility. Each included macroeconomic, cyclic and market-level control variable is statistically significant across at the 1% level of significance.

For robustness a final regression is performed based on an autoregressive process that captures lagged conditional volatility variables into the regression. This method is aimed to control for serial correlation remaining in the model's monthly conditional volatility observations. The HFT volatility fundamental association is tested intra-day applying an EGARCH (1,1) as suggested by Nelson (1991). These results are presented in Appendix 4. Appendix 4 however indicates that within short term contemporary periods across one trading day the two factors are not fundamentally linked, the association is rather a more long term lead lag association.

In summary, for 38 large-capitalization stocks that are listed on the OMX Helsinki Exchange across the period of January 2008 to September 2009 I find that a fundamental association between HFT and stock market volatility exists and it is negative and statistically significant. This association is also economically significant as a decrease in volatility of 0.484% with each 1% increase in HFT activity translates into 1.1 billion € per day fall in the variation in market capitalization across the sample stocks. During the investigated period HFT participation varies between about 20% and 30%.

5.3 The impact of different investor category trades on stock price volatility

The extent to which a specific group of investors impact upon volatility is of significant importance to the participation levels of other players in the market (Groth, 2011). I am specifically interested in the effect of HFT's on price volatility before and after the co-location of HFT servers on the OMXH. Such a central shift in market structure and trading execution processes presents a relevant technological 'shock' vantage point from which to view the mutation of the HFT-Volatility association in the context of technological change. Figure 2 indicates that the level of HFT activity as well as volatility is different after the co-location. This section builds a regression model to test for the causal impact of HFT activity on price volatility when market structures are fundamentally altered.

Following Kirilenko et al. (2011) I estimate a regression model that attempts to model how contemporary HFT drive future price changes and market volatility. Prior period HFT activity is modeled as an explanatory variable for price change computations, weighted by each individual stocks contemporary volatility level. HFT activity is weighted through an Aggressiveness Imbalance indicator (AI) value which captures whether liquidity is being removed from the market based on trading direction. In order to test the association, the following regression equations are estimated:

$$\frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1} * \sigma_{i,t-1}} = \alpha + \left[\sum_{C=1}^{5} \beta_C * \frac{AI_{C,i,t}}{CLASS_{C,i,t-1} * 100{,}000} \right] + \varepsilon_t \qquad (14)$$

where $S_{i,t}$ is the price of stock i at time t pre Co-location, $\sigma_{i,t-1}$ is the volatility of stock i at time t-1 (as defined in Equation 1). The independent variable $AI_{C,i,t}$ is an aggressiveness imbalance indicator for stock i at time t for investor class C while $CLASS_{C,i,t-1}$ is the fraction of total value traded by investor Class C in stock i at time $t-1$. β_C is the coefficient of determination for each of the five separate investor classes, α is the regression intercept term and ε_t is the error term. In order to ensure there is scalable and relevant data and to avoid confounding effects, intervals of one week prior and post the introduction of Co-locations servers on the Nasdaq OMXH is tested using trade by trade data.

The dependent variable in the regression represents the realized price return for the current period scaled by the previous period volatility level. The Aggressiveness Imbalance weighting is calculated as the difference between the number of aggressive buy transactions during the period and the number of aggressive sell transactions by stock. The AI value indicates how investors behave during each period, hence their trade direction in periods of low and high volatility. The independent variable estimates weights the specific investor class's aggressiveness imbalance level during the current period by the previous period's level of investor trading activity in the market scaled by 100,000 to avoid large coefficients.

The results for the regressions testing the association between prices changes and HFT activity pre co-location are computed in Table 8 Panel A, and indicate that HFT's have a positive and significant impact on price change's during the week prior to co-location. This result is significant the 1% level with a p-value of 0.007. The coefficient for HFT of 10.2 is of most interest. A positive value indicates that HFTs levels in the current period are actually positively driving price changes and volatility in the next period. Thus, HFT's are having a positive impact and are in fact driving volatility in the market. These results are in line with the order imbalance analysis in previous sections, while it to some extent contradicts the negative fundamental association between HFT activity and volatility intra-day and also between monthly levels of HFT activity and conditional volatility.

The results for the regressions testing the association between prices changes and HFT activity post co-location are computed in Table 8 Panel B, and indicate that HFT's have a negative and weakly significant impact on price change's during the week post co-location. The findings indicate that the null hypotheses, that levels of HFT activity in prior periods do not drive volatility levels in current periods, can be rejected at the 10% significance level, with the resulting t-stats measured indicating a p-value of 0.088 for the Gamma coefficient. The parameter estimated is negative and quite large with a value of -24.1. An interpretation of this result is that in the period post co-location, levels of HFT activity had an impact of a stronger magnitude on price changes than in the week before co-location. Furthermore, this association with price changes is determined to be negative in correlation for HFT's trading indicating that increasing levels of HFT in during the current period are significantly related to a decrease in price volatility in future periods. Essentially, increasing levels of HFT activity dampen price volatility, and at a stronger rate, in the period post co-location.

These results imply beneficial impact of technological advancement on lowering overall market volatility through the inducement of HFT's to supply more trade on the market. Lower latency times and transaction costs may lead to a rise in the level of order-flow from HFT investors. By trading through more scalable liquidity provision strategies such as those defined as passive or structural, HFTs are contributing to overall market quality.

5. Conclusions

While the literature is more or less in agreement that algorithmic trading and high frequency trading is generally beneficial for market quality, it has not been able to alleviate the concern that the machines may cause volatility. This paper contributes with new evidence that allowing low latency traders in, can improve the way the market corresponds to volatility shocks. In the Nasdaq OMX Helsinki case the timing of co-location could not have been better, just as the global credit crisis of 2007 and 2008 hit Europe.

An analysis of the general, statistical and fundamental association between HFT activity and market volatility is conducted by applying an investor classification model to our data of 38 stocks across 441 trading days. The association is evaluated based on a unique dataset of second-stamped transactions conducted on the OMX Helsinki stock Exchange over the period of January 2008 to September 2009.

Evidence presented indicates that there is a general fundamental association between HFT activity and realized stock price volatility of statistical and economic significance. In the intra-day space with lags up to minutes high HFT driven order imbalances increase volatility. This effect appears to reverse in longer lags. Higher HFT participation measured as their fraction of total value traded is consistently negatively related to realized volatility. An analysis of the association at a monthly level also indicates that, when controlling for third variables, the association between HFT participation and volatility is statistically significant and negative. The results assert a negative correlation between stock market volatility and HFT participation to a greater extent than those of and Hendershott, Jones and Menkveld (2011) and Brogaard, Hendershott and Riordan (2013). One interpretation in line with previous literature (Viljoen, Westerholm and Zheng (2013) is that HFT enter the market during periods of low volatility and while they contribute to some short term volatility in near future periods, the overall long term effect is to dampen volatility as a result of the liquidity provision HFT provides.

