Need for speed: Hard information processing in a high-frequency world

DOI:
10.1002/fut.21861

Document Version
Accepted author manuscript

Citation for published version (APA):
Zhang, S. S. (2017). Need for speed: Hard information processing in a high-frequency world. *Journal of Futures Markets*, *38*(1), 3-21. https://doi.org/10.1002/fut.21861

Published in:
Journal of Futures Markets

Need for Speed: Hard Information Processing in a High-frequency World

S. Sarah Zhang*

Abstract

I study the role of high-frequency traders (HFTs) and non-high-frequency traders (nHFTs) in transmitting hard price information from the futures market to the stock market using an index arbitrage strategy. Using intraday transaction data with HFT identification, I find that HFTs process hard information faster and trade on it more aggressively than nHFTs. In terms of liquidity supply, HFTs are better at avoiding adverse selection than nHFTs. Consequently, HFTs enhance the linkage between the futures and stock markets, and significantly contribute to information efficiency in the stock market by reducing the delay between the stock and the futures markets.

JEL Classification: G10, G14.

Keywords: High-frequency Trading, Information Processing, Index Arbitrage, Price Discovery, Efficiency.

I thank Frank Hathaway and Jeff Smith at NASDAQ OMX for providing data and comments. I am also grateful to Torben Andersen, Kevin Aretz, Michael Bowe, Michael Brennan, Todd Doersch, Terrence Hendershott, Jianfeng Hu, Eirini Konstantinidi, Alex Kostakis, Albert S. (Pete) Kyle, Albert Menkveld, Ryan Riordan, Elvira Sojli, Gunther Wuyts, Chen Yao, Mao Ye, Jean-Pierre Zigrand, and seminar participants at the doctoral tutorials at the EFA 2012 and the 5th Erasmus Liquidity Conference of the LSE Systemic Risk Centre, HU Berlin; NASDAQ Brownbag Seminar, University of Maryland; CQA Conference 2012; FMA Annual Meeting 2012, University of Toronto; University of Zurich; KU Leuven; and the International Conference of the French Finance Association 2012 for insightful discussions and helpful comments and suggestions. Part of the paper was written during a visit to the University of Toronto hosted by Andreas Park and Katya Malinova, whose hospitality is gratefully acknowledged. Financial support from the IME Graduate School funded by Deutsche Forschungsgemeinschaft (DFG) and the Karlsruhe House of Young Scientists is gratefully acknowledged. All errors are mine.

*Assistant Professor of Finance, E-mail: sarah.zhang@manchester.ac.uk, University of Manchester, Alliance Manchester Business School, Booth Street East, Manchester M15 6PB, United Kingdom, Phone: +44 (0) 161 306 3452, Fax +44 (0) 161 275 4023.

1. Introduction

In recent years, the volume and complexity of information accessible to participants in financial markets has grown to exceed human information processing capacity. As a result, computer algorithms are now used to process large amounts of information more quickly. One specific group of computer trading algorithms is that used by high-frequency traders (HFTs). HFTs distinguish themselves from other groups of traders through their use of high speed trading and information processing, their high trading volume, as well as their sophisticated algorithms.[1] A major concern of regulatory authorities, such as the U.S. Securities and Exchange Commission (SEC) and the U.S. Commodity Futures Trading Commission (CFTC), is the influence of HFTs on market quality and price discovery (cf. the call for comments of the SEC, 2010). HFTs contribute to price discovery through the application of different information processing strategies. Common information processing strategies of HFTs include arbitrage trading strategies. Index arbitrage focuses on mispricings between an index (such as the S&P 500) and its components. It can thus be categorized as "hard" quantitative information processing.

To address the regulatory concerns relating to price discovery, I analyze the role of HFTs and non-high-frequency traders (nHFTs) in interpreting "hard" futures price information for index arbitrage strategies, and study the implications for information efficiency. This paper provides an empirical test of hard information processing strategies used by HFTs, specifically index arbitrage strategies between the E-mini futures and the U.S. stock markets.

The results show that HFTs use their competitive advantage to react to hard quantitative information shocks faster and more strongly than nHFTs. Specifically, they trade in the direction of hard information shocks within the first few seconds and quickly start selling off their trading position in order to realize their trading profits. This trading behavior translates into HFTs increasing information efficiency by reducing the delay between the stock and futures markets' price adjustment. Consequently, they function as "messengers" between the futures and stock markets, and create a stronger interlinkage between both markets.

The results have important implications for price discovery in the U.S. stock market. This study is the first to document this type of HFT behavior and analyze its effects on information efficiency.

First, the high-frequency dynamics of an index arbitrage strategy conducted by HFTs are illustrated. I show that liquidity-demanding HFTs trade on hard and short-lived information more aggressively and in higher volumes than nHFTs. After about 10 seconds, HFTs start selling off their trading position, arguably in order to quickly realize their trading profits. Roşu (2016) makes predictions concerning the aversion of fast traders to holding inventory. I support this finding as I observe the reversion behavior of HFTs after 10 seconds. However, liquidity supplying HFTs are able to prevent being adversely selected in contrast to liquidity supplying nHFTs who provide a substantial amount of liquidity against the direction of information shocks. This indicates that HFTs are better at avoiding adverse selection costs compared to nHFTs and that nHFT liquidity supply becomes more important after information shocks.

Second, I show that higher HFT activity is further related to a decrease in delay between futures and stock markets. The lead–lag relationship between the futures and stock markets has been well-documented, such as by Hasbrouck (2003). In this context, my contribution is to show that HFTs use futures price information to trade on stock markets. They thereby contribute to a shorter delay of price adjustment on both markets as they enable information to travel faster from the futures market to the stock market. While a shorter delay implies higher efficiency, other papers raise concerns that higher exchange speed and faster traders might also create negative externalities, such as higher adverse selection and a decrease in liquidity and total welfare (cf. Budish, Cramton, & Shim, 2015; Menkveld & Zoican, 2017).

Empirical evidence shows that HFTs generally improve information efficiency. Jovanovic and Menkveld (2015) find a positive correlation between middlemen/HFT activity and hard information days. Further insight is provided into the high-frequency dynamics of hard information processing by HFTs and nHFTs, and the relationship between information processing and price discovery. The latter is further in line with the theoretical predictions of Jovanovic and Menkveld (2015). My results are consistent with the findings of Brogaard, Hendershott, and Riordan (2014) that HFTs contribute to short-term price discovery, and I relate their findings to hard information shocks. Brogaard, Hendershott, and Riordan (2014) use a state space model to decompose the market return time series into a transitory component (i.e., pricing errors) and a permanent component (i.e., permanent price changes). They show that liquidity-demanding trades are positively related to permanent price changes, while HFT liquidity-supplying trades are positively related to pricing errors. As mentioned by Hendershott, Jones, and Menkveld (2011), "Algorithmic Trading (AT) can also improve linkages between markets." Chaboud, Chiquoine, Hjalmarsson, and Vega (2014) find that AT increases price efficiency in FX markets by applying triangular arbitrage strategies. Boehmer, Fong, and Wu (2014) show that AT improves liquidity and efficiency but increases volatility. Menkveld (2013) further finds that HFTs play a major role in the interlinkages between different stock markets in Europe, specifically NYSE–Euronext and Chi–X. My results complement the findings of previous studies by showing that HFTs also engage in index arbitrage strategies and thus create a more efficient linkage between the futures and stock markets.

Other papers also demonstrate the negative effects of HFTs on price discovery, market volatility, and total welfare. Kirilenko, Kyle, Samadi, and Tuzun (2014) find that HFTs did not trigger the "Flash Crash" on May 6th, 2010[3], it exacerbated market volatility during this extreme event. Hirschey (2013) discovers that HFTs are better at anticipating buying and selling pressures and order flow. Based on their theoretical model, Biais, Foucault, and Moinas (2015) demonstrate that HFTs increase adverse selection costs for slower traders, which subsequently lowers welfare. To overcome the decrease in welfare, they propose separating fast and slow traders on different platforms and introducing Pigovian taxes on fast technology investment. Predatory HFT strategies, including quote stuffing, are analyzed empirically by Gai, Yao, and Ye (2012) and Egginton, Van Ness, and Van Ness (2013). In my analyses, I concentrate on index arbitrage trading strategies and their effects on information efficiency.

The remainder of the paper is structured as follows. Section 2 describes the data sample underpinning the analyses and provides the correlation results. Section 3 presents results of price and trading responses to hard information shocks. Section 4 discusses the relationship of the speed of price discovery with hard information shocks and the implications for information efficiency. Section 5 concludes.

2. Data and Sample Selection

I match a proprietary HFT dataset provided by NASDAQ with E-mini futures price data. The HFT identification of the transaction dataset is based on NASDAQ's knowledge of their customers and analysis of firms' trading such as order duration and order-to-trade ratio (For further information on the dataset see Brogaard, Hendershott, and Riordan, 2014). Trade data are tick-by-tick data time-stamped to milliseconds, covering the years 2008 and 2009. For each trade, the liquidity demander and liquidity supplier of a trade is identified as an HFT or nHFT. The same dataset is used by Brogaard, Hendershott, and Riordan (2014), O'Hara, Yao, and Ye (2014), Hirschey (2013), and Carrion (2013). Datasets used in research often employ proxies for HFT and AT (such as Hendershott, Jones, & Menkveld (2011), Hasbrouck & Saar (2013)). Only a few available datasets directly identify AT and HFT. Hendershott and Riordan (2013) use data from Deutsche Börse in Germany, which identifies ATs. Furthermore, several papers based on proprietary datasets with trader identification conduct HFT identification based on trader characteristics, for example, Kirilenko, Kyle, Samadi, and Tuzun (2014), who use E-mini futures data, and Hagströmer and Nordén (2013), who study proprietary data from NASDAQ-OMX Stockholm Exchange.

