# 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 also 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

## 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 perhaps the most preeminent issue in the contemporary regulatory discourse. Current debates in the regulatory space are asking the question of how markets have come to be dominated by computer-driven algorithms and what probability of success does human cognitive induced decision-making have 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.

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 Brogaard, 2010, and Hendershott, Jones and Menkveld, 2011). However, limited access to relevant data and the lack of a universal definition of HFT have stalled the development of an overall consensus. Previous studies have used HFT data such as research samples provided by the exchange, raw trading data, and regulatory, industry or self identification methodologies (For example see, Brogaard, Hendershott and Riordan (2013), Kirilenko, Kyle, Samadi, Tuzan (2011), and Hendershott and Rioridan (2012) respectively). Breckenfelder (2013) and Frino et al. (2013) have access to information about which trading channel investors use the access the market and use this to identify algorithmic trading.

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 future markets that trade a single security.

The association between HFT 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 activity 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. Analysis of the association between abnormal volatility and HFT activity, causality tests, and regressions models of market volatility versus HFT activity and period lead and lag volatility vs. HFT activity, are employed to confirm these findings. 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 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 on the 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-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 has traditionally been simple at 0.01 EURO 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. The exchange is home to companies in the technology sector like Nokia, Telia-Sonera, Elcoteq, Vaisala, Comptel and Suunto, and their presence may also have alerted international investors to the other largest companies in the exchange, typically in the industries forestry, resources and engineering. 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 Nokia stock was held by 13f registered US institutional investors during the period and most of the 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 have 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 starkly 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 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. Methods and Results

### 4.1. Definition of HFT activity

HFT's conduct operations through a hyper-active algorithmic based trading 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 and Analysis

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

Volatility measures in the empirical literature are typically based on actual security price levels attained during the time period analyzed. Volatility estimations based on the logarithmic function of high and low price values has been shown to be more accurate than measures that use daily price returns (Parkinson, 1980). Furthermore, the dispersion of measurement errors is materially reduced when using logarithmic returns which results in a more efficient method of calculating volatility (Alizadeh, 2002). In consideration of the literature on realized volatility (see Andersen and Bollerslev (1998) and Andersen, Bollerslev, Diebold, and Labys (2003)), I also apply the realized range measure suggested by Martens and van Dijk (2007) and implemented by Kalev and Duong (2008), which combines aspects of both range based and realized volatility metrics suitable for intraday data analysis (not reported in this version). When volatility metrics are computed using intraday observations these may be affected by microstructure noise and infrequent trading. Martens and van Dijk (2007) propose a bias-correction procedure where the realized range is scaled by the ratio of the average level of daily range and the average level of the realized range over previous trading days. This section identifies daily and intra-day stock volatility measures and equations to calculate their realized values.

Volatility for individual stocks is measured across time-periods of 10 seconds, 30 seconds, 1 minute, 15 minutes, 1 hour, 3 hours, and 1 trading day. The OMXH25 exchange-wide volatility is measured at 10 second, 1 minute, 15 minute, 1 hour, and 1 day intervals. These measures of estimating volatility are calculated through the method of period-length logarithmic high-low levels. Period volatility is defined by the equation:

$$VOL_{s,t} = \ln\left(\frac{s_{t,max}}{s_{t,min}}\right) \qquad (1)$$

Where $VOL_{s,t}$ is the period t volatility measure for security s, $s_{t,max}$ is the maximum value for security s in time period t, and $s_{t,min}$ is the minimum value for security s in time period t. This measure captures the level of volatility across daily and intra-day periods.

Monthly volatility is based on the average daily logarithmic high and low range values. This measure is used to calculate Monthly volatility for the OMHX25 Index and several macroeconomic control variables that are reported on an intra-day basis. Monthly volatility is calculated as:

$$VOL_{I,m} = \left(\frac{1}{n}\right)\sum_{d=1}^{n}\left(\ln\left(I_{d,max}\right) - \ln\left(I_{d,min}\right)\right) \qquad (2)$$

Where $VOL_{I,m}$ is the monthly volatility measure for index I which is measured based on the average of the logarithm of the daily, d, maximum value, $\ln\left(I_{d,max}\right)$, minus the daily minimum value, $\ln\left(I_{d,min}\right)$. This measure captures the level of volatility across monthly periods based on intra-day data. All measures of monthly stock market volatility are then annualized by multiplying $VOL_{I,m}$ by the square root of 12.

Market volatility is calculated based on the OMX Helsinki 25 market index (OMXH25) which is the HEX’s gross value weighted index consisting of the 25 most actively traded common stocks. The components of this index as officially stated at semi-annual intervals are identified in Appendix II. This index captures the broad movements by the primary components of the OMXH market and is the primary large cap Finnish index that is tracked by traders. Tick-by-tick second-level data is available for the OMXH25 Index. As a result, intra-day market volatility is calculated using Equation 1 and monthly volatility is calculated using Equation 2.

Individual stock and market wide volatility are measured across the period of January 2008 to September 2009. This period accounts for both the market's sharp contraction leading into March 2009, as well as the partial recovery in the latter half of 2009. This period represents historically high levels of volatility for the OMXH. Figure 1 depicts the level of monthly market volatility across the period superimposed on the OMXH25 index level. Table 3 provides summary statistics for the annualized volatility estimates of stock market volatility on a daily and monthly basis. These statistics reinforce the highly volatile nature of the period being analyzed.

### 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 analysed 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 smooths out 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 manifested 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, with its largest trading currency, the US Dollar (ERU), and the universally excepted internationally weighted instrument, SDR's (ERS). In order to account for 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, see Appendix III.

