Information Flow Between Spot and Futures Market - The Role of Algorithmic Traders

Anirban Banerjee* Ashok Banerjee†

January 2, 2020

Abstract

We inspect the intraday lead-lag relationship between the spot and stock futures market for a sample set of 160 stocks in National Stock Exchange (NSE) of India using six months of data during Jan-Jun 2015. Instead of looking at the relationship between the price movements of the two markets, we look at the order imbalance in these two markets to establish the direction of information flow. We show that information in the futures market leads the spot market by one minute. We split the order imbalance due to the different category of traders. We find that algorithmic and high-frequency traders are not informed, and the information flow is primarily established through non-algorithmic traders. The results are consistent even during periods of extreme price movements.

1 Introduction

We show that high-frequency algorithmic traders are not informed, and non-algorithmic traders generate information. We test for the flow of information between the stock and

*Assistant Professor, Finance, Accounting and Control Group, Indian Institute of Management Kozhikode, Email: anirban@iimk.ac.in

†Professor, Finance and Control Group, Indian Institute of Management Calcutta

stock-futures market for a sample set of 160 stocks listed in National Stock Exchange of India and find that the information in futures market leads the spot market by one minute. Instead of looking into the co-movement of prices in the spot and futures market to determine the lead-lag nature, we attempt to look into the relationship through order flow. Though prior research has used information from order book, we segregate this flow of information from various category of traders, primarily distinguishing between algorithmic and non-algorithmic human trader categories - a novel attempt.

Information flow between the cash and futures market is well established. The much studied lead-lag relationship of price movement between these two markets indicates how quickly new information is reflected in one market relative to the other (Chan, 1992). In an ideal frictionless environment, price movements across markets should be contemporaneously correlated and not cross-correlated. A situation where one market assimilates information faster compared to the other gives rise to a lead-lag relationship between price movements.

Order imbalance may arise due to inventory pressure of the market maker as well as information asymmetry. In an order-driven market without designated market-maker, information is expected to be the primary source of order imbalance. For an individual stock, information, both public or private, can induce order imbalance. If information is acted upon earlier in one market compared to another, it is likely to be reflected first in the order book. We look at the information flow between the spot and futures market by inspecting how individual stock returns are impacted by order imbalance of these two markets.

Empirical research regarding information flow between markets has largely been restricted between the cash (spot) or index and the index futures markets. Problem of using index data is that though aggregate market-level information is captured in the index, it cannot be used to infer any stock-specific information flow. In this study, we overcome this difficulty by using a proprietary dataset obtained from the National Stock Exchange of India (NSE). Unlike many other markets, single stock futures are heavily traded in NSE, with NSE ranked as the second largest exchange for trading of single stock futures both in terms of traded volume as well as number of contracts traded [1]. Stock futures are important financial instruments, especially for any informed investor. For an investor with price sensitive information, spot and futures markets provide two options to utilize that information. The benefit of leveraging suggests that the investor is better off acting on that information in the futures market compared to the cash (spot) market. If that information is utilized for placing orders first in the futures market, information will flow from the futures to the spot market. Restrictions on short-selling in the spot market, as in the case of Indian equity markets [2], provides strong motivation that the futures market would be first impacted by new information.

With the advancement of technology, exchanges have witnessed significant growth in algorithmic trading activities, whereby orders are placed automatically from computer terminals using proprietary algorithms. The growth of algorithmic trading has been met with skepticism from other market participants as it is often feared that algorithmic traders may be able to manipulate the market through their advantage of speed. Frequent events of flash crashes like that on May 6th, 2010 in the US equity market or October 5th, 2012 in the Indian market have not helped the cause for algorithmic traders and forced the market regulators to intervene. In the Indian market, where algorithmic traders provide close to half of the trading volume [3], the market regulator SEBI (Securities and Exchange Board of India) has been forced to intervene to strengthen algo trading regulations [4]. Despite criticism from various market participants, academic literature however provides evidence that algorithmic traders improve liquidity (Hendershott, Jones, & Menkveld, 2011) and the much maligned high-frequency traders continue to provide liquidity even during periods of stress (Nawn & Banerjee, 2018).

[1] World Federation of Exchanges Report,2015

[2] https://www.sebi.gov.in/legal/circulars/dec-2007/short-selling-and-securities-lending-and-borrowing_9463.html

[3] https://economictimes.indiatimes.com/markets/stocks/news/sebi-framing-algo-trading-rules-for-retail-investors/articleshow/61841163.cms

[4] https://economictimes.indiatimes.com/markets/stocks/news/full-text-of-key-decisions-at-sebis-board-meet/articleshow/63516953.cms

Machines interpret information much faster compared to human beings. As such it might be expected that with the growth of algo trading activity, the flow of information between interconnected financial markets would be much faster. That should result in reduction in the order of time to which one market leads or lags a connected market. We inspect if algorithmic trading plays any role in the lead-lag relationship of financial markets and whether algorithmic trades are informed.

We make a number of contributions to the existing literature on information flow and algorithmic trading. We provide evidence of lead-lag relationship at the level of stock and stock futures. Though a number of studies have looked into the lead-lag relationship at the market level, we do not come across studies looking at the phenomenon at the level of individual stocks. Prior studies have looked at the issue of information flow at the aggregate trade level. We also look at this issue at category of trader level because information flow is not expected to be uniform across different trader groups.

We find that the single stock-futures market leads the spot market by one minute. When we split the order imbalance posed by different category of traders, we find that algorithmic traders are not informed and the non-algorithmic traders are responsible for this flow of information between the futures and spot markets. The results hold even during periods of extreme price movements, when the advantage of utilizing sensitive information is the most. Segmenting the algorithmic traders into the level of proprietary [5] and agency [6] algorithmic traders do not change our results. Therefore, we try to put to rest the allegation that high-frequency traders have informational advantage.

The rest of the paper is arranged as follows - in Section 2, we briefly discuss the relevant literature. Section 3 talks about algorithmic trading activities in the Indian market and Section 4 describes our dataset. We explore the relationship between order imbalance and intraday-returns in Section 5 and then categorize the order imbalance due to various trader groups in Section 6. We conclude in Section 7.

[5] trading on own account, primarily high-frequency traders

[6] executing on behalf of someone else

2 Related Literature

A number of studies have looked into the temporal relationship between the cash index and the futures market (Chan, 1992; Finnerty & Park, 1992; Kawaller, Koch, & Koch, 1987; Harris, 1989; Stoll & Whaley, 1990). Empirical evidence suggest futures market returns lead cash index returns, although there is weak evidence pointing towards predictive ability of cash index returns for futures market returns. Kawaller et al. (1987) report that the S&P 500 futures lead the cash market by 20-45 minutes, while the lead from cash to futures is rarely more than a minute. Similar findings by Stoll and Whaley (1990) show that the S&P 500 and Major Market Index (MMI) futures lead the spot by 5-10 minutes, while the feedback is of much shorter duration. Chan (1992) reports that futures market leads the spot to a greater degree in presence of market-wide information.

The relationship between trading activity, and more specifically trading volume (which it can be argued is a good proxy for liquidity) and stock returns has been extensively studied in the literature (Gallant, Rossi, & Tauchen, 1992; Hiemstra & Jones, 1994; Lo & Wang, 2000). But intuitively, order imbalance should have a more pronounced effect on market returns compared to aggregate trading volume. According to the Kyle (1985) model of price formation, price change is related to net (pooled) order flow. Empirical investigations have looked into the relationship of order imbalance and individual stock returns (Cushing & Madhavan, 2000; Stoll, 2000) and more specifically into institutional order imbalances (Lakonishok, Shleifer, & Vishny, 1992; Kraus & Stoll, 1972; Sias, 1997; Wermers, 1999). Chordia, Roll, and Subrahmanyam (2002) and Chordia and Subrahmanyam (2004) show that the relationship holds for NYSE stocks in longer term (1988-1998) both for index returns and individual stock returns. Though the model was proposed for intermediated markets, similar results were obtained from inspection of order-driven markets (Handa, Schwartz, & Tiwari, 2003; Huang & Chou, 2007).

3 Algorithmic Trading and Indian Markets

Derivative products were introduced in the Indian market by NSE and BSE (Bombay Stock Exchange) during 2000. The relatively newer exchange NSE (setup in 1992) has surpassed the incumbent BSE (setup in 1875) in terms of volume of trade. Presently NSE represents the largest share of the equity as well as derivative market activities in India. NSE is a completely order driven market with no designated market-maker. Trading session starts from 9:15 AM IST (GMT + 5:30) and continues till 3:30 PM IST without any breaks in between. The annual turnover in the NSE equity derivatives for the year 2014-15 was 556 trillion INR, which is equivalent to approximately 8.8 trillion USD [7]. The turnover in the single stock futures segment for the same period was close to 83 trillion INR.

