Pricing information: Evidence from transaction-level data in financial markets

Anirban Banerjee* — *IIMA*
Arnab Chakrabarti† — *FLAME*
Anindya S. Chakrabarti‡ — *IIMA*

Abstract

High-frequency trading is known to facilitate incorporating information into asset prices. Do such traders utilize aggregate market-wide information or idiosyncratic stock-specific information? In this paper, we exploit novel transaction-level data from the National Stock Exchange in India, to construct a measure of information by order imbalance i.e. excess demand for assets. Upon filtering out aggregate factors, we show that the traders act in accordance to the direction of information – importantly, magnitude of information is barely related to high frequency trades, the direction matters a lot more. Market-wide and idiosyncratic information are used almost equally – even during times of extreme fluctuations. All of our results appropriately control for aggregate market-index return, stock specific liquidity and they are robust to a range of fixed effects accounting for stock- , day- and minute-level unobserved heterogeneity.

JEL Codes: G14, G12, G10
Keywords: High-frequency traders, Order imbalance, Asset prices, Market-wide and stock-specific information.

Word Count: 7270

*Finance and Accounting Area, Indian Institute of Management Ahmedabad, Gujarat, India. Email: anirbanb@iima.ac.in
†Department of Economics, FLAME University, Maharashtra, India. Email: arnab.chakrabarti@flame.edu.in
‡Economics Area, Indian Institute of Management Ahmedabad, Gujarat, India. Email: anindyac@iima.ac.in

1 Introduction

What moves stock prices? Known contenders are information, news and noise. Brogaard *et al.* (2022) recently have shown that 8% of the return variance can be explained by market-wide information whereas around 61% of the return variance can be attributed to firm-specific information. Starting from information and ending with pricing fluctuations leave one of the intermediate steps unclear. The pricing of information happens due to trades executed by the traders in the market. How do they exactly carry out these trades in response to different kinds of information? A complete answer to this question necessitates observing individual traders in the market and analyzing their trading history. However, such datasets are practically non-existent. In this paper, we take up the question and pursue it in a different form. How do traders, in particular High-Frequency Traders (HFTs henceforth), exploit information which are market-wide or stock-specific for carrying out their trades? We exploit a novel transaction-level dataset from the National Stock Exchange in India (NSE henceforth) to answer this question.

We consider HFTs as the focal point for our analysis due to the prominent role they play in modern financial markets to assimilate information into asset prices. High frequency trading is a relatively new and one of the most important technological disruptions in the recent times. The last decade has seen an exponential growth in algorithmic trading activity around the globe. By definition, algorithmic trading refers to the phenomenon of computer programmes automatically making trading decisions (Menkveld, 2016). HFTs belong to the group of traders who engage in algorithmic trading. These traders carry out extremely fast trading, with trading decisions executed in fraction of seconds. Due to the inherent advantage of speed and the potential to manipulate financial markets, such traders have garnered significant academic and policy interests in the recent past. Through their role in provisioning of continuous liquidity to financial markets, HFTs have often been ascribed the role of modern market makers (Hagströmer & Nordén, 2013; Menkveld, 2013).

A large stream of the literature on HFT is focused on their use of information. HFT trading activity is correlated to public information (Brogaard *et al.*, 2014; Zhang, 2012; Chakrabarty *et al.*, 2020). They are known to react faster than other groups of traders to publicly disclosed corporate announcements (Frino *et al.*, 2017) as well as macroeconomic announcements (Scholtus *et al.*, 2014). Due to these trading activities, HFTs enhance the speed of information getting assimilated to security prices, improving price efficiency. At the same time, due to the diffusion of aggregate market-level information, HFT activity also result in significant co-movements of security returns and liquidity (Malceniece *et al.*, 2019). There are complementary channels through which HFT trading strategies may affect stock returns. One of the possible channels is HFTs' market-making strategies which require optimal inventory control. HFT market-making also requires the provisioning of liquidity during periods of stress. As this liquidity-provisioning is typically related to an increase in expected earnings (Nagel, 2012), HFT trading strategies may be positively related to stock returns. This association may also stem from HFT's front-running and back-running algorithms. Front-running on stock-specific information results in predatory-type trading behavior (Brunnermeier & Pedersen, 2005). Back-running stems from these traders trying to identify orders split from larger informed (possibly institutional) orders (Yang & Zhu, 2020). While the literature has developed an understanding of timing of HFTs trading activities with respect to information and its broad correlations with aggregate information, there is a prominent gap in understanding how HFTs utilize information which are market-wide or stock-specific.

In this paper, we analyze how HFTs use information which can be specific to stocks or the market as a whole. Let us elaborate on the novelty of this decomposition of information in terms of its scope. The present literature primarily focuses on HFT activity immediately after public information disclosures. The focus is on whether the source of the information is public or private. However, they do not provide any evidence if HFTs primarily use market-wide or stock-specific information to price individual securities. Both of these types of information can have a public or a private source. Thus the current literature cannot differentiate between the two types of effects – is it related to individual stocks or the market as a whole? We take a direct approach to decompose information into stock-specific and market-wide components and examine which component has higher impact on stock pricing via HFT trades.

As information can not be directly observed, we use order imbalance (OIB) as an ex-post measure of information. We know from the work of Chordia *et al.* (2002) and Chordia & Subrahmanyam (2004) that order imbalance is positively related to security prices. In presence of positive (negative) information in the market, we anticipate the order imbalance to be also positive (negative). This motivation comes from the observation of Chordia & Subrahmanyam (2004) that order imbalance can arise in intermediated markets either due to inventory pressure of the market-maker or asymmetric information. NSE is completely driven by limit orders with no designated market-maker. Also, HFTs are known not to carry inventory, even on an intraday basis. Therefore, the source of order imbalance at NSE for HFTs is new information. Building on this insight, we create a measure of order imbalance specifically for the HFTs and decompose it to construct stock-specific and market-wide components.

We find that HFT order imbalance is positively associated with individual stock returns at an intraday level. The HFTs rely on the direction of the information (e.g. positive or negative) rather than the magnitude of the information. The informativeness of the HFTs order imbalance can arise from either stock-specific or market-wide factors. To differentiate between the two, we employ a factor model to decompose the order imbalance into components due to stock-specific or market-wide information. Empirically, both of these components have significant impact on the asset returns – with heterogeneous effects over time. Most of the time during our period of analysis, the impact of stock-specific information seem to be more than market-wide information. However, during periods of extreme price movements, both these components result in very similar magnitudes of effects.

Does order imbalance capture information? In short, we argue that - it does. In an order-driven market, Investors can participate in a trade in two ways-either as an active participant, where they act as liquidity-taker; or as a passive participant, where they act as liquidity providers. Due to time-sensitive nature of financial market information, informed investors are more likely to participate in trades as active participants, rather than passive liquidity-providers. This allows us to use order imbalance as a proxy of information, in absence of better alternatives.

An intuitive way to see why is that we can interpret order imbalance for an asset as excess demand for that asset. Note that we are not calculating order imbalance of an individual investor. We are instead computing it at the level of an asset aggregating over all investors. Since order imbalance is effectively excess demand for an asset, order imbalance would be positive (negative) in presence of positive (negative) information. Therefore, changes in information available at the level of traders will likely be captured by the order imbalance.

However, one can now flip the question and ask - is information the only source of fluctuations in order imbalance? In practice, order imbalance may also arise out of liquidity demand of traders, hedging needs and need to manage their inventory. However, these effects would be zero-mean and would not affect our analysis for the following reason. As described above, we do not compute the order imbalance of individual traders/investors. Instead, we estimate the order imbalance of an entire trader group aggregating over all traders belonging to that group. Though individual traders at any point in time may execute trades that are not motivated by information, and goes against the direction of prevalent information (based on their liquidity or hedging needs), it is extremely improbable that a large group of investors in a trading group will do so impacting the order imbalance aggregated at the level of the market.

To summarize the above discussion, information is the most likely source that can synchronize trading positions for a trading group as a whole. Therefore, we consider asset-level order imbalance to reflect information.

The rest of the paper is arranged as follows. Section 2 describes the related literature.

Section 3 describes the data and institutional setting. In section 4, we carry out and describe the main empirical analysis on information and prices. Section 5 provides robustness checks through panel estimation with fixed effects to account for unobserved heterogeneity. Section 6 summarizes and concludes.

2 Related Literature

Our paper relates to three strands in the literature. First, we review the current understanding of HFT behavior and the resulting impact on the market microstructure. The extant literature on HFTs has primarily focused on their impact on market quality. The evidence on the impact on market quality has been mostly positive. HFTs have been shown to improve market quality in terms of bid-ask spread and price efficiency (Menkveld, 2013; Hagströmer & Nordén, 2013; Brogaard *et al.*, 2014). HFTs monitor liquidity more closely compared to human traders and help the market by acting as the liquidity provider when it is expensive and consume liquidity when it is cheap (Hendershott & Riordan, 2011). The quotes provided by HFTs are seen to be more efficient (Hendershott *et al.*, 2011). HFTs trade in the direction of permanent price changes and against transitory changes (Brogaard *et al.*, 2014). More generally, algorithmic traders improves informational efficiency through faster price discovery, though it imposes adverse selection cost to human traders (Chaboud *et al.*, 2014).

There is a large literature on market volatility and agent behavior (Schroeder *et al.*, 2020). The behavior of HFTs are typically endogenous to the state of the market. It is not a priori clear exactly what causes large prices changes resulting in volatility. Farmer *et al.* (2004) attribute large price changes to changes in supply of liquidity. Boudt & Petitjean (2014) find evidence consistent with this liquidity-oriented view on price jumps. Jiang *et al.* (2011) on the other hand, show that macroeconomic news announcements account for a large fraction of price jumps. How do HFTs respond to such jumps? Possibly they exacerbate the volatility by adversely reacting through liquidity withdrawal or by not providing the liquidity to counteract the trades of non-HFTs. Notably, Kirilenko *et al.* (2017) found such behavior during the 2010 flash crash in the Dow Jones Index, an event that prompted policy-makers to be worried about the role of HFTs in contributing to market instability.[1] Golub *et al.* (2012) suggests that that stock-specific flash crashes may have an origin from the trading nature of HFTs. This behavior is consistent with the empirical findings on a larger class of endogenous liquidity providers (Anand & Venkataraman, 2016).