An important market level control variable 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. These nine control variables are measured based on their volatility and used as control variables when analyzing the HFT and market volatility association.

### 4.4 Are HFT and Volatility dynamically related?

HFT is an important component of modern financial markets with the empirical literature showing their role in positive liquidity provision (Hendershott, 2011) and contribution to price discovery (Brogaard, Hendershott and Riordan 2013). These investors generally operate during regular trading periods by providing liquidity at the bid-ask spread. During periods of extreme volatility however, markets could be severely distressed if HFT activity evaporates as their trading strategies are rendered redundant through the extant high risk environment of the market. Events such as the 2010 'Flash Crash' and more recently the wide fluctuation in opening prices for Blue Chip stocks on the ASX show the risks inherent in fully electronic market trading systems when liquidity evaporates and a cascade develops. Whilst several studies have rebuked the role of HFT's in facilitating the Flash Crash (Cvitanic, 2010) there does exist a systematic risk to the integrity of financial markets due to HFT behavior during periods of sustained volatility.

#### 4.4.1 Model definition

Volatility levels are computed across the entire dataset based on the fluctuation in individual stock prices across equally spaced time period intervals. The method of testing the association between HFT and volatility has been extended from that utilized by Brogaard et.al (2013). The measure used to calculate intra-day period volatility is set out in Equation 1. Observations that record nil stock volatility during a certain period t are erased from the dataset. These final observations are based on separate time interval analyses based on activity during periods of 10 and 30 seconds, 1 and 15 minutes, 1 and 3 hours. Abnormal Volatility values are then calculated as follows:

$$ABN\_VOL_{s,t} = \frac{VOL_{s,t} - \overline{VOL}_s}{\overline{VOL}_s} * \frac{1}{\sigma_s} \qquad (3)$$

$VOL_{s,t}$ is the volatility of stock $s$ across period $t$ as measured by the lognormal function of the maximum price range of stock $s$ over period $t$. $\overline{VOL}_s$ is the average price volatility for stock $s$ across all time periods. $\sigma_s$is calculated as the standard deviation of all $VOL_{s,t}$ values for stock $s$ across all time periods. $ABN\_VOL_{s,t}$ is thus computed for each individual stock and time interval where a value of $VOL_{s,t}$ is positive. Value's of $ABN_VOL_{s,t}$are then ranked from smallest to largest and placed into 20 groups ($p = 20$) of similar levels of abnormal volatility.

For each stock and time period that attains a positive Abnormal-Volatility value, the anomalistic level of HFT activity in that stock during the period is calculated. In this context HFT activity is defined as the percentage fraction of total value traded in stock $s$ during period $t$ by HFT Classed traders. Given that the aim is to determine the association between high and low proportions of HFT activity in varying levels of market volatility, a transformation is created to express the abnormal HFT activity as follows:

$$ABN\_HFT_p = \sum_{ABN\_VOL_{s,t} \cap p} \frac{1}{N_p} * \left( \frac{HFT_{s,t} - \overline{\overline{HFT}}_s}{\overline{HFT}_s} \right) \qquad (4)$$

$HFT_{s,t}$ is the level of activity of HFT in stock $s$ during period $t$ as a percentage of value traded. $\overline{HFT}_s$ is the expected level of HFT activity in stock $s$ across all time periods $t$. $N_p$is the number of $ABN\_HFT_{s,t}$ observations in bin *p*. $ABN\_HFT_p$ explains the level of deviation in the fraction of shares that HFT's trade in a specific stock during period $t$ from the mean expected HFT fraction over the period.

The results are presented graphically plotting values of $ABN\_HFT_p$ on the Y-axis pertaining to their corresponding $ABN\_VOL_{s,t} \cap p$ bin, which is plotted on the X-axis. Essentially this visual representation will plot abnormal levels of HFT activity that is prevalent during periods of low to high stock volatility.

#### 4.4.2 Results and Analysis

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. The abnormal association between market volatility and HFT activity is analyzed through two dimensions – trade direction and liquidity provision. Figure 3 compares the association ABN-HFT before and after 'Proximity Services' were introduced on the HEX. Abnormal HFT and price volatility are calculated over the following time intervals – 10 seconds, 30 seconds, 1 minute, 15 minutes, 1 hour, and 3 hours. The results indicate that as volatility increases there is a general decrease in HFT activity across all time intervals. This analysis is most acute for longer time intervals as the slope for the longer horizon graphs is greater. The results are interesting from one perspective. For all the periods analyzed, 1 Hour, 1 Minute, 10 Seconds, the results for periods after Co-location are slightly 'flatter' than those for periods before Co-location. That is, HFT-deviations from their expected levels are lower in periods of low and high volatility when testing after co-location. These results indicate that Co-location technologies impact on volatility levels.

### 4.5. Does HFT precede Volatility? Bi-Directional Granger Causality analysis

The causal statistical association between HFT activity and stock market volatility is analyzed by testing for Granger causality (1969) between the two variables. Granger causality between

HFT activity and stock market volatility is tested in both directions with the impact from certain underlying variables being controlled for. The association is tested over time interval periods of 1 day, 1 hour, 15 minutes, 1 minute and 10 seconds. Furthermore, causality is tested on a market-wide basis as well as at the stock-level. Thus, if lagged variables of HFT activity help predict volatility levels at time t, then HFT Granger cause's volatility. The opposite association, that volatility Granger causes HFT activity is also tested.