Regulators and policy makers seeking to curtail the level of HFT activity in the market must weigh up the benefits that this class of investors brings to overall market quality. HFT's have been shown to significantly decrease price deficiencies in the bid-ask spread

(Hendershott, Jones and Menkveld, 2011). Our results indicate a fundamental and negative association between HFT participation levels and both stock specific and market wide volatility. These benefits exist in light of the systematic risk that HFTs pose to the efficient and robust operation of financial markets if their order flow evaporates during times of severe market distress. Given that these new market participants fulfil the role of the modem day market maker, but without the fiduciary and legal obligation to trade during periods of market stress, their participation is critical. Evidence presented by Kirilenko et al. (2011) though seems to detract from this notion as they find HFTs compete for liquidity in periods of extreme volatility.

The implications of this study in informing future regulations shows in the findings related to how HFT's change their behavior after periods of technological innovation. HFTs dampen volatility at a negative and stronger level after Co-location servers were made available to HFT investors. Lines of debate against HFT participation in markets based on the argument that technology provides an advantage of HFT's over other investors, particularly retail but also slow frequency buy side traders, must be weighed against the benefits of lower volatility and higher liquidity these traders contribute to the market. One option to lower potential negative implications of HFT would be to enforce their participation also during high volatility episodes as a requirement for access to low latency platforms. Ultimately, it is essential that market participants, regulators and industry garner a clear understanding of the role and risks that HFT's can contribute to financial markets in periods of significant leaps in technology.

REFERENCES

Aldridge I., 2012, *High Frequency Trading - A Practical Guide to Algorithmic Strategies and Trading Systems,* John Wiley & Sons: Hoboken, New Jersey

Alexander C., Lazar E., 2009, Modelling regime-specific stock price volatility, *Oxford Bulletin of Economics and Statistics,* Volume 71, Issue 6, pp. 761-769.

Alizadeh, S., Brandt, M., Diebold, F., 2002, Range-Based Estimation of Stochastic Volatility Models, *Journal of Finance*, Colume 7, Issue 3, pp. 1047-1090.

Andersen, T. G., and Bollerslev, T., 1998, Answering the sceptics: Yes, standard volatility models do provide accurate forecasts, *International Economic Review*, 39, 885–905.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P., 2003, Modelling and forecasting realized volatility, *Econometrica*, 75, 579–626.

Baillie R., Bollerslev., 1989, The message in daily exchange rates: a conditional variance tale, *Journal of Business and Economic Statistics*, Volume 7, pp. 297-305

Bank of Finland, 2012, Helsinki, Available: http://www.suomenpankki.fi/en/Pages/default.aspx. Accessed from 1/08/12

Bercich, Justin, High Frequency Trading and Market Volatility in an Environment of Technological Innovation, Honours thesis, University of Sydney, Australia.

Berkman H., Koch P., Westerholm J., 2012, Informed Trading Through the Accounts of Children, *Journal of Finance*.

Bollerslev, T., 1986, Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, Volume 31, pp. 307-327

Brandt, M., Kang, Q., 2004, On the Relationship between the Conditional Mean and Volatility of Stock Returns: a Latent VAR Approach, *Journal of Financial Economics,* Volume 72, pp. 217-257

Brogaard, J. A., 2010, High Frequency Trading and its Impact on Market Quality, North-western University, Working Paper

Brogaard, J. A., 2011a, High Frequency Trading and Market Quality, University of Washington, Working Paper

Brogaard, J. A., 2011ba, High Frequency Trading and Volatility, University of Washington, Working Paper

Brogaard, J. A., T. Hendershott and R. Riordan, 2013, High Frequency Trading and Price Discovery. Working paper UC Berkeley, Forthcoming in Review of Financial Studies.

Burnham K., Anderson R., 2002, *Model Selection and Multimodal Inference: A Practical Information-Theoretical Approach,* Second Edition, Springer-Verlag, New York

Castura, J., Dwivedi, Y., Gorelick, R., Litzenberger, R., 2010, *Market Efficiency and Microstructure Evolution in U.S. Equity Markets: A High Frequency Perspective,* RGM Advisors, Working Paper

CESR., 2010, Micro-structural issues of the European equity markets, published 1 April 2010, viewed on 3/9/12 at http://www.esma.europa.eu/data/document/10_142.pdf.

Chaboud A., Chiquoine B., Hjalmarsson E., Vega C., 2009, *Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market,* International Finance Discussion Papers – Board of Governors of the Federal Reserve System, #980

Chordia, T. and Subrahmanyam A., 2004, Order imbalance and individual stock returns: Theory and evidence, *Journal of Financial Economics*, 72, 485–518

Corradi, V., Distaso, W., Mele, A., 2009, *Macroeconomic Determinants of Stock Market Volatility and Volatility Risk Premia,* Working Paper

Cvitanic, J., Kirilenko, A. A., 2010, *High Frequency Traders and Asset Prices*, California Institute of Technology, Working Paper

Drakos G., Kouretas P., Zarangas L., 2010, Forecasting financial volatility of the Athens Stock Exchange daily returns: An application of the asymmetric normal mixture GARCH model, *International Journal of Finance and Economics*, Volume 1, Issue 4, pp. 1-19

Engle, R., 1982, Autoregressive Conditional Heteroscedasticity With Estimates of the Variance of United Kingdom Inflation, *Econometrica*, Volume 50, Issue 4, pp. 987-1007

Foucault T., 1999, Order Flow Composition and Trading Costs in a dynamic limit order market, *Journal of Financial Markets*, Volume 2, Issue 2, pp. 99-134

Godfrey L., 1996, Misspecification tests and their uses in econometrics, *Journal of Statistical Planning and Inference*, Volume 49, Issue 2, pp. 241-260

Granger C., 1969, Investigating causal relations by econometric models and cross-spectral Methods, *Journal of the Econometric Society*, Volume 4, pp. 424-438

Granger, C., 1969, Investigating Causal Relations by Economic Models and Cross Spectral Methods, *Econometrica,* Volume 37, pp. 424-438

Groth, S., 2011, *Does Algorithmic Trading Increase Volatility? Empirical Evidence from the Fully-Electronic Trading Platform Xetra*, Goethe University Frankfurt, Working Paper

Gsel M., 2008, *Assessing the Impact of Algorithmic Trading on Markets: A Simulation Approach*, Center for Financial Studies, Working Paper

Harris, L., 2003, *Trading and Exchanges: Market Microstructure for Practitioners*, Oxford University Press, Oxford

Hasbrouck J., Saar G., 2009, Technology and liquidity provision: The blurring of traditional definitions, *Journal of Financial Markets*, Volume 12, Issue 2, pp. 143-172