Based on the sample selection criteria in Brogaard, Hendershott, and Riordan (2014), the data sample is restricted to 80 randomly selected stocks in the Russell 2000. The complete list of sample stocks and relative portion of HFT activity can be found in Appendix A, Table A.1. The HFT dataset contains the timestamp, symbol, volume, price, and buy–sell indicators for each transaction. It also has an indicator called "type," which can assume the values {HH, HN, NH, and NN}. For each trade, the first character indicates whether the liquidity-demanding marketable order that initiates the trade is classified as an HFT (H) or nHFT (N). The second character indicates whether the passive liquidity-supplying limit order against which the marketable order is executed is classified as either an HFT or nHFT. Thus, I am able to distinguish between the HFT groups demanding liquidity (HFT_d, which combines the HH and HN trades) and supplying liquidity (HFT_s, which combines the HH and NH trades), and the corresponding nHFTs demanding liquidity ($nHFT_d$) and supplying liquidity ($nHFT_s$). This is in the spirit of the demand and supply classification by Brogaard, Hendershott, and Riordan (2014). However, Hagströmer and Nordén (2013) show that HFT strategies are diverse and can involve both liquidity demand and supply. Chordia, Goyal, Lehmann, and Saar (2013) further argue that one algorithm might be involved in liquidity-demanding and -supplying activity. I do not exclude the possibility that the activity of the same HFT firm can be represented in the HFT demand and supply group. However, I argue that liquidity demand and supply distinguish different levels of immediacy that are especially important around extreme events, and that the differences in results for liquidity demand and supply are economically meaningful in the context of my analysis.

Only continuous trading is considered in order to measure the direct intra-day reaction after an information event. The first and last five minutes of each trading day are omitted in order to exclude trading on overnight information and biases associated with market opening and closing. Thus, the data span the time interval from 9:35 a.m. to 3:55 p.m.

2.1. Data and Sample Selection

Descriptive statistics of the data sample are shown in Table 1.

[INSERT TABLE 1 HERE]

The table presents daily averages of the sample of 80 stocks, specifically the average market capitalization (*MCap*), the average stock price over the sample (*Price*), and the average total number of daily shares traded (*Shares*), all sourced from Compustat. The quoted spread at the best bid and ask level (*Qspread*) is obtained from one-second snapshots of the National Best Bid and Offer (NBBO). The second part of Table 1 shows descriptive statistics based on the NASDAQ HFT dataset. HFT_d and HFT_s ($nHFT_d$ and $nHFT_s$) are the total net (buyer-initiated minus seller-initiated) trade volume of the liquidity-demanding and -supplying HFT (nHFT) group, respectively. *RelHinit* and *RelHpass* are the participation rates of HFTs on the NASDAQ dollar trading volume.

Sample stocks are capitalized with an average of $24 billion, a price of $42.9, and possess moderate liquidity with average transaction costs measured by a quoted spread of 8.92 bps. The trading volume in *Shares* shows that the NASDAQ trading volume in the sample represents around 30% of the total trading volume of sample stocks (2.57 million shares per stock traded daily on the NASDAQ as compared to 8.83 million shares overall).

The net trading variables show that both HFTs and nHFTs buy and sell in equal amounts, whereas HFTs have a slightly higher buying volume (more were bought than sold by $80,000 and $10,000 for liquidity demand and supply, respectively) than nHFTs. The participation rates show that HFTs are responsible for around 40% of the liquidity-demanding and around 30% of all liquidity-supplying trades. While it is often assumed that a major activity component of HFTs is related to their involvement in market making, it is interesting to note that liquidity-demanding transactions outweigh liquidity supply in this sample.

2.2. Index Futures Arbitrage and Price Delay

The lead–lag relationship of the stock index futures and stock markets is well documented in the finance literature. While the cost-of-carry relationship should lead to a contemporaneous correlation of returns in a frictionless market, differences in market microstructure and other frictions can lead to a lag between both markets. Chan (1992) documents the lead–lag relationship between the returns of the Major Market cash index and those of the Major Market index futures and S&P 500 futures. Grossman and Miller (1988) argue that the futures market is more liquid and offers lower transaction costs and higher immediacy than stock markets, resulting in the futures market leading the stock market. Hasbrouck (2003) finds that for the S&P500 and Nasdaq-100 indexes, most of the price discovery occurs in the E-mini futures market. This lag between the futures and stock markets is often exploited by arbitrage strategies. These strategies both profit from short-term price differences between the markets and increase efficiency by transmitting information from the futures market to the stock markets. In a high-frequency world, the time lag between the futures and stock markets has decreased significantly, and high speed is nowadays essential in order to profit from arbitrage strategies. I argue that HFTs play a major role in decreasing the documented gap between the futures and stock markets by executing an index arbitrage strategy, that is, using futures price information to trade on the stock market, and thus transmit information between both markets.

For this analysis, I focus on extreme futures price returns and use measures of price delay in the stock market. Specifically, in the spirit of Jovanovic and Menkveld (2015), I utilize extreme market futures returns as a major source of information for index arbitrage strategies in the intraday analysis. S&P 500 futures prices on a tick-by-tick basis from Thomson Reuters Tick History are collected.[4] The first and last 5 minutes of the trading day are excluded to eliminate market opening and closing effects, and determine the 1% and 99% percentiles of S&P 500 futures one-second returns over the whole observation period. Returns above the 99% (6.11bps) and below the 1% (-6.04bps) level are considered to be hard information shocks. As a robustness check, I further include returns above the 95% (3.52 bps) and below the 5% (-3.47 bps) level.

Price delay as a measure of price efficiency is proposed by Mech (1993) and Hou and Moskowitz (2005) and applied by Carrion (2013), who utilize measures of price delay across different stocks to analyze the effect of speed of information diffusion (as a measure of market friction) or price adjustment. Mech (1993) finds that transaction costs cause portfolio autocorrelation by delaying price adjustments. Hou and Moskowitz (2005) find that a large delay commands a large return premium that is not explained by size, liquidity, or microstructure effects. While Mech (1993) and Hou and Moskowitz (2005) use weekly returns, Carrion (2013) employ an intraday adaption based on 1-minute and 5-minute returns. In a similar fashion, I run the following regression of the stock's return per second on contemporaneous E-mini one-second returns (Model 1) and 10 seconds of lagged E-mini one-second returns (Model 2):

$$r_{i,t} = \alpha_{i.d} + \beta_{i,d} r_{m,t} + \varepsilon_{i,t} \quad (1)$$

$$r_{i,t} = \alpha_{i,d} + \beta_{i,d} r_{m,t} + \sum_{n=1}^{10} \delta_{i,d}^{(-n)} r_{m,t-m} + \varepsilon_{i,t} \quad (2)$$

If the information is impounded instantly, then $\beta_{i,d}$, but none of the $\delta_j^{(-n)}$s will be different from zero. The price delay can be characterized by the significance of the $\delta_j^{(-n)}$. I derive the following measure for price delay according to Hou and Moskowitz (2005) and Mech (1993):

$$\text{DELAY} = 1 - \frac{R^2 from\ Model\ 1}{R^2 from\ Model\ 2} \tag{3}$$

The measure provides an estimate of how much of the price discovery occurs after the first second. The larger this measure, the more the return variation explained by lagged returns and the stronger the effects of any delay in the response to return innovations. I hypothesize that HFTs significantly decrease the delay in price discovery and have a negative relationship with this measure.

Daily correlation results of trading variables and the proposed measures of hard information and price delay are reported in Table 2.

[INSERT TABLE 2 HERE]

Trading variables are the total number of trades per stock and day (*Trades*) and the relative HFT liquidity-demanding (HFT_d) and -supplying (HFT_s) trades divided by the total number of trades per stock and day. *Delay* is defined as in Equation 3. The results in Panel A show positive contemporaneous correlations of relative HFT liquidity demand (HFT_d) and supply (HFT_s) with each other. This indicates that HFTs are both active liquidity demanders and suppliers on HFT-intensive days. Furthermore, both HFT variables have a negative correlation with *Delay*. This means that higher HFT activity days indicate a lower delay of price discovery in the stock market, thereby a higher connection of the stock market to the futures market. The correlation results give the first indication of the role of HFTs for the transmission of hard information in the stock market and their beneficial effect in reducing delay in information transmission.

3. Intraday Responses to Information Shocks

3.1. Response of Returns and Overall Net Trading

A central assumption in market microstructure is that trade and quote changes convey information in markets with asymmetrically informed traders. Based on this assumption, one can model price changes as a function of past quote changes and contemporaneous and past trades. Vector auto-regressions (VARs) have been used in a number of settings to estimate this relationship. I estimate three VAR models with exogenous shocks (VARX models): one for total HFT net trading and two, separately, for both the liquidity demand and supply of HFTs and nHFTs.

Following Hasbrouck (1991a), I model the relationship between returns and trades as a system of equations. In order to apply this methodology to the HFT and futures data set, I include exogenous information shocks and adapt the used timeframes and lags to a high-frequency setting. All variables are aggregated into one-second intervals. I further incorporate exogenous information shocks represented by extreme positive and negative futures returns into the model (transforming the VAR model into a VARX model). The VARX model is applied to directed net trading (buy minus sell volume) and stock returns in the spirit of Tookes (2008) and is implemented as follows (estimated separately by OLS):

$$V_{i,t}^{h} = \alpha_i^h + \sum_{j=1}^{k} \beta_{i,j}^{h} V_{i,t-j}^{h} + \sum_{j=1}^{k} \delta_{i,j}^{h} r_{i,t-j} + \sum_{w=0}^{W} \phi_{i,w}^{h} D_{i,t-w} + \varepsilon_{i,t}^{h}$$

$$r_{i,t} = \alpha_i^r + \sum_{j=0}^{k} \beta_{i,j}^{r} V_{i,t-j}^{h} + \sum_{j=1}^{k} \delta_{i,j}^{r} r_{i,t-j} + \sum_{w=0}^{W} \phi_{i,w}^{r} D_{i,t-w} + \varepsilon_{i,t}^{r} \quad (4)$$

where $r_{i,t}$ denotes the return, and $V_{i,t}$, the signed net trading ($-volume bought minus $-volume sold) of HFTs (superscript h). The subscripts i and j denote the stock and lags after an information shock, respectively, while t is the respective one-second interval. The relevant lags after an information shock are denoted by W, which may be 1, 10, or 30 one-second intervals. $D_{i,t-j}$ specifies the direction of the information shock and equals 1 if a positive information shock and -1 if a negative information shock occurs in t or in less than W before t, and 0 otherwise. The coefficients are β_i and δ_i, where superscripts h and r denote HFTs and return, respectively. The α_i are intercepts, and $s_{i,t}$ refers to the error terms. The model is applied to overall HFT net trading.[5] I control for k=30 lags of stock return and past HFT net trading (chosen based on the Bayesian information criterion). The coefficients of interest are φ^r and φ^h, which represent the responses of stock returns and HFT net trading after exogenous information shocks, respectively.