### 4.5.1 Equations

To test for Granger Causality between HFT activity and volatility the following two sets of equations are estimated:

$$VOL_{s,t} = \alpha_0 + \sum_{i=1}^{n} \alpha_1 HFT_{s,t-i} + \sum_{i=1}^{n} \alpha_2 VOL_{s,t-i} + \alpha_3 \delta + \varepsilon_{s,t} \qquad (5)$$

$$HFT_{s,t} = \beta_0 + \sum_{i=1}^{n} \beta_1 HFT_{s,t-i} + \sum_{i=1}^{n} \beta_2 VOL_{s,t-i} + \beta_3 \delta + \mu_{s,t} \qquad (6)$$

Where $VOL_{s,t}$ is the volatility logarithmic function of the high and low price for stock $s$ in period $t$ as derived from Equation 1. $HFT_{s,t}$ is the fraction of total stock $s$ turnover during period $t$ in which HFT investors are involved. $\delta$ is representative of the control variables utilised in the model. The two control variables include the log of the total value of trades in stock $s$ during the trading day that period $t$ occurs in, and the EURO STOXX 50 Volatility Index level at time period $t$. $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ and $\beta_0, \beta_1, \beta_2, \beta_3$ 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 which reflect the variations in VOL and HFT that are not a result of lagged independent and control variables.

$$VOL_{m,t} = \gamma_0 + \sum_{i=1}^{n} \gamma_1 HFT_{m,t-i} + \sum_{i=1}^{n} \gamma_2 VOL_{\mathrm{m},t-i} + \gamma_3 \tau + \varepsilon_{m,t} \tag{7}$$

$$HFT_{m,t} = \theta_0 + \sum_{i=1}^{n} \theta_1 HFT_{m,t-i} + \sum_{i=1}^{n} \theta_2 VOL_{m,t-i} + \theta_3 \tau + \mu_{m,t} \tag{8}$$

Where $VOL_{m,t}$ is the logarithmic volatility of the high and low value for the OMXH25 market index $m$ for period $t$ as derived from Equation 1. $HFT_{m,t}$ is the value of trades conducted by HFT's as a fraction of total turnover across the sample dataset during period $t$. $\tau$ represents a set of control variables used in the model including the total log value of trades on the Helsinki Exchange market $m$ during the trading day that period $t$ occurs in, and the EURO STOXX 50 Volatility Index level at time period $t$. $\gamma_0, \gamma_1, \gamma_2, \gamma_3$ and $\theta_0, \theta_1, \theta_2, \theta_3$ are the coefficients tested in the Volatility and HFT dependent variable regression respectively. $\varepsilon_{m,t}$ and $\mu_{m,t}$ are error disturbance terms which reflect the variations in VOL and HFT respectively that are not a result of lagged independent and control variables.

The first set of equations looks at the association between dependent variables and lagged values of the independent variables at the stock-level. The second set of equations takes accounts for HFT activity and Volatility based on a market-wide interpretation of the two variables. The causal association between HFT activity and market volatility is analysed through the Granger causality equations estimated above. Each equation is tested across different evenly spaced lags of 3, 5, and 10 ($n$=3,5,10) period lags for each variable. The model also independently and implicitly determines the most accurate and efficient number of lags using an Akaike-Schwarz information criterion (see Burnham and Anderson, 2002)

Wald tests are conducted on each of these equations to determine evidence against the null hypothesis that $x$ does not Granger cause $y$. That the parametric coefficient's of lagged variables are statistically different from zero is tested in an F-test.

### 4.5.2 Results and Analysis

Table's 4 and 5 show the respective stock-level and market-level Granger Causality results for time interval periods of 10 seconds, 1 minute, 15 minutes, 1 hour and 1 trading day. Symbol V->HFT denotes results for testing whether volatility Granger causes HFT, and HFT->V denotes results regarding whether HFT Granger causes volatility. The first column for each time interval tested indicates the p-value derived from the F-statistic measured by the regression (Fpval1). Similarly, the implied p-value from the Wald test Chi-square statistics expressed in the adjacent columns (Wpval2). The null hypothesis of no Granger Causality is rejected at the 1% level ($p<0.01$), 5% level ($p<0.05$), and 10% level ($P<0.01$) and emphasised within the results.

Results at the stock-level expressed in Table 4 provide very strong evidence for Granger causality in both directions across all periods and number of time lags used in the model. These results are generally in line with those derived by Brogaard (2011a).

The Market-level results in Table 5 provide further evidence of Granger causality in both directions. HFT Granger causing volatility and volatility granger causing HFT is strongly supported over the 10 second, 1 minute and 15 minute time frames. This finding is particularly relevant in the context of the short-horizon investment strategies driven by the low-latency environment in which HFT's operate on the HEX during this period. However, there is an evident weakening of the statistical causality in both directions over time frames longer than 1 hour. An interpretation of this result is that whilst HFT trading strategies are tuned acutely to holding positions of stocks over ultra-short periods of seconds and milliseconds, lagged volatility levels of 10 seconds and upto 15 minutes are statistically relevant when analysing current HFT levels. Thus, this asserts that criticism levelled at HFT's that their large volume of order-flow tends to 'evaporate' as soon as volatility levels increase may not be essentially correct. HFT's tend to take into account volatility over a longer period when deciding whether to execute a particular trading strategy or not.

The bi-directional causality is weaker during longer time-periods there is stronger evidence to support the assertion that Volatility Granger causes HFT, than in the opposite direction of HFT Granger causing Volatility. These results can be understood within a framework of varying volatility. Given that HFT's seem extremely adverse to operate in periods of high volatility, and the finding that there is a 'preceding' association of HFT to volatility, the significant impact of volatility on HFT activity becomes apparent. HFT's may exit the market during sustained periods of high volatility. The following sections will attempt to determine whether the variables efficaciously explain the behaviour of the other, or whether the association is best explained by an underlying third variable that drives both processes.