Hasbrouck J., Saar G., 2010, *Low-Latency Trading,* New York University, Working Paper

Hendershott, T., Jones, C., Menkveld, A., 2011, Does Algorithmic Trading Improve Liquidity? *Journal of Finance,* Volume 66, Issue 1

Hendershott, T., Riordan, R., 2012, *High Frequency Trading and Price Discovery,* Forthcoming Journal of Financial and Quantitative Analysis

Hoorn D., Nilsson M., 2012, *The relationship between High Frequency Trading and Stock Market Volatility*, Lund University, Working Paper

International Organization of Securities Commissions (IOSCO), 2011, Regulatory Issues Raised by the Impact of Technological Changes on Market Integrity and Efficiency – Consultation Report, Available: http://www.iosco.org/library/pubdocs/pdf/IOSCOPD354.pdf. Accessed 4/10/12

Ito T., Lyden S., 2012, *Towards an Understanding of High Frequency Trading*, Instinet, Working Paper

Jarnecic E., Snape M., 2010, *An Analysis of Trades by High Frequency Participants on the London Stock Exchange*, Working Paper

Jovanovic, B., Menkveld, A., 2010, *Middlemen in Limit Order Markets,* VU University Amsterdam, Working Paper

Kalev and Duong, 2008, *A test of the Samuelson Hypothesis using Realized Range*, The Journal of Futures Markets, Vol. 28, No. 7, pp. 680–696.

Kaminska I., 2011, Algo trading and the Nymex, *Financial Times Alphaville Blog, Available*: http://ftalphaville.ft.com/blog/2011/03/04/505021/algo-trading-and-the-nymex/. Accessed: 3/10/12

Kirilenko, A., Kyle, S., Samadi, M., Tuzun, T., 2011, *The Flash Crash: The Impact of High Frequency Trading on an Electronic Market*, Commodity Futures Trading Commission, Working Paper

Latteman C., 2012, High Frequency Trading: Costs and Benefits in Securities Trading and its Necessity of Regulations, *Business and Information Systems Engineering*, Volume 4, Issue 2, pp. 93-108

Liljeblom, E., Stenius, M., 1997, Macroeconomic Volatility and Stock Market Volatility: Empirical Evidence on Finnish Data, *Applied Financial Economics*, Volume 7, Issue 4, pp. 419 426

Martens, M., & van Dijk, D., 2007, Measuring volatility with the realized range. Journal of Econometrics, Volume 138, pp. 181–207

Menkveld A., 2010, High Frequency Trading and New-Market Makers, University of Amsterdam, Working Paper

NASDAQ, 2008, *OMX Proximity Services – Service Summary,* OMX, New York

Nelson, D., 1991, Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica,* Volume 59, pp. 347-370

Newey W., 1994, Kernel Estimation of Partial Means and a General Variance Estimator, *Econometric Theory*, Volume 10, pp. 233-253

Olsen R., 1998, Behavioral finance and its implications for stock price volatility, *Financial Analysts Journal*, Volume 54, Issue 2, pp. 10-18

Parkinson, M., 1980, The Extreme Value Method for Estimating the Variance of the Rate of Return, *Journal of Business,* Volume 53, pp. 61-65

Parlour C., 1998, Price Dynamics in Limit Order Markets, *Review of Financial Studies*, Volume 11, pp. 789–816
Rigobon R. and Sack B., Measuring The Reaction of Monetary Policy to the Stock Market, Quarterly Journal of Economics, 118, 639-669
Roll R., 1984, A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market, *The Journal of Finance*, Volume 39, Issue 4, pp. 1127-1139
Savani, R., 2012, High-Frequency Trading: The Faster, the Better?, *Intelligent Systems*, Volume 27, Issue 4, pp. 70-73
Statistics Finland, 2012, Helsinki, Available: http://tilastokeskus.fi/index_en.html. Accessed from 1/08/12
Stephens M., 1974, EDF Statistics for Goodness of Fit and Some Comparisons. *Journal of the American Statistical Association*, Volume 69, pp. 730-737
Tahir, M., Keung, W., *Linkage between Stock Market Prices and Exchange Rate: A Causality Analysis for Pakistan*, National University of Singapore, Working Paper
Thomson Retuers, 2012, Thomson Reuters Tick Hoitroyt, New York, Available: https://tickhistory.thomsonreuters.com. Accessed from 1/08/12
US Securities & Exchange Commission, 2010, Concept release on equity market structure, *Release No. 34-61458; File No. S7-02-10*, Available: http://www.sec.gov/rules/concept/2010/34-61358.pdf. Accessed 5/10/12
US Securities & Exchange Commission, 2012, Self-Regulatory Organizations; The NASDAQ Stock Market LLC; NYSE Arca, Inc - Order Instituting Proceedings to Determine Whether to Approve or Disapprove Proposed Rule Changes Relating to Market Maker Incentive Programs for Certain Exchange- Traded Products *Release No. 34-67411; File Nos. SR-NASDAQ-2012-043: SR-NYSE Arca-2012-37)*. Available: http://www.sec.gov/rules/sro/nasdaq/2012/34-67411.pdf. Accessed: 2/10/12
Viljoen, T., Westerholm P.J.and Zheng H., 2013, Algorithmic Trading, Liquidity and Price Discovery: An Intraday Analysis of the SPI 200 Futures, forthcoming *Financial Review*.
Walid, C., Chaker, A., Masood, A., 2011, Stock Market Volatility and Exchange Rates in Emerging Countries: A Markov-State Switching Approach, *Emerging Markets Review*, Volume 12, pp. 272-293
Wang, X., 2010, The Relationship between Stock Market Volatility and Macroeconomic Volatility: Evidence from China, *International Research Journal of Finance and Economics*, Volume 49, pp. 159-160
Xiao, J., Brooks, R., 2009, GARCH and Volume Effects in the Australian Stock Markets, *Journal of Financial Economics*, Volume 5, pp. 79-105
Zhang F., 2010, High-Frequency Trading, Stock Volatility, and Price Discovery, Working Paper