In order to differentiate whether the response in net trading is due to liquidity demand or supply, I include two trade time series (HFT and nHFT) in the model and estimate it separately for the liquidity demand and supply variables. HFT and nHFT liquidity demand (HFT_d and $nHFT_d$) is the sum of all trades that were initiated by an HFT and nHFT, respectively (the trader who submitted the marketable order). Supply variables (HFT_s and $nHFT_s$) are defined according to the passive side of the trades. The VARX model including both HFT and nHFT net trading is implemented as follows (estimated separately by OLS):

$$V_{i,t}^h = \alpha_i^h + \sum_{j=1}^{k} \beta_{i,j}^h V_{i,t-j}^h + \sum_{j=0}^{k} \gamma_{i,j}^h V_{i,t-j}^n + \sum_{j=1}^{k} \delta_{i,j}^h r_{i,t-j} + \sum_{w=0}^{W} \phi_{i,w}^h D_{i,t-w} + \varepsilon_{i,t}^h$$

$$V_{i,t}^n = \alpha_i^n + \sum_{j=0}^{k} \beta_{i,j}^n V_{i,t-j}^h + \sum_{j=1}^{k} \gamma_{i,j}^n V_{i,t-j}^n + \sum_{j=1}^{k} \delta_{i,j}^n r_{i,t-j} + \sum_{w=0}^{W} \phi_{i,w}^n D_{i,t-w} + \varepsilon_{i,t}^n \qquad (5)$$

$$r_{i,t} = \alpha_i^r + \sum_{j=0}^{k} \beta_{i,j}^r V_{i,t-j}^h + \sum_{j=0}^{k} \gamma_{i,j}^r V_{i,t-j}^n + \sum_{j=1}^{k} \delta_{i,j}^r r_{i,t-j} + \sum_{w=0}^{W} \phi_{i,w}^r D_{i,t-w} + \varepsilon_{i,t}^r$$

The model adopts notations identical to those of the VARX model in Equation 5, and also incorporates coefficients and variables with superscript *n*, denoting nHFT. Panel B of Table 3 presents the aggregated coefficients of the VARX model for HFT and nHFT liquidity demand and supply and the difference between HFT and nHFT. The results for liquidity-demanding trades of HFTs (HFT_d) and nHFTs ($nHFT_d$) and their difference (*Diff*) appear on the left-hand side, and those for liquidity-supplying trades appear on the right-hand side of Table 3.

Subsequent to the occurrence of an information event, net trading in the same direction as this information event is interpreted as informed trading activity based on the event. Thus, the cumulative net trading can be interpreted as the cumulative impact of a directed information shock on net trading in the same direction. The rationale for this interpretation is that a higher amount of positive net trading, that is, more buy than sell orders, should be expected after positive information shocks, and more negative net trading, after negative information shocks. The results are reported for the aggregated coefficients within one second, ($\sum_{j=0}^{1} \phi_{i,j}^{r}$ and $\sum_{j=0}^{1} \phi_{i,j}^{h}$); 10 seconds, ($\sum_{j=0}^{10} \phi_{i,j}^{r}$ and $\sum_{j=0}^{10} \phi_{i,j}^{h}$); and 30 seconds, ($\sum_{j=0}^{30} \phi_{i,j}^{r}$ and $\sum_{j=0}^{30} \phi_{i,j}^{h}$), after the shock.

Table 3 presents the aggregated coefficients of the VARX model in Equation (5) for return and trading responses.

[INSERT TABLE 3 HERE]

As indicated in Panel A, there is an immediate response in stock returns and HFT net trading 1 second after the information shock. My interpretation of the magnitude of the response is that an average information shock induces a 2.03-basis-point response in returns after the first second and a response of 0.37 in net trading by HFTs. Given that the daily HFT net trading averages around $90,000, 0.37 equates to HFT net trading of $3,700, accounting for 4% of the total daily HFT net trading position within 1 second. Furthermore, most of the net trading response occurs during the first second. In comparison, after 10 seconds, the cumulated net trading is 0.51, while the coefficient has increased by a further 0.14, which is cumulatively much less than that in the first second. After 30 seconds, cumulative net trading is only 0.26, which suggests that HFTs begin to sell off their trading position and trade in the opposite direction.

In summary, the results document the immediate response of HFTs' net trading to information shocks and consistent trading in the direction of the shock within the first 10 seconds. Subsequently, they quickly reverse their trading behavior and start selling off their acquired trading position.

Panel B shows that this reversion behavior is mainly driven by liquidity-demanding HFTs. They reduce their position to less than half of what it was after 30 seconds, while liquidity-demanding nHFTs, on the other hand, trade continuously in the direction of the information shock. The difference between HFT and nHFT net trading becomes insignificant after 10 seconds, with nHFTs gaining a significantly larger position after 30 seconds, while HFT positions are now insignificantly different from 0. Furthermore, HFTs react much faster and more strongly to the shock: liquidity-demanding HFTs are able to build a position of 0.41 bps, while it is significantly less for nHFTs.

The inference that can be drawn from the analysis of overall trading response is that HFTs exhibit an aversion to holding inventory, which is consistent with the theoretical model of Roşu (2014). In contrast, they have an incentive to sell off their accumulated trading position in order to quickly realize their trading profits. As pointed out by Roşu (2016), HFTs appear to avoid maintaining a large inventory over time, as is common for market makers in order to avoid the inventory risks and capital constraints arising from a high level of inventory. Thus, speed is important for HFTs, not only in order to react to new information, but also for initiating a quick mean-reversion in their inventory position.

Liquidity-supplying HFTs are adversely selected, that is, they supply liquidity in the direction opposite to the shock and build a position against it, as shown by the negative coefficients in Panel B of Table 3. However, the coefficient for liquidity-supplying HFTs is only significant for the first one-second interval. This can be interpreted as HFTs being able to withdraw their liquidity supply from the market quickly after a shock occurs. Liquidity-supplying nHFTs do not withdraw from the market as quickly, and suffer significantly more from adverse selection than HFTs. However, coefficients for both liquidity-supplying HFTs and nHFTs increase after 30 seconds, thus showing a slightly reversion behavior of nHFT liquidity suppliers within 30 seconds after a shock. These results indicate that liquidity suppliers suffer from adverse selection as they build a trading position in the direction opposite to extreme hard information shocks, but they are able to reverse their positions after 30 seconds.

The different results on liquidity demand and supply shed further light on their distinct functions. Liquidity demanders require immediacy as they have to be the first to trade on new information in order to make a significant profit, while liquidity suppliers on the other side are more patient but suffer from adverse selection when there are large price changes.

I conduct another run of the analysis with net trading and return variables standardized by subtracting their mean and dividing the difference by their standard deviation in order to remove any biases due to the respective trading day. The results of the model coefficients using standardized coefficients are presented in Table 4.

[INSERT TABLE 4 HERE]

Compared to the non-standardized coefficients, the magnitudes of the coefficients for overall HFT net trading are higher, as indicated by Panel A, and liquidity demand and supply coefficients remain stable. T-statistics are higher for all coefficients, pointing toward less variation of standardized coefficients compared to non-standardized coefficients. The reversion effect of liquidity-demanding HFTs becomes less distinct after 10 seconds with standardization. Thus, the reversion effect becomes weaker but is still present. The adverse selection of liquidity-supplying HFTs (their position is in the direction opposite to the shock) is still significant and lower than that for nHFTs.

Figures 1 and 2 give further insights into the high-frequency response of stock returns and net trading after information shocks.

[INSERT FIGURE 1 HERE]

Figure 1 illustrates that the return response (green dotted line) to hard futures shocks is highest within the first second after the shock (lag = 0 in the graph), and the majority of the response occurs within 5 seconds. The trading response to hard information shows that HFTs trade in the direction of the information shock within the first 10 seconds following the shock (solid black line) and subsequently sell off their positions by trading in the direction opposite to the information shock. Thus, HFTs perfectly capture the majority of the return reaction by their trading, but then reverse their trading behavior, presumably in order to realize trading profits quickly. As all trades are classified as either HFT or nHFT on both sides of a trade, the cumulative net trading of nHFTs is the negative of HFT net trading.

Figure 2 further illustrates the net trading response of liquidity-demanding and –supplying HFTs and nHFTs separately.

[INSERT FIGURE 2 HERE]

Figure 2 reveals that liquidity-demanding HFTs build up a trading position within 8 seconds after a hard information shock. Subsequently, they reverse their trading behavior, representing a liquidation of the trading position acquired within the first 8 seconds. In contrast, nHFTs exhibit a consistent reaction in the direction of the information shock for a period of up to 30 seconds.

As discussed above, liquidity-supplying HFTs and nHFTs are both adversely selected within the first second after information shocks. HFTs and nHFTs both exhibit net trading patterns in the direction opposite to the information shock, as shown in Table 3. However, HFTs are able to maintain a neutral position. NHFTs consistently provide liquidity against the direction of the information shock for the first 6 seconds, and are only able to avoid further adverse selection 6 seconds after the shock.

On the one hand, the results show the clear speed advantage of HFTs on both the supply and demand sides. They exhibit a faster reaction to shocks and build a larger position with their liquidity demand. Thus, liquidity-demanding HFTs directly trade on the information shocks, and lead price discovery by bridging the lag between the futures and spot markets. In the long run, HFTs trade in the direction opposite to the hard information shock. This behavior implies a reduction of their trading positions and a realization of their short-term profits. On the supply side, HFTs are able to quickly withdraw liquidity and suffer less from adverse selection that nHFTs. This trading behavior, however, might increase short-term volatility due to the high level of directed trading, and potentially leads to instability in the market. I further discuss the implications for short- and long-term price discovery in the next section.

On the other hand, nHFTs trade consistently in the direction of the shock and tend to keep a stable position. Possibly due to their slower reaction, they provide liquidity against the directions of the shock and the price correction for a longer time after the shock. The availability of liquidity supply against the direction of the shock leads to an increase in the importance of nHFT liquidity supply after information shocks.