### 4.6 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 periods. 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 controls for macroeconomic, cyclical and market factor structures are introduced.

#### 4.6.1 Monthly Statistics

Monthly statistics for the 11 variables used in the regressions are conveyed 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 monthly time-series of market volatility and HFT depicted in Figure 4. This Figure demonstrates that HFT activity follows a contrapositive pattern to market volatility, with high levels of HFT activity in the first half of 2008, corresponding to lower volatility levels. As the period studied progresses into the latter half of 2009 there is a discernible decrease in HFT activity as volatility levels generally tend to increase.

### 4.6.2 Model

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 utilises 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. The estimated equations are:

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

$$VOL_t = \alpha_1 + \delta_1 ERU_t + \delta_2 ERS_t + \delta_3 IREA_t + \delta_4 IREU_t + \vartheta_n MACRO_t + \epsilon_t \tag{10}$$

$$VOL_t = \alpha_1 + \gamma_1 MKT_t + \rho_n CYCLICAL_t + \vartheta_n MACRO_t + \tau_t \tag{11}$$

$$VOL_t = \alpha_1 + \beta_1 HFT_t + \gamma_1 MKT_t + \rho_n CYCLICAL_t + \vartheta_n MACRO_t + \mu_t \tag{12}$$

The Volatility, HFT activity, Market-level, Cyclical and Macroeconomic variables in Table 6, with their observed values at time *t* included in the regression. $MACRO_t$, $CYCLICAL_t$, $MKT_t$ 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 which tests the association between HFT activity and stock market volatility.

The regressions are estimated using three methodologies when calculating HFT activity. HFT activity is defined as the participation of HFT investors in a) any transaction across the period, b) transactions where HFT's demand liquidity and c) transactions where HFT's supply liquidity. 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.

For robustness regressions (9) to (12) are also estimated using the realized range measure of volatility, Equation 3, instead of the range based volatility, Equations 1 and 2. These results are reported in Appendix 5.

### 4.6.3 Results and Analysis

The final regression model in this section, Equation (12) tests the dependent conditional market volatility variable against the fraction of HFT monthly explanatory variable in addition to the SDR and US Exchange rate volatility, the 3-month Eurepa Interest Rate and the log of market turnover as control variables. HFT activity is defined across three dimensions based on whether these investors are transacting on the market (Panel A), demanding liquidity (Panel B), or supplying liquidity (Panel C). Regression analyses employ GMM model transformations with results reported in Table 7 Panels A, B and C:

The results provide strong evidence to assert a fundamental association between HFT activity and stock market volatility. Across all three Panels the HFT coefficient parameter estimate $\beta_1$ is significantly different from 0 at the with a p-value less than 0.05. When testing all HFT activity on volatility the estimated coefficient parameter of -0.48 perpetuates a p-value of 0.029 presenting strong evidence against the null hypothesis of no fundamental association between HFT and market volatility. This finding indicates that based on observations through the period, conditional market volatility has been on average 0.48% lower for every 1% increase in HFT activity. The results from HFT demand and supply driven activity indicate similar coefficients of -0.51 and -0.46 which are both significant at the 5% level. 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. This is indicative of the increase in R-squared values from 0.69 in regression 3 to 0.74 in regression 4 (Panel A). There is also clear evidence that the four control variables are statistically related to market volatility which asserts their relevance as control variables when regressing against market volatility. Each control variable in the model is statistically significant across all three Panels at the 1% level of significance.

By testing the association with relevant macroeconomic, cyclic and market-level control variables incorporated, the evidence is further strengthened. Adjusted R-squared values for all levels of HFT activity in Regression 4 are greater than 0.73 which dictate the strength of the model and the weighting given to evaluating the results. The findings indicate that there is a strong and fundamental association between HFT and volatility across the sample. 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-capitalisation 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 capitalisation across the sample stocks. During the investigated period HFT participation varies between about 20% and 30%.

### 4.7. Does HFT activity drive Volatility? The impact of HFT's on stock prices

Volatility is both a natural occurrence and an accepted risk of investor's participation in financial markets. 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.

### 4.7.1 Model

Following Kirilenko et al. (2011) I estimate a regression model that attempts to model how contemporary HFT investors drive future price changes, and market volatility, through the execution of their trading strategies. Prior period HFT activity is modeled as an explanatory variable for price change computations, weighted by each individual stocks contemporary volatility level. Furthermore, 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:

$$\frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1} * \sigma_{i,t-1}} = \alpha + \left[\sum_{C=1}^{5} \beta_C * \frac{AI_{C,i,t}}{CLASS_{C,i,t-1} * 100{,}000}\right] + \mu_t \qquad (13)$$

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

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 $t-1$. $\beta_C$ is the coefficient of determination for each of the five separate investor classes, $\alpha$ is the regression intercept term and $\mu_t$ is the error term. Equation (13) is estimated pre and Equation (14) is estimated post Co-location.

The dependent variable in the regression represents the realised 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.

### 4.7.2 Results and Analysis

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 computed. The Beta coefficient 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 not in line with the fundamental and negative association that was determined in section 6.3 in analysis of the association between monthly levels of conditional volatility and HFT activity.

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 prior to 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 -24.13. 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 can be interpreted as evident of the 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 market volatility of statistical and economic significance. Such a association is most pronounced during periods of very short intervals. However, an analysis of the association at a monthly level also indicates that, when controlling for third variables, the association is statistically significant and negative. The results strongly assert a negative correlation between stock market volatility and HFT to a greater extent than those of and Hendershott, Jones and Menkveld (2011) and Brogaard, Hendershott and Riordan (2013).