APPENDICES

Appendix 1 – Summary of the 38 large capitalization stocks included in dataset

Company	ISIN	Ticker	GICS	Price	NoTrans	Value Traded	Market Cap	% BuyInit
Kesko Corporation B	FI0009000202	KESBV	30101030	$21.95	932925	$5,534,669,412	$1,449,253,814	64.34%
Nokia Corporation	FI0009000681	NOK1V	45201020	$13.19	25420056	$538,822,207,015	$51,919,033,411	78.18%
Uponor Oyj	FI0009002158	UNR1V	20102010	$11.21	460564	$2,068,799,954	$820,348,686	69.44%
Raisio	FI0009002943	RAIVV	30202030	$1.76	45655	$131,463,492	$229,797,733	75.04%
Finnair	FI0009003230	FIA1S	20302010	$5.78	97057	$577,867,864	$741,170,394	73.71%
Rautaruukki K	FI0009003552	RTRKS	15104050	$19.28	1788481	$11,359,037,246	$2,703,149,826	65.43%
Finnlines	FI0009003644	FLG1S	20303010	$11.06	11894	$124,267,704	$449,926,539	67.20%
Nokian Tyres Plc	FI0009005318	NRE1V	25101020	$18.49	1683604	$12,966,772,689	$2,281,170,440	66.13%
Konecranes Plc	FI0009005870	KCR1V	20106020	$19.23	1046774	$6,696,272,824	$1,170,067,216	64.88%
Stora Enso	FI0009005953	STEAV	15105020	$6.31	9190	$27,386,129	$1,120,745,383	70.56%
Stora Enso Oyj R	FI0009005961	STERV	15105020	$6.10	2351307	$21,668,431,685	$3,734,973,112	76.59%
UPM-Kymmene	FI0009005987	UPM1V	15105020	$9.49	3008589	$24,129,754,015	$5,019,621,646	75.36%
HKSCAN	FI0009006308	HKSAV	30202030	$7.30	57363	$223,339,128	$247,674,363	71.61%
Atria Group	FI0009006548	ATRAV	30202030	$12.43	28529	$139,309,593	$236,987,400	66.68%
Poyry	FI0009006696	POY1V	20201030	$11.75	93069	$440,005,165	$688,429,281	66.21%
Sponda OYJ	FI0009006829	SDA1V	40403010	$4.25	359276	$1,552,905,872	$472,183,100	76.54%
Fortum Corporation	FI0009007132	FUM1V	55101010	$21.39	4303674	$37,654,193,234	$19,079,992,090	71.51%
Metso Corporation	FI0009007835	MEO1V	20106020	$19.07	2602561	$18,321,541,387	$2,702,873,945	69.57%
Elisa Corporation	FI0009007884	ELI1V	50101020	$13.77	1537085	$10,413,918,701	$2,290,112,954	72.56%
Kesko Corporation A	FI0009007900	KESAV	30101030	$25.65	10027	$68,035,661	$814,050,210	72.60%
Comptel	FI0009008221	CTL1V	45103010	$0.99	25250	$79,996,748	$106,277,020	74.89%
Tekla	FI0009008833	TLA1V	45103010	$7.84	15568	$83,031,414	$177,146,825	59.70%
Okmetic	FI0009009054	OKM1V	45301020	$2.58	7542	$32,024,834	$43,621,295	60.79%
Aprman B	FI0009009377	CPMBV	40203010	$1.56	23986	$50,637,966	$115,027,774	82.61%
Suominen Group	FI0009010862	SUY1V	30301010	$1.13	3251	$12,469,118	$26,875,741	44.26%
Suomen Tresvo	FI0009012413	SUT1V	35102015	$1.50	14971	$51,773,920	$101,801,833	72.90%
Alma Media	FI0009013114	ALN1V	25401040	$8.18	84840	$925,023,959	$610,336,163	74.45%
Neste oil	FI0009013296	NES1V	10102030	$14.72	2421099	$15,649,821,913	$3,774,496,285	73.27%
Affecto	FI0009013312	AFE1V	45102010	$2.70	12290	$51,410,669	$58,022,018	67.41%
Cargotec Oyj	FI0009013429	CGCBV	20106020	$15.69	754687	$3,122,354,775	$857,032,786	69.25%
Oriola A	FI0009014344	OKDAV	35102010	$2.67	14086	$37,610,484	$136,913,927	68.57%
Oriola B	FI0009014351	OKDBV	35102010	$2.49	139865	$391,543,646	$224,059,884	82.68%
Orion A	FI0009014369	ORNAV	35202010	$12.52	26355	$92,647,578	$661,489,063	75.01%
Orion B	FI0009014377	ORNBV	35202010	$12.33	703160	$2,478,338,606	$1,090,814,162	74.97%
Outotec Oyj	FI0009014575	OTE1V	20103010	$22.76	1266280	$8,143,082,128	$955,770,938	64.87%
YIT Corporation	FI0009800643	YTY1V	20103010	$8.63	336035	$1,296,483,206	$1,095,090,145	72.78%
F Secure	FI0009801310	FSC1V	45103020	$2.36	69966	$298,814,855	$366,164,458	78.12%
Olvi A	FI0009900401	OLVAS	30201010	$19.38	11901	$69,088,219	$165,028,249	63.20%

Appendix 2 – Summary of OMXH25 Market Index components across sample period

Jan-08			Jul-09			
AMEAS	FI0009000285	Amer Sports Corporation	RTRKS	FI0009003552	MRLBV	FI0009000665
CGCBV	FI0009013429	Cargotec Oyj				
ELI1V	FI0009007884	Elisa Corporation	Jan-09			
FUM1V	FI0009007132	Fortum Corporation				
KCR1V	FI0009005870	Konecranes Plc	POH1S	FI0009003222	AMEAS	FI0009000285
KESBV	FI0009000202	Kesko Corporation B	RMR1V	FI0009007066	RTRKS	FI0009003552
KNEBV	FI0009013403	KONE Corporation	SAA1V	FI0009007694	SWS1V	FI0009007694
MEO1V	FI0009007835	Metso Corporation				
MRLBV	FI0009000665	M-real Corporation B	Jul-09			
NDA1V	FI0009902530	Nordea Bank AB				
NES1V	FI0009013296	Neste Oil Corporation	ORNBV	FI0009014377	RMR1V	FI0009007066
NOK1V	FI0009000681	Nokia Corporation	TLV1V	FI0009014716	SWS1V	FI0009007694
NRE1V	FI0009005318	Nokian Tyres Plc				
OTE1V	FI0009014575	Outotec Oyj				
OUT1V	FI0009002422	Outokumpu Oyj				
RTRKS	FI0009003552	Rautaruukki Corporation K				
SAMAS	FI0009003305	Sampo Plc A				
STERV	FI0009005961	Stora Enso Oyj R				
SWS1V	FI0009007694	SanomaWSOY Corporation				
TIE1V	FI0009000277	TietoEnator Oyj				
TLS1V	SE0000667925	TeliaSonera AB				
UNR1V	FI0009002158	Uponor Oyj				
UPM1V	FI0009005987	UPM-Kymmene				
WRTV	FI0009003727	Wärtsilä Corporation B				
YTY1V	FI0009800643	YIT Corporation				

Appendix 3 –period EWMA and GARCH – Conditional Volatility analysis

Several statistical models have been developed to account for issues of multicolinearity, stationarity, and non-normality that is manifest in financial time series studies. In order to account for these statistical limitations, monthly measures of volatility – stock market, macroeconomic and cyclical - are estimated by fitting an EWMA and alternatively a GARCH (1,1) model to the data. The economic cyclical factors – exchange rates and interest rates – have intra-day observable values. First, monthly volatility for these variables is estimated using Equation 2. Macroeconomic indicators – prices and output – are computed on a monthly basis based on the official monthly index values. Two comparable conditional volatility models are employed to estimate the univariate volatility for the time series returns. Firstly, an Exponentially Weighted Moving Average (EWMA) model is employed to estimate monthly volatility. Secondly, volatility is estimated by fitting a Generalised Autoregressive Conditional Heteroscadacity (GARCH) model to the monthly returns, with the computed coefficients fitted to the data to determine monthly volatility.