The Indian capital market regulator SEBI started providing DMA (Direct Market Access) services to the investors in 2008, which can be considered as the first step towards introduction of algorithm based trading activities in India. This service enabled the investors to directly access the exchange trading system through the broker's network without any manual intervention of the broker. Further stimulus was provided through the introduction of co-location facility in 2010, which allowed the brokers to place their servers at the exchange premises. Since then, Indian markets have witnessed a significant growth of algorithmic trading activities. During 2015, more than 40% of the trading volume in the cash segment as well as single stock futures was provided by algorithmic traders.

4 Data

We use a unique dataset obtained from National Stock Exchange of India (NSE). The dataset contains intraday order and trade level data for both the cash and the derivatives

[7] Based on the nominal USD-INR exchange rate as on 31st March 2015

market. Number of companies listed in the NSE cash market [8] is 1511. A subset of these stocks are permitted to be traded in the derivatives market. The dataset provides the entire set of order messages received by the exchange. There are flags provided to identify if the orders were generated from algorithmic trading terminals or otherwise. There are also indicators to identify if the order was placed by a proprietary trader (Prop), a custodian (Cust) or a non-proprietary non-custodian trader(NCNP). Custodial services are primarily availed by financial institutions who are legally not allowed to have a direct trading account with the exchange [9].

We consider six months of intraday trading data from 1st January 2015 to 30th June 2015. Our sample dataset consists of 160 stocks whose single stock futures are traded on NSE. The companies have to fulfill certain criteria in order for their derivatives to be traded on the exchange. These are the stocks with highest liquidity and largest market capitalization. The average market capitalization of these 160 stocks is 467 Billion INR (7.05 Billion USD) and median market capitalization being 208 Billion INR (3.14 Billion USD) [10]. NSE offers futures contract with 3 expiry dates at any point of time, contracts expiring end of current month (near-month contract), expiring at the end of next month (middle-month contract) and expiring at the end of next to next month (far-month contract). Contract expiry date is usually the last Thursday of a month. For the sake of liquidity, we only consider near-month contracts for our analysis.

For the measure of order imbalance, we require the trades to be classified as buyer-initiated or seller-initiated. The NSE dataset gives us a unique benefit of having the entire sequence of order messages received by the exchange. As such, we may use this information directly to classify trades as buyer-initiated or seller-initiated instead of using algorithms. As a measure of order imbalance, we use two alternate definitions similar to Chordia and Subrahmanyam (2004).

$OIBNUM$: number of buyer-initiated minus the number of seller-initiated trades

[8] as on 31st March 2015

[9] https://www.nseindia.com/products/content/derivatives/equities/cp_deals.htm

[10] Market capitalization as on 31 Dec 2015. USD figures have been computed using USD-INR exchange rate as on 31 Dec 2015.

scaled by the total number of trades over a period.

$OIBVOL$: INR volume of buyer-initiated minus the INR volume of seller-initiated trades scaled by the total INR volume of trades over a period.

We use the prefix CM or FUT to identify if the variable is defined for the spot or futures market respectively.

5 Order Imbalance and Intraday Returns

Chordia and Subrahmanyam (2004) suggests that daily stock returns can be explained by contemporaneous and lagged order imbalance. Discretionary liquidity traders split their order across periods, resulting in the order imbalances to be positively autocorrelated. Presence of any positive information in a particular period translates to positive order imbalance and vice-versa. Due to the autocorrelation of order imbalances, contemporaneous and lagged order imbalances are positively related to price change and, therefore, returns. They also showed that if the price change is regressed on the contemporaneous and lagged order imbalances, the co-efficient of the contemporaneous order imbalance is expected to be positive while the co-efficient of the lagged order imbalances are expected to be negative. Though the model (Chordia & Subrahmanyam, 2004) was proposed for explaining daily returns, we use the same model for explaining intraday returns. The highly autocorrelated nature of order imbalances remain valid for intraday measures also. Instead of using daily time series regressions (Chordia & Subrahmanyam, 2004), we use fixed effect panel regression.

$$R_{it} - R_{mt} = a + \sum_{k=0}^{5} b_k CM\_OIB_{i,t-k} + \delta_i + e_{it} \tag{1}$$

We regress the excess return for any particular stock on the contemporaneous and lagged order imbalances. Market return is proxied by the Nifty 50 Index [11] returns. The serial autocorrelation of intraday order imbalances is quite high. For our sample dataset,

[11] The NIFTY 50 is a diversified 50 stock index for the National Stock Exchange of India

the first order autocorrelation of spot market order imbalance measures (CM_OIBNUM) computed at both five minute and one minute intervals, are statistically significant (at 1% level) for all 160 stocks, with the cross sectional average being 0.544 and 0.367 respectively. As such, we expect the coefficient of the contemporaneous order imbalance to be positive while the same for lagged order imbalances to be negative.

$$R_{it} - R_{mt} = a + \sum_{k=0}^{5} b_k CM\_OIB_{i,t-k} + \sum_{k=0}^{5} c_k FUT\_OIB_{i,t-k} + \delta_i + e_{it} \quad (2)$$

In order to check for flow of information from the futures market to the spot market, we introduce futures market order imbalances in the model to explain spot market returns (Eqn.2). Trading volume in the spot and futures market are expected to be correlated. As such, in absence of any lead-lag relationship between the markets, the sign of the coefficients of the futures market order imbalances are expected to be similar to the spot market order imbalances. But if information in the futures market leads the spot market, we can expect not just the coefficient of the contemporaneous futures market order imbalance to be positive, but also the coefficients of the lagged futures market order imbalances to be positive.

5.1 Results

We compute our measures of order imbalances - $OIBNUM$ and $OIBVOL$, for the spot and futures market at five minute and one minute intervals. Table 1 reports descriptive statistics of our data. Apart from the order imbalance measures and returns, we also report summary statistics for the unscaled measures of order imbalances. As it can be seen, trading is more frequent in the spot market, whereas trade volume (INR) is much more for the futures market. This is primarily due to the presence of minimum trading units in the futures market - a phenomenon absent in the spot market.

Table 2 proves the cross-sectional average of individual time-series correlations for the order imbalance measures and the excess spot market returns. The correlation values for the measures computed at five minute and one minute intervals are quite similar. We do not report the correlation values for the unscaled order imbalance measures, as we do not use them in our final model. The correlation between the two order imbalance measures, $OIBNUM$ and $OIBVOL$ is extremely high, both for spot and futures market, indicating that we might expect similar results using these two alternate definitions in our model.

Returns have been computed as logarithmic returns over the last traded price for the specified time intervals, both for spot as well as the futures market.

Table 1: The table presents the summary statistics for the NSE stocks traded in the futures market during Jan-Jun 2015. Number of stocks included in the sample is 160. Variables have been computed for 5 min and 1 min intervals.

Variable	5 Min Interval		1 Min Interval	
	Mean	Std. Dev.	Mean	Std. Dev.
CM_OIBNUM	-0.0186	0.3688	-0.0115	0.4951
CM_OIBVOL	-0.0151	0.3911	-0.0088	0.5499
CM_OIBNUM (Unscaled)	-5.3854	243.4568	-1.0079	69.7841
CM_OIBVOL (Unscaled) (INR million)	-1.3320	21.2451	-0.0506	36.2470
FUT_OIBNUM	-0.0371	0.4017	-0.0305	0.5525
FUT_OIBVOL	-0.0431	0.4198	-0.0328	0.5639
FUT_OIBNUM (Unscaled)	-2.2609	33.3387	-0.5203	11.1568
FUT_OIBVOL (Unscaled) (INR million)	-1.3320	18.0152	-0.3044	6.6334
Spot Trade Count	410.3087	577.8125	86.3000	136.7881
Spot Trade Volume (INR million)	11.5594	36.1484	2.4887	38.6544
Futures Trade Count	62.8276	124.5347	15.6600	32.9394
Futures Trade Volume (INR million)	24.2399	53.8856	6.0117	14.6809
Spot Return ($\times 10^4$)	-0.2834	24.3463	-0.0682	18.0121
Futures Return ($\times 10^4$)	-0.2720	25.4459	-0.0325	54.0708
Index Return ($\times 10^4$)	-0.1141	8.7042	-0.0527	4.2304
Excess Spot Return ($\times 10^4$)	-0.1693	22.7030	-0.0180	17.6726
Excess Futures Return ($\times 10^4$)	-0.1477	23.4507	0.0343	54.1141

First, we run a fixed effect panel model, regressing the excess spot market returns computed at five minute intervals on contemporaneous and 5 lagged order imbalances. Next, we include futures market order imbalances and re-run the model. We use both the definitions of order imbalance- $OIBNUM$ and $OIBVOL$. The results (Table 3) suggest that, similar to earlier daily return models (Chordia & Subrahmanyam, 2004), the coefficient of the contemporaneous order imbalance is positive, while the coefficients of the

Table 2: Cross-sectional average of individual time-series correlations for order imbalance measures computed at 5 min and 1 min interval.