3.2. Robustness Checks

In order to check the robustness of the results, I perform the analysis for a wider set of information shocks, for different time periods and distinguish between the directions of information shocks. Overall, the results are stable across a wider range of information shocks, time periods, and the direction of information shocks.

Table 5 presents the results for futures returns greater than 95% or less than 5% of the overall sample in order to test the robustness of the results for a wider set of information shocks.[6]

[INSERT TABLE 5 HERE]

Similar to the results for 1% and 99% percentile shocks, the results confirm that liquidity demanding HFTs trade quickly in the direction of the shock and reverse their trading behavior after more than 10 seconds, while nHFTs consistently trade in the direction of the shock after the information shock. Comparing the coefficients, the magnitude of the coefficients is smaller for 5%/95% shocks than for 1%/99% shocks, however they are still statistically and economically significant. Figure 3 provides details of the HFT behavior within the 30 seconds period after the shock.

[INSERT FIGURE 3 HERE]

As shown in Figure 3, the reaction to the wider range of shocks is smaller on average. For liquidity demanding HFTs, the impact on net trading is 0.4 compared to 0.5 in Figure 2. For liquidity supplying nHFTs, adverse selection is lower as the coefficient is -0.4 compared to -0.6. Thus, the results still hold for a wider range of information shocks, although the magnitude of the effects is smaller compared to more extreme shocks.

For different time periods, the analysis is performed using standardized coefficients, separately for time periods of high volatility (during the financial crisis from September 2008 to June 2009) and low volatility (pre- and post-financial crisis), as well as positive and negative information shocks. I choose time periods according to the VIX value, which increases to above 30 in September 15, 2008, and decreases again to below 30 after July 10, 2009. The results are shown in Appendix B, Table B.1. The results are qualitatively similar, and the coefficients are generally higher in the pre-crisis period. This finding is interesting as HFTs are sometimes blamed for increasing volatility during the financial crisis. I can refute these concerns in the context of hard information shocks, as HFTs' reaction to hard information is no stronger during the crisis than during non-crisis periods.

Furthermore, the results for hard information shocks are consistent for both positive and negative shocks. The results for positive and negative shocks are presented separately in Appendix C, Table C.1. The magnitudes of the responses to positive and negative hard information shocks are similar, indicating that there is no asymmetric reaction to quantitative information shocks. As HFTs and algorithms in general should not be subject to any behavioral biases, there should be no significant difference in reactions for positive and negative shocks.

4. Implications for Information Efficiency

Information efficiency involves different dimensions, one of which is the timeliness of price adjustment. In this section, I analyze the degree of price discovery measured by a measure of price delay, as applied by Hou and Moskowitz (2005), Comerton-Forde and Putniņš (2015), and Carrion (2013).

Delay is defined as $1 - \frac{R^2 of\ model\ without\ lagged\ futures\ returns}{R^2 of\ model\ with\ lagged\ futures\ returns}$ in the spirit of Comerton-Forde and Putniņš (2015) and Carrion (2013), as described in Section 2.2. In order to analyze the influence of HFT activity on the speed of price discovery after hard information shocks, I implement the following cross-sectional regression:

$$Delay_{i,d} = a_d + b_{1,d} HFT_{i,d} + b_{2,d} logMC_{i,d} + b_{3,d} D_{i,d}^{nyse}$$
$$+ b_{4,d} \times precrisis_{i,d} + b_{5,d} \times postcrisis_{i,d} + \varepsilon_{i,d} \qquad (9)$$

where subscripts *i* and *d* denote stock *i* and day *d*, respectively. $HFT_{i,d}$ is the participation rate of HFTs in total trading volume (HFT dollar volume divided by total dollar volume), with HFT_{all} representing the participation rate for overall HFT activity and HFT_d and HFT_s representing participation rates for HFT liquidity demand and supply. The control variables are the same as in Equation 8 ($logMC_{i,d}$ is the natural logarithm of the market capitalization (in billions of dollars), $D_{i,d}^{nyse}$ is a dummy for NYSE-listed stocks, and $precrisis_{i,d}$ and $crisis_{i,d}$ denote the dummies for the pre-crisis and crisis periods, respectively). The results of the daily regressions with and without controls are presented in Table 6.

[INSERT TABLE 6 HERE]

Table 6 presents negative and significant coefficients for HFT_{all}, HFT_d, and HFT_s (except for HFT_s with controls), showing that higher overall HFT activity and higher HFT participation rate in liquidity demand is negatively related to the delay measure. As described in Section 2.2, a decrease in *Delay* points towards a higher R^2 of contemporaneous futures returns on stock returns compared to a model which includes lagged futures returns. For overall HFT activity HFT_{all}, the coefficient is around -0.31 before and -0.14 after including control variables. Although including control variables lowers the absolute value of the coefficient, it is still statistically significant and economically significant compared to the coefficients of the control variables. Analyzing HFT liquidity demand and supply separately illustrates that the main effect stems from the liquidity demanding activities rather than liquidity supply as the latter shows ambiguous results. Thus, HFTs generally improve price efficiency of stock markets by reducing the delay of price discovery between the stock and futures markets with their liquidity demanding activities.

While the *HFT* variable in the previous section analyzes general HFT activity during the day, I apply another measure $HFThard$ in order to analyze the influence of daily HFT activity due to hard information[7]. $HFThard$ is the regression coefficient derived from the 1 second regression of HFT net trading on futures returns: $HFT_{i,t} = \alpha_{i.d} + HFThard_{i,d} r_{m,t} + \varepsilon_{i,t}$. It represents HFT activity in the direction of futures price changes in contrast to overall HFT activity and thus serves as a measure of stock-specific HFT net trading due to market price changes in the same direction. The following regression model is applied:

$$Delay_{i,d} = a_d + b_{1,d} HFThard_{i,d} + b_{2,d} logMC_{i,d} + b_{3,d} D_{i,d}^{nyse}$$
$$+ b_{4,d} \times precrisis_{i,d} + b_{5,d} \times postcrisis_{i,d} + \varepsilon_{i,d} \quad (10)$$

The results of the daily regressions with and without controls are presented in Table 7.

[INSERT TABLE 7 HERE]

Table 7 shows consistently negative and significant coefficients, except for HFT liquidity supply HFT_s. The magnitude of the coefficients is generally larger than in the model in Equation 9 and the magnitude does not deteriorate with the inclusion of control variables. For overall HFT activity HFT_{all}, the coefficient is around -3.07 before and -3.17 after including control variables. Similar to above, the decreasing effect on delay stems primarily from the liquidity demanding activities of HFTs rather than their liquidity supply.

These results support the previous conclusion that HFTs play an important role in short-term price discovery by incorporating hard futures price information in their trading strategies. Thereby, HFTs decrease the delay between the futures and stock markets and strengthen the interlinkage between the futures and stock markets.

However, price delay is not the only dimension of efficiency; academics argue that an increase in speed might also create negative externalities. Biais, Foucault, and Moinas (2015) find that fast traders might generate higher adverse selection. Budish, Cramton, and Shim (2015) argue that arbitrage rents that can only be achieved at a frequency of milliseconds actually harm liquidity provision and decrease total welfare. Thus, it is important to conduct a more detailed study of the effects of liquidity and market stability in the context of increase in exchange speed and trading frequency.

5. Conclusion

Information processing speed matters especially for "hard" quantitative information shocks and arbitrage trading. HFTs react more quickly and aggressively to hard information shocks than nHFTs and later reverse their trading behavior quickly afterwards. The reverting HFT behavior shortly after the information shock can be interpreted as a strategy to realize profits from this shock and as a form of "inventory aversion" or "level aversion," as argued by Roşu (2014). As a consequence, HFTs play an important role in short-term price discovery, especially on days with a high amount of hard information, and they significantly decrease the delay in price discovery between stock and futures markets.

The results have several implications for the public discussion relating to HFTs. The different results for liquidity demand and supply emphasize the importance of differentiating between different types of HFT strategies, particularly arbitrage strategies and market making strategies. As shown by Biais, Declerck, and Moinas (2015), traders who are fast and/or proprietary exhibit different trading behavior, trading profits, and influence on market quality and price efficiency. This helps to clarify the notion of HFTs: rather than one homogenous group of traders, it serves as an umbrella term for different types of automated trading strategies, with varying effects on market quality and welfare. In this paper, I concentrate on index arbitrage strategies, one of the most commonly applied HFT strategies relying on HFTs' ability to process a large amount of machine-processable and quantitative information at high speed. I find that this type of HFT is associated with higher price efficiency and shorter delay in price discovery. More research is required to analyze the effects of other types of HFTs and their effects on long-term price discovery.

Concerns can also be raised in the context of stronger linkage between the futures and stock markets: While a higher degree of comovement indicates higher information efficiency of the stock market, this effect might also cause a "domino effect" in times of financial distress, such as during the "Flash Crash" between different markets. A stronger linkage also enables erroneous information to travel faster between markets, with the potential to worsen market disruptions. These findings strengthen regulatory discussions on stricter market-wide circuit breaker rules, that is, trading halts that are triggered by abnormally high volatility within a certain time period. These rules have recently been revised, for example, by the SEC (cf. SEC, 2012, "Market-Wide Circuit Breaker Approval Order"). More research is needed to explore HFT behavior in extreme market situations and whether they make a net contribution to market stability.

References

Biais, B., F. Declerck, & S. Moinas (2015). Who supplies liquidity, how and when?. Working Paper.

Biais, B., T. Foucault, & S. Moinas (2015). Equilibrium Fast Trading. Journal of Financial Economics, 116, 292–313.

Boehmer, E., K. Fong, & J. Wu (2014). International Evidence on Algorithmic Trading. Working Paper.

Brogaard, J., T. Hendershott, & R. Riordan (2014). High Frequency Trading and Price Discovery. Review of Financial Studies, 27, 2267–2306.

Budish, E., P. Cramton, & J. Shim (2015). The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response. Quarterly Journal of Economics, 130, 1547–1621.

Carrion, A. (2013). Very fast money: High-frequency trading on the NASDAQ. Journal of Financial Markets, 16, 680–711.

Chaboud, A., B. Chiquoine, E. Hjalmarsson, & C. Vega (2014). Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market. Journal of Finance, 69, 2045–2048.

Chan, K. (1992). A Further Analysis of the Lead-Lag Relationship Between the Cash Market and Stock Index Futures Market. Review of Financial Studies, 5, 123–152.