EWMA and GARCH models take into account the effect of serial correlation between both short-term return's and variance levels. When analyzing volatility it is essential to account for this correlation which is a pervasive trait of time-series macroeconomic variables, including inflation and output levels. Both models include the squared error of lagged return variables as well as lagged unconditional variance values. This influences the models assessment of current volatility by incorporating the effect of persistent levels of volatility across the time series. The EWMA model uses a normal distributive process to weight lagged squared return and variance values to estimate current period volatility. The model can be defined by the following function:

$$\sigma_{i,t}^2 = (1-\gamma)r_{i,t-1}^2 + \gamma\sigma_{i,t-1}^2 \qquad (A1)$$

$\sigma_{i,t}^2$ is the variance at time t, $r_{i,t-1}^2$ is the squared return of variable I at time t-1, $\sigma_{i,t-1}^2$ is variance of variable I at time *t-1*, and γ is is a weight given to each variable.

Equation 3 is used to estimate current period volatility. The equation is solved by setting the weight parameters, γ and $(1-\gamma)$, to maximise the log of the Maximum Likelihood value as defined by:

$$\max_{\sigma_{i,t}} \sum_{t=1}^{n} \left(-ln\sigma_{i,t} - \frac{r_{i,t}^2}{\sigma_i} \right) \qquad (A2)$$

Monthly volatility is estimated by solving equation (A2) for volatility.

The GARCH model encompasses a similar method of estimating monthly volatility values. The model was first introduced by Bollerslev (1986) and operates as an extension to the ARCH model developed by Engle (1982). The primary difference is that the GARCH model includes lagged values of conditional variance. In order to estimate monthly volatility of macroeconomic factor's a GARCH (1,1) model is used, which includes 1 lag of squared returns and conditional variance estimates. The data for output and prices is analyzed over a period of 21 months with lagged variables for 12 months prior also incorporated into the model. The model can be defined as:

$$\sigma_{i,t}^2 = \delta + \alpha r_{i,t-1}^2 + \beta \sigma_{i,t-1}^2 \qquad (A3)$$

Where α and β are weights assigned to lagged return and variance variables respectively, and δ is a constant calculated as the third weight, ρ, multiplied by the long-run variance rate, V_L. The weights α, β and ρ are solved for, calibrated for efficiency, and then inserted into the model from which final monthly volatility measures are calculated.

After final volatility measures are estimated it is necessary to apply consistency tests on the two models outputs to determine which one is the most accurate and relevant reflection of the macroeconomic variable volatility. Correlogram's for both the EWMA and GARCH (1,1) models are constructed to determine the level of autocorrelation that each model has accounted for. A Box-Pierce and Ljung-Box test is then conducted over 10 lags at the 95% confidence interval. Results indicate that the notion that both models have completely removed all levels of autocorrelation from the model time series can be rejected at the 95% confidence interval. However, the GARCH (1,1) does in fact attain slightly better results, which as a result is used to estimate monthly volatility of the macroeconomic factors.

Appendix 4 - HFT-Volatility Fundamental association –Daily analysis

The association between HFT activity and market volatility is also tested based on daily data across the period of analysis. The fraction of total value within a trading day that HFT's are counter-party to serves as the basis for calculating HFT activity. Volatility is more complex in nature as levels of volatility tend to be persistent over short periods of time elucidating the issue of serial correlation between daily observations. It is hence pertinent to apply a form of Autoregressive conditional heteroscadacity (ARCH) model, the Generalised ARCH (GARCH), to serially correlated volatility levels in order to more accurately and dynamically model volatility in financial time-series. The primary shortcoming of the GARCH (1,1) is that it only captures a portion of the data's skewness and leptokurtosis (Drakos, 2010). This results incorrect conditional volatility estimates if the observed volatility conditional densities are not normally distributed (Ballie, 1989).

Due to the limitations of the GARCH model, Nelson (1991) proposed an extension to the model that accounts for the effect of asset prices on conditional volatility based on their directional movement. The resulting EGARCH (1,1) conditional variance model accounts for the asymmetric responses of volatility regardless of the direction of returns. The EGARCH (1,1) model has been shown to be more efficient in modelling the volatility in returns for a large portion of financial instruments (Alexander, 2009).

An EGARCH (1,1) model is fitted to the daily volatility and return data over the 432 observations and a relevant regression model to test the fundamental daily association between HFT activity and market volatility is developed. The conditional market variance is estimated through the following regression:

$$\ln(\sigma_t^2) = \alpha_1 + \beta_1 HFT_t + \gamma_1 \ln(\sigma_{t-1}^2) + \delta_1 \frac{r_{t-1}}{\sqrt{\sigma_{t-1}^2}} + \theta_1 \frac{|r_{t-1}|}{\sqrt{\sigma_{t-1}^2}} + \vartheta_1 MKT_t + \mu_t \qquad (A4)$$

Where σ_t^2 is the daily conditional variance estimation of the EGARCH (1,1) model, HFT_t is the level of HFT activity during day *t*, r_{t-1} is the daily return for the market on day *t-1*, MKT_t is the logarithm of daily turnover value, and μ_t is the regression error term. The six coefficients of the dependent variables are estimated by applying a GMM model to the OLS regression to compute more accurate test statistics.

The daily impact of HFT on market volatility is tested assuming that at the outset HFT activity does not contemporaneously impact on conditional market volatility over the short-term. That is, there is no fundamental association between HFT and conditional volatility across the daily-level.

The variables composing the conditional variance equation are estimated through application of an EGARCH (1,1) model to the daily return and volatility raw data. The resulting regression seeks to test the HFT-volatility association through the inclusion of lagged conditional variance terms to account for any serial correlation in volatility values. Furthermore, both lagged returns and asymmetric measures are captured by the model. Finally, the logarithm of the total market daily turnover is also included as a control variable. Together these variables consummate an applicable model that tests the HFT-volatility association based on daily observations and in the absence of more efficient control measures such as those used in the monthly analysis.