Panel A : 5 Min Interval

	CM_OIBVOL	FUT_OIBNUM	FUT_OIBVOL	Excess Spot Returns
CM_OIBNUM	0.7196	0.2620	0.2363	0.2540
CM_OIBVOL	1.0000	0.3080	0.2799	0.2670
FUT_OIBNUM		1.0000	0.9566	0.2724
FUT_OIBVOL			1.0000	0.2527

Panel B : 1 Min Interval

	CM_OIBVOL	FUT_OIBNUM	FUT_OIBVOL	Excess Spot Returns
CM_OIBNUM	0.7832	0.2570	0.2465	0.2976
CM_OIBVOL	1.0000	0.2727	0.2614	0.2914
FUT_OIBNUM		1.0000	0.9826	0.2337
FUT_OIBVOL			1.0000	0.2256

CM_OIBNUM - Spot market order imb. in number of transactions scaled by total trans.
CM_OIBVOL - Spot market order imb. in INR scaled by total INR volume
FUT_OIBNUM - Fut. market order imb. in number of transactions scaled by total trans.
FUT_OIBVOL - Futures market order imb. in INR scaled by total INR volume

lagged order imbalances are negative and significant. The inclusion of futures market order imbalances does not change the signs for the spot order imbalances and the signs of futures market order imbalances are identical to spot market order imbalances. Both spot and futures market contemporaneous order imbalances are positively related to spot market returns, possibly due to strong correlation of trading volume in these two markets. There does not seem to be any evidence of the futures market leading the spot, at least in the order of five minutes or more. The results are consistent for both definitions of order imbalance - $OIBNUM$ and $OIBVOL$.

We run the same model for variables computed at one minute intervals. Results (Table 4) are similar when we use only spot market order imbalances as regressors. We find that only the contemporaneous spot order imbalance has positive and significant coefficient. The coefficients for lagged order imbalances are negative and significant. However, when we include the futures market order imbalances in the model, we find that not only the coefficient of contemporaneous futures market order imbalance is positive and significant, the coefficient of first lag of futures market order imbalance is also positive and significant. This indicates that spot returns are positively influenced by one minute lagged futures market order imbalance, or information in the futures market lead the spot market in the order of one minute. The results are consistent for both definitions of order imbalance.

As a robustness check, we run separate panel regressions for each month, using one minute spot returns as the dependent variable and both spot and futures market order imbalances as independent variables. Aggregate results showed that coefficient corresponding to the first lag of the futures market order imbalance remains positive, indicating information flow from futures to spot market. As such, we are interested in the coefficient corresponding to the first lag of the futures market only for the monthly regression models. Results using $OIBNUM$ as the measure of order imbalance (Table 5) shows that the coefficient for the first lag of the futures market order imbalance remains positive and significant throughout all the six months. [12]

To test if the flow of information is unidirectional in nature, i.e., from futures to spot but not the other way around, we use excess futures market as the dependent variable in the fixed effect panel regression model. Our first model uses only futures market order imbalances and the second model includes the measures of spot order imbalances also.

$$R_{ft} - R_{mt} = a + \sum_{k=0}^{5} c_k FUT\_OIB_{i,t-k} + +\delta_i + e_{it} \tag{3}$$

$$R_{ft} - R_{mt} = a + \sum_{k=0}^{5} b_k CM\_OIB_{i,t-k} + \sum_{k=0}^{5} c_k FUT\_OIB_{i,t-k} + \delta_i + e_{it} \tag{4}$$

Results of the panel data analysis (Table 6) using excess futures market returns as dependent variable indicates that coefficients for only the contemporaneous order imbalances are positive and significant. Coefficients for lagged order imbalances are negative, indicating that the information flow is unidirectional from futures to spot market and

[12] We also run similar test using $OIBVOL$ as the measure of order imbalance and find that the sign of the coefficient for the first lag of futures market order imbalance remains positive across all six months, but the level of significance is not consistent throughout.

Table 3: Panel Data Regressions of 5 min interval cash market excess returns on contemporaneous and lagged order imbalances of the cash and futures market. Panel A of the table defines Order imbalance(OIB) as the estimated number of buyer-initiated minus seller-initiated trades scaled by the total number of trades (OIBNUM) over five minute interval. Panel B of the table defines Order imbalance(OIB) as the estimated buyer-initiated minus seller-initiated INR volume of transactions scaled by total INR volume of trade (OIBVOL) over five minute interval.

Dependent Variable: Excess Return				
	(1)		**(2)**	
Variable	**Estimate**	**t Stat**	**Estimate**	**t Stat**
Panel A : OIBNUM				
Constant	-0.071***	(-11.93)	0.198***	(13.50)
CM_OIBNUM	24.670***	(27.89)	20.770***	(27.99)
L1_CM_OIBNUM	-9.565***	(-22.14)	-8.798***	(-21.59)
L2_CM_OIBNUM	-3.674***	(-22.88)	-3.077***	(-21.28)
L3_CM_OIBNUM	-2.604***	(-18.63)	-2.044***	(-16.26)
L4_CM_OIBNUM	-2.321***	(-17.32)	-1.799***	(-15.46)
L5_CM_OIBNUM	-2.310***	(-15.66)	-1.791***	(-13.57)
FUT_OIBNUM			10.140***	(30.55)
L1_FUT_OIBNUM			-0.478***	(-3.80)
L2_FUT_OIBNUM			-0.424***	(-6.16)
L3_FUT_OIBNUM			-0.539***	(-8.06)
L4_FUT_OIBNUM			-0.473***	(-8.05)
L5_FUT_OIBNUM			-0.440***	(-7.66)
Adj. R^2		0.101		0.129
Panel B : OIBVOL				
Constant	-0.065***	(-12.51)	0.205***	(14.63)
CM_OIBVOL	18.410***	(28.32)	15.320***	(28.98)
L1_CM_OIBVOL	-4.763***	(-24.80)	-4.259***	(-27.06)
L2_CM_OIBVOL	-2.417***	(-22.34)	-2.009***	(-21.67)
L3_CM_OIBVOL	-1.767***	(-19.68)	-1.361***	(-16.68)
L4_CM_OIBVOL	-1.773***	(-21.53)	-1.410***	(-18.79)
L5_CM_OIBVOL	-1.701***	(-19.95)	-1.367***	(-17.54)
FUT_OIBVOL			9.432***	(31.58)
L1_FUT_OIBVOL			-0.939***	(-8.47)
L2_FUT_OIBVOL			-0.538***	(-10.04)
L3_FUT_OIBVOL			-0.529***	(-9.71)
L4_FUT_OIBVOL			-0.386***	(-7.79)
L5_FUT_OIBVOL			-0.373***	(-7.05)
Adj. R^2		0.081		0.107
No. of obs.		1,353,730		1,353,730

Table 4: Panel Data Regressions of 1 min interval cash market excess returns on contemporaneous and lagged order imbalances of the cash and futures market. Panel A of the table defines Order imbalance(OIB) as the estimated number of buyer-initiated minus seller-initiated trades scaled by the total number of trades (OIBNUM) over one minute interval. Panel B of the table defines Order imbalance(OIB) as the estimated buyer-initiated minus seller-initiated INR volume of transactions scaled by total INR volume of trade (OIBVOL) over one minute interval.