Chordia, T., A. Goyal, B. N. Lehmann, & G. Saar (2013). High-frequency trading. Journal of Financial Markets, 16, 637–645.

Comerton-Forde, C., & T. J. Putniņš (2015). Dark trading and price discovery. Journal of Financial Economics, 118, 70–92.

Easley, D., M. M. L. de Prado, & M. O'Hara (2011). The Microstructure of the Flash Crash: Flow Toxicity, Liquidity Crashes and the Probability of Informed Trading. Journal of Portfolio Management, 37, 118–128.

Egginton, J., B. Van Ness, & R. Van Ness (2013). Quote Stuffing. Financial Management, 45, 583–608.

Foucault, T., J. Hombert, & I. Roşu (2015). News Trading and Speed. Journal of Finance, 71, 335–382.

Gai, J., C. Yao, & M. Ye (2012). The Externalities of High Frequency Trading. Working Paper.

Grossman, S. J., & M. H. Miller (1988). Liquidity and Market Structure. The Journal of Finance, 43, 617–633.

Hagströmer, B., and L. Nordén (2013). The Diversity of High Frequency Traders. Journal of Financial Markets, 16, 741–770.

Hasbrouck, J. (1991a). Measuring the Information Content of Stock Trades. Journal of Finance, 46, 179–207.

Hasbrouck, J. (1991b). The summary informativeness of stock trades: an econometric analysis. Review of Financial Studies, 4(3), 571–595.

Hasbrouck, J. (2003). Intraday Price Formation in U.S. Equity Index Markets. Journal of Finance, 58, 2375–2399.

Hasbrouck, J., and G. Saar (2013). Low-Latency Trading. Journal of Financial Markets, 16, 646–679.

Hendershott, T., C. M. Jones, and A. J. Menkveld (2011). Does Algorithmic Trading Improve Liquidity? Journal of Finance, 66, 1–33.

Hendershott, T., and A. J. Menkveld (2014). Price pressures, Journal of Financial Economics. 114, 405–423.

Hendershott, T., and R. Riordan (2013). Algorithmic Trading and the Market for Liquidity. Journal of Financial and Quantitative Analysis, 48, 1001–1024.

Hirschey, N. H. (2013). Do High-Frequency Traders Anticipate Buying and Selling Pressure ? Working Paper.

Hou, K., and T. J. Moskowitz (2005). Market Frictions, Price Delay, and the Cross- Section of Expected Returns. Review of Financial Studies, 18, 981–1020.

Jovanovic, B., and A. J. Menkveld (2015). Middlemen in Limit-Order Markets. Working Paper.

Kirilenko, A., Kyle, A. S., Samadi, M., & Tuzun, T. (2016). The Flash Crash: High Frequency Trading in an Electronic Market. Journal of Finance, forthcoming.

Kyle, A. S. (1985). Continuous Auctions and Insider Trading. Econometrica, 53, 1315–1335.

Mech, T. S. (1993). Portfolio return autocorrelation. Journal of Financial Economics, 34, 307–344.

Menkveld, A. J. (2013). High Frequency Trading and The New-Market Makers. Journal of Financial Markets, 16, 712–740.

Menkveld, A. J., and M. A. Zoican (2017). Need for Speed? Exchange Latency and Liquidity. Review of Financial Studies, forthcoming.

O'Hara, M., C. Yao, and M. Ye (2014). What's Not There: The Odd-Lot Bias in TAQ Data. Journal of Finance, 69, 2199–2236.

Roşu, I. (2016). Fast and Slow Informed Trading. Working Paper.

SEC (2010). Concept Release on Equity Market Structure, Release No. 34-61358, January 14th, 2010.

SEC (2012). Market-Wide Circuit Breaker Approval Order, Release No. 34-67090, May 31st, 2012.

Thompson, S. B. (2011). Simple Formulas for Standard Errors that Cluster by Both Firm and Time. Journal of Financial Economics, 99, 1–10.

Tookes, H. (2008). Information, Trading and Product Market Interactions: Cross-Sectional Implications of Insider Trading. Journal of Finance, 63, 379–413.

6. Tables

Table 1 Descriptive Statistics

This table provides descriptive statistics of the final sample of 80 stocks for the years 2008 and 2009 based on daily averages per stock. *MCap* denotes the average market capitalization of the stocks and *Price* the average stock price. *Shares* is the average total number of shares traded. *Qspread* is the quoted spread at the best bid and ask level. The second part of the table provides daily statistics based on the NASDAQ dataset. *Traded Shares* is the average number of shares traded on NASDAQ and *Trades* the average number of transactions on NASDAQ. HFT_d and $nHFT_d$ is net trading (buyer-initiated minus seller-initiated trade volume) of liquidity demanding HFT and nHFT, while HFT_s and $nHFT_s$ is net trading of liquidity supplying HFT and nHFT. *RelHinit* and *RelHpass* are the participation rates of HFT on the NASDAQ dollar trading volume.

Variable		**Units**	**Mean**	**Std**	**Min**	**Median**	**Max**
		Stock Characteristics					
Mcap	Compustat	$ 1 b	24.44	43.41	0.23	5.23	409.46
Price	Compustat	$	42.88	57.34	0.90	31.29	685.33
Shares	Compustat	1m shares	8.83	18.93	0.02	2.46	752.40
Qspread	NBBO	1 bps	8.92	6.87	0.90	7.08	93.00
		Daily Trading Variables					
Traded Shares	Nasdaq	1m shares	2.57	5.35	0.00	0.68	145.22
Trades	Nasdaq	1,000 trades	13.27	19.97	0.01	5.58	348.23
HFT_{all}	Nasdaq	$ 1m	0.09	6.35	-124.88	0.00	146.87
HFT_d	Nasdaq	$ 1m	0.08	5.32	-123.56	0.01	133.64
$nHFT_d$	Nasdaq	$ 1m	-0.06	10.38	-480.09	-0.01	264.96
HFT_s	Nasdaq	$ 1m	0.01	4.43	-95.62	-0.01	84.65
$nHFT_s$	Nasdaq	$ 1m	-0.03	11.84	-360.74	0.01	487.83
RelHinit	Nasdaq		0.39	0.12	0.02	0.39	0.83
RelHpass	Nasdaq		0.29	0.17	0.00	0.26	0.75

Table 2 Daily Correlations

This table presents daily correlations of trade, and market variables. Trade variables are *Trades* (number of daily transactions per stock), HFT_d (number of HFT liquidity demanding trades divided by total number of trades) and HFT_s (number of HFT liquidity supplying trades divided by total number of trades). *Delay* is the measure of price delay as proposed by Hou and Moskowitz (2005) (see Section 2.2 for details). Superscripts ***, **, and * denote significance at the 1%, 5%, and 10% level respectively.

	Trades	HFT_d	HFT_s	*Delay*
Trades	1	0.11***	0.58***	-0.21***
HFT_d		1	0.17***	-0.30***
HFT_s			1	-0.18***
Delay				1

Table 3 Differential Return and Trading Responses to Futures Shocks

This table presents the intraday trading response of HFTs and nHFTs within 1, 10, and 30 seconds after futures shocks. Coefficients in Panel A are estimated using the following VARX model:

$$V_{i,t}^{h} = \alpha_i^h + \sum_{j=1}^{k} \beta_{i,j}^{h} V_{i,t-j}^{h} + \sum_{j=1}^{k} \delta_{i,j}^{h} r_{i,t-j} + \sum_{w=0}^{W} \phi_{i,w}^{h} D_{i,t-w} + \varepsilon_{i,t}^{h}$$

$$r_{i,t} = \alpha_i^r + \sum_{j=0}^{k} \beta_{i,j}^{r} V_{i,t-j}^{h} + \sum_{j=1}^{k} \delta_{i,j}^{r} r_{i,t-j} + \sum_{w=0}^{W} \phi_{i,w}^{r} D_{i,t-w} + \varepsilon_{i,t}^{r} \quad (4)$$

where $r_{i,t}$ denotes the return and $V_{i,t}$ the net trading of HFT (superscript h). All variables are aggregated into one-second intervals. $D_{i,t-w}$ equals one for a positive and -1 for a negative shock occurring in t or less than W one-second intervals before t, and 0 otherwise. The cumulative responses are the aggregated coefficients $\varphi_{i,w}$ for 1, 10, and 30 seconds. In Panel B, the model is estimated separately for HFT and nHFT liquidity demand (HFT_d and $nHFT_d$) and supply (HFT_s and $nHFT_s$). Coefficients are aggregated per stock-day and tested using robust standard errors clustered by stock and trading day (cf. Thompson, 2011). t-statistics are in parentheses. Superscripts ***, **, and * denote significance at the 1%, 5%, and 10% level respectively.

Panel A: Return and Trading Responses		
Cum. Response	Return	HFT_{all}
1 second	2.03***	0.37***
	(22.53)	(6.29)
10 seconds	3.95***	0.51***
	(18.15)	(5.22)
30 seconds	4.58***	0.26***
	(14.85)	(2.96)

Panel B: Response of Liquidity Demand and Supply						
Cum. Response	HFT_d	$nHFT_d$	*Diff*	HFT_s	$nHFT_s$	*Diff*
1 second	0.41***	0.28***	0.13***	-0.07***	-0.58***	0.51***
	(6.03)	(5.77)	(3.35)	(-2.58)	(-6.51)	(6.52)
10 seconds	0.41***	0.48***	-0.07	0.03	-0.85***	0.89***
	(4.24)	(5.16)	(-0.98)	(0.85)	(-5.40)	(5.85)
30 seconds	0.17	0.71***	-0.54***	0.06	-0.73***	0.79***
	(1.61)	(4.69)	(-3.98)	(1.37)	(-4.01)	(4.69)