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. Firstly, HFT's have been shown to decrease price deficiencies in the bid-ask spread significantly (Hendershott, Jones and Menkveld, 2011). Secondly, evidence of a fundamental and negative association between HFT activity and market volatility is presented. So as HFT's enter into the market volatility levels tend to dampen. 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) seems to detract from this notion as HFTs actually seek to rebalance positions by competing for liquidity at periods of high volatility.

The implications of this study in informing future regulations shows in the findings related to how HFT's change their behaviour 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, 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

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

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

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

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

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
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

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 GARCH & EWMA – 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 $\gamma$ 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, $\gamma$ and $(1-\gamma)$, to maximise the log of the Maximum Likelihood value as defined by:

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

By solving the equation above, and calculating the parameters, monthly volatility is estimated.

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 analysed 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 $\alpha$ and $\beta$ are weights assigned to lagged return and variance variables respectively, and $\delta$ is a constant calculated as the third weight, $\rho$, multiplied by the long-run variance rate, $V_L$. The weights $\alpha$, $\beta$ and $\rho$ 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. As identified in sections 5.3 it is 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:

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

Where $\sigma_t^2$ 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 $\mu_t$ 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 & 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($\sigma_t^2$) | 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 evident 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. This finding does not contribute evidence to reject the null hypothesis that coefficients of daily HFT are statistically different from zero when regressing against volatility. An interpretation of this result is that there is no fundamental association between HFT and market volatility when analyzing over medium-term time intervals. This result is in line with the previous Granger causality findings which indicate that within contemporary periods across one trading day the two factors are not fundamentally linked, most likely due to that HFT do not hold positions overnight. In the Granger causality lagged variables of both HFT and market volatility help forecast current levels when analyzing very short-term intervals however.

Table 1 Descriptive Statistics

| | 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 | | |

Table 3 - Volatility summary statistics for OMXH25 and individual Stocks

The table provides summary statistics for the annualised volatility estimates of stock market volatility on a daily and monthly basis.

| Period | Average | N | Maximum | Minimum | Standard Deviation |
|---|---|---|---|---|---|
| OMXH25 (Daily) | 45.9% | 439 | 159.2% | 11.2% | 24.1% |
| OMXH25 (Monthly) | 45.8% | 21 | 90.0% | 24.8% | 15.7% |
| ISIN (Daily) | 65.0% | 38 | 87.9% | 41.4% | 14.0% |

Figure 1 OMXH25 Market index and monthly volatility levels across the period.