Dependent Variable: Excess Return				
	(1)		**(2)**	
Variable	**Estimate**	**t Stat**	**Estimate**	**t Stat**
Panel A : OIBNUM				
Constant	0.001	(0.77)	0.094***	(25.56)
CM_OIBNUM	8.983***	(28.31)	8.078***	(26.51)
L1_CM_OIBNUM	-3.354***	(-13.68)	-3.397***	(-13.42)
L2_CM_OIBNUM	-0.991***	(-24.49)	-0.888***	(-24.69)
L3_CM_OIBNUM	-0.808***	(-24.35)	-0.689***	(-23.27)
L4_CM_OIBNUM	-0.752***	(-26.32)	-0.629***	(-25.06)
L5_CM_OIBNUM	-0.765***	(-25.79)	-0.640***	(-24.14)
FUT_OIBNUM			3.284***	(38.41)
L1_FUT_OIBNUM			0.318***	(7.18)
L2_FUT_OIBNUM			-0.084***	(-6.28)
L3_FUT_OIBNUM			-0.122***	(-9.31)
L4_FUT_OIBNUM			-0.102***	(-9.72)
L5_FUT_OIBNUM			-0.126***	(-12.02)
Adj. R^2		0.103		0.123
Panel B : OIBVOL				
Constant	0.000	(0.32)	0.090***	(25.03)
CM_OIBVOL	6.848***	(35.06)	6.079***	(33.85)
L1_CM_OIBVOL	-1.854***	(-17.17)	-1.891***	(-16.74)
L2_CM_OIBVOL	-0.611***	(-22.40)	-0.541***	(-23.33)
L3_CM_OIBVOL	-0.547***	(-25.19)	-0.465***	(-24.20)
L4_CM_OIBVOL	-0.503***	(-26.33)	-0.420***	(-24.84)
L5_CM_OIBVOL	-0.518***	(-27.32)	-0.435***	(-25.49)
FUT_OIBVOL			3.261***	(36.58)
L1_FUT_OIBVOL			0.135***	(3.94)
L2_FUT_OIBVOL			-0.150***	(-12.22)
L3_FUT_OIBVOL			-0.153***	(-12.60)
L4_FUT_OIBVOL			-0.124***	(-12.43)
L5_FUT_OIBVOL			-0.142***	(-14.73)
Adj. R^2		0.082		0.102
No. of Obs.		7,085,452		7,085,452

Table 5: Panel Data Regressions of 1 min interval cash market excess returns on contemporaneous and lagged order imbalances of the cash and futures market on a monthly basis from Jan-2015 to Jun-2015. Order imbalance(OIB) defined as the estimated number of buyer-initiated minus seller-initiated trades scaled by the total number of trades (OIBNUM) over one minute interval.

Dependent Variable: Excess Return						
Variable	**Jan-15**	**Feb-15**	**Mar-15**	**Apr-15**	**May-15**	**Jun-15**
Constant	-0.0121***	0.255***	0.147***	0.0591***	0.0568***	0.0314***
	(-12.43)	(23.41)	(27.14)	(30.28)	(18.47)	(7.59)
CM_OIBNUM	7.627***	8.731***	7.572***	8.549***	7.951***	8.344***
	(26.75)	(24.82)	(26.77)	(29.47)	(25.63)	(19.93)
L1_CM_OIBNUM	-3.066***	-3.246***	-3.366***	-3.334***	-3.398***	-3.576***
	(-14.26)	(-15.30)	(-14.95)	(-13.38)	(-12.09)	(-10.19)
L2_CM_OIBNUM	-0.741***	-0.874***	-0.969***	-0.847***	-0.790***	-0.820***
	(-17.08)	(-18.36)	(-21.27)	(-18.38)	(-15.74)	(-15.83)
L3_CM_OIBNUM	-0.552***	-0.663***	-0.862***	-0.679***	-0.515***	-0.607***
	(-16.33)	(-12.80)	(-18.44)	(-18.14)	(-15.57)	(-13.91)
L4_CM_OIBNUM	-0.564***	-0.533***	-0.745***	-0.588***	-0.480***	-0.568***
	(-16.68)	(-13.77)	(-16.47)	(-16.15)	(-13.48)	(-15.47)
L5_CM_OIBNUM	-0.512***	-0.623***	-0.761***	-0.605***	-0.492***	-0.522***
	(-18.58)	(-15.60)	(-18.08)	(-14.76)	(-15.20)	(-13.12)
FUT_OIBNUM	2.907***	3.603***	3.171***	3.351***	3.256***	3.477***
	(36.26)	(34.73)	(35.42)	(31.04)	(31.25)	(32.11)
L1_FUT_OIBNUM	0.294***	0.188***	0.202***	0.405***	0.347***	0.479***
	(7.19)	(4.01)	(4.73)	(7.91)	(7.06)	(6.47)
L2_FUT_OIBNUM	0.0442*	-0.188***	-0.179***	-0.0202	-0.0726**	-0.0672
	(2.15)	(-7.81)	(-8.69)	(-0.99)	(-2.99)	(-1.93)
L3_FUT_OIBNUM	-0.0616***	-0.222***	-0.211***	-0.0342	-0.101***	-0.0772*
	(-3.52)	(-9.62)	(-9.78)	(-1.59)	(-5.33)	(-2.10)
L4_FUT_OIBNUM	-0.0413*	-0.241***	-0.186***	-0.0280	-0.0390	-0.0578
	(-2.29)	(-11.44)	(-8.96)	(-1.58)	(-1.66)	(-1.89)
L5_FUT_OIBNUM	-0.0362*	-0.262***	-0.224***	-0.0440*	-0.0830***	-0.0832**
	(-2.07)	(-12.06)	(-12.09)	(-2.08)	(-3.51)	(-2.80)
No. of obs.	1,223,628	1,163,623	1,222,540	1,101,492	1,155,387	1,219,428
$Adj.R^2$	0.125	0.124	0.107	0.139	0.137	0.115

not the other way around. Results are consistent for both definitions of order imbalance.

Table 6: Panel Data Regressions of 1 min interval futures market excess returns on contemporaneous and lagged order imbalances of the cash and futures market. Panel A of the table defines Order imbalance(OIB) as the estimated number of buyer-initiated minus seller-initiated trades scaled by the total number of trades (OIBNUM) over one minute interval. Panel B of the table defines Order imbalance(OIB) as the estimated buyer-initiated minus seller-initiated INR volume of transactions scaled by total INR volume of trade (OIBVOL) over one minute interval.

Dependent Variable: Excess Futures Return				
	(1)		**(2)**	
Variable	**Estimate**	**t Stat**	**Estimate**	**t Stat**
Panel A : OIBNUM				
Constant	0.148***	(31.28)	0.164***	(30.90)
CM_OIBNUM			6.535***	(29.94)
L1_CM_OIBNUM			-0.485***	(-5.39)
L2_CM_OIBNUM			-0.985***	(-19.28)
L3_CM_OIBNUM			-0.841***	(-16.98)
L4_CM_OIBNUM			-0.829***	(-14.06)
L5_CM_OIBNUM			-0.895***	(-20.92)
FUT_OIBNUM	6.248***	(29.41)	5.031***	(26.85)
L1_FUT_OIBNUM	-0.656***	(-5.52)	-0.763***	(-6.24)
L2_FUT_OIBNUM	-0.536***	(-12.56)	-0.355***	(-8.40)
L3_FUT_OIBNUM	-0.520***	(-17.30)	-0.311***	(-10.64)
L4_FUT_OIBNUM	-0.353***	(-11.76)	-0.125***	(-3.93)
L5_FUT_OIBNUM	-0.487***	(-17.32)	-0.245***	(-8.81)
Adj. R^2		0.011		0.016
Panel B : OIBVOL				
Constant	0.153***	(30.11)	0.156***	(30.67)
CM_OIBVOL			5.184***	(32.90)
L1_CM_OIBVOL			-0.041	(-0.54)
L2_CM_OIBVOL			-0.541***	(-13.27)
L3_CM_OIBVOL			-0.537***	(-15.81)
L4_CM_OIBVOL			-0.529***	(-11.41)
L5_CM_OIBVOL			-0.529***	(-14.74)
FUT_OIBVOL	5.815***	(28.45)	4.734***	(25.35)
L1_FUT_OIBVOL	-0.516***	(-4.70)	-0.668***	(-5.82)
L2_FUT_OIBVOL	-0.498***	(-11.81)	-0.401***	(-8.97)
L3_FUT_OIBVOL	-0.479***	(-16.43)	-0.335***	(-12.04)
L4_FUT_OIBVOL	-0.321***	(-11.06)	-0.157***	(-5.02)
L5_FUT_OIBVOL	-0.449***	(-16.50)	-0.279***	(-9.78)
Adj. R^2		0.010		0.015
No. of Obs.		5,608,792		5,608,792

6 Categorized Order Imbalance

We find evidence of unidirectional information flow from the futures market to the spot market. But it does not provide any indication about who is responsible for this flow of information. The question regarding the informational role of certain trader groups is even more interesting considering the substantial growth of algorithmic and high-frequency trading activities. By observing how order imbalance of various trader categories relate to spot market returns, we intend to contribute to the ongoing debate on the informational role of algorithmic and high-frequency traders.