The second stream of literature our work relates to is on order imbalance and its relationship

[1] US Securities and Exchange Commission's investigation indicated a possible 'algorithmic' origin of this crash. While HFTs by themselves did not initiate the crash, they were thought of to be net contributors to it.

with information. In one of the earliest papers related to order imbalance, Chordia *et al.* (2002) show that market returns are positively related to contemporaneous and lagged order imbalances. The same holds true for individual stock returns (Chordia & Subrahmanyam, 2004). Using a similar notion of order imbalance, Cushing & Madhavan (2000) show that trading in the last five minutes of any day explains a large fraction of stock variation. Chan & Fong (2000) shows that order imbalance explains stock returns more for large stocks compared to small stocks. There are other related notions which, in principle, are very similar to the idea of order-imbalance as mentioned by Chordia *et al.* (2002), such as marketable imbalance (Hirschey, 2021), trade-imbalance (Hsieh & Lee, 2021), or buying-selling pressure (Bollen & Whaley, 2004; Kang & Park, 2008).

Hirschey (2021) shows that through their analysis of past patterns in orders and trades, HFTs are able to predict order flow from other class of traders. He argues that through this anticipatory channel, HFTs increase trading cost for non-HFTs. Using trade imbalance as measure of informativeness, Hsieh & Lee (2021) show that mutual funds and foreign institutions are the primary source of information in the market.

The idea of buying/selling pressure is more prevalent in the derivatives literature. Bollen & Whaley (2004) show that the changes in implied volatility is directly related to buying pressure arising from public order flow. Kang & Park (2008) further show that the net buying pressure of calls (puts) increases the implied volatility of calls (puts), while the net buying pressure of puts (calls) lowers the implied volatility of calls (puts).

Third, our paper relates to the literature on information usage in trading. It is understood for a very long time that traders act on information which gets reflected in the asset prices. In fact, this idea itself forms the basic tenet of efficient market hypothesis. One strand of the current literature attempts to understand the modes of analyzing such information (Bernales *et al.*, 2022). We are interested in a complementary strand of literature which analyzes the sources of such information – does it have stock-specific or market-wide origin?

Sequential trade models (Copeland & Galai, 1983; Glosten & Milgrom, 1985; Easley & O'hara, 1987) in traditional market microstructure theory explains how public information gets incorporated into security prices. Using Dow Jones announcements as proxis, Mitchell & Mulherin (1994) show that public information is directly related to market activity. Berry & Howe (1994) show a moderate positive relationship between public information (news releases by Reuter's) and trading volume.

Kyle (1985) explains the process through which private information that gets assimilated into prices through the process of continuous trading. In follow-up empirical work, Ito *et al.* (1998) show the presence of private information in the Tokyo FX market. In more recent work, Kacperczyk & Pagnotta (2019) show that on the days of private information based trading, the indicators for asymmetric information display abnormal values. Volume and volatility are unusually high, whereas illiquidity is low.

3 Data and Institutional Background

3.1 List of Variables

Let us first introduce some notations. We denote the number of observation on any day by n_0 and the total number of days by n_1. The total number of observations, therefore, is $n_0 \times n_1$. Throughout the paper, we denote this as n. We denote the spot return of the ith stock at day t and intraday period τ as $r_{i,t,\tau}$, where $i = 1,\ 2,\ ...,\ p$, $t = 1,\ 2,\ ...,\ n_1$ and $\tau = 1,\ ,2,\ ...,\ n_0$. The NIFTY50[2] index return at date t and intraday period τ is denoted by $r^m_{t,\tau}$. We denote the measure of liquidity as $L_{i,t,\tau}$. Return and liquidity measures are standard in the literature. So we will not elaborate on them.

Here is a list of order imbalance-related variables we use throughout the paper.

$OIBNUM_{i,t,\tau}$: the number of buyer-initiated trades less the number of seller-initiated trades during the period τ on the t-th date for the i-th stock scaled by the total number of trades in the said period.

$OIBVOL_{i,t,\tau}$: the INR aggregate volume of buyer-initiated trades less the INR aggregate volume of seller-initiated trades during the period τ on the t-th date for the i-th stock scaled by the total INR volume of the trades during the said period.

These two measures of order imbalance (OIBNUM and OIBVOL) expresses the aggregate information in the market. We may construct a similar measure specific to the HFTs.

$O^{HFT}_{i,t,\tau}$: the number of HFT buyer-initiated trades less the number of HFT seller-initiated trades during the period τ on the t-th date for the i-th stock scaled by the total number of trades in the said period.[3]

This measure of order imbalance is representative of the ex-post information utilized by HFTs to execute their trades.

[2]NIFTY50 is the benchmark value-weighted index of NSE consisting of the 50 largest firms.

[3]Please note that though we have defined the order imbalance measure for the HFTs based on the trade counts, a similar measure may be developed using the value of trades as well.

Table 1: Descriptive Statistics

Variable	No. of Obs.	Mean	SD	Minimum	Maximum
OIBNUM	14,512,125	-0.0141	0.4957	-1.00	1.00
OIBVAL	14,512,125	-0.0098	0.5562	-1.00	1.00
HFT OIBNUM	14,512,125	-0.0036	0.1831	-1.00	1.00
HFT OIBVAL	14,512,125	-0.0019	0.2129	-1.00	1.00
Spot Return ($\times 10^4$)	14,325,459	-0.0585	14.0320	-1,823.22	1,083.80
Index Return ($\times 10^4$)	14,370,449	-0.0442	4.0113	-74.78	123.59
Spot Trade Count	14,325,459	83.3779	133.5942	1.00	26,620.00
Spot Traded Value (INR)	14,325,459	2.30E+06	2.74E+07	5.45	7.05E+10
Liquidity ($\times 10^{-8}$)	11,963,003	76.4488	833.4658	5.91E-06	1.33E+06

Note: The table presents the summary statistics for the sample stocks traded in NSE during the year 2015. The variables are computed at one minute intervals for each of the 166 sample stocks during the entire calendar year of 2015 which corresponds to 246 trading days. Here, we have scaled spot and index returns and liquidity for the ease of comparison of magnitudes. None of these scaling affect the results.

3.2 Institutional Background

The Market: For our study, we use an unique dataset obtained from the National Stock Exchange (NSE) of India. NSE is a completely limit order driven market without any traditional market-maker (or specialist). The exchange was established in 1992 and within three decades it has surpassed the incumbent Bombay Stock Exchange (BSE) in terms of market share. Presently the Indian stock market operates as a duopoly with NSE having more than 92.5% of the equity market trading volume for FY 2021-22.[4] At present there are more than 2000 firms listed in NSE. The exchange operates on a price-time priority. The market operates from 9:15 AM IST (GMT + 5:30) to 3:30 PM IST without any breaks. There is a pre-open period for the spot market (from 9:00 AM to 9:15 AM) which is used to discover daily opening price through a call-auction. Presently NSE stands as one of the largest stock exchange in the world with an annual turnover of more than 2.18 trillion USD (FY 2021-22) in the spot market.[5][6]

In our dataset, we have exact identifiers for algorithmic orders and subsequent trades. This implies that we do not need to use proxies for identification of algorithmic traders, which has been one of the majors limitations in this literature. The exchange had 1,794 stocks listed during the year 2015. In our analysis, we consider trading data for all the stocks listed in NSE that are permitted to issue derivative securities. This criteria ensures that the selected stocks are large-cap stocks with sufficient liquidity.

[4]For FY 2015-16, NSE had a market share of 85%.

[5]For FY 2015-16, the turnover was 639 billion USD.

[6]World Federation of Exchanges website – *statistics.world-exchanges.org* accessed on 25 Sep 2022.

Consistent with academic (Hendershott *et al.*, 2011) as well as regulatory conventions (SEC, 2010), we use proprietary algorithmic traders as our proxy for high-frequency traders. Existing literature suggests that HFT contribute significant proportion of the order messages, have high order-to-trade ratio and relatively low traded volume. For our dataset, we find that HFTs contribute 76.37% of the message traffic and participate in 13.18% of the trades.[7]

Order Imbalance: We use the framework provided by Chordia *et al.* (2002) and Chordia & Subrahmanyam (2004) to use order imbalance as a measure of aggregate ex-post information. In a limit order driven market without any designated intermediary (market-maker), information asymmetry is likely to drive order-imbalance. In presence of positive (negative) information in the market, we expect more trades in a particular time interval to be initiated by buyers (sellers) who expect price appreciation (depreciation) in the future.

By definition, the normalized order imbalance measures $OIBNUM_{i,t}$ and $OIBVOL_{i,t}$ take a value between -1 to +1, where -1 indicates all trades initiated by sellers and +1 indicates all trades initiated by buyers. In the context of intermediated markets, order imbalance can arise from information as well as the inventory pressure of the market-maker. However, in the context of a completely limit order driven electronic market, order imbalance can be conceptualized as an ex-post measure of aggregate information in the market. It is well established that the contemporaneous order imbalance measures are positively correlated with the stock returns in that period.

How well does order imbalance reflect active trade? In short, the answer is - quite well. Traditionally, high frequency traders are known to be liquidity providers, primarily participating in trades as passive participants. However, with time the number of high frequency traders have increased, resulting in reduced profitability of passive market making trades. We observe that almost half of the trades executed by HFTs in the Indian market is active or liquidity-demanding. Our estimation of order imbalance considers only the trades where HFTs participated as active participants.

Theoretically it may be possible that informed investors place non-marketable orders that would not reflect in the order imbalance. But such orders are unlikely to reflect price-sensitive information. In the case of HFTs, it becomes even more difficult to track informativeness of orders that are not executed. Especially because HFTs modify their orders very frequently

[7]Over the years, NSE has evolved significantly in its trader composition. Though it is difficult to pinpoint an exact time when algorithmic trading started in NSE, the introduction of Direct Market Access (DMA) in 2008 was possibly the first step, followed by the provision of co-location services in 2010. Since then, the market had witnessed an exponential growth in algorithmic trading activity. The market consists of other types of traders as well. Proprietary non-algorithmic traders, who execute trades using their own funds contribute 7.45% of the overall traded volume during the period. Institutional traders contribute 32.74% while other traders (high net individuals, corporate entities, retail traders etc.) contribute 46.63% of the traded volume.

leading to a high order to trade ratio. As such, we only focus on possible price sensitive information, that is acted upon by HFTs through marketable orders.

3.3 Data Description

For our study, we use a novel dataset of order and trade level data obtained from the National Stock Exchange of India. Our dataset consists of 166 stocks which are permitted to be traded in the derivative market. These are usually stocks with highest liquidity. In this dataset, we can exactly identify the algorithmic traders – in particular, HFT traders.[8] The dataset spans across the calendar year of 2015 consisting of 246 trading days.