Results –Daily HFT and Conditional Volatility association

The fundamental association between HFT and market volatility is tested on a daily level by estimating conditional volatility through an EGARCH (1,1) process. These volatility measures are regressed against daily HFT activity and lagged conditional variance, returns, absolute returns, and the log of daily market turnover which acts as a market-level control proxy. The results are computed in Table A1

Table A1 –

The table presents results for regression Equation A4.1:

$$\ln(\sigma_t^2) = \alpha_1 + \beta_1 HFT_t + \gamma_1 \ln(\sigma_{t-1}^2) + \delta_1 \frac{r_{t-1}}{\sqrt{\sigma_{t-1}^2}} + \theta_1 \frac{|r_{t-1}|}{\sqrt{\sigma_{t-1}^2}} + \vartheta_1 MKT_t + \mu_t$$

***, **, and * denote statistically significant p-values at 1%, 5% and 10% levels respectively

Dependent	DF Model	DF Error	SSE	MSE	R-Sq	Adj R-Sq
LN(σ_t^2)	6	432	444.4	1.0146	0.1984	0.1892

Parameter	Estimate	Std Error	t-value	Pr > |t|
α_1	-16.259	2.046	-7.950	<.0001***
β_1	-0.224	0.230	-0.970	0.330
γ_1	0.331	0.062	5.380	<.0001***
δ_1	-0.002	0.003	-0.530	0.595
θ_1	0.007	0.006	1.120	0.265
ϑ_1	0.567	0.095	5.950	<.0001***

From the results it appears that there is no congenial association between HFT and stock market volatility on an intra-day level. The regression took into account computations of conditional variance and HFT in the period January 2008 to September 2009. The coefficient of HFT in the regression outputs is not statistically significant at the 10% level. One interpretation of this result is that there is no fundamental association between HFT and market volatility when analyzing over medium-term time intervals. More importantly these results show that the results in this paper are not altered (nor confirmed) by introducing more sophisticated models of volatility such as EWMA and EGARCH.

Table 1 Descriptive Statistics

This table presents descriptive statistics for the investor account level transaction data for all Finnish stocks listed on the large cap board of the Nasdaq OMX during the investigated period 2008 and 2009.

	Total-All Periods	Average Monthly	Std Dev Monthly	Average Daily	Std Dev Daily
# ISIN's Traded	38	37.29	0.46	36.94	0.66
Trading Days	440	20.95	1.28	1	0
Value Traded	$725,786,332,806	$34,561,253,943	$15,245,046,528	$1,649,514,393	$1,077,515,191
# Transactions	51,778,812	2,465,658	484,094	117,679	51,053
# Unique Buy Accounts	195714	28520	9067	2587	1115
# Unique Sell Accounts	132747	15312	4392	1465	618
% Buyer Initiated Trade	74.52%	74.21%	5.77%	73.84%	6.36%
% Seller Initiated Trade	25.48%	25.79%	5.77%	26.16%	6.36%

Table 2 Activity by Trader Category

Descriptive statistics for each of the trading categories identified by the investor classification model of Kirilenko et al. (2011) are produced in the table representing the behavior of each group across the 21 analyzed months.

Class	% Value	% Transactions	% Unique Accounts	% Trades Initiated	Average Trade Value
Fundamental Buyers	14.6%	16.3%	28.6%	63.3%	$12,450
Fundamental Sellers	16.0%	15.9%	17.6%	38.7%	$14,112
HFT	31.8%	28.6%	0.09%	49.3%	$15,580
Intermediaries	0.62%	0.66%	3.6%	49.0%	$13,212
Opportunistic	37.0%	36.8%	1.8%	48.6%	$14,082
Small	0.19%	1.8%	48.4%	68.7%	$1,457
	Value	# Transactions	# Unique Accounts		
TOTAL	€1,451,572 Million	103,557,624	372,184		

Figure 1 Large capitalization stock price volatility over the investigated period.

The figure describes the daily average realized volatility for the sample of 38 large capitalization stocks from OMXH. The realized

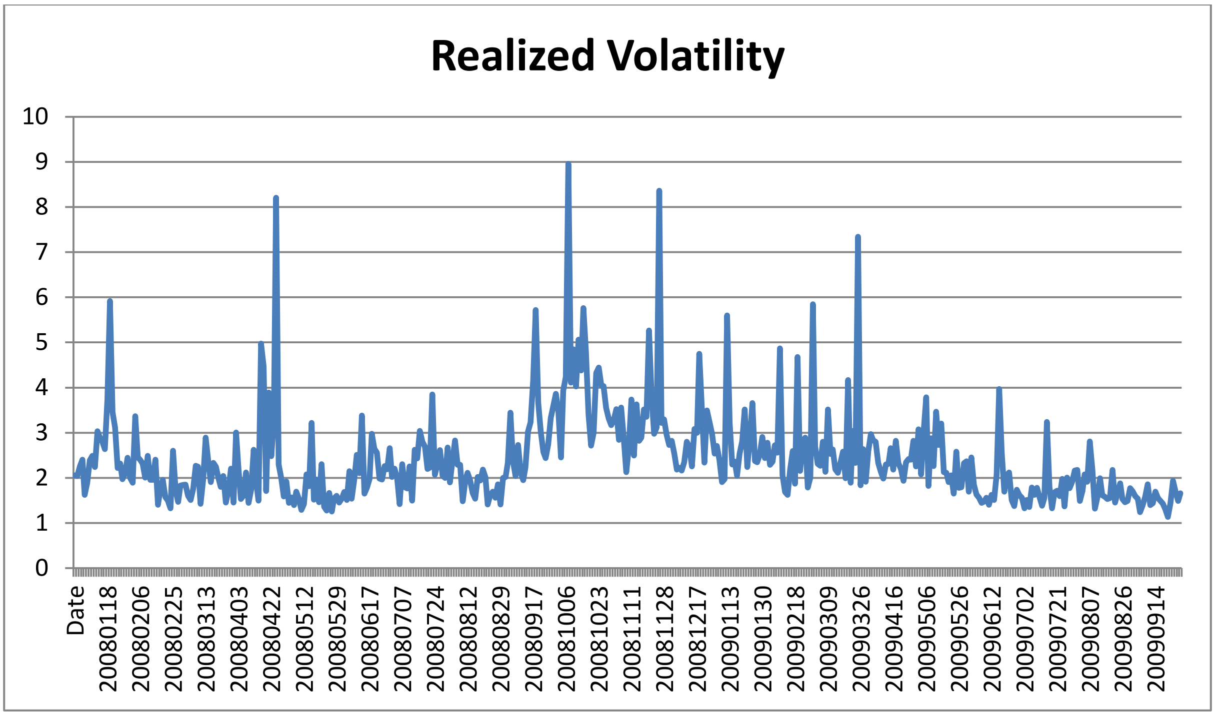

Figure 2 Daily HFT activity and volatility around co-location

The figure describes the development in the daily HFT participation in the sample of 38 large capitalization companies around the date of introduction of co-locations services, October 1, 2008. The figure also shows the corresponding daily standard deviation of the return on the OMXH25 share and futures index.