A number of existing studies have looked into the role of these machine enabled trading activities and informational efficiency. Results indicate that algorithmic and high-frequency trading activity is correlated with public information (Brogaard, Hendershott, & Riordan, 2014) and they improve the speed of assimilation of this information into security prices. Algorithms are used to quickly process information contained in order flow to determine when security prices deviate from efficient prices (Hendershott et al., 2011). But most of the existing literature on algorithmic trading use proxies in the absence of direct identification to categorize an order coming from high frequency traders. By using the unique dataset from NSE of India, we are able to overcome that limitation - the dataset identifies algo and non-algo orders.

High-frequency traders are subset of algorithmic traders, who submit and revise orders to the exchange at extremely high speed, with round trip execution time in the order of microseconds. High-frequency traders (HFT) are primarily proprietary algorithmic traders. Apart from HFT, algorithmic trading is also undertaken by agency algorithmic traders, who use algorithms to execute trades on someone else's (primarily institutional investors) behalf. We separately study if HFTs and agency algorithmic trades are informed.

We split the aggregate order imbalance into components of order imbalance of different categories of traders. In each of the categorized models, the categorization is carried out in such a way that the sum of the categorized order imbalance equals the market order imbalance. For all of the following models, we use fixed-effect panel regression using spot market returns at 1-minute interval as the dependent variable. Consistent with the earlier results, we analyze six months of intraday data (1st January 2015 to 30th June 2015). Earlier we did not find any evidence of information flow beyond one minute. Hence, we use contemporary and one-minute lagged order imbalance measures as independent variables.

The aspect of information flow is more prominent in periods of extreme price movements and consequently in periods at the extreme ends of the excess spot-market returns distribution. As such, we separately run models where we regress the excess spot market returns on the categorized order imbalance measures, where the excess return belongs to the top 10% (positive extreme returns) or bottom 10% (negative extreme returns) of the excess return distribution for each stock.

Each trade in our dataset has identifiers for the class (Proprietary, Custodian [13] and others [14]) and nature (Algorithmic and non-algorithmic) of traders. We define order imbalance (OIBNUM) for any type of trader-group as the number of buyer-initiated trades where that particular trader-group is mentioned as the buyer, minus the number of seller-initiated trades where that particular trader-group is mentioned as the seller, scaled by the total number of trades within that one-minute period. For the alternate definition of order imbalance (OIBVOL), we use INR volume of trades in place of number of trades.

We carry out the categorization of order imbalance in three different ways. The first set of models (Table 7 and Table 8) categorize order imbalance into components due to high-frequency traders (HFT) and non high-frequency traders. As our dataset does not provide exact identifiers for HFT, we use proprietary algorithmic traders as our proxy for HFT. The second set of models (Table 9 and Table 10) club proprietary and agency algorithmic traders into a single group of algorithmic traders and categorize order-imbalance into components due to algorithmic and non-algorithmic traders. The final set of mod-

[13] Custodians trades are primarily executed for institutional investors

[14] Non-Proprietary Non-Custodian traders

els (Table 11 and Table 12) break down the order imbalance into components due to all the identifiable trader categories as available in the dataset - Proprietary Algorithmic (PA), Custodian[15] Algorithmic (CA), Non-Proprietary Non-Custodian Algorithmic (NCNPA), Proprietary Non-Algorithmic (PNA), Custodian Non-Algorithmic (CNA) and Non-Proprietary Non-Custodian Non-Algorithmic (NCNPNA) traders. We use the prefix CM to define the variables for the spot market and the prefix FUT for the single stock futures market.

Table 7: Results of panel-data regression models of cash market excess returns at one minute interval on contemporaneous and lagged order imbalances of the cash and futures market split across categories of HFT and non-HFT. Order imbalance (OIB) for this panel is measured as the number of buyer-initiated minus seller-initiated trades scaled by the total number of trades (OIBNUM) over one minute interval. Model 1 uses only contemporaneous spot and futures order imbalance measures. Model 2 is the full model using both contemporaneous and lagged (one minute) spot and futures market order imbalance measures. Model 3 uses the full model with the subset with positive extreme returns, while Model 4 uses the subset with negative extreme returns.

Dependent Variable: Excess Spot Return				
Variable	**(1)**	**(2)**	**(3)**	**(4)**
Constant	0.152***	0.129***	19.85***	-19.73***
	(32.40)	(31.84)	(291.58)	(-216.97)
CM_OIBNUM (HFT)	6.685***	6.914***	-0.0858	-0.445
	(18.37)	(18.80)	(-0.26)	(-1.33)
CM_OIBNUM (Non-HFT)	5.965***	7.738***	1.363***	0.978***
	(27.58)	(24.17)	(6.42)	(4.93)
FUT_OIBNUM (HFT)	3.122***	2.754***	-3.358***	-2.787***
	(18.92)	(16.74)	(-11.42)	(-9.71)
FUT_OIBNUM (Non-HFT)	3.576***	3.491***	0.755***	0.322*
	(41.91)	(40.70)	(6.47)	(2.17)
L1_CM_OIBNUM (HFT)		-1.639***	0.143	0.169
		(-15.77)	(0.58)	(1.31)
L1_CM_OIBNUM (Non-HFT)		-4.619***	0.731***	0.214*
		(-14.95)	(7.83)	(2.28)
L1_FUT_OIBNUM (HFT)		0.152*	-0.746*	0.117
		(2.23)	(-2.07)	(1.29)
L1_FUT_OIBNUM (Non-HFT)		0.283***	0.654***	0.298***
		(7.64)	(14.11)	(4.74)
No. of Obs.	7,162,806	7,162,806	709,038	706,888
Adj. R^2	0.050	0.062	0.001	0.002

[15]institutional

Table 8: Results of panel-data regression models of cash market excess returns at one minute interval on contemporaneous and lagged order imbalances of the cash and futures market split across categories of HFT and non-HFT. Order imbalance (OIB) for this panel is measured as the INR volume of buyer-initiated minus seller-initiated trades scaled by the total INR volume of trades (OIBVOL) over one minute interval. Model 1 uses only contemporaneous spot and futures order imbalance measures. Model 2 is the full model using both contemporaneous and lagged (one minute) spot and futures market order imbalance measures. Model 3 uses the full model with the subset with positive extreme returns, while Model 4 uses the subset with negative extreme returns.

Dependent Variable: Excess Spot Return				
Variable	**(1)**	**(2)**	**(3)**	**(4)**
Constant	0.130***	0.118***	20.04***	-19.83***
	(32.46)	(31.33)	(208.84)	(-240.56)
CM_OIBVOL (HFT)	5.514***	5.538***	-1.283***	-1.323***
	(22.64)	(22.68)	(-5.45)	(-5.32)
CM_OIBVOL (Non-HFT)	5.281***	5.941***	0.797*	0.847***
	(31.58)	(31.42)	(2.15)	(5.42)
FUT_OIBVOL (HFT)	2.628***	2.459***	-3.743***	-3.397***
	(16.10)	(14.97)	(-12.25)	(-11.58)
FUT_OIBVOL (Non-HFT)	3.409***	3.420***	0.782***	0.291*
	(39.06)	(38.40)	(6.82)	(2.07)
L1_CM_OIBVOL (HFT)		-0.885***	0.503***	0.295**
		(-15.36)	(5.38)	(3.06)
L1_CM_OIBVOL (Non-HFT)		-2.492***	0.909***	0.533***
		(-19.39)	(15.64)	(8.12)
L1_FUT_OIBVOL (HFT)		-0.0788	-0.929**	-0.122
		(-1.32)	(-2.65)	(-1.31)
L1_FUT_OIBVOL (Non-HFT)		0.0697*	0.583***	0.189**
		(2.06)	(13.67)	(3.14)
No. of Obs.	7,162,806	7,162,806	709,038	706,888
Adj. R2	0.047	0.052	0.001	0.003

For each of the three aforementioned categorization models, we use four set of regression specifications. The first set of regression specifications, uses only the contemporaneous order imbalance measures due to various trader categories, while for the second we use the full model having both contemporaneous and lagged categorized order imbalance measures. The third and fourth set of models use the full model as in case of model 2 but for the subset of data for positive and negative extreme returns respectively. Continuing from the earlier section, we are interested about the sign and significance of the coefficient corresponding to the lagged futures market order imbalance of the different trader categories to infer about their informational role. For each of the models, we use both definitions of order imbalance - OIBNUM and OIBVOL.