For all the stocks, we create a time series of intraday order imbalance and returns computed at one minute interval. A one-minute interval allows us to capture enough trades so that the price series is not flat, and at the same time, it is granular enough to capture high-frequency behavior of the traders.

Table 1 provides descriptive statistics of the minute-level trading data regarding the order imbalance measures and individual stock returns. The market in general was slightly bearish for the period. The benchmark NIFTY50 stock index declined from a level of 8284.00 at the closing of 1st Jan 2015 to 7946.30 at the closing on 31st Dec 2015. The mean value of the returns and imbalance measures in this period are both negative. As the trading activity differs significantly across the selected stocks, we observe high degree of dispersion for the number of trades (spot trade count) and the traded volume (in INR). Stock and index returns are estimated as log returns over the one-minute periods.[9] Our liquidity measure is constructed as the inverse of the illiquidity measure proposed by Amihud (2002).

4 How is Information Priced?

HFTs are generally thought of to be informed traders (Brogaard *et al.*, 2014) in the line of using information efficiently as described by Kyle (1985). Here we empirically analyze the informational component of the HFT trades.

[8] The algorithmic identifier flag in the dataset is provided by the exchange (National Stock Exchange India). We combine this flag with the type of trader (Proprietary) flag, to identify Proprietory-Algorithmic (PA) traders.

[9] In general, this is estimated as close-to-close returns over the one-minute intervals, but for the first trading minute of the day, this is estimated as open to close returns.

4.1 Constructing Aggregate and Stock-specific Factors

To analyze the relationship between stock returns and information, we have to look at the idiosyncratic component of returns after controlling for known factors that explain return variability. There are well-recognized factors that explain stock returns at a daily level. Since the introduction of the Capital Asset Pricing Model (Sharpe, 1964; Lintner, 1969), the aggregate market return has been used as one of the primary factors to explain individual stock returns. Further, Amihud & Mendelson (1986) show that even after accounting for aggregate market risk, illiquidity demands a premium for investors. To account for these two channels, we use a simple two-factor model to regress the intraday returns for a particular stock on the market returns and stock-specific liquidity as described in equation 1. We use this model to extract the return residuals after controlling for systematic components arising from fundamentals and liquidity –

$$r_{i,t,\tau} = \alpha_i + \underbrace{\beta_i^{mkt} r_{t,\tau}^m}_{\substack{\text{fundamentals-driven} \\ \text{systematic component}}} + \underbrace{\beta_i^{liq} L_{i,t,\tau}}_{\substack{\text{liquidity-driven} \\ \text{systematic component}}} + \underbrace{\epsilon_{i,t,\tau}}_{\text{return residual}} . \tag{1}$$

The intraday return for the i-th stock at day t and intraday period τ is denoted as $r_{i,t,\tau}$ are is calculated as log returns over one-minute periods. The market return on date t and intraday period τ is denoted by $r_{t,\tau}^m$ and is calculated as log returns over one-minute periods of the NIFTY50 index – the market index for NSE. Liquidity $L_{i,t,\tau}$ is estimated through the inverse of the illquidity measure proposed by Amihud (2002). In this model, the β^{mkt} coefficient is analogous to the market model beta. The term $\epsilon_{i,t,\tau}$ denotes the return residuals. We expect this residual to be related to the positive or negative information arrival in the said period. We run individual time-series regression for each stock to extract the residuals. The summary of the results of the time-series regression is tabulated in table A.2.[10] We also estimated the model using a panel setup using fully specified fixed effects and we discuss the corresponding results in the robustness section.

While we are not interested in the estimated coefficients as only the residuals from this estimation after purging out systematic factors would contain information, it is useful to discuss them for a sanity check. The coefficient for the market returns is positive and significant for all the regressions. The coefficient for the liquidity parameter is positive –

[10] As the returns are computed at one-minute intervals, the magnitude of the return variables (spot returns and index returns) are very small. As the illiquidity measure takes the return component in the denominator, the computed liquidity estimate blows up making the interpretation of the measure cumbersome. For the ease of interpreting the magnitude of the coefficients, we scale down the liquidity variable by a factor of 10^12. This scaling only improves the readability of the corresponding co-efficient and does not impact the statistical significance of the results. In the following, we work only with the residuals after normalization. Hence, this scaling is automatically cancelled out.

the positiveness is explained by the fact that we employ the measure of illiquidity which is inverse of liquidity. However, they are not statistically significant in most of the cases. The stock-level residuals from this regression, form the idiosyncratic component of the returns free of the market return and stock-specific liquidity effects. We use these residuals for our following analysis as they are free of the aggregate index return and effects due to liquidity.

4.2 HFT-specific Information and Stock Returns

After constructing the idiosyncratic component of the returns, we explore how HFT-specific component of the information impacts the idiosyncratic component of the returns. Chordia *et al.* (2002) and Chordia & Subrahmanyam (2004) show that order-imbalance is positively associated with stock returns. If HFTs are informed, the HFT-specific component of order imbalance should also be positively associated to idiosyncratic returns in the said period. In order to test the statement, we set up our first hypothesis as follows.

Hypothesis 1: *HFT specific information is associated with market returns.*

As both the residuals and information parameters may differ significantly in size across individual stocks, we first normalize these parameters. We denote the normalized residual vector obtained from equation 1 as $\hat{\epsilon}$:

$$\hat{\epsilon} = \frac{\epsilon - \bar{\epsilon}}{SD(\epsilon)} \tag{2}$$

where we represent sample average and standard deviation of a variable X by $\bar{X}$ and $SD(X)$ respectively. Similarly, we also normalize the HFT information vector O^{HFT} as $\hat{O}^{HFT}$:

$$\hat{O}^{HFT} = \frac{O^{HFT} - \bar{O}^{HFT}}{SD(O^{HFT})}. \tag{3}$$

The normalized residual is used as the explanatory variable in the next stage. We regress the residual return on normalized information acted upon by high frequency traders. If our null hypothesis is to be accepted, the beta coefficient β_i^{info} should be statistically significant.

Model 1:

$$\hat{\epsilon}_{i,t,\tau} = \alpha_i + \beta_i^{info}\hat{O}_{i,t,\tau}^{HFT} + \xi_{i,t,\tau} \tag{4}$$

In terms of Model 1 (equation 4), hypothesis 1 can be written as $H_{0,i}^1 : \beta_i^{info} = 0$ vs $H_{1,i}^1 : \beta_i^{info} \neq 0$ for all i.

Table 2: Summary of Model 1

	Mean	% positive	%+ve & significant	%-ve & significant
Intercept	0.0002	66.27%	0.00%	0.00%
β^{info}	0.1547	100.00%	100.00%	0.00%
Average R^2	0.0281			

Note: The table presents the summary of the individual time series regressions as in Model 1 (equation 4). The model uses normalized idiosyncratic returns as the dependent variable and normalized HFT order imbalance is used as the independent variable. The dataset consists of 166 sample stocks traded in the NSE during the period 01-Jan-2015 till 31-Dec-2015. The statistical significance reported are at 1% level.

The HFT-information variable $HFT_{i,t,\tau}$ is created after aggregating the order imbalances of different HFT traders within in a fixed period of time - one-minute intervals in our case. This aggregation captures only the effect of information and not inventory control by HFT market-making strategies.[11]

We run individual time-series regressions for every stock on a time series of intraday observations. Both the idiosyncratic returns (normalized) and the HFT order imbalance (normalized) are estimated on a one-minute interval. The results of the individual time-series regressions are presented in table 2. The stock-wise regressions show that all the β_i^{info}s turn out to be significant (p val < 0.01). Average β_i^{info} across the stocks is 0.15. The results clearly show that HFT information is positively associated with the contemporaneous stock returns supporting hypothesis 1.

Figure A.1 reports the estimated β^{info} and the corresponding confidence intervals. For ease of visualization, we have rank-ordered the stocks on the x-axis in ascending order of the estimated coefficient values. The narrow confidence band around the estimated beta values indicate that the β^{info} values are statistically significant for most of the individual companies.

4.3 Does Direction of Information Matter or the Magnitude?

HFTs' trading strategies are built around their advantage of speed of reaction to information. This happens as they extract information from limit orderbook (LOB) rather than searching for private information. Informed high frequency traders may rely more on the direction of the information rather than the actual magnitude of information. This idea forms our next

[11] One can think of liquidity dislocation affecting returns and hence, it may introduce bias in our estimation. We note two points here. One, our dependent variable in equation 1 is constructed after controlling for systematic factors and hence accounts for changes in the return through liquidity provisioning to begin with. Two, as long as the β_i^{info} coefficient is statistically significant, we know that the HFT- information variable has significant impact on return via the residual. We confirm that this is indeed empirically the case and this observation allows us to proceed with the rest of the analysis.

hypothesis which we split into two interrelated parts.

Due to the mode of constructing the measure, order imbalance can be either positive or negative. In presence of positive information, we expect OIB to be positive and vice-versa. Due to the inherent nature of trade executions by HFTs, it is likely the HFTs find it easier to interpret information in a binary form - positive or negative (+1/-1) rather than judging the degree of positiveness or negativeness of the information. We set up the next hypothesis arguing that the sign of the information in terms of positive or negative matter, but the magnitude of the information does not matter.

Hypothesis 2a: *Direction of information matters.*

Hypothesis 2b: *Magnitude of information doesn't matter.*

First, we define $\text{sign}(\hat{O}^{HFT}_{i,t,\tau})$ as

$$\text{sign}(\hat{O}^{HFT}_{i,t,\tau}) = \begin{cases} +1 & \text{if } \hat{O}^{HFT}_{i,t,\tau} > 0 \\ 0 & \text{if } \hat{O}^{HFT}_{i,t,\tau} = 0 \\ -1 & \text{if } \hat{O}^{HFT}_{i,t,\tau} < 0 \end{cases}.$$

We define $\text{abs}(\hat{O}^{HFT}_{i,t,\tau})$ as

$$\text{abs}(\hat{O}^{HFT}_{i,t,\tau}) = \begin{cases} \hat{O}^{HFT}_{i,t,\tau} & \text{if } \hat{O}^{HFT}_{i,t,\tau} \geq 0 \\ -\hat{O}^{HFT}_{i,t,\tau} & \text{if } \hat{O}^{HFT}_{i,t,\tau} < 0 \end{cases}.$$

We use the following empirical model 2 to test Hypothesis 2a and 2b.