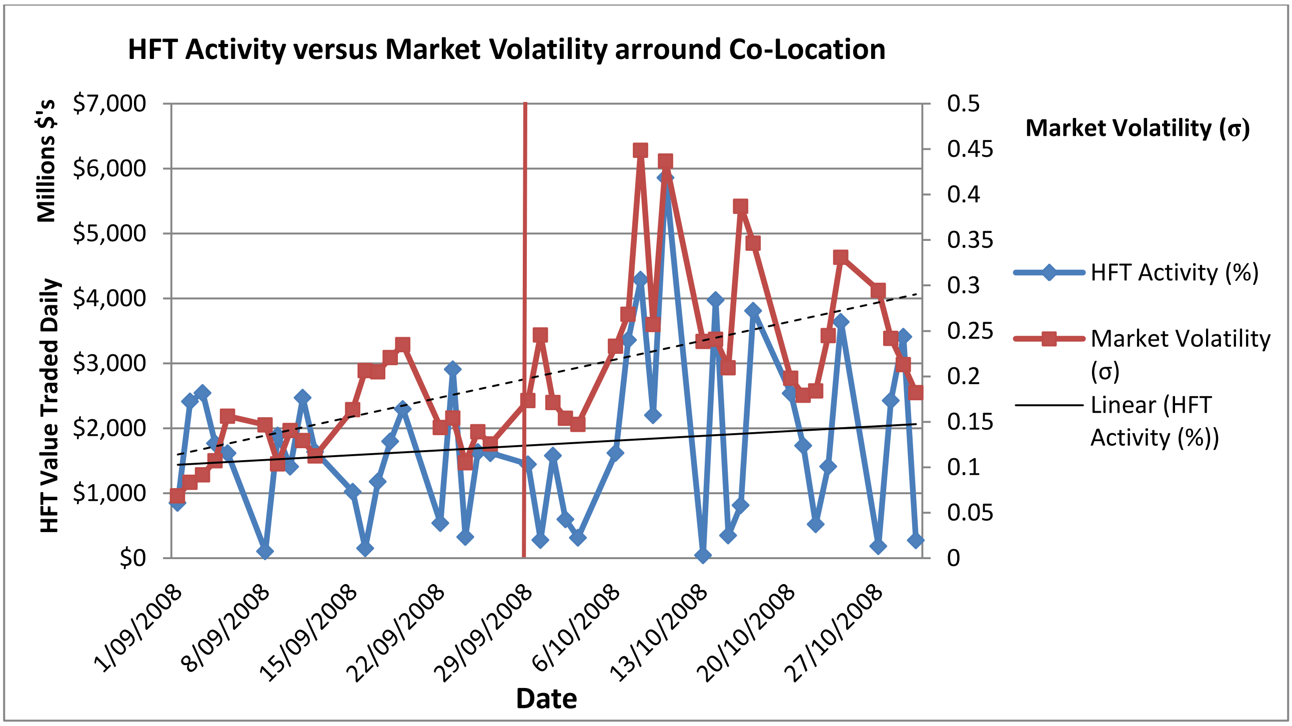

Table 3 Analysis of dynamic association volatility vs. HFT participation

This table presents the results of an analysis of the associations between realized stock volatility and HFT participation measured as fraction of value traded where a HFT identified account is a counterparty to the trade. Observations are computed for each 5 minute period during the trading day starting at 10 am and ending at 6.30 pm for the period 2008 to 2009. The following two equations are estimated:

$$VOL_{s,t} = \alpha_0 + \sum_{i=1}^{n} \alpha_1 HFT_{s,t-i} + \varepsilon_{s,t} \qquad (2)$$

$$HFT_{s,t} = \beta_0 + \sum_{i=1}^{n} \beta_2 VOL_{s,t-i} + \mu_{s,t} \qquad (3)$$

Firm fixed effects and cluster robust standard errors are reported in parenthesis. ***, ** and * indicates statistical significance of the coefficients at the 1%, 5% and 10% level, respectively.

Dependent	Realized Volatility $_t$	HFT Participation $_t$
Intercept	7.231***	0.619***
	(1.128)	(0.00004)
HFT Participation $_t$	-5.537***	
	(1.628)	
HFT Participation $_{t-1}$	-6.180***	
	(0.899)	
HFT Participation $_{t-2}$	-1.046***	
	(0.347)	
HFT Participation $_{t-3}$	-0.208	
	(0.155)	
HFT Participation $_{t-4}$	0.284*	
	(0.321)	
HFT Participation $_{t-5}$	1.679***	
	(0.321)	
Realized Volatility $_t$		-0.000248***
		(0.00008)
Realized Volatility $_{t-1}$		0.0000429
		(0.000045)
Realized Volatility $_{t-2}$		-0.000089**
		(0.000041)
Realized Volatility $_{t-3}$		-4.17E-06
		(0.000026)
Realized Volatility $_{t-4}$		-0.0000233
		(0.000019)
Realized Volatility $_{t-5}$		-0.0000179
		(0.000018)
Obs	754,338	679,403

Table 4 Analysis of dynamic association volatility vs. HFT participation

This table presents the results of an analysis of the associations between realized stock volatility and HFT participation around the introduction of co-location services on Oct 1, 2008. The following equation is estimated:

$$VOL_{s,t} = \theta_0 + \delta_1 \tau + \sum_{i=1}^{n} \theta_1 HFT_{m,t-i} + \sum_{i=1}^{n} \theta_2 VOL_{m,t-i} + \mu_{m,t} \qquad (6)$$

Firm fixed effects and cluster robust standard errors are reported in parenthesis. ***, ** and * indicates statistical significance of the coefficients at the 1%, 5% and 10% level, respectively..

Dependent	Realized Volatility $_t$		
Intercept	3.950***		
	(0.314)		
Co-location Indicator $_t$	0.069***		
	(0.0214)		
HFT Participation $_t$	-4.859***		
	(0.8682)		
HFT Participation $_{t-1}$	-4.050***	Realized Volatility $_{t-1}$	0.706***
	(0.4808)		(0.0991)
HFT Participation $_{t-2}$	1.184***	Realized Volatility $_{t-2}$	-0.409***
	(0.4237)		(0.1393)
HFT Participation $_{t-3}$	0.232	Realized Volatility $_{t-3}$	0.238*
	(0.3337)		(0.1301)
HFT Participation $_{t-4}$	0.316	Realized Volatility $_{t-4}$	-0.126
	(0.2839)		(0.0951)
HFT Participation $_{t-5}$	2.163***	Realized Volatility $_{t-5}$	0.056
	(0.2466)		(0.0485)
HFT Participation $_{t-6}$	0.510**		
	(0.2103)		
HFT Participation $_{t-7}$	-0.188		
	(0.1815)		
HFT Participation $_{t-8}$	-0.387**		
	(0.1613)		
HFT Participation $_{t-9}$	-0.860***		
	(0.1335)		
HFT Participation $_{t-10}$	-0.190*		
	(0.1129)		
Obs	681,347		
R^2	0.374		