For the first set of categorization models, we find that the coefficient corresponding to Non-HFT lagged futures market order imbalance measure is positive and significant for all the models, while the same for HFT traders is not true. The results are similar for both measures of order imbalance, both using number of trades (Table 7), and size of trades (Table 8). The results indicate that HFTs are not informed.

For the second set of models, we club proprietary and agency algorithmic traders. From the analysis of algorithmic trading, we find that the coefficients corresponding to lagged futures market order imbalance for non-algo traders are continuously positive and significant for all the models (Table 9 & 10). But the coefficients corresponding to the lagged futures market order imbalance due to the algo traders are not significant and positive throughout the models. The results from the analysis of HFT and algo traders are consistent with our understanding of machine traders. Machines interpret public information in order flow much faster than human traders. But they do not seem to have any access to private information. As such, it is logical to expect them not to have any role in informational flow between spot and futures market.

In our final set of models, we split the market order imbalance into components due to all six trader categories. Consistent with our other results, we find that the coefficient corresponding to the lagged futures market order imbalance due to all three non-algorithmic trader groups (PNA, CNA & NCNPNA) are positive and significant (Table 11 & 12), while none of them are positive and significant for the algorithmic trader groups (PA, CA & NCNPA). These results provide further evidence to our observation that algorithmic traders are not informed traders- their activity does not explain the flow of information between the futures market and the spot market.

Table 9: Results of panel-data regression models of cash market excess returns at one minute interval on contemporaneous and lagged order imbalances of the cash and futures market split across categories of algorithmic and non-algorithmic traders. Order imbalance (OIB) for this panel is measured as the number of buyer-initiated minus seller-initiated trades scaled by the total number of trades (OIBNUM) over one minute interval. Model 1 uses only contemporaneous spot and futures order imbalance measures. Model 2 is the full model using both contemporaneous and lagged (one minute) spot and futures market order imbalance measures. Model 3 uses the full model with the subset with positive extreme returns, while Model 4 uses the subset with negative extreme returns.

Dependent Variable: Excess Spot Return				
	(1)	**(2)**	**(3)**	**(4)**
Constant	0.148***	0.137***	19.87***	-19.70***
	(30.37)	(32.00)	(317.92)	(-224.43)
CM_OIBNUM (Algo)	5.349***	6.405***	0.414*	0.451*
	(22.64)	(24.52)	(2.18)	(2.12)
CM_OIBNUM (Non-Algo)	6.703***	8.480***	1.556***	0.780**
	(25.55)	(22.04)	(6.32)	(3.16)
FUT_OIBNUM (Algo)	3.058***	2.809***	-2.002***	-1.624***
	(29.67)	(28.25)	(-11.16)	(-8.39)
FUT_OIBNUM (Non-Algo)	3.942***	3.919***	1.959***	1.160***
	(43.24)	(41.57)	(15.59)	(7.04)
L1_CM_OIBNUM (Algo)		-2.837***	0.654***	0.591***
		(-25.47)	(8.05)	(7.29)
L1_CM_OIBNUM (Non-Algo)		-5.314***	0.546***	-0.143
		(-13.29)	(4.99)	(-0.95)
L1_FUT_OIBNUM (Algo)		0.0577	-0.0845	0.108
		(1.49)	(-1.49)	(1.81)
L1_FUT_OIBNUM (Non-Algo)		0.341***	0.665***	0.245**
		(8.77)	(8.93)	(2.76)
No. of Obs.	7,162,806	7,162,806	709,038	706,888
Adj. R^2	0.051	0.063	0.001	0.003

Table 10: Results of panel-data regression models of cash market excess returns at one minute interval on contemporaneous and lagged order imbalances of the cash and futures market split across categories of algorithmic and non-algorithmic traders. Order imbalance (OIB) for this panel is measured as the INR volume of buyer-initiated minus seller-initiated trades scaled by the total INR volume of trades (OIBVOL) over one minute interval. Model 1 uses only contemporaneous spot and futures order imbalance measures. Model 2 is the full model using both contemporaneous and lagged (one minute) spot and futures market order imbalance measures. Model 3 uses the full model with the subset with positive extreme returns, while Model 4 uses the subset with negative extreme returns.

Dependent Variable: Excess Spot Return				
	(1)	**(2)**	**(3)**	**(4)**
Constant	0.135***	0.127***	20.05***	-19.81***
	(33.58)	(34.24)	(205.59)	(-244.36)
CM_OIBVOL (Algo)	4.546***	4.973***	-0.320*	0.000140
	(26.50)	(28.67)	(-2.31)	(0.00)
CM_OIBVOL (Non-Algo)	5.970***	6.547***	0.912	0.663**
	(30.17)	(29.39)	(1.77)	(3.21)
FUT_OIBVOL (Algo)	2.743***	2.659***	-2.141***	-1.835***
	(26.56)	(25.87)	(-12.19)	(-9.88)
FUT_OIBVOL (Non-Algo)	3.792***	3.830***	2.007***	1.158***
	(41.25)	(40.07)	(16.05)	(7.44)
L1_CM_OIBVOL (Algo)		-1.652***	0.666***	0.727***
		(-31.71)	(7.65)	(11.43)
L1_CM_OIBVOL (Non-Algo)		-2.748***	0.962***	0.310**
		(-16.86)	(10.15)	(2.83)
L1_FUT_OIBVOL (Algo)		-0.167***	-0.184**	-0.0338
		(-5.46)	(-3.33)	(-0.61)
L1_FUT_OIBVOL (Non-Algo)		0.157***	0.625***	0.146
		(3.83)	(10.84)	(1.68)
No. of Obs.	7,162,806	7,162,806	709,038	706,888
Adj. R^2	0.048	0.053	0.001	0.004

Table 11: Results of panel-data regression models of cash market excess returns at one minute interval on contemporaneous and lagged order imbalances of the cash and futures market split across six trader categories. Order imbalance (OIB) for this panel is measured as the number of buyer-initiated minus seller-initiated trades scaled by the total number of trades (OIBNUM) over one minute interval. Model 1 is the full model using both contemporaneous and lagged (one minute) spot and futures market order imbalance measures. Model 2 uses the full model with the subset with positive extreme returns, while Model 3 uses the subset with negative extreme returns.

Explanatory Variable: Excess Spot Returns						
	Full Sample		Positive Extremes		Negative Extremes	
Constant	0.134***	(36.50)	19.84***	(368.43)	-19.65***	(-226.12)
CM_CA_OIBNUM	6.040***	(39.71)	0.725*	(2.07)	1.456***	(5.57)
CM_PA_OIBNUM	6.963***	(20.06)	-0.354	(-1.11)	-0.575	(-1.72)
CM_NCNPA_OIBNUM	6.851***	(14.22)	0.894***	(3.46)	0.777*	(2.59)
CM_CNA_OIBNUM	7.899***	(35.18)	1.584***	(3.63)	2.817***	(8.01)
CM_PNA_OIBNUM	9.376***	(8.80)	3.191*	(2.18)	2.280***	(7.70)
CM_NCNPNA_OIBNUM	8.257***	(29.14)	1.212***	(4.02)	0.217	(0.64)
FUT_CA_OIBNUM	3.289***	(26.11)	0.113	(0.81)	0.241	(1.24)
FUT_PA_OIBNUM	2.906***	(17.84)	-2.898***	(-10.52)	-2.448***	(-9.03)
FUT_NCNPA_OIBNUM	2.333***	(19.08)	-3.002***	(-13.89)	-2.570***	(-11.04)
FUT_CNA_OIBNUM	1.645***	(19.98)	-1.107***	(-3.97)	-1.237***	(-5.32)
FUT_PNA_OIBNUM	4.807***	(27.93)	2.582***	(11.63)	0.973***	(4.39)
FUT_NCNPNA_OIBNUM	3.976***	(40.02)	2.164***	(19.33)	1.650***	(8.49)
L1_CM_CA_OIBNUM	-3.480***	(-31.88)	0.962***	(4.03)	0.718***	(5.18)
L1_CM_PA_OIBNUM	-2.255***	(-22.01)	-0.0143	(-0.06)	0.0914	(0.71)
L1_CM_NCNPA_OIBNUM	-1.981***	(-12.33)	0.791***	(6.45)	0.446**	(3.24)
L1_CM_CNA_OIBNUM	-3.840***	(-30.18)	0.473*	(1.99)	0.841***	(3.82)
L1_CM_PNA_OIBNUM	-8.315***	(-8.19)	-2.050***	(-4.44)	-2.707***	(-9.36)
L1_CM_NCNPNA_OIBNUM	-4.801***	(-19.23)	1.309***	(9.22)	0.416	(1.87)
L1_FUT_CA_OIBNUM	-0.0902	(-1.96)	0.200	(1.62)	0.248*	(2.40)
L1_FUT_PA_OIBNUM	0.115	(1.70)	-0.773	(-1.90)	0.0813	(0.92)
L1_FUT_NCNPA_OIBNUM	0.0743	(1.49)	-0.0352	(-0.13)	-0.184*	(-2.48)
L1_FUT_CNA_OIBNUM	0.394***	(6.80)	1.665**	(2.97)	0.565***	(4.33)
L1_FUT_PNA_OIBNUM	0.332***	(6.46)	1.034***	(10.45)	0.294***	(3.47)
L1_FUT_NCNPNA_OIBNUM	0.335***	(7.31)	0.519**	(2.94)	0.313**	(2.65)
No. of Obs		7,162,806		709,038		706,888
Adj. R^2		0.065		0.001		0.006