Model 2:

$$\hat{\epsilon}_{i,t,\tau} = \alpha_i + \beta_i^{sign}\text{sign}(\hat{O}^{HFT}_{i,t,\tau}) + \beta_i^{\text{abs}}\text{abs}(\hat{O}^{HFT}_{i,t,\tau}) + \xi_{i,t,\tau} \tag{5}$$

In terms of the Model 2 (equation 5), Hypothesis 2a implies β_i^{sign} is significant for all i i.e. $H^2_{0,i} : \beta_i^{sign} = 0$ vs $H^2_{1,i} : \beta_i^{sign} \neq 0$ for all i. Hypothesis 2b implies β_i^{abs} is not significant, i.e. $H^3_{0,i} : \beta_i^{abs} = 0$ vs $H^3_{1,i} : \beta_i^{abs} \neq 0$ for all i. We also subset the model to estimate the effects individually –

Model 2a:

$$\hat{\epsilon}_{i,t,\tau} = \alpha_i + \beta_i^{sign}\text{sign}(\hat{O}^{HFT}_{i,t,\tau}) + \xi_{i,t,\tau} \tag{6}$$

and

Model 2b:

$$\hat{\epsilon}_{i,t,\tau} = \alpha_i + \beta_i^{\text{abs}}\text{abs}(\hat{O}^{HFT}_{i,t,\tau}) + \xi_{i,t,\tau} \tag{7}$$

Table 3: Summary of Model 2

Panel A: Model 2a

	Mean	**% positive**	**%+ve & significant**	**%-ve & significant**
Intercept	-0.0191	25.90%	24.10%	67.47%
β^{sign}	0.1399	100.00%	98.80%	0.00%
Average R^2	0.0198			

Panel B: Model 2b

	Mean	**% positive**	**%+ve & significant**	**%-ve & significant**
Intercept	-0.0009	46.99%	0.00%	1.20%
β^{abs}	0.0014	53.01%	6.63%	0.00%
Average R^2	1.75×10^{-5}			

Panel C: Model 2 joint estimation

	Mean	**% positive**	**%+ve & significant**	**%-ve & significant**
Intercept	-0.0323	25.90%	24.70%	71.69%
β^{sign}	0.1393	100.00%	97.59%	0.00%
β^{abs}	0.0185	74.10%	69.88%	24.10%
Average R^2	0.0176			

Note: The table presents the summary of the individual time series regressions as specified in Model 2 (equation 5). The models use normalized idiosyncratic returns as the dependent variable while abs(O^{HFT}) and sign(O^{HFT}) are used as the independent variables respectively. The dataset consists of 166 sample stocks traded in the NSE during the period 01-Jan-2015 till 31-Dec-2015. The statistical significance reported are at 1% level.

The results for the individual time series regressions is presented in table 3. Let us first focus on the individual regression models (panels A and B). The results indicate that the β^{sign} is positive and significant for 164 out of the 166 individual firms wheres the β^{abs} is positive and significant for only 11 firms. The average estimate for β^{sign} is much higher compared to β^{abs}. The average model explanatory power for Model 2a is also much higher compared to Model 2b. Thus, both hypotheses 2a and 2b are empirically supported.

Figure A.2 plots the estimated β^{sign} and β^{abs} from Models 2a and 2b with corresponding confidence intervals. For Model 2a, the band of confidence interval is much narrow whereas for Model 2b, the band is much wider.

The joint estimation in Panel C retains very similar feature where β^{sign} is significant almost for all stocks whereas β^{abs} has much lesser frequency of significant relationships and it is also directionally inconsistent (both positive and negative effects are seen across stocks although negatives are fewer in number – 24.1% < 69.8%). The directional inconsistency appears due to the fact that large downswings are associated with price drops (downward movement implies negative excess demand or more simply, excess supply) but that is picked up by the variable as a positive upswing due to conversion to absolute values. To summarize, direction of order imbalance is much more consistent and has a significant relationship in explaining return residuals than the magnitude of order imbalance.

4.4 Scope of Information: Stock-specific and Market-wide

Information may have two different scopes of effect – stock-specific and market-wide. Such information may originate from public or private sources. Both types of information have well-documented effects on the market. Macroeconomic news is known to induce real time effects on price discovery (Andersen *et al.*, 2003). Order book imbalance is another well known source of information (Cao *et al.*, 2009). While all traders use such information, HFTs most likely exploit such information with a lead of a few seconds than the rest of the traders (Brogaard *et al.*, 2014).

We split the aggregate information into idiosyncratic and market-wide components. As noted above, HFTs are known to react faster to any market-level information (e.g., macroeconomic announcements). In the absence of any specific instrument to use in the spot market for these market-level information, these traders are likely to interpret the implication of such information for individual stocks and take positions accordingly. The aggregate information acted upon by HFTs for individual stocks may arise from either stock specific information or market-wide information. In order to decompose this aggregate information $O^{HFT}_{i,t,\tau}$ into market-wide and stock specific components, we use a unobserved factor model with a single latent factor. Ex-ante we hypothesise that both these type of information for HFTs will be significantly related to the idiosyncratic component of the return as HFTs exploit both types of information. Our next hypothesis is set up accordingly.

Hypothesis 3: *Both market-wide information and idiosyncratic information affect idiosyncratic component of returns.*

Let us first set up an unobserved factor model to carry out the decomposition of OIB into two components - a common factor and the residual. Consider a single factor model for a p-variable data set Z with a latent factor F and factor loading $\mathbf{L}$ in the following form:

$$Z^{p\times 1} = \mu^{p\times 1} + \mathbf{L}^{p\times 1}F^{1\times 1} + \psi^{p\times 1}. \tag{8}$$

Our intent is to replace Z by the OIB and carry out the above decomposition. In our implementation, we will consider p number of OIB series. Since we have already normalized OIB as part of pre-processing as per equation 3, we set the constant vector equal to zero i.e., $\mu = 0$. Therefore, the resulting model is

$$Z^{1\times p} = \mathbf{L}^{1\times p}F^{1\times 1} + \psi^{1\times p} \tag{9}$$

after taking a transpose. This operation of taking a transpose helps explaining the estimation of the model with stacked variables as explained below.

The model for OIB described by equation 9, holds at every point of time- and day-index. Therefore, to carry out estimation of the factor jointly across OIB indexed with respsct to both time and day, we stack them up and write it in the matrix form in the following way. Let n_0 denote the number of minutes in a trading day. Consider $Z^{n\times p}$ to denote the data matrix where the $\{i, j\}$-th element represents OIB information of jth stock on day t and at time τ such that $(t-1)\times n_0 + \tau = i$, n denotes the total number of observations. We use the following one factor model for Z to extract market-wide information and idiosyncratic information,

$$Z^{n\times p} = F^{n\times 1}\mathbf{L}^{1\times p} + \Psi^{n\times p}. \tag{10}$$

We denote the estimate of F as $\hat{F}$ and the vectorized (row-wise) estimate of Ψ as $\hat{\psi}$. Specifically,

$$\hat{\psi}_{i,t,\tau} = \big((t-1)\times\tau, i\big)^{th} \text{ element of } \hat{\Psi}. \tag{11}$$

$\hat{F}$ gives us the estimated common factor capturing market-wide information and $\hat{\Psi}$ gives us the idiosyncratic stock-specific information.[12]

[12]In terms of actual implementation, the number of stocks $p = 166$, the number of minute within the

Table 4: Summary of Model 3

	Mean	**% positive**	**%+ve & significant**	**%-ve & significant**
Intercept	0.0024	73.49%	4.22%	0.00%
β^{mw}	0.1630	99.40%	97.59%	0.00%
β^{ss}	0.6919	100.00%	100.00%	0.00%
Average R^2	0.0579			

Notes: The table presents the summary of the individual time series regressions as specified in Model 3 (equation 12). The model uses normalized idiosyncratic returns as the dependent variable while shock-specific and market-wide components of HFT-specific information are used as the independent variables. The dataset consists of 166 sample stocks traded in the NSE during the period 01-Jan-2015 till 31-Dec-2015. The statistical significance reported are at 1% level.

Our factor analysis to decompose information has a parallel with the principal component decomposition implemented in Boehmer *et al.* (2018) although our objective is very different. They used principal components decomposition of the total number of messages that HFT firms send to the market to initiate changes in their positions, with the purpose of identifying different underlying HFT trading strategies. Our objective is to capture information directly via OIB rather than the number of messages which captures at best, a possible indication of trade rather than capturing the trade itself.

To study the impact of market-wide information and stock-specific information, we split these two components of Z and gauge their relative effects by plugging their estimated values into the same regression model. Specifically, we estimate the following model:

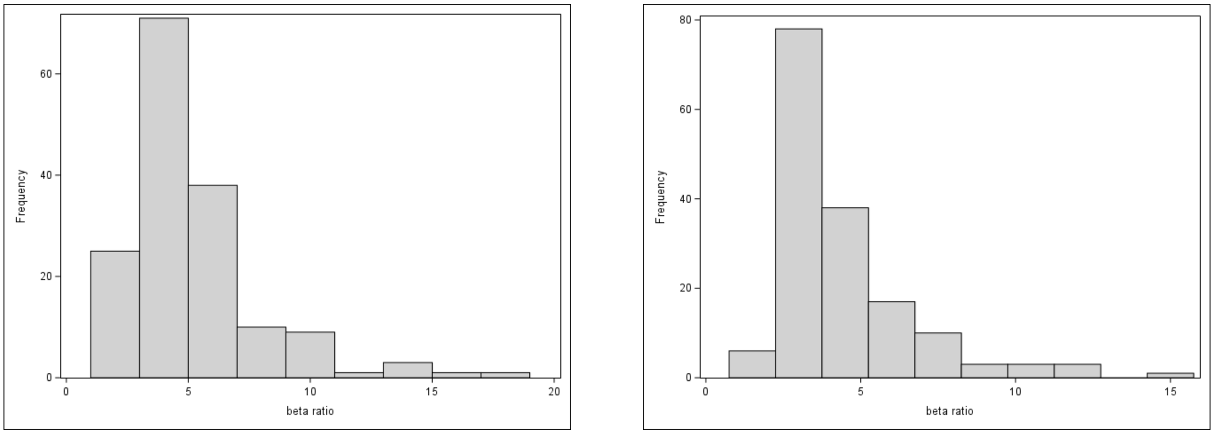
Model 3:

$$\hat{\epsilon}_{i,t,\tau} = \alpha_i + \beta_i^{ss}\hat{\psi}_{i,t,\tau} + \beta_i^{mw}\hat{F}_{t,\tau} + \xi_{i,t,\tau}. \tag{12}$$

In terms of equation 12, hypothesis 3 implies that both β_i^{ss} and β_i^{mw} are statistically significant.