Table 4 Differential Return and Trading Responses - Standardized Variables

This table presents the intraday trading response of HFTs and nHFTs within 1, 10, and 30 seconds after futures shocks. Coefficients in Panel A are estimated using the following VARX model:

$$V_{i,t}^{h} = \alpha_i^h + \sum_{j=1}^{k} \beta_{i,j}^{h} V_{i,t-j}^{h} + \sum_{j=1}^{k} \delta_{i,j}^{h} r_{i,t-j} + \sum_{w=0}^{W} \phi_{i,w}^{h} D_{i,t-w} + \varepsilon_{i,t}^{h}$$

$$r_{i,t} = \alpha_i^r + \textstyle\sum_{j=0}^{k} \beta_{i,j}^{r} V_{i,t-j}^{h} + \sum_{j=1}^{k} \delta_{i,j}^{r} r_{i,t-j} + \sum_{w=0}^{W} \phi_{i,w}^{r} D_{i,t-w} + \varepsilon_{i,t}^{r} \qquad (4)$$

where $r_{i,t}$ denotes the return and $V_{i,t}$ the net trading of HFT (superscript h). All variables are aggregated into one-second intervals and standardized by their mean and standard deviation of the respective trading day. $D_{i,t-w}$ equals one for a positive and -1 for a negative shock occurring in t or less than W one-second intervals before t, and 0 otherwise. The cumulative responses are the aggregated coefficients $\varphi_{i,w}$ for 1, 10, and 30 seconds. In Panel B, the model is estimated separately for HFT and nHFT liquidity demand (HFT_d and $nHFT_d$) and supply (HFT_s and $nHFT_s$). Coefficients are aggregated per stock-day and tested using robust standard errors clustered by stock and trading day (cf. Thompson, 2011). *t*-statistics are in parentheses. Superscripts ***, **, and * denote significance at the 1%, 5%, and 10% level respectively

Panel A: Return and Trading Responses

Cum. Response	*Return*	HFT_{all}
1 second	1.50***	0.59***
	(21.80)	(18.95)
10 seconds	2.78***	0.94***
	(21.33)	(12.74)
30 seconds	3.14***	0.69***
	(21.75)	(15.84)

Panel B: Response of Liquidity Demand and Supply

Cum. Response	HFT_d	$nHFT_d$	*Diff*	HFT_s	$nHFT_s$	*Diff*
1 second	0.38***	0.21***	0.17***	-0.06***	-0.53***	0.47***
	(16.92)	(13.28)	(8.33)	(-5.66)	(-19.01)	(19.11)
10 seconds	0.52***	0.39***	0.13***	-0.01	-0.86***	0.85***
	(10.42)	(13.01)	(3.01)	(-0.70)	(-14.12)	(15.97)
30 seconds	0.36***	0.52***	-0.16***	0.01	-0.81***	0.82***
	(5.06)	(10.54)	(15.95)	(0.58)	(-8.40)	(15.91)

Table 5 Differential Return and Trading Responses to 5% and 95% Information Shocks

This table presents the intraday trading response of HFTs and nHFTs within 1, 10, and 30 seconds after 5% and 95 % futures return shocks. Coefficients in Panel A are estimated using the following VARX model:

$$V_{i,t}^{h} = \alpha_i^h + \sum_{j=1}^{k} \beta_{i,j}^{h} V_{i,t-j}^{h} + \sum_{j=1}^{k} \delta_{i,j}^{h} r_{i,t-j} + \sum_{w=0}^{W} \phi_{i,w}^{h} D_{i,t-w} + \varepsilon_{i,t}^{h}$$

$$r_{i,t} = \alpha_i^r + \sum_{j=0}^{k} \beta_{i,j}^{r} V_{i,t-j}^{h} + \sum_{j=1}^{k} \delta_{i,j}^{r} r_{i,t-j} + \sum_{w=0}^{W} \phi_{i,w}^{r} D_{i,t-w} + \varepsilon_{i,t}^{r} \quad (4)$$

where $r_{i,t}$ denotes the return and $V_{i,t}$ the net trading of HFT (superscript h). All variables are aggregated into one-second intervals and standardized by their mean and standard deviation of the respective trading day. $D_{i,t-w}$ equals one for a positive and -1 for a negative shock occurring in t or less than W one-second intervals before t, and 0 otherwise. The cumulative responses are the aggregated coefficients $\varphi_{i,w}$ for 1, 10, and 30 seconds. In Panel B, the model is estimated separately for HFT and nHFT liquidity demand (HFT_d and $nHFT_d$) and supply (HFT_s and $nHFT_s$). Coefficients are aggregated per stock-day and tested using robust standard errors clustered by stock and trading day (cf. Thompson, 2011). t-statistics are in parentheses. Superscripts ***, **, and * denote significance at the 1%, 5%, and 10% level respectively

Panel A: Return and Trading Responses		
Cum. Response	Return	HFT_{all}
1 second	0.82***	0.37***
	(28.23)	(25.68)
10 seconds	1.34***	0.52***
	(28.27)	(14.73)
30 seconds	1.54***	0.40***
	(30.38)	(19.43)

Panel B: Response of Liquidity Demand and Supply						
Cum. Response	HFT_d	$nHFT_d$	*Diff*	HFT_s	$nHFT_s$	*Diff*
1 second	0.33***	0.12***	0.21***	-0.04***	-0.35***	0.31***
	(17.70)	(20.05)	(11.42)	(-4.18)	(-27.18)	(32.49)
10 seconds	0.41***	0.26***	0.15***	-0.02	-0.56***	0.54***
	(11.91)	(23.22)	(4.62)	(-1.08)	(-18.83)	(27.74)
30 seconds	0.31***	0.41***	-0.11***	-0.03	-0.59***	0.56***
	(6.71)	(21.51)	(22.88)	(-1.17)	(-12.94)	(22.67)

Table 6 Influence of HFT Activity on Price Delay

This table presents the coefficients of a daily regression of price delay as proposed by Hou and Moskowitz (2005) on the participation rate of HFT in total trading volume (*HFT*) or liquidity demand (HFT_d) or supply (HFT_s). The estimation applies a regression without and with controls (market capitalization $logMC_{i,d}$, NYSE dummy D^{nyse}, and dummies for pre-crisis and crisis periods):

$$Delay_{i,d} = a_d + b_{1,d}HFT_{i,d} + b_{2,d}logMC_{i,d} + b_{3,d}D_{i,d}^{nyse}$$

$$+b_{4,d} \times precrisis_{i,d} + b_{5,d} \times crisis_{i,d} + \varepsilon_{i,d} \quad (9)$$

t-statistics are in parentheses. Superscripts ***, **, and * denote significance at the 1%, 5%, and 10% level respectively.

Dependent Variable: Price Delay						
HFT_{all}	-0.31***	-0.14***				
	(-63.61)	(-22.07)				
HFT_d			-0.57***	-0.41***		
			(-64.10)	(-46.24)		
HFT_s					-0.25***	0.11***
					(-36.70)	(12.99)
logmcap		-0.04***		-0.04***		-0.06***
		(-44.44)		(-65.40)		(-65.43)
NYSE		-0.04***		-0.03***		-0.05***
		(-20.84)		(-15.72)		(-24.53)
Pre		-0.08***		-0.07***		-0.05***
		(-26.70)		(-26.51)		(-17.93)
Crisis		-0.10***		-0.08***		-0.09***
		(-35.22)		(-31.41)		(-33.38)

Table 7 Influence of HFT Hard Information Activity on Price Delay

This table presents the coefficients of a daily regression of price delay as proposed by Hou and Moskowitz (2005) on the total HFT activity due to hard information (*HFThard*) or liquidity demand (*HFThard*$_d$) or supply (*HFThard*$_s$). The estimation applies a regression without and with controls (market capitalization $logMC_{i,d}$, NYSE dummy, and dummies for pre-crisis and crisis periods):

$$Delay_{i,d} = a_d + b_{1,d}HFThard_{i,d} + b_{2,d}logMC_{i,d} + b_{3,d}D_{i,d}^{nyse}$$
$$+b_{4,d} \times precrisis_{i,d} + b_{5,d} \times crisis_{i,d} + \varepsilon_{i,d} \quad (9)$$

t-statistics are in parentheses. Superscripts ***, **, and * denote significance at the 1%, 5%, and 10% level respectively.

	Dependent Variable: Price Delay					
HFT_{all}	-3.07***	-3.17***				
	(-91.01)	(-93.89)				
HFT_d			-2.70***	-2.40***		
			(-91.44)	(-68.92)		
HFT_s					0.57***	-0.36***
					(8.95)	(-6.05)
logmcap		-0.04***		-0.03***		-0.05***
		(-72.12)		(-38.75)		(-80.29)
NYSE		-0.04***		-0.05***		-0.05***
		(-22.81)		(-25.89)		(-24.39)
Pre		0.00***		-0.01***		-0.06***
		(0.76)		(-4.69)		(-23.64)
Crisis		-0.12***		-0.11***		-0.10***
		(-49.70)		(-44.13)		(-38.31)

7. Figures

Figure 1 Overview of Return and Trading Response

This figure plots the cumulative response of stock returns (right y-axis) and net trading of HFTs and nHFTs (left y-axis) to hard information shocks estimated using the VARX model in Equation 4. All variables are aggregated into one-second intervals and standardized using mean and standard deviation of the stock-day. Coefficients are aggregated per stock-day. The x-axis represents the 30 lags after an information shock in seconds.

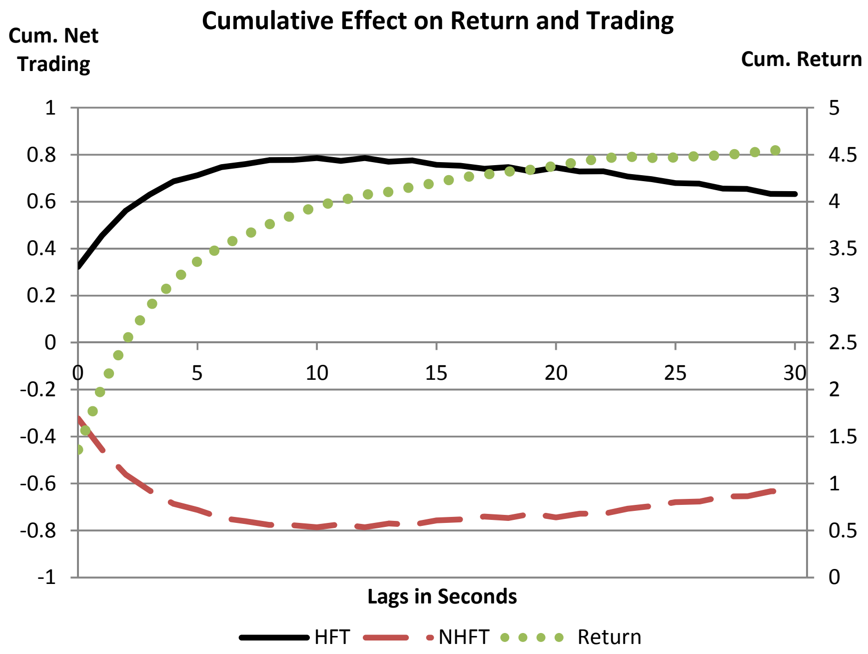


Figure 2 Response of Liquidity Demand and Supply to Information Shocks

This figure plots the trading response of HFT liquidity demand and supply (*HFT_d* and *HFT_s*, black solid line) and nHFT liquidity demand and supply (*NHFT_d* and *NHFT_s*, red dashed line) after an information shock (as a result of the VARX model in Equation 5). All trade variables are standardized by mean and standard deviation of the stock-day. The x-axis represents the lags after an information shock in seconds.