Table 12: Results of panel-data regression models of cash market excess returns at one minute interval on contemporaneous and lagged order imbalances of the cash and futures market split across six trader categories. Order imbalance (OIB) for this panel is measured as the INR volume of buyer-initiated minus seller-initiated trades scaled by the total INR volume of trades (OIBVOL) over one minute interval. Model 1 is the full model using both contemporaneous and lagged (one minute) spot and futures market order imbalance measures. Model 2 uses the full model with the subset with positive extreme returns, while Model 3 uses the subset with negative extreme returns.

Explanatory Variable: Excess Spot Returns						
	Full Sample		Positive Extremes		Negative Extremes	
Constant	0.130***	(36.85)	20.04***	(203.62)	-19.76***	(-246.80)
CM_CA_OIBVOL	4.977***	(45.93)	0.706***	(4.11)	1.571***	(7.62)
CM_PA_OIBVOL	5.603***	(24.10)	-1.487***	(-6.21)	-1.427***	(-5.56)
CM_NCNPA_OIBVOL	4.720***	(16.04)	0.184	(1.04)	0.296	(1.56)
CM_CNA_OIBVOL	5.644***	(36.44)	1.077**	(2.97)	2.031***	(9.05)
CM_PNA_OIBVOL	8.683***	(8.89)	1.961	(0.72)	1.991***	(4.43)
CM_NCNPNA_OIBVOL	6.223***	(39.65)	0.717***	(4.09)	0.241	(0.95)
FUT_CA_OIBVOL	3.216***	(25.28)	0.0616	(0.50)	0.255	(1.45)
FUT_PA_OIBVOL	2.618***	(16.19)	-3.237***	(-11.50)	-3.001***	(-10.98)
FUT_NCNPA_OIBVOL	2.182***	(17.31)	-3.153***	(-14.48)	-2.811***	(-12.55)
FUT_CNA_OIBVOL	1.410***	(19.47)	-0.835***	(-3.77)	-0.931***	(-4.89)
FUT_PNA_OIBVOL	4.775***	(27.20)	2.577***	(12.01)	1.007***	(4.61)
FUT_NCNPNA_OIBVOL	3.963***	(38.76)	2.301***	(18.64)	1.684***	(9.35)
L1_CM_CA_OIBVOL	-2.517***	(-35.77)	0.484*	(2.46)	0.883***	(8.51)
L1_CM_PA_OIBVOL	-1.232***	(-20.72)	0.418***	(3.89)	0.240*	(2.48)
L1_CM_NCNPA_OIBVOL	-0.961***	(-11.04)	0.646***	(6.61)	0.435***	(4.91)
L1_CM_CNA_OIBVOL	-1.749***	(-22.59)	0.816***	(3.58)	0.837***	(5.28)
L1_CM_PNA_OIBVOL	-3.596***	(-9.30)	-0.665*	(-2.27)	-1.386***	(-6.54)
L1_CM_NCNPNA_OIBVOL	-2.762***	(-18.79)	1.301***	(15.22)	0.515**	(3.26)
L1_FUT_CA_OIBVOL	-0.308***	(-7.76)	0.0374	(0.31)	0.0792	(0.83)
L1_FUT_PA_OIBVOL	-0.0555	(-0.89)	-0.928*	(-2.22)	-0.136	(-1.49)
L1_FUT_NCNPA_OIBVOL	-0.163***	(-3.96)	-0.146	(-0.57)	-0.340***	(-4.84)
L1_FUT_CNA_OIBVOL	0.259***	(5.05)	1.438**	(2.85)	0.416***	(3.63)
L1_FUT_PNA_OIBVOL	0.190***	(3.57)	0.933***	(9.74)	0.181*	(2.05)
L1_FUT_NCNPNA_OIBVOL	0.122**	(2.75)	0.485**	(3.13)	0.227*	(1.98)
No. of Obs.		7162806		709038		706888
Adj. R^2		0.054		0.001		0.007

7 Conclusion

We look at the problem of determining lead-lag relationship between financial markets through order imbalance. Using six months (Jan-Jun 2015) of intraday data for 160 stocks traded in both the spot and futures market of National Stock Exchange (NSE), we find that information in the futures market leads the spot market in the order of one minute. Using the unique dataset, we also categorize the order imbalance due to different trader categories. We find that none of the algorithmic trader groups are informed. We also find that information flow from the futures to the spot market, even during extreme market volatility, happen due to non-algorithmic traders only.

A Robustness Check

In order to ascertain if our findings related to categorized order imbalances are robust, we re-run our tests following different classification schemes. The dataset allows us to classify traders as prop traders, custodians and non-custodian non-prop (NCNP) categories. It also allows us to distinguish algo and non-algo traders. Non-algo NCNP category is primarily made up of retail investors [16]. As such in this analysis, we exclude the NCNP trader group, both algo and non-algo from the consideration.

For the first set of models, we classify traders into four categories - prop algo, institutional (custodian) algo, prop non-algo and institutional non-algo. Next, similar to earlier section, we regress excess spot market returns at one-minute intervals on spot and futures market order imbalance due to these four trader categories. We use both definitions of order imbalance - OIBNUM and OIBVOL. We use the full sample of returns as well as subset of positive and negative extreme returns as independent variables.

Consistent with earlier results, we find that the co-efficients of the lagged futures market order imbalance for both non-algo trader categories remain positive and significant, suggesting non-algo traders are responsible for flow of information from futures to spot market. The co-efficient of the lagged futures market order imbalance for the algo institutional traders is positive and significant during periods of extreme returns, but negative and significant for the full sample. This result suggests that institutional trades may be informed only during extreme price movements, not always.

[16] It should be noted that retail traders are not permitted to use algorithms to execute their trades in the Indian markets.

Table 13: Panel Data Regressions of 1 min interval cash market excess returns on contemporaneous and lagged order imbalances of the cash and futures market split across different trader categories. Order imbalance (OIB) for this panel is measured as the estimated number of buyer-initiated minus seller-initiated trades scaled by the total number of trades (OIBNUM) over one minute interval. The first model uses the full sample of returns, while the second and third model uses the subset with positive and negative extreme returns.