In table 4, we present the resulting coefficient estimates. Both β_i^{ss} and β_i^{mw} are positive and significant for almost all stocks. These findings are clearly in favour of hypothesis 3. For analysing the relative size of the two coefficients for individual stocks, we plot a histogram of the ratio of these beta coefficients – $\beta_i^{ss}/\beta_i^{mw}$ in panel (a) of figure 1. As can be observed, the ratio is larger than 1 for most of the individual stocks, indicating that β^{ss} is larger than β^{mw}.

active trading time in a typical day $n_0 = 375$, and the number of trading days is 246. Thus, $n = 166 \times 246 \times 375$.



(a) Histogram of $\beta_i^{ss}/\beta_i^{mw}$ (b) Histogram of $\bar{\beta}_i^{ss}/\bar{\beta}_i^{mw}$

Figure 1: The two figures plot the histogram of the beta ratios for the overall data (model 3) and for the subset of data during periods of extreme idiosyncratic returns. Figure (a) plots the ratio during for the overall data while figure (b) plots the same for the subset of data during extreme returns. For both of the scenarios, we observe that the probability mass is spread on the right side of 1, indicating β^{ss} is more than β_i^{mw}. We also observe that on a relative basis, the ratios are closer to 1 during periods of extreme movements.

4.5 HFT Responses during Extreme Price Movements

HFTs are known to stabilize extreme price movements (Brogaard *et al.*, 2018). Notably, Grossman & Miller (1988) posited that traders behaving as endogenous liquidity providers choose to supply liquidity during times of order imbalance. HFTs on an average, behave like this which in turn, contributes to stabilization of prices. Thus the role of HFTs during extreme price movements is understood in the literature quite well.

What is not so well understood is exactly which kind of information do they use during times of extreme price movements. Building on the above results, we check the informativeness of HFTs during such periods of extreme movements. For each stock we run a model similar to model 3 (equation 12), but only during periods where the returns were either in the top or bottom ten percentile of the returns[13] individually for each stock. HFTs are machine enabled traders. Possibly, they would be less exposed to behavioral biases compared to human traders. Ex-ante we expect the results corresponding to hypothesis 3 to be same as hypothesis 4 where both the stock-specific component as well the market-wide component of the HFTs' information significantly affects the idiosyncratic component of the return.

Hypothesis 4: *Both market-wide information and idiosyncratic information affects extreme idiosyncratic return.*

[13]The choice of tenth percentile to describe extreme returns is not unique. We run robustness tests with different cutoffs. All results hold.

Table 5: Summary of Model 4

	Mean	**% positive**	**%+ve & significant**	**%-ve & significant**
Intercept	0.0201	83.73%	15.66%	0.00%
$\bar{\beta}^{mw}$	0.5720	96.99%	96.39%	2.41%
$\bar{\beta}^{ss}$	2.2469	100.00%	100.00%	0.00%
Average R^2	0.0614			

Note: The table presents the summary of the individual time series regressions as specified in Model 4 (equation 14). The model uses normalized idiosyncratic returns as the dependent variable while shock-specific and market-wide components of HFT-specific information are used as the independent variables, but only for the subset of periods of extreme returns. The dataset consists of 166 sample stocks traded in the NSE during the period 01-Jan-2015 till 31-Dec-2015. The statistical significance reported are at 1% level.

Suppose $F(.)$ denotes the empirical distribution function of $\hat{\epsilon}$. For $0 < p < 1$, define $Q(\hat{\epsilon}, p) = F^{-1}(p)$, where $F^{-1}(p) = \inf\{u : F(u) \leq p\}$. We also define

$$\bar{\epsilon} = \{\hat{\epsilon}_{i,t,\tau} : \text{ either } \hat{\epsilon}_{i,t,\tau} > Q(\hat{\epsilon}, 0.9) \text{ or } \hat{\epsilon}_{i,t,\tau} < Q(\hat{\epsilon}, 0.1)\}. \tag{13}$$

$\bar{\epsilon}$ represents the extreme parts of the residual vector. The corresponding market-wide information and idiosyncratic information are denoted by $\bar{F}$ and $\bar{\gamma}$.

To study the same hypothesis for extreme idiosyncratic return we propose the following model -

Model 4:

$$\hat{\epsilon}_{i,t,\tau} = \alpha_i + \bar{\beta}_i^{ss}\hat{\psi}_{i,t,\tau} + \bar{\beta}_i^{mw}\hat{F}_{t,\tau} + \xi_{i,t,\tau}. \tag{14}$$

We want to test whether both market-wide information and idiosyncratic information affect extreme idiosyncratic return i.e. the alternative hypothesis suggests $\bar{\beta}_i^{ss} \neq 0$ & $\bar{\beta}_i^{mw} \neq 0$ for all i.

Our analysis suggests that $\bar{\beta}_i^{ss}$ is positive and significant for 56% of stocks and $\bar{\beta}_i^{mw}$ are significant for 71% stocks. Average $\bar{\beta}_i^{ss}$ turns out to be 0.035 where average $\bar{\beta}_i^{mw}$ is 0.069. Panel (b) of figure 1 plots the histogram of beta ratio – $\beta_i^{ss}/\beta_i^{mw}$ during periods of extreme price movements. As it can be observed, similar to panel (a), most of the coefficients are larger than one, indicating that the β^{mw} is larger than β^{ss} during periods of extreme movements also. However, the probability mass is shifted towards the left for panel (b), indicating that the relative importance of stock-specific component over the market-wide component decreases during these periods of extreme movements.

Table 6: Split sample analysis – summary of Model 3 separating periods of positive and negative returns.

	Mean	**% positive**	**%+ve & significant**	**%-ve & significant**
Panel A : Periods of positive idiosyncratic returns				
Intercept	0.6873	100.00%	100.00%	0.00%
β^{mw}	0.0497	89.16%	80.72%	7.83%
β^{ss}	0.2562	100.00%	99.40%	0.00%
Average R^2	0.0080			
Panel B : Periods of negative idiosyncratic returns				
Intercept	-0.6786	0.00%	0.00%	100.00%
β^{mw}	0.0568	89.76%	84.94%	7.23%
β^{ss}	0.2060	99.40%	96.99%	0.00%
Average R^2	0.0082			

Note: The table presents the summary of the individual time series regressions as specified in Model 3 (equation 12) split into periods of positive and negative idiosyncratic returns. The model uses normalized idiosyncratic returns as the dependent variable while stock-specific and market-wide components of HFT-specific information are used as the independent variables. The dataset consists of 166 sample stocks traded in the NSE during the period 01-Jan-2015 till 31-Dec-2015. The statistical significance reported are at 1% level.

4.6 Do HFTs exhibit Asymmetric Responses to Return Fluctuations?

Accounting for asymmetric response to positive and negative expected returns, Glosten *et al.* (1993) show that positive (negative) unanticipated returns result in downward (upward) revision of the conditional volatility. Given that HFTs respond to information that drive prices up or down, it is a natural extension of the above analysis to check how they respond to positive as opposed to negative information. Given that generally HFTs act very fast on information and as we have seen above they respond to the direction rather than magnitude of information, it is a prioi unlikely that they would respond asymmetrically to return fluctuations. We formally test this hypothesis by estimating model 3 (equation 12) separately on positive and negative idiosyncratic returns.

Table 6 presents the results of the individual time series regressions split into periods of positive and negative idiosyncratic returns. For each stock, we split the sample into positive and negative idiosyncratic returns. As the idiosyncratic return measure is normalized, this splits the samples into two exact halved for each stock. The results corresponding to periods of positive idiosyncratic returns (normalized) is presented in Panel A, whereas the same corresponding to negative idiosyncratic returns (normalized) is presented in Panel B. We observe that the coefficients corresponding to the market-wide and stock-specific component of HFT information are consistent across the two scenarios, indicating that HFTs are not exposed to the mentioned behavioral biases.

5 Robustness

5.1 Panel Estimation with Fully Specified Fixed Effects

In the preceding analysis, we use time series regression coupled with a factor model approach. One may have a concern that the empirical results do not account for unobserved heterogeneity at the levels of stocks being traded, days of trading and intra-day time of trading. Here we take the previous models to a panel data format and conduct joint estimation with additional fixed effects to account for all such possible unobserved heterogeneity.

In the earlier sections, we had estimated the basic model coeffcients through individual time-series models, which can be be expressed in the following generalized form –

$$y_{i,t,\tau} = \alpha_i + \beta_i x_{i,t,\tau} + \xi_{i,t,\tau}. \tag{15}$$

Here $y_{i,t,\tau}$ represents the dependent variable, which is our case is the the idiosyncratic individual time-series of intraday stock returns for the i-th stock in the t-th day for the τ-th period. $x_{i,t,\tau}$ represents the model-specific independent variables. For Model 1, the independent variable was HFT-specific information. For the two models 2a and 2b, the independent variables were the sign and absolute value of HFT-specific information respectively. For models 3 and 4, the independent variables were both market-wide and stock-specific component of information.

Next, we test for the robustness of the existing models (1-4) in a panel setup, which can be expressed in the following generalized form –

$$y_{i,t,\tau} = \beta x_{i,t,\tau} + \delta_i + \gamma_t + \omega_\tau + \delta_i \times \gamma_t + \xi_{i,t,\tau}. \tag{16}$$

The dependant variable $y_{i,t,\tau}$ and the independent variable $x_{i,t,\tau}$ carry the same meanings as described in the earlier paragraph. δ_i, γ_t, and ω_τ are used to control firm, day and intraday fixed-effects respectively. The detailed results of the panel regressions models are provided in the appendix.

Table A.3 provides the results of robustness test for model 1. Consistent with our findings reported in table 2, where the coefficient for β^{info} term was positive and significant for all the individual stocks, we find that in the panel form (table A.3), the same coefficient is positive and significant across all the different panel specifications. The values of the β^{info} coefficient in the panel specification (table A.3) is also similar to the average value of the β^{info} coefficient (0.1547) in table 2.

Similarly, for models 2a and 2b, we also find that the results of the individual time-series regressions reported in table 3 are consistent with the panel regression results reported in table A.4. The coefficient corresponding to β^{sign} was positive and significant for 164 of the 166 sample stocks 3. In contrast, the coefficient corresponding to β^{abs} was positive and significant for only 6.63% of the stocks. In the panel specifications (table A.4), we find that though both of these coefficients are positive and significant across specifications, the magnitude of the β^{sign} component is approximately 100 times the magnitude of the β^{abs} component.