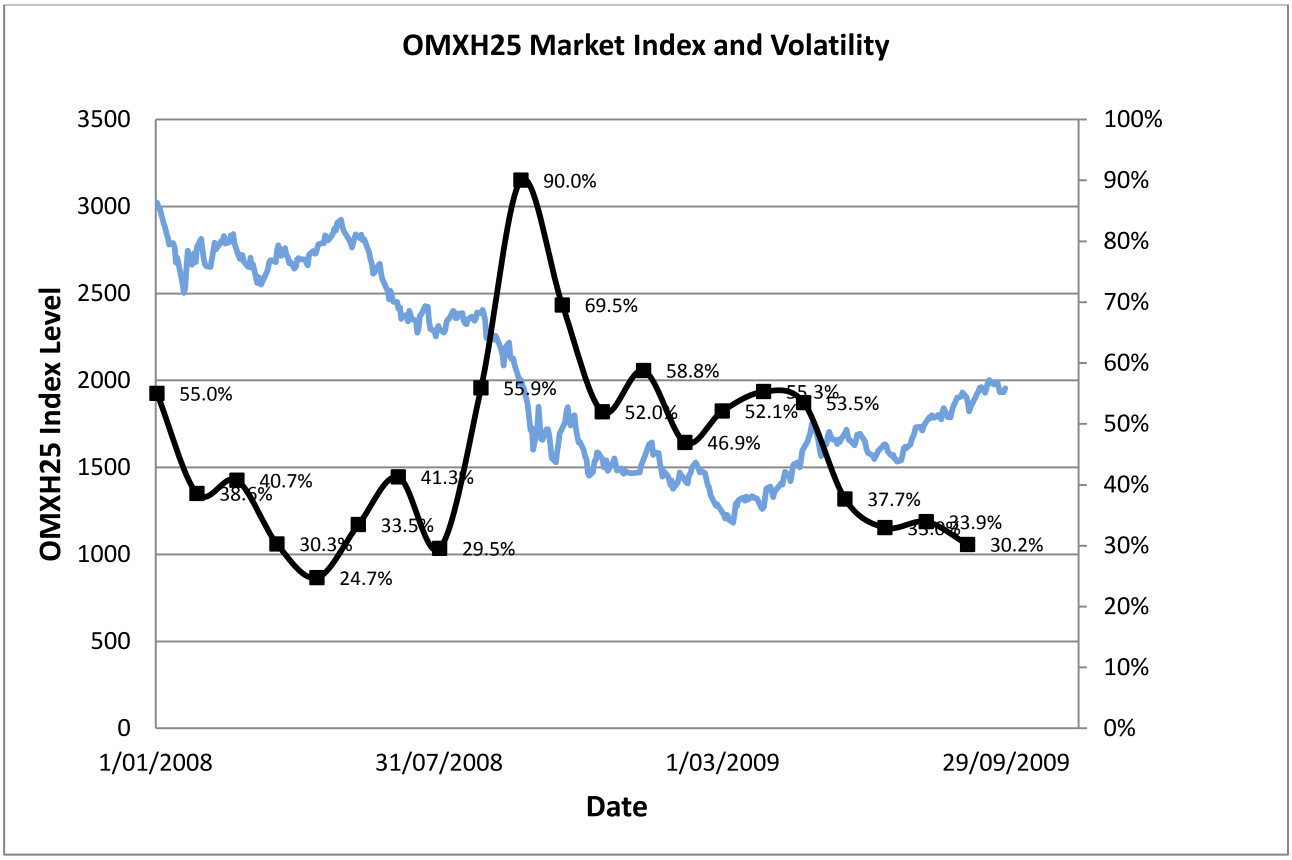


Figure 2 Daily HFT activity and volatility around co-location

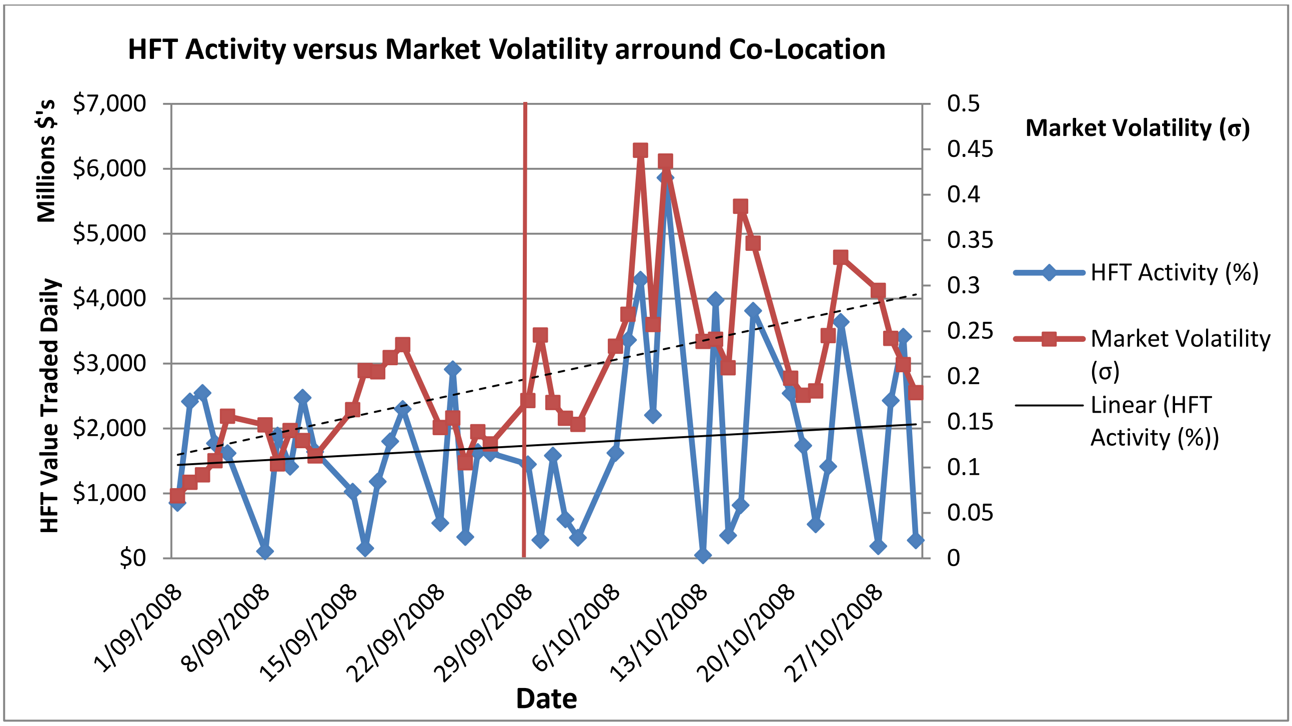


Figure 3 Abnormal HFT activity and abnormal stock volatility around co-location

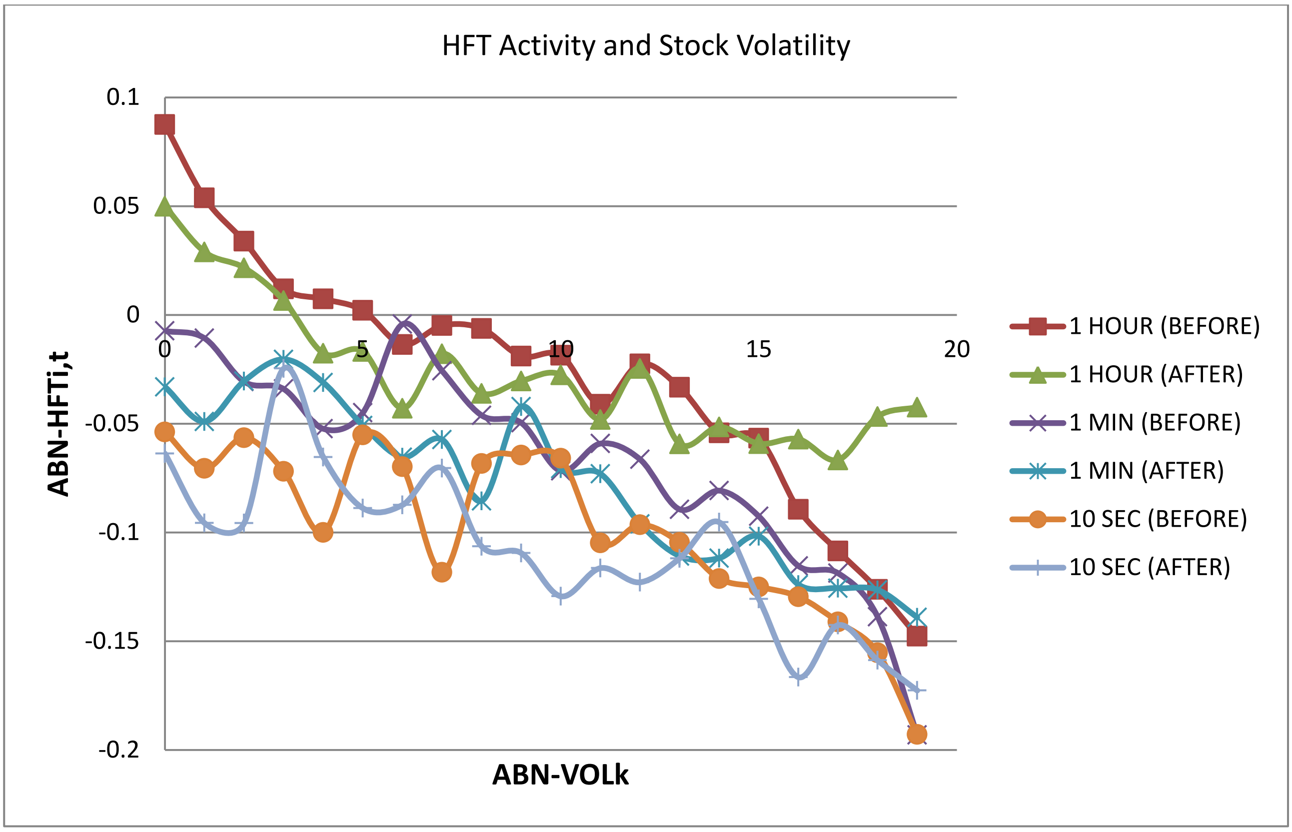


Table 4 Stock level causality tests

Granger causality results at the stock-level for HFT to V (Range based volatility) and V to HFT across different time intervals. The p-value for the Wald test statistic is reported. ***, **, and * denote statistically significant p-values at 1%, 5% and 10% levels respectively.

| Stock Level | Day<br>*Wald Test p-value* | Hour 1<br>*Wald Test p-value* | Min 15<br>*Wald Test p-value* | Min 1<br>*Wald Test p-value* | Sec 10<br>*Wald Test p-value* |
|---|---|---|---|---|---|
| *10 Time Lags* | | | | | |
| V to HFT | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** |
| HFT to V | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** |
| *5 Time Lags* | | | | | |
| V to HFT | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** |
| HFT to V | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** |