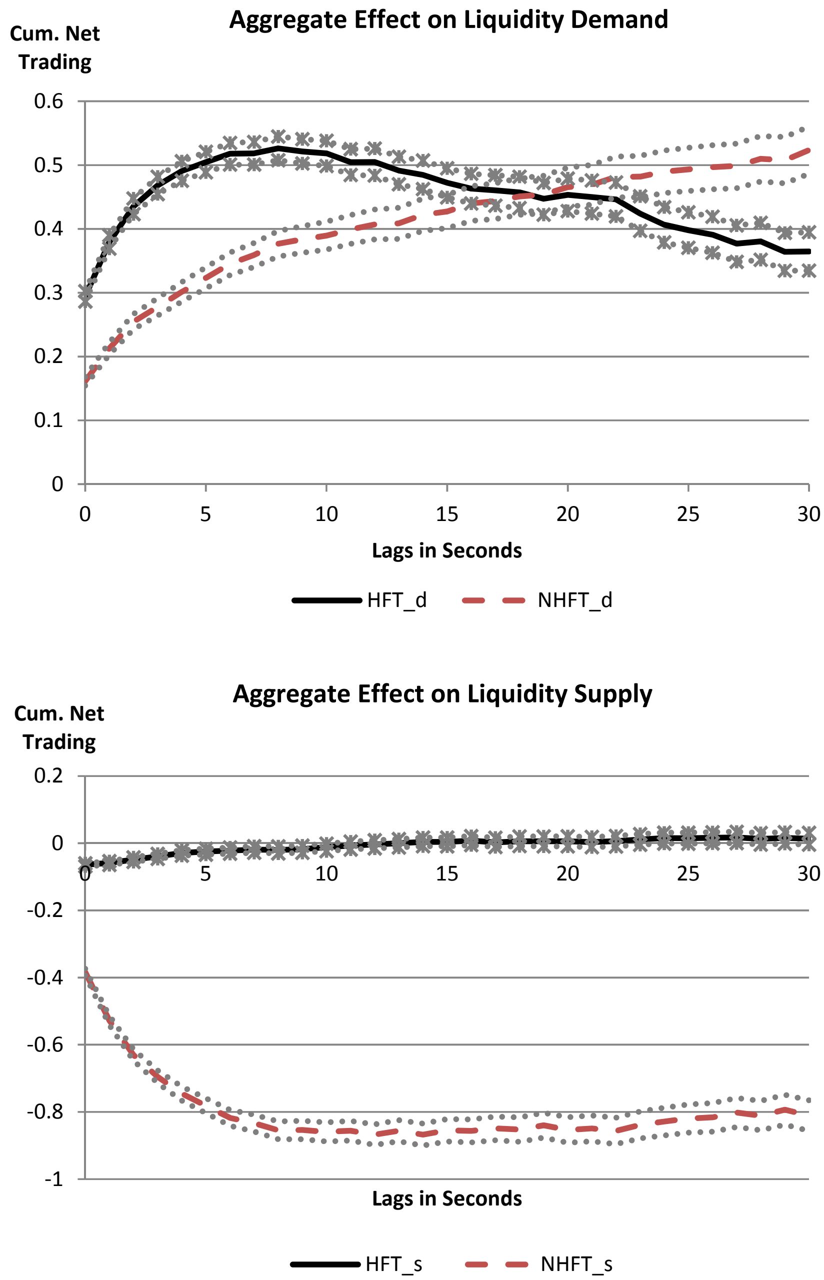
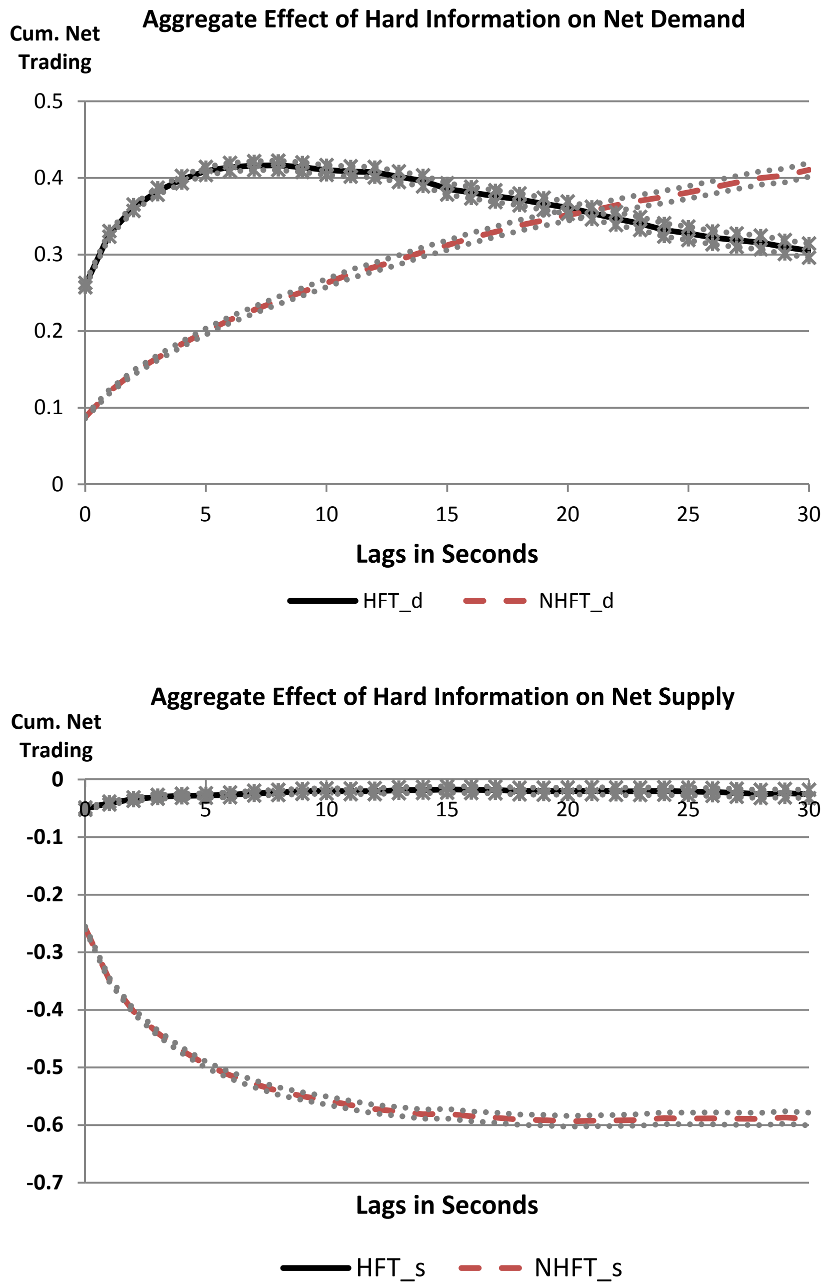


Figure 3 Response of Liquidity Demand and Supply to 5%/95% Information Shocks

This figure plots the trading response of HFT liquidity demand and supply (*HFT_d* and *HFT_s*, black solid line) and nHFT liquidity demand and supply (*NHFT_d* and *NHFT_s*, red dashed line) after a 5%/95% information shock (as a result of the VARX model in Equation 5). All trade variables are standardized by mean and standard deviation of the stock-day. The x-axis represents the lags after an information shock in seconds.



Appendix A List of Sample Stocks

Table A.1 Sample Statistics

This table presents the descriptive statistics of 80 sample stocks and the absolute and relative HFT activity. *Total* denotes the average total number of trades per stock day, *HFTrades* the number of trades involving an HFT. *Abs.HFTd* and *Abs.HFTs* denote the absolute number of trades by liquidity supplying and demand HFT respectively, while *Rel.HFTd* and *Rel.HFTs* is the relative percentage of HFT demand and supply.

Ticker	*Total*	*HFTrades*	*Abs.HFTd*	*Abs.HFTs*	*Rel.HFTd*	*Rel.HFTs*
AA	52,558	27,712	11,474	16,238	0.45	0.61
AAPL	169,252	83,803	40,882	42,921	0.48	0.48
ADBE	43,152	18,529	9,025	9,504	0.43	0.44
AGN	8,823	2,771	1,718	1,053	0.40	0.24
AINV	7,556	2,280	1,066	1,214	0.29	0.33
AMAT	66,867	34,405	14,037	20,367	0.42	0.61
AMED	6,105	1,708	1,105	604	0.35	0.19
AMGN	49,972	18,732	8,725	10,007	0.35	0.40
AMZN	56,634	22,506	14,393	8,114	0.51	0.28
ARCC	4,752	1,379	620	759	0.24	0.29
AXP	56,661	28,795	13,603	15,192	0.49	0.52
AYI	2,072	555	433	121	0.42	0.13
BARE	4,753	956	529	427	0.21	0.20
BHI	25,017	12,857	7,481	5,376	0.61	0.42
BIIB	22,860	7,461	4,707	2,754	0.41	0.24
BRCM	62,542	29,984	13,932	16,052	0.45	0.51
BRE	4,325	1,894	1,288	606	0.58	0.26
BXS	3,470	1,384	993	392	0.57	0.23
CB	14,581	6,471	3,954	2,517	0.53	0.34
CBT	2,203	700	509	192	0.45	0.16
CELG	30,968	10,264	6,072	4,192	0.39	0.27
CETV	4,034	935	612	323	0.31	0.16
CHTT	3,282	822	592	230	0.36	0.15
CKH	1,281	488	366	122	0.57	0.19
CMCSA	78,308	41,952	17,492	24,459	0.44	0.63
CNQR	6,125	1,454	1,088	366	0.35	0.12
COO	2,018	480	346	134	0.35	0.13
COST	38,738	16,532	9,630	6,902	0.50	0.36
CPWR	12,880	5,634	2,341	3,293	0.35	0.52
CR	1,481	382	301	81	0.41	0.11
CRI	2,603	673	448	226	0.35	0.17
CSCO	124,424	65,480	27,151	38,329	0.43	0.61
CSE	8,656	3,541	1,602	1,939	0.39	0.47
CSL	1,770	470	343	127	0.38	0.15
CTSH	29,615	12,013	6,092	5,921	0.42	0.40
DELL	74,994	37,223	15,122	22,102	0.40	0.59
DIS	39,664	20,885	9,017	11,868	0.46	0.59
DOW	35,979	18,129	8,171	9,958	0.46	0.53
EBAY	59,286	27,726	11,507	16,219	0.39	0.55
ERIE	1,253	321	246	75	0.38	0.12