In model 3, we decompose the HFT-specific information is decomposed into stock-specific and market-wide components. We find that the individual time series results (table 4) are consistent with panel regression results (table A.5). The mean parameter estimates for the coefficient of the stock-specific component is almost four times that of the market-wide component for either specification.

In model 4, we run model 3 for the subset of periods belonging to either positive or negative extreme returns. For individual time-series regressions, we find that the stock-specific component is positive and significant for approximately half of the stocks while same is true for close to three-fourth of the stocks for the market-wide component (table 5). In the panel specifications, both the components are statistically significant (table A.6). Though the coefficient values differ across the two specifications, the coefficient for the stock-specific component is almost twice that of the market-wide component, in either specification.

5.2 Alternate Definition of Extreme Events

So far we have constructed extreme events on the basis of normalized idiosyncratic returns rather than raw individual stock returns. To show robustness of our result, we estimate the same models only for extreme periods signified by top and bottom ten percentile of individual stock returns. All results continue to hold (details can be found in the online appendix). Both the individual time-series (table OA.1) as well as the panel regression models (table OA.2) during the extreme return periods suggest that the both the coefficients of the stock specific component and the market-wide component are statistically significant across specifications and the coefficients of the stock-specific component is twice in magnitude compared to the market-wide component. We further segregate the periods into that for only either the positive or negative extreme returns. Both the individual time-series (table OA.3) and panel (table OA.4) are consistent with the earlier results with a difference that while the coefficient for the stock-specific component is positive, the coefficient for the market-wide component is negative. This can be explained by the observation that the extreme events are constructed conditional on the individual raw returns, i.e. without normalization. Therefore, it is likely that the extreme fluctuations in them would be correlated with the stock-specific component. The market-wide component moderates the effect possibly due to inducing common direction of fluctuations in returns across stocks.

6 Summary and Discussion

In this paper, we analyze the role of HFTs in terms of pricing information. HFTs are known to help in price discovery (Brogaard *et al.*, 2014). What is not understood well is how do they exploit information affecting the market as a whole or information specific to stocks, for that purpose? An answer to this question will shed light on how the joint movement of prices change with respect to shocks (Lan *et al.*, 2023; Raunig, 2023).

Using a dataset obtained from the National Stock Exchange (NSE) of India with identifiers for algorithmic traders, we show that HFT-specific information is positively associated with idiosyncratic stock returns. Empirically, the direction of information matters – sign of the information is correlated with the idiosyncratic component of returns much more than the magnitude of the information. Using a factor model, we decompose the HFT-specific information into stock-specific and market-wide components. We find that both these components are positively associated with idiosyncratic component of returns, though the importance of stock-specific component is more than that of the market-wide component. We also find that the relative importance of the stock-specific component compared to market-wide component is lower during periods of extreme price movements. The findings are consistent across individual time-series as well panel specifications after accounting for stock-, trading day- and trading time-level unobserved heterogeneity.

Our work sheds light on how traders engage with the market. This has two sides to it. One, traders get influenced by market states. Two, traders themselves may influence the market. While we do not attempt the disentangle this endogeneity in the relationship between the traders and the states of the market, our approach is to focus on the statistical nature of how traders trade on the basis of the aggregate market-wide information and stock-specific information. Our work should be seen as depicting the equilibrium relationship between trading behavior and the market behavior.

We end the paper with some caveats and thoughts on future work. One limitation of our dataset is that we cannot differentiate between stock- or more generally, company-specific information available publicly and company-specific information obtained privately. Therefore, our factor model decomposition allows us to keep track of market-wide information and stock-specific information where the nomenclature emphasizes the scope of the effect of the information rather than whether the origin of the information in public or private. However, observing the source of information is an extremely difficult task. Our approach is akin to inferring information based only on *equilibrium* prices which has an obvious limitation that we cannot differentiate between the supply side and the demand side. The second limitation is that the order imbalance appearing from the HFTs is aggregated over all HFTs. Thus, we lose granular trader-level heterogeneity. Again, data limitations stop us from exploiting trader-level information. We leave it to future work with more granular data to explore the direct relationship between sources of information and how it gets priced via individual traders.

Competing Interests: The authors have no financial or proprietary interests in any material discussed in this article.

References

Amihud, Yakov. 2002. Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, **5**(1), 31–56.

Amihud, Yakov, & Mendelson, Haim. 1986. Liquidity and stock returns. *Financial Analysts Journal*, **42**(3), 43–48.

Anand, Amber, & Venkataraman, Kumar. 2016. Market conditions, fragility, and the economics of market making. *Journal of Financial Economics*, **121**(2), 327–349.

Andersen, Torben G, Bollerslev, Tim, Diebold, Francis X, & Vega, Clara. 2003. Micro effects of macro announcements: Real-time price discovery in foreign exchange. *American Economic Review*, **93**(1), 38–62.

Bernales, Alejandro, Valenzuela, Marcela, & Zer, Ilknur. 2022. Effects of Information Overload on Financial Markets: How Much Is Too Much? *Available at SSRN 3904916*.

Berry, Thomas D, & Howe, Keith M. 1994. Public information arrival. *The Journal of Finance*, **49**(4), 1331–1346.

Boehmer, Ekkehart, Li, Dan, & Saar, Gideon. 2018. The competitive landscape of high-frequency trading firms. *Review of Financial Studies*, **31**(6), 2227–2276.

Bollen, Nicolas PB, & Whaley, Robert E. 2004. Does net buying pressure affect the shape of implied volatility functions? *The Journal of Finance*, **59**(2), 711–753.

Boudt, Kris, & Petitjean, Mikael. 2014. Intraday liquidity dynamics and news releases around price jumps: Evidence from the DJIA stocks. *Journal of Financial Markets*, **17**, 121–149.

Brogaard, Jonathan, Hendershott, Terrence, & Riordan, Ryan. 2014. High-Frequency Trading and Price Discovery. *Review of Financial Studies*, **27**(8), 2267–2306.

Brogaard, Jonathan, Carrion, Allen, Moyaert, Thibaut, Riordan, Ryan, Shkilko, Andriy, & Sokolov, Konstantin. 2018. High frequency trading and extreme price movements. *Journal of Financial Economics*, **128**(2), 253–265.

Brogaard, Jonathan, Nguyen, Thanh Huong, Putnins, Talis J, & Wu, Eliza. 2022. What Moves Stock Prices? The Roles of News, Noise, and Information. *The Review of Financial Studies*, **35**(9), 4341–4386.

Brunnermeier, Markus K, & Pedersen, Lasse Heje. 2005. Predatory trading. *The Journal of Finance*, **60**(4), 1825–1863.

Cao, Charles, Hansch, Oliver, & Wang, Xiaoxin. 2009. The information content of an open limit-order book. *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, **29**(1), 16–41.

Chaboud, Alain P., Chiquoine, Benjamin, Hjalmarsson, Erik, & Vega, Clara. 2014. Rise of the machines: Algorithmic trading in the foreign exchange market. *The Journal of Finance*, **69**(5), 2045–2084.

Chakrabarty, Bidisha, Moulton, Pamela C., & Wang, Xu. 2020. *Limited Attention and Market Pricing of Earnings in a High Frequency World.*

Chan, Kalok, & Fong, Wai-Ming. 2000. Trade size, order imbalance, and the volatility–volume relation. *Journal of Financial Economics*, **57**(2), 247–273.

Chordia, Tarun, & Subrahmanyam, Avanidhar. 2004. Order imbalance and individual stock returns: Theory and evidence. *Journal of Financial Economics*, **72**(3), 485–518.

Chordia, Tarun, Roll, Richard, & Subrahmanyam, Avanidhar. 2002. Order imbalance, liquidity, and market returns. *Journal of Financial economics*, **65**(1), 111–130.

Copeland, Thomas E, & Galai, Dan. 1983. Information effects on the bid-ask spread. *the Journal of Finance*, **38**(5), 1457–1469.

Cushing, David, & Madhavan, Ananth. 2000. Stock returns and trading at the close. *Journal of Financial Markets*, **3**(1), 45–67.

Easley, David, & O'hara, Maureen. 1987. Price, trade size, and information in securities markets. *Journal of Financial economics*, **19**(1), 69–90.

Farmer, J. Doyne, Gillemot, Laszlo, Lillo, Fabrizio, Mike, Szabolcs, & Sen, Anindya. 2004. What really causes large price changes? *Quantitative Finance*, **4**(4), 383–397.

Frino, Alex, Prodromou, Tina, Wang, George H.K., Westerholm, P. Joakim, & Zheng, Hui. 2017. An empirical analysis of algorithmic trading around earnings announcements. *Pacific Basin Finance Journal*, **45**, 34–51.

Glosten, Lawrence R, & Milgrom, Paul R. 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of financial economics*, **14**(1), 71–100.

Glosten, Lawrence R, Jagannathan, Ravi, & Runkle, David E. 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, **48**(5), 1779–1801.

Golub, Anton, Keane, John, & Poon, Ser-Huang. 2012. High frequency trading and mini flash crashes. *arXiv preprint arXiv:1211.6667*.

Grossman, Sanford J, & Miller, Merton H. 1988. Liquidity and market structure. *Journal of Finance*, **43**(3), 617–633.

Hagströmer, Björn, & Nordén, Lars. 2013. The diversity of high-frequency traders. *Journal of Financial Markets*, **16**(4), 741–770.

Hendershott, Terrence, & Riordan, Ryan. 2011. *Algorithmic Trading and Information.*

Hendershott, Terrence, Jones, Charles M, & Menkveld, Albert J. 2011. Does Algorithmic Trading Improve Liquidity? *The Journal of Finance*, **66**(1), 1–33.

Hirschey, Nicholas. 2021. Do high-frequency traders anticipate buying and selling pressure? *Management Science*, **67**(6), 3321–3345.

Hsieh, Wen-liang Gideon, & Lee, Chin-Shen. 2021. Who reacts to what information in securities analyst reports? Direct evidence from the investor trade imbalance. *Pacific-Basin finance journal*, **65**, 101492.

Ito, Takatoshi, Lyons, Richard K, & Melvin, Michael T. 1998. Is there private information in the FX market? The Tokyo experiment. *The Journal of Finance*, **53**(3), 1111–1130.