| *3 Time Lags* | | | | | |
| V to HFT | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** |
| HFT to V | 0.00*** | 0.00*** | 0.00*** | 0.00*** | 0.00*** |

Table 5 Market Level Causality Tests

Granger causality results at the market-level for HFT to V (Range based volatility) and V to HFT across different time intervals. The p-value for the Wald test statistic is reported. ***, **, and * denote statistically significant p-values at 1%, 5% and 10% levels respectively.

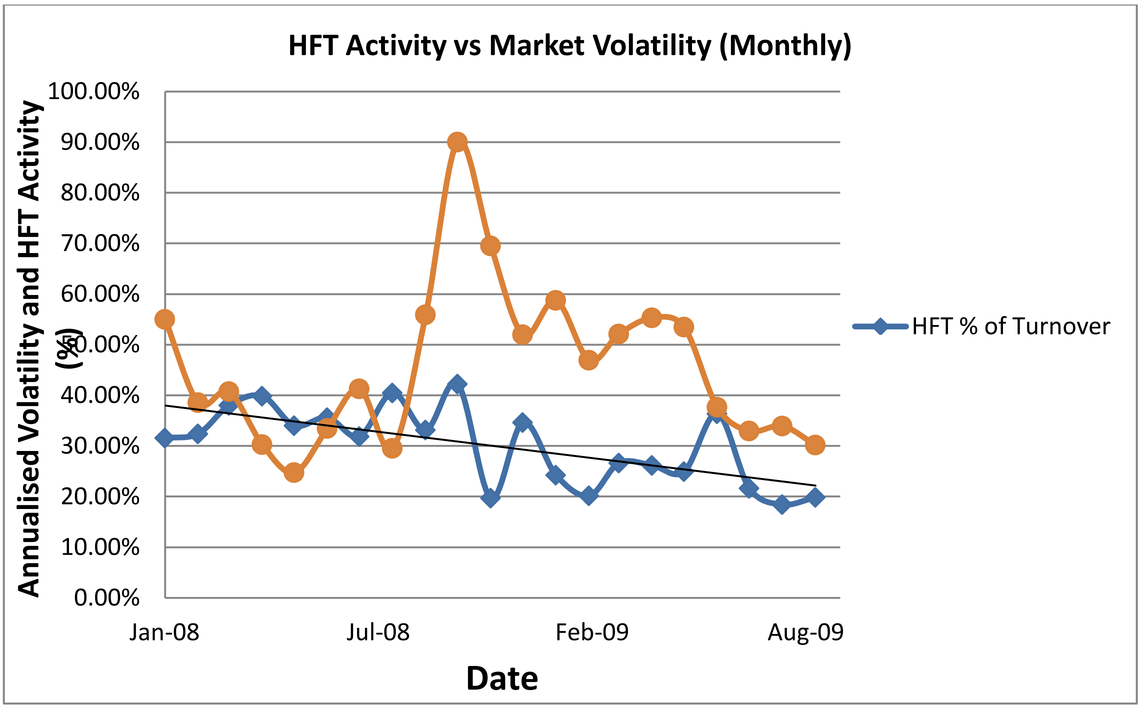
| Market Level | Day<br>*Wald Test p-value* | Hour 1<br>*Wald Test p-value* | Min 15<br>*Wald Test p-value* | Min 1<br>*Wald Test p-value* | Sec 10<br>*Wald Test p-value* |
|---|---|---|---|---|---|
| *10 Time Lags* | | | | | |
| V->HFT | 0.09* | 0.02** | 0.00*** | 0.00*** | 0.00*** |
| HFT->V | 0.35 | 0.99 | 0.00*** | 0.00*** | 0.00*** |
| *5 Time Lags* | | | | | |
| V->HFT | 0.04** | 0.39 | 0.01*** | 0.00*** | 0.00*** |
| HFT->V | 0.14 | 0.29 | 0.00*** | 0.00*** | 0.00*** |
| *3 Time Lags* | | | | | |
| V->HFT | 0.04** | 0.18 | 0.00*** | 0.00*** | 0.00*** |
| HFT->V | 0.08* | 0.13 | 0.00*** | 0.00*** | 0.00*** |