Table 5 Analysis of dynamic association volatility vs. HFT order imbalance

This table presents the results of an analysis of the associations between realized stock volatility and HFT order imbalance measured as the difference in value between HFT buyer and HFT seller initiated trades scaled by the total number of HFT trades.

$$VOL_{s,t} = \alpha_0 + \sum_{i=1}^{n} \alpha_1 HFT\_OI_{s,t-i} + \varepsilon_{s,t} \qquad (7)$$

$$HFT\_OI_{s,t} = \beta_0 + \sum_{i=1}^{n} \beta_2 VOL_{s,t-i} + \mu_{s,t} \qquad (8)$$

Newey-West (1987) standard errors are reported in parenthesis. ***, ** and * indicates statistical significance of the coefficients at the 1%, 5% and 10% level, respectively.

Dependent	ln(Realized Volatility$_t$)	ln(HFT Order Imbalance $_t$)
Intercept	-4.015***	-4.32157
	(0.0303)	(0.0507)
ln(HFT Order Imbalance $_t$)	0.013***	
	(0.00214)	
ln(HFT Order Imbalance $_{t-1}$)	0.007***	
	(0.00209)	
ln(HFT Order Imbalance $_{t-2}$)	-0.030***	
	(0.002)	
ln(HFT Order Imbalance $_{t-3}$)	-0.037***	
	(0.002)	
ln(HFT Order Imbalance $_{t-4}$)	-0.067***	
	(0.002)	
ln(HFT Order Imbalance $_{t-5}$)	-0.092***	
	(0.00202)	
ln(Realized Volatility $_t$)		0.051***
		(0.0126)
ln(Realized Volatility $_{t-1}$)		0.026***
		(0.015)
ln(Realized Volatility $_{t-2}$)		0.001*
		(0.0104)
ln(Realized Volatility $_{t-3}$)		-0.00653
		(0.013)
ln(Realized Volatility $_{t-4}$)		0.030172
		(0.0146)
ln(Realized Volatility $_{t-5}$)		-0.022*
		(0.0127)
Obs	563,160	174,614

Table 6 Descriptive Statistics of Regression Variables

The Table presents monthly summary statistics based on observations of independent, dependent and control variables across the sample period of January 2008 to September 2009.

Variables	Frequency	Mean	Max	Min	StDev
Independent Variables					
Market Volatility					
Stock Market Volatility % *(V)*	21	45.83%	90.03%	24.75%	0.157
Dependent Variables					
HFT Investor Activity					
Value % traded by HFT's *(HFT)*	21	30.09%	42.22%	18.42%	0.076
Control Variables					
Market Factors					
Market Turnover logarithm *(MKT)*	21	24.18%	24.95%	23.59%	0.004
Cyclical Factors (% Volatility)					
Exchange Rate USD Volatility *(ERU)*	21	7.62%	22.23%	3.57%	0.051
Exchange Rate SDR Volatility *(ERS)*	21	4.96%	24.26%	1.51%	0.050
Interest Rate 3Month Eurepa Volatility *(IREA)*	21	23.07%	42.16%	5.64%	0.128
Interest Rate 3Month Euribor Volatility *(IREU)*	21	19.12%	37.00%	2.65%	0.101
Macroeconomic Factors (% Volatility)					
CPI Inflation Volatility *(InfH)*	21	5.70%	6.04%	3.13%	0.006
Producer Price Inflation Volatility *(InfP)*	21	8.34%	10.22%	4.97%	0.016
Economic Output *(OutO)*	21	19.01%	24.30%	7.15%	0.047
Industrial Output *(OutP)*	21	23.95%	28.16%	20.18%	0.032

Table 7 Estimating the fundamental association between HFT and volatility

The table presents the results for the estimation of Equation (13) using bootstrapped GMM: $VOL_t = \alpha_1 + \beta_1 HFT_t + \gamma_1 MKT_t + \rho_1 ERU_t + \rho_2 ERS_t + \rho_3 IREA_t + \mu_t$ where HFT is the proportion of HFT trades in month t. All other variables measure conditional monthly volatility estimated by fitting a EWMA process for market volatility (VOL) and a GARCH(1,1) process for market share (MKT), Euro vs. USD (ERU), Euro vs. a currency basket (ERS) and the 3 month interbank interest rate (IREA). See Appendix 3 for details. ***, **, and * denote statistically significant p-values at 1%, 5% and 10% levels respectively.

Variable	Estimate
α_1	-5.952***
	(7.14)
β_1 (HFT)	-0.484**
	(2.41)
γ_1 (MKT)	25.700***
	(7.31)
ρ_1 (ERU)	3.129***
	(4.88)
ρ^2(ERS)	-1.591***
	(3.02)
ρ^3(IREA)	0.794***
	(5.35)
Obs	21
Adj R^2	0.739

Table 8 Does HFT activity drive Volatility? The impact of HFT's on stock prices

Panel A: Pre co-location

The table reports results for the regression: $\frac{S_{i,t}-S_{i,t-1}}{S_{i,t-1}*\sigma_{i,t-1}} = \alpha + \left[\sum_{C=1}^{5} \beta_C * \frac{AI_{C,i,t}}{CLASS_{C,i,t-1}*100{,}000}\right] + \varepsilon_t$

***, **, and * denote statistically significant p-values at 1%, 5% and 10% levels respectively.

Variable	Estimate
β_FUNDAMENTAL_BUY	0.147
	(1.27)
β_FUNDAMENTAL_SELL	- 0.345
	(0.63)
β_INTERMEDIATE	- 4.33
	(1.06)
β_OPPORTUNISTIC	- 0.040**
	(2.20)
β_HFT	10.166***
	(2.70)
Obs	564,050

Panel B: Post co-location

The table reports results for the regression: $\frac{S_{i,t}-S_{i,t-1}}{S_{i,t-1}*\sigma_{i,t-1}} = \alpha + \left[\sum_{C=1}^{5} \beta_C * \frac{AI_{C,i,u}}{CLASS_{C,i,u-1}*100{,}000}\right] + \varepsilon_u$

***, **, and * denote statistically significant p-values at 1%, 5% and 10% levels respectively.

Variable	Estimate
β_FUNDAMENTAL_BUY	3.07
	(0.38)
β_FUNDAMENTAL_SELL	- 1.31
	(0.35)
β_INTERMEDIATE	-68.5
	(0.91)
β_OPPORTUNISTIC	- 4.33***
	(4.48)
β_HFT	- 24.1*
	(1.71)
Obs	564,050