Table A.1 Sample Statistics

Ticker	*Total*	*HFTrades*	*Abs.HFTd*	*Abs.HFTs*	*Rel.HFTd*	*Rel.HFTs*
ESRX	19,741	5,945	3,999	1,946	0.41	0.20
EWBC	9,162	3,208	1,856	1,352	0.42	0.29
FCN	3,413	907	667	240	0.39	0.13
FL	9,880	4,972	2,655	2,317	0.51	0.46
FMER	8,742	3,281	2,167	1,115	0.48	0.26
FULT	10,175	4,279	2,198	2,081	0.43	0.41
GAS	2,593	845	608	237	0.47	0.19
GE	122,616	70,809	28,545	42,264	0.47	0.67
GENZ	23,165	7,403	4,755	2,648	0.41	0.23
GILD	49,144	18,135	9,011	9,124	0.37	0.37
GLW	37,611	19,123	7,952	11,171	0.44	0.59
GOOG	44,131	20,420	12,154	8,267	0.55	0.37
GPS	33,076	18,258	8,321	9,937	0.50	0.59
HON	23,230	11,094	5,607	5,487	0.48	0.47
HPQ	57,979	28,473	11,547	16,926	0.41	0.57
INTC	131,109	69,357	27,759	41,598	0.42	0.63
ISIL	17,478	7,212	3,644	3,567	0.42	0.41
ISRG	10,501	4,070	2,627	1,444	0.50	0.27
JKHY	6,006	1,809	999	810	0.33	0.28
KMB	8,865	3,372	1,910	1,462	0.43	0.33
KR	26,584	13,731	6,273	7,458	0.47	0.55
LANC	1,773	375	264	111	0.29	0.13
LECO	3,402	1,072	793	280	0.45	0.17
LPNT	6,448	1,559	1,042	518	0.32	0.16
LSTR	7,240	2,569	1,909	660	0.51	0.19
MANT	2,824	739	537	201	0.35	0.15
MELI	4,834	1,115	725	391	0.29	0.15
MMM	19,126	8,626	4,832	3,794	0.50	0.37
MOS	26,226	12,610	7,175	5,434	0.55	0.39
NSR	2,656	588	404	183	0.31	0.15
NUS	1,106	244	157	86	0.26	0.14
PFE	69,450	37,313	13,525	23,787	0.39	0.68
PG	48,559	23,052	9,395	13,657	0.38	0.54
PNC	23,068	10,704	6,524	4,180	0.57	0.35
PNY	1,459	341	266	75	0.33	0.12
PTP	1,806	451	346	105	0.38	0.12
ROC	2,456	629	429	199	0.35	0.15
SF	1,419	460	344	116	0.46	0.17
SFG	1,924	658	491	166	0.49	0.16
SWN	21,104	9,727	5,965	3,763	0.56	0.34
All Stocks	26,545	12,498	5,867	6,630	0.42	0.33

Appendix B Robustness over Time

Table B.1 Differential Trading Responses - Robustness over time

This table presents the impact of information shocks on the cumulative response of HFT and nHFT liquidity demand and supply for different time periods of high and low volatility, taking a VIX value of 30 as the limit for high/low volatility before and after the financial crisis. The VARX model is the same as in Equation 5. Panel A reports the cumulative response 1, 10, and 30 seconds after hard information shocks on liquidity demand of HFT (HFT_d) and nHFT ($nHFT_d$), liquidity supply of HFT (HFT_s) and nHFT ($nHFT_s$) and their difference (*Diff*). Variables are aggregated per stock-day and differences are tested using double clustered standard errors on stock and trading day (cf.Thompson, 2011). *t*-statistics are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level respectively.

Panel A: Effects on Liquidity Demand									
	Low Volatility (Pre-Crisis)			High Volatility			Low Volatility (Post-Crisis)		
	HFT_d	$nHFT_d$	*Diff*	HFT_d	$nHFT_d$	*Diff*	HFT_d	$nHFT_d$	*Diff*
1 second	0.52***	0.34***	0.18***	0.30***	0.14***	0.16***	0.39***	0.22***	0.17***
	(11.27)	(8.52)	(4.16)	(13.18)	(10.75)	(7.53)	(8.30)	(7.35)	(4.23)
10 seconds	0.74***	0.56***	0.18**	0.40***	0.27***	0.13***	0.52***	0.45***	0.07
	(7.76)	(8.66)	(1.97)	(9.62)	(11.86)	(3.33)	(5.68)	(6.97)	(0.94)
30 seconds	0.51***	0.70***	-0.19	0.31***	0.40***	-0.09*	0.30**	0.60***	-0.29**
	(3.47)	(6.31)	(-1.45)	(5.58)	(10.31)	(-1.69)	(2.03)	(5.75)	(-2.33)

Panel B: Effects on Liquidity Supply									
	Low Volatility (Pre-Crisis)			High Volatility			Low Volatility (Post-Crisis)		
	HFT_s	$nHFT_s$	*Diff*	HFT_s	$nHFT_s$	*Diff*	HFT_s	$nHFT_s$	*Diff*
1 second	-0.09***	-0.79***	0.70***	-0.04***	-0.39***	0.35***	-0.06***	-0.52***	0.46***
	(-4.29)	(-12.71)	(13.21)	(-4.87)	(-15.27)	(15.69)	(-3.25)	(-9.04)	(9.19)
10 seconds	0.04	-1.36***	1.39***	-0.04***	-0.60***	0.55***	0.00	-0.86***	0.86***
	(1.09)	(-11.14)	(12.51)	(-3.13)	(-12.47)	(13.20)	(-0.05)	(-7.11)	(8.05)
30 seconds	0.10**	-1.32***	1.41***	-0.04**	-0.58***	0.54***	0.03	-0.73***	0.76***
	(1.98)	(-6.47)	(7.64)	(-2.03)	(-7.86)	(8.44)	(0.57)	(-3.68)	(4.38)

Appendix C Trading Responses to Positive and Negative Shocks

Table C.1 Differential Trading Responses to Positive and Negative Information Shocks

This table presents the intraday trading response of HFTs and nHFTs within 1, 10, and 30 seconds after and positive and negative futures shocks, estimated using the same VARX model as in Table 3/ Equation 5. The model is applied to liquidity-demanding HFTs and nHFTs (*HFTd* and *nHFTd*) and liquidity-supplying HFTs and nHFTs (*HFTs* and *nHFTs*). All variables are aggregated into one-second intervals and standardized using mean and standard deviation of the stock-day. Variables are aggregated per stock-day and tested using robust standard errors clustered by stock and trading day (cf. Thompson, 2011). *t*-statistics are in parentheses. Superscripts ***, **, and * denote significance at the 1%, 5%, and 10% level respectively.

Panel A: Positive Shocks						
Cum. Response	*HFTd*	*nHFTd*	*Diff*	*HFTs*	*nHFTs*	*Diff*
1 second	0.36***	0.20***	0.16***	-0.05***	-0.51***	0.45***
	(14.41)	(11.25)	(7.02)	(-4.97)	(-16.10)	(16.21)
10 seconds	0.42***	0.34***	0.08*	0.00	-0.73***	0.73***
	(8.48)	(10.59)	(1.81)	(-0.01)	(-11.88)	(13.53)
30 seconds	0.29***	0.39***	-0.10***	0.01	-0.62***	0.63***
	(4.05)	(7.91)	(11.95)	(0.43)	(-6.40)	(11.81)

Panel B: Negative Shocks						
Cum. Response	*HFTd*	*nHFTd*	*Diff*	*HFTs*	*nHFTs*	*Diff*
1 second	-0.33***	-0.17***	-0.16***	0.05***	0.44***	-0.39***
	(-14.42)	(-12.16)	(-7.93)	(5.40)	(16.67)	(-17.19)
10 seconds	-0.49***	-0.31***	-0.18***	0.03	0.75***	-0.72***
	(-10.25)	(-11.44)	(-4.76)	(1.57)	(12.53)	(-13.90)
30 seconds	-0.36***	-0.45***	0.10***	0.00	0.72***	-0.72***
	(-5.42)	(-9.55)	(-12.83)	(0.15)	(8.09)	(-12.77)

Footnotes

[1] High-frequency trading is a subcategory of algorithmic trading (AT), which is commonly defined as the use of computer algorithms to support the trading process (cf. Hendershott, Jones, & Menkveld, 2011).

[2] Price pressures are transitory price changes that are not related to permanent price changes because of new information.

[3] On May 6th, 2010, the Dow Jones Industrial Average dropped rapidly by 10% followed by a similarly rapid recovery within half an hour. Easley, de Prado, and O'Hara (2011) find that while order flow toxicity measured by the volume-synchronized probability of informed trading (VPIN) increased before the "Flash Crash" took place.

[4] I thank SIRCA for providing access to the Thomson Reuters DataScope Tick History.

[5] As the total trading volume of liquidity-demanding and -supplying trades is zero sum, total nHFT net trading (buy minus sell volume) is the negative value of HFT net trading. Thus, I omit the nHFT variable as it is complementary to the HFT results.

[6] I greatly appreciate this suggestion provided by an anonymous referee.

[7] I thank an anonymous referee for suggesting this since it allows for a more robust analysis.