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 | FREQ | AVE | 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 3-M Eurepa Volatility *(IREA)* | 21 | 23.07% | 42.16% | 5.64% | 0.128 |
| Interest Rate 3-M 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 |

Figure 4 Monthly HFT fraction of turnover and Market Volatility (January 2008 - September 2009)



## Table 7 Estimating the fundamental association between HFT and volatility

The table presents the results for the estimation of Equation (12) using bootstrapped GMM:
$VOL_t = \alpha_1 + \beta_1 HFT_t + \gamma_1 MKT_t + \rho_1 ERU_t + \rho_2 ERS_t + \rho_3 IREA_t + \mu_t$

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

Panel A - MODEL Procedure – ALL HFT TRADES

| Dependent | DF Model | DF Error | SSE | MSE | R-Sq | Adj R-Sq |
|---|---|---|---|---|---|---|
| VOLATILITY | 6 | 15 | 0.0962 | 0.00458 | 0.804 | 0.739 |

| Parameter | Estimate | Std Error | t-value | Pr > \|t\| |
|---|---|---|---|---|
| α_1 | -5.952 | 0.834 | -7.140 | <.0001*** |
| β_1 | -0.484 | 0.201 | -2.410 | 0.029** |
| γ_1 | 25.700 | 3.518 | 7.310 | <.0001*** |
| ρ_1 | 3.129 | 0.641 | 4.880 | 0.000*** |
| ρ_2 | -1.591 | 0.527 | -3.020 | 0.009*** |
| ρ_3 | 0.794 | 0.148 | 5.350 | <.0001*** |

Panel B - MODEL Procedure - HFT DEMAND

| Dependent | DF Model | DF Error | SSE | MSE | R-Sq | Adj R-Sq |
|---|---|---|---|---|---|---|
| VOLATILITY | 6 | 15 | 0.0953 | 0.00454 | 0.806 | 0.742 |

| Parameter | Estimate | Std Error | t-value | Pr > \|t\| |
|---|---|---|---|---|
| α_1 | -5.891 | 0.821 | -7.180 | <.0001*** |
| β_1 | -0.513 | 0.209 | -2.460 | 0.027** |
| γ_1 | 25.472 | 3.457 | 7.370 | <.0001*** |
| ρ_1 | 3.107 | 0.631 | 4.920 | 0.000*** |
| ρ_2 | -1.590 | 0.523 | -3.040 | 0.008*** |
| ρ_3 | 0.778 | 0.146 | 5.330 | <.0001*** |

Panel C - MODEL Procedure - HFT SUPPLY

| Dependent | DF Model | DF Error | SSE | MSE | R-Sq | Adj R-Sq |
|---|---|---|---|---|---|---|
| VOLATILITY | 6 | 15 | 0.097 | 0.00462 | 0.803 | 0.737 |

| Parameter | Estimate | Std Error | t-value | Pr > \|t\| |
|---|---|---|---|---|
| α_1 | -6.004 | 0.847 | -7.090 | <.0001*** |
| β_1 | -0.457 | 0.194 | -2.350 | 0.033** |
| γ_1 | 25.890 | 3.576 | 7.240 | <.0001*** |
| ρ_1 | 3.146 | 0.651 | 4.830 | 0.000*** |
| ρ_2 | -1.591 | 0.531 | -3.000 | 0.009*** |
| ρ_3 | 0.808 | 0.151 | 5.350 | <.0001*** |

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 regression: $\frac{S_{i,t}-S_{i,t-1}}{S_{i,t-1}*\sigma_{i,t-1}} = \alpha + \left[\sum_{C=1}^{5} \beta_C * \frac{AI_{C,i,t}}{CLASS_{C,i,t-1}*100{,}000}\right] + \mu_t$

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

| Parameter | Estimate |
|---|---|
| *b*_FUNDAMENTAL_BUY | 0.147 |
| | (1.27) |
| *b*_FUNDAMENTAL_SELL | - 0.345 |
| | (0.63) |
| *b*_INTERMEDIATE | - 4.33 |
| | (1.06) |
| *b*_OPPORTUNISTIC | - 0.040** |
| | (2.20) |
| *b*_HFT | 10.166*** |
| | (2.70) |

Panel B: Post co-location

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

| Parameter | Estimate |
|---|---|
| *g*_FUNDAMENTAL_BUY | 3.07 |
| | (0.38) |
| *g*_FUNDAMENTAL_SELL | - 1.31 |
| | (0.35) |
| *g*_INTERMEDIATE | -68.5 |
| | (0.91) |
| *g*_OPPORTUNISTIC | - 4.33*** |
| | (4.48) |
| *g*_HFT | - 24.1* |
| | (1.71) |