Jiang, George J, Lo, Ingrid, & Verdelhan, Adrien. 2011. Information shocks, liquidity shocks, jumps, and price discovery: Evidence from the US Treasury market. *Journal of Financial and Quantitative Analysis*, **46**(2), 527–551.

Kacperczyk, Marcin, & Pagnotta, Emiliano S. 2019. Chasing private information. *The Review of Financial Studies*, **32**(12), 4997–5047.

Kang, Jangkoo, & Park, Hyoung-Jin. 2008. The information content of net buying pressure: Evidence from the KOSPI 200 index option market. *Journal of Financial Markets*, **11**(1), 36–56.

Kirilenko, Andrei, Kyle, Albert S, Samadi, Mehrdad, & Tuzun, Tugkan. 2017. The flash crash: High-frequency trading in an electronic market. *The Journal of Finance*, **72**(3), 967–998.

Kyle, Albert S. 1985. Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, 1315–1335.

Lan, Tingting, Shao, Liuguo, Zhang, Hua, & Yuan, Caijun. 2023. The impact of pandemic on dynamic volatility spillover network of international stock markets. *Empirical Economics*, 1–30.

Lintner, John. 1969. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets: A reply. *Review of Economics and Statistics*, 222–224.

Malceniece, Laura, Malcenieks, KĂrlis, & PutniÅš, TĂlis J. 2019. High frequency trading and comovement in financial markets. *Journal of Financial Economics*, **134**(2), 381–399.

Menkveld, Albert J. 2013. High frequency trading and the new market makers. *Journal of Financial Markets*, **16**(4), 712–740.

Menkveld, Albert J. 2016. The Economics of High-Frequency Trading. *Annual Review of Financial Economics*, **8**, 1–24.

Mitchell, Mark L, & Mulherin, J Harold. 1994. The impact of public information on the stock market. *The Journal of Finance*, **49**(3), 923–950.

Nagel, Stefan. 2012. Evaporating liquidity. *The Review of Financial Studies*, **25**(7), 2005–2039.

Raunig, Burkhard. 2023. Using causal graphs to test for the direction of instantaneous causality between economic policy uncertainty and stock market volatility. *Empirical Economics*, 1–20.

Scholtus, Martin, Van Dijk, Dick, & Frijns, Bart. 2014. Speed, algorithmic trading, and market quality around macroeconomic news announcements. *Journal of Banking and Finance*, **38**(1), 89–105.

Schroeder, Florian, Lepone, Andrew, Leung, Henry, & Satchell, Stephen. 2020. Flash crash in an OTC market: trading behaviour of agents in times of market stress. *The European Journal of Finance*, **26**(15), 1569–1589.

SEC. 2010. *Concept release on market structure.*

Sharpe, William F. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, **19**(3), 425–442.

Yang, Liyan, & Zhu, Haoxiang. 2020. Back-running: Seeking and hiding fundamental information in order flows. *The Review of Financial Studies*, **33**(4), 1484–1533.

Zhang, S. Sarah. 2012. *Need for Speed - An Empirical Analysis of Hard and Soft Information in a High Frequency World.*

APPENDIX

A Additional Figures

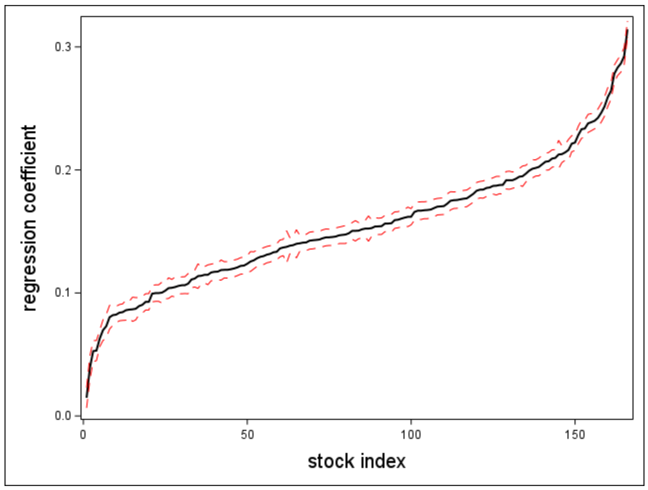
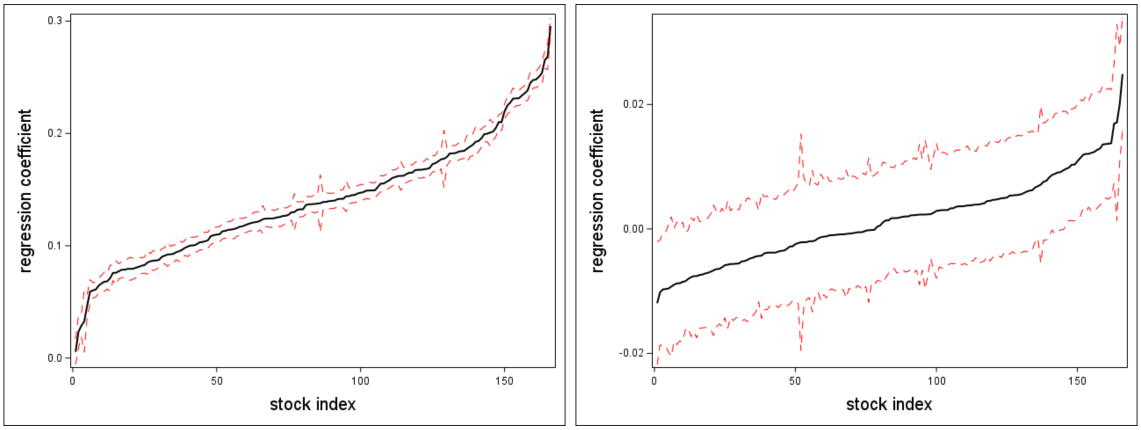


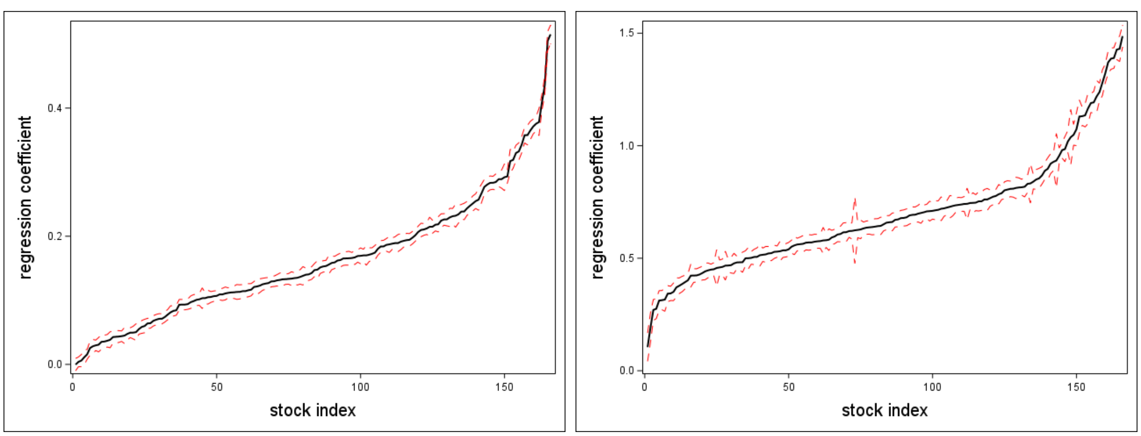
Figure A.1: The figure plots the regression coefficient and the corresponding confidence intervals for β_i^{info} estimated from Model 1. The figures illustrates that the most of the β^{info} are statistically significant. For ease of visualization, we have ordered the stocks indices on the x-axis in (weakly) ascending order of the estimated coefficient – β_i^{info} in this case.



(a) Model 2: Single variable estimation of β_i^{sign} and its CI.

(b) Model 2: Single variable estimation of β_i^{abs} and its CI.

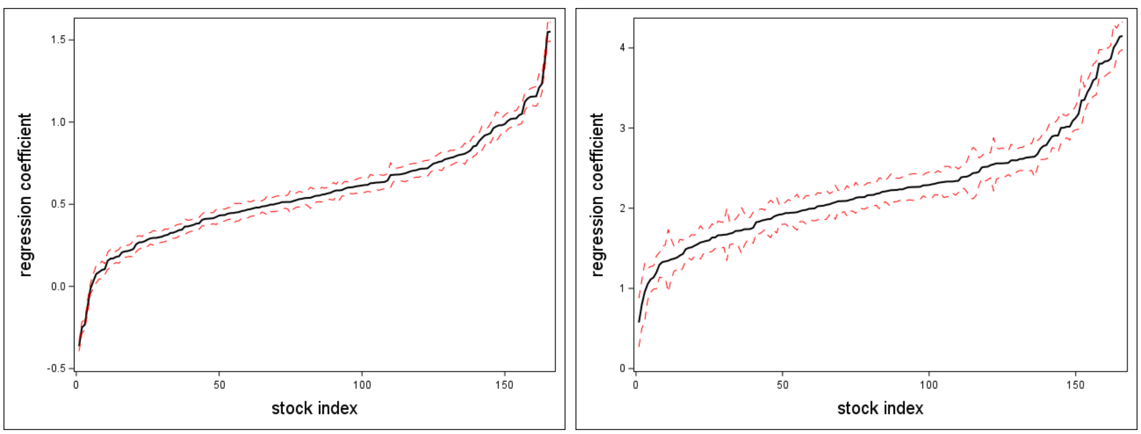
Figure A.2: Plot for estimated regression coefficient for the two hypothesis 2a and 2b. The confidence intervals for each estimate is also shown. For hypothesis 2a, most of the βs are significant. However, for hypothesis 2b, the value 0 is contained in the confidence intervals for most of the βs. For ease of visualization, we have ordered the stocks indices on the x-axis in (weakly) ascending order of the estimated coefficients – β_i^{sign} and β_i^{abs} in panels (a) and (b) respectively.



(a) Model 3: β^{mw} and its CI.

(b) Model 3: β^{ss} and its CI.

Figure A.3: Plot for estimated regression coefficient for the parameters β^{mw} and β^{ss} estimated from model 3. The confidence intervals for each estimate is also shown. For model 2a most of the βs are significant. However, for model 2b, the value 0 is contained in the confidence intervals for most of the βs. For ease of visualization, we have ordered the stocks indices on the x-axis in (weakly) ascending order of the estimated coefficient – β_i^{mw} and β_i^{ss} in panels (a) and (b) respectively.