Explanatory Variable: Excess Spot Returns						
	Full Sample		Positive Extremes		Negative Extremes	
Constant	0.0661***	(22.54)	20.07***	(311.90)	-19.82***	(-429.48)
CM_CA_OIBNUM	6.440***	(38.69)	0.338	(0.92)	1.282***	(5.10)
CM_PA_OIBNUM	9.649***	(22.42)	0.362	(1.05)	-0.0424	(-0.12)
CM_CNA_OIBNUM	7.840***	(35.42)	1.136*	(2.57)	2.528***	(7.42)
CM_PNA_OIBNUM	9.928***	(9.81)	3.155*	(2.12)	2.189***	(7.40)
FUT_CA_OIBNUM	3.525***	(24.24)	-0.320*	(-2.11)	-0.143	(-0.70)
FUT_PA_OIBNUM	3.653***	(19.83)	-3.302***	(-11.22)	-2.855***	(-10.07)
FUT_CNA_OIBNUM	1.659***	(19.11)	-1.250***	(-4.40)	-1.355***	(-5.75)
FUT_PNA_OIBNUM	5.561***	(29.10)	2.791***	(11.77)	0.992***	(4.47)
L1_CM_CA_OIBNUM	-3.957***	(-33.32)	1.103***	(4.65)	0.782***	(5.65)
L1_CM_PA_OIBNUM	-1.374***	(-14.61)	0.832***	(3.38)	0.563***	(5.04)
L1_CM_CNA_OIBNUM	-3.569***	(-29.08)	0.726**	(2.97)	0.977***	(4.53)
L1_CM_PNA_OIBNUM	-7.946***	(-7.71)	-1.624**	(-3.30)	-2.464***	(-8.79)
L1_FUT_CA_OIBNUM	-0.389***	(-8.37)	0.379**	(3.03)	0.390***	(3.64)
L1_FUT_PA_OIBNUM	0.0109	(0.18)	-0.516	(-1.36)	0.196*	(2.27)
L1_FUT_CNA_OIBNUM	0.193**	(3.25)	1.760**	(3.14)	0.630***	(4.82)
L1_FUT_PNA_OIBNUM	0.436***	(8.36)	1.375***	(12.74)	0.489***	(5.18)
No. of Obs.		7,162,806		709,038		706,888
Adj. R^2		0.033		0.001		0.003

Table 14: Panel Data Regression of 1 min interval cash market excess returns on contemporaneous and lagged order imbalances of the cash and futures market split across different trader categories. Order imbalance (OIB) for this panel is measured as the estimated buyer-initiated minus seller-initiated INR volume of transactions scaled by total INR volume (OIBVOL) over one minute interval. The first model uses the full sample of returns, while the second and third model uses the subset with positive and negative extreme returns.

Explanatory Variable: Excess Spot Returns						
	Full Sample		Positive Extremes		Negative Extremes	
Constant	0.051***	(18.38)	20.200***	(231.92)	-19.910***	(-457.75)
CM_CA_OIBVOL	5.220***	(43.91)	0.464**	(2.78)	1.466***	(7.22)
CM_PA_OIBVOL	7.516***	(26.70)	-1.048***	(-4.16)	-1.065***	(-3.98)
CM_CNA_OIBVOL	5.669***	(36.43)	0.771*	(2.24)	1.819***	(8.38)
CM_PNA_OIBVOL	9.589***	(10.01)	2.144	(0.79)	2.009***	(4.42)
FUT_CA_OIBVOL	3.398***	(23.79)	-0.399**	(-2.83)	-0.156	(-0.84)
FUT_PA_OIBVOL	3.381***	(18.27)	-3.677***	(-12.26)	-3.447***	(-11.90)
FUT_CNA_OIBVOL	1.460***	(18.82)	-0.947***	(-4.27)	-1.036***	(-5.35)
FUT_PNA_OIBVOL	5.416***	(28.78)	2.755***	(12.29)	0.997***	(4.58)
L1_CM_CA_OIBVOL	-2.759***	(-37.64)	0.607**	(3.01)	0.915***	(8.77)
L1_CM_PA_OIBVOL	-0.698***	(-11.74)	0.944***	(8.96)	0.579***	(6.90)
L1_CM_CNA_OIBVOL	-1.761***	(-22.22)	0.915***	(4.05)	0.897***	(5.72)
L1_CM_PNA_OIBVOL	-3.079***	(-7.94)	-0.301	(-1.03)	-1.200***	(-6.05)
L1_FUT_CA_OIBVOL	-0.531***	(-12.99)	0.234	(1.94)	0.229*	(2.34)
L1_FUT_PA_OIBVOL	-0.101	(-1.82)	-0.637	(-1.63)	-0.005	(-0.06)
L1_FUT_CNA_OIBVOL	0.101	(1.93)	1.535**	(3.05)	0.474***	(4.10)
L1_FUT_PNA_OIBVOL	0.387***	(7.90)	1.280***	(12.26)	0.389***	(3.82)
No. of Obs.		7,162,806		709,038		706,888
Adj. R^2		0.028		0.001		0.004

References

References

Brogaard, J., Hendershott, T., & Riordan, R. (2014). High-Frequency Trading and Price Discovery. *Review of Financial Studies*, *27*(8), 2267–2306. doi: 10.1093/rfs/hhu032

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*(1), 123–152. doi: 10.1017/CBO9781107415324.004

Chordia, T., Roll, R., & Subrahmanyam, A. (2002). Order imbalance, liquidity, and market returns. *Journal of Financial Economics*, *65*(1), 111–130. doi: 10.1016/S0304-405X(02)00136-8

Chordia, T., & Subrahmanyam, A. (2004). Order imbalance and individual stock returns: Theory and evidence. *Journal of Financial Economics*, *72*(3), 485–518. doi: 10.1016/S0304-405X(03)00175-2

Cushing, D., & Madhavan, A. (2000). Stock returns and trading at the close. *Journal of Financial Markets*, *3*(1), 45–67. doi: 10.1016/S1386-4181(99)00012-9

Finnerty, J. E., & Park, H. Y. (1992). Stock Index Futures: Does the Tail Wag the Dog? *Financial Analysts Journal*, *43*(2), 57–61.

Gallant, A. R., Rossi, P. E., & Tauchen, G. (1992). Stock Prices and Volume. *The Review of Financial Studies*, *5*(2), 199–242. doi: 10.1093/poq/nfi058

Handa, P., Schwartz, R., & Tiwari, A. (2003). Quote setting and price formation in an order driven market. *Journal of Financial Markets*, *6*(4), 461–489. doi: 10.1016/S1386-4181(02)00041-1

Harris, L. (1989). The October 1987 S & P 500 Stock-Futures Basis. *The Journal of Finance*, *44*(1), 77–99.

Hendershott, T., Jones, C. M., & Menkveld, A. J. (2011). Does Algorithmic Trading

Improve Liquidity? *The Journal of Finance*, *66*(1), 1–34. doi: 10.1111/j.1540-6261.2010.01624.x

Hiemstra, C., & Jones, J. D. (1994). Testing for Linear and Nonlinear Granger Causality in the Stock Price- Volume Relation. *The Journal of Finance*, *49*(5), 1639–1664.

Huang, Y. C., & Chou, J. H. (2007). Order imbalance and its impact on market performance: Order-driven vs. quote-driven markets. *Journal of Business Finance and Accounting*, *34*(9-10), 1596–1614. doi: 10.1111/j.1468-5957.2007.02038.x

Kawaller, I. G., Koch, P. D., & Koch, T. W. (1987). The Temporal Price Relationship Between S&P 500 Futures and the S&P 500 Index. *The Journal of Finance*, *42*(5), 1309–1329.

Kraus, A., & Stoll, H. R. (1972). Parallel Trading by Institutional Investors. *The Journal of Financial and Quantitative Analysis*, *7*(5), 2107–2138.

Kyle, A. S. (1985). Continuous Auctions and Insider Trading. *Econometrica*, *53*(6), 1315–1335. doi: 10.3982/ECTA6822

Lakonishok, J., Shleifer, A., & Vishny, R. W. (1992). The impact of institutional trading on stock prices. *Journal of Financial Economics*, *32*, 23–42.

Lo, A. W., & Wang, J. (2000). Trading Volume : Definitions , Data Analysis , and Implications of Portfolio Theory. *The Review of Financial Studies*, *13*(2), 257–300.

Nawn, S., & Banerjee, A. (2018). Do Proprietary Algorithmic Traders Withdraw Liquidity during Market Stress? *Financial Management*, 1–36. doi: 10.1111/fima.12238

Sias, R. W. (1997). Price pressure and the role of institutional investors in closed-end funds. *Journal of Financial Research*, *20*(2), 211–229. doi: 10.1111/j.1475-6803.1997.tb00245.x

Stoll, H. R. (2000). Friction. *The Journal of Finance*, *55*(4), 1479+1478+1480—-1514.

Stoll, H. R., & Whaley, R. E. (1990). The Dynamics of Stock Index and Stock Index Futures Returns. *The Journal of Financial and Quantitative Analysis*, *25*(4), 441–468.

Wermers, R. (1999). Mutual Fund Herding and the Impact on Stock Prices. *The Journal of Finance*, *54*(2), 581–622.