(a) Model 4: estimated $\bar{\beta}_i^{ss}$ and its CI.

(b) Model 4: estimated $\bar{\beta}_i^{mw}$ and its CI.

Figure A.4: Estimated regression coefficients and confidence intervals for Model 4. For ease of visualization, we have ordered the stocks indices on the x-axis in (weakly) ascending order of the estimated coefficients – $\bar{\beta_i^{mw}}$ and $\bar{\beta_i^{ss}}$ in panels (a) and (b) respectively.

B Additional Tables

Table A.1: Correlation table

Variable	OIBNUM	OIBVAL	HFT OIBNUM	HFT OIBVAL	Index Ret.	Spot Ret.
OIBVAL	0.7696					
HFT OIBNUM	0.4477	0.3516				
HFT OIBVAL	0.3513	0.4422	0.7994			
Index Ret.	0.1677	0.1648	0.1701	0.1663		
Spot Ret.	0.3395	0.3332	0.1952	0.1783	0.2983	
Liquidity	0.0043	0.0031	0.0032	0.0020	-0.0015	0.0044

Note: The table provides a cross-sectional average of individual time-series correlations for order imbalance measures computed at one minute interval.

Table A.2: Summary of the filtering model

	Mean	% positive	%+ve & significant	%-ve & significant
Intercept	-0.0507	31.33%	0.00%	3.61%
β^{mkt}	1.0454	100.00%	100.00%	0.00%
β^{liq}	0.0083	89.16%	13.25%	0.00%
Average R^2	0.1132			

Note: The table presents the summary of the individual time series regressions as illustrated in equation 1. The model uses individual stock returns as the dependent variable while market return and market liquidity are used as the independent variables. The dataset consists of 166 sample stocks traded in the NSE during the period 01-Jan-2015 till 31-Dec-2015. The statistical significance reported are at 1% level.

Table A.3: Panel regression models with HFT order imbalance (Model 1)

	(1)	(2)	(3)	(4)	(5)
(Intercept)	0.0002				
	(0.58)				
β^{info}	0.1555***	0.1555***	0.1557***	0.1558***	0.1639***
	(555.00)	(555.00)	(555.19)	(555.80)	(572.38)
Num. obs.	11,917,377	11,917,377	11,917,377	11,917,377	11,917,377
R^2	0.0252	0.0252	0.0254	0.0259	0.0292
Adj. R^2	0.0252	0.0252	0.0253	0.0258	0.0260
Firm FE		Yes	Yes	Yes	
Day FE			Yes	Yes	
Intraday FE				Yes	
Firm*Day FE					Yes

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Note: The table presents the results of the panel regression models using normalized idiosyncratic returns as the dependent variable and normalized HFT order imbalance as the independent variable. The dataset consists of 166 sample stocks traded in the NSE during the period 01-Jan-2015 till 31-Dec-2015. Column 1 reports the results of a a pooled regression model without any fixed-effect. Column 2 reports results with firm fixed-effects. Column 3 uses both firm and day fixed effects. Column 4 uses firm, day and intraday period fixed-effects. Column 5 uses an interaction of firm and day fixed-effects.

Table A.4: Panel regression models with sign and absolute value of HFT order imbalance (Models 2a and 2b)

PanelA : Model2a					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
(Intercept)	-0.0171***				
	(-59.02)				
β^{sign}	0.1292***	0.1406***	0.1407***	0.1412***	0.1474***
	(445.87)	(465.34)	(465.42)	(466.94)	(478.55)
Num. obs.	11,917,377	11,917,377	11,917,377	11,917,377	11,917,377
R^2	0.0164	0.0178	0.0180	0.0186	0.0213
Adj. R^2	0.0164	0.0178	0.0180	0.0186	0.0181
Firm FE		Yes	Yes	Yes	
Day FE			Yes	Yes	
Intraday FE				Yes	
Firm*Day FE					Yes

PanelB : Model2b					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
(Intercept)	-0.0008**				
	(-2.16)				
β^{abs}	0.0013***	0.0013***	0.0012***	0.0012***	0.0015***
	(3.49)	(3.51)	(3.27)	(3.30)	(3.88)
Num. obs.	11,917,377	11,917,377	11,917,377	11,917,377	11,917,377
R^2	0.0000	0.0000	0.0001	0.0007	0.0024
Adj. R^2	0.0000	-0.0000	0.0001	0.0006	-0.0008
Firm FE		Yes	Yes	Yes	
Day FE			Yes	Yes	
Intraday FE				Yes	
Firm*Day FE					Yes

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Note: The table presents the results of the panel regression models using normalized idiosyncratic returns as the dependent variable and sign(O^{HFT}) as the independent variable. The dataset consists of 166 sample stocks traded in the NSE during the period 01-Jan-2015 till 31-Dec-2015. Column 1 reports the results of a a pooled regression model without any fixed-effect. Column 2 reports results with firm fixed-effects. Column 3 uses both firm and day fixed effects. Column 4 uses firm, day and intraday period fixed-effects. Column 5 uses an interaction of firm and day fixed-effects.

Table A.5: Panel regression models with market-wide and stock specific components of HFT order imbalance (Model 3)

	(1)	(2)	(3)	(4)	(5)
(Intercept)	0.0022***				
	(7.74)				
β^{mw}	0.1404***	0.1405***	0.1406***	0.1415***	0.1457***
	(357.64)	(357.77)	(357.84)	(359.28)	(367.77)
β^{ss}	0.5528***	0.5532***	0.5536***	0.5539***	0.5803***
	(401.04)	(401.19)	(401.20)	(401.50)	(413.06)
Num. obs.	11,917,377	11,917,377	11,917,377	11,917,377	11,917,377
R^2	0.0146	0.0146	0.0148	0.0154	0.0179
Adj. R^2	0.0146	0.0146	0.0147	0.0153	0.0147
Firm FE		Yes	Yes	Yes	
Day FE			Yes	Yes	
Intraday FE				Yes	
Firm*Day FE					Yes

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Note: The table presents the results of the panel regression models using normalized idiosyncratic returns as the dependent variable and market-wide and stock specific components of HFT information as the independent variable. The dataset consists of 166 sample stocks traded in the NSE during the period 01-Jan-2015 till 31-Dec-2015. Column 1 reports the results of a a pooled regression model without any fixed-effect. Column 2 reports results with firm fixed-effects. Column 3 uses both firm and day fixed effects. Column 4 uses firm, day and intraday period fixed-effects. Column 5 uses an interaction of firm and day fixed-effects.

Table A.6: Panel regression models with market-wide and stock specific components of HFT order imbalance during periods of extreme idiosyncratic returns (Model 4)

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
(Intercept)	0.0285***				
	(27.57)				
$\bar{\beta}^{mw}$	0.6395***	0.6397***	0.6403***	0.6451***	0.6527***
	(480.13)	(480.19)	(479.91)	(481.81)	(482.87)
$\bar{\beta}^{ss}$	1.2881***	1.2884***	1.2885***	1.2905***	1.3450***
	(262.62)	(262.59)	(262.52)	(263.04)	(268.56)
Num. obs.	2,851,818	2,851,818	2,851,818	2,851,818	2,851,818
R^2	0.0776	0.0777	0.0789	0.0804	0.0912
Adj. R^2	0.0776	0.0777	0.0788	0.0801	0.0787
Firm FE		Yes	Yes	Yes	
Day FE			Yes	Yes	
Intraday FE				Yes	
Firm*Day FE					Yes

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Note: The table presents the results of the panel regression models using normalized idiosyncratic returns as the dependent variable and market-wide and stock specific components of HFT information as the independent variable during periods of extreme idiosyncratic returns. The dataset consists of 166 sample stocks traded in the NSE during the period 01-Jan-2015 till 31-Dec-2015. Column 1 reports the results of a a pooled regression model without any fixed-effect. Column 2 reports results with firm fixed-effects. Column 3 uses both firm and day fixed effects. Column 4 uses firm, day and intraday period fixed-effects. Column 5 uses an interaction of firm and day fixed-effects.

Table A.7: Panel regression models with market-wide and stock specific components of HFT order imbalance (split into periods of positive and negative idiosyncratic returns)

Panel A : Periods of positive idiosyncratic returns					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**
(Intercept)	0.6805***				
	(2242.54)				
β^{mw}	0.0467***	0.0438***	0.0433***	0.0542***	0.0413***
	(112.36)	(105.45)	(105.22)	(136.67)	(100.55)
β^{ss}	0.2007***	0.1983***	0.1987***	0.2021***	0.2069***
	(137.53)	(136.06)	(137.63)	(145.67)	(141.95)
Num. obs.	5,923,472	5,923,472	5,923,472	5,923,472	5,923,472
R^2	0.0033	0.0086	0.0286	0.1035	0.0602
Adj. R^2	0.0033	0.0086	0.0285	0.1034	0.0540
Firm FE		Yes	Yes	Yes	
Day FE			Yes	Yes	
Intraday FE				Yes	
Firm*Day FE					Yes

	(1)	**(2)**	**(3)**	**(4)**	**(5)**
Panel B : Periods of negative idiosyncratic returns					
(Intercept)	-0.6718***				
	(-2238.85)				
β^{mw}	0.0553***	0.0533***	0.0532***	0.0461***	0.0526***
	(136.09)	(131.16)	(132.16)	(118.40)	(130.79)
β^{ss}	0.1503***	0.1506***	0.1525***	0.1568***	0.1649***
	(104.97)	(105.23)	(107.68)	(114.71)	(115.25)
Num. obs.	5,993,905	5,993,905	5,993,905	5,993,905	5,993,905
R^2	0.0032	0.0081	0.0307	0.0973	0.0613
Adj. R^2	0.0032	0.0081	0.0306	0.0972	0.0552
Firm FE		Yes	Yes	Yes	
Day FE			Yes	Yes	
Intraday FE				Yes	
Firm*Day FE					Yes

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Note: The table presents the results of the panel regression models using normalized idiosyncratic returns as the dependent variable and market-wide and stock specific components of HFT information as the independent variable during periods of positive (Panel A) and negative (Panel B) idiosyncratic returns. The dataset consists of 166 sample stocks traded in the NSE during the period 01-Jan-2015 till 31-Dec-2015. Column 1 reports the results of a a pooled regression model without any fixed-effect. Column 2 reports results with firm fixed-effects. Column 3 uses both firm and day fixed effects. Column 4 uses firm, day and intraday period fixed-effects. Column 5 uses an interaction of firm and day fixed-effects.