Order Imbalance and Return Predictability: Evidence from Korean Index Futures

Hyoung-Goo Kang
Department of Finance
Hanyang University
hyoungkang@hanyang.ac.kr

Soo-Hyun Kim[1]
School of Business administration
Soongsil University
soo_hyun.kim@ssu.ac.kr

Kyuseok Lee
College of Business
KAIST
kslee2@business.kaist.ac.kr

Abstract

The higher the ratio of buy-order quantities minus sell-order quantities to the entire limit order book quantity, the higher the future returns in the next 5- to 20-minute intervals. The order imbalance predicts the size of future returns as well. In addition, the ratio of trading volume to open interests, predicts the direction and size of future returns. These novel findings are obtained from analyzing all the open limit-order book information provided in a real-time basis for the KOSPI200 futures. The findings suggest that traders can use aggregate quantities of buy versus sell orders in order to optimize their trades.

Key-words: order book, limit order, order imbalance, high frequency data, futures contract

JEL classification: G12, G14

[1] Corresponding author.
Address: 417 Kyungsang-Kwan, Soonsil University, 369 Sangdoro Dongjak-Gu, Seoul, Korea.
Tel: +82-2-820-0544, Fax: +82-2-824-4384

Introduction

The information content of an order book is an important subject, which is the focus of a debate on whether and how much order book information enhances the process of price discovery. As an "open order book" makes headway, such debates become even more significant. Indeed, major exchanges, including the CME□ Globex, NYMEX, and Singapore Exchange,disclose order book information to the public on a real-time basis ("open order book"), although they provide the order book information up to first five bid and ask orders.

□ □

Baruch (2005) and Boehmer, Saar, and Yu (2005) find that the open order book improves market liquidity and informational efficiency in transacted prices in the New York Stock Exchange. On the other hand, Madhavan, Porter, and Weaver (2005) find that the open order book increases volatility and bid-ask spreads in the Toronto Stock Exchange.

Recently, Cao,□ Hansch□ and Wang (2008, 2009) test whether the open order book information contributes to price discovery and whether it can predict future short-term returns. Analyzing the 100 most actively traded stocks in the Australian Stock Exchange, they find that the order book information contributes to price discovery, although its extent is moderate (22% in terms of the Hasbrouck's (1995) information share), and is helpful in predicting future short-term returns.

We contribute to the literature on the information content of an order book by analyzing the order book information provided in a real-time basis for the KOSPI200 futures, Korean stock index futures known as one of the most actively traded financial products in the world. The KOSPI200 futures market offers a unique environment for analyzing the information content of an order book for three reasons. First, the market provides all the order book information to public traders. In comparison, in most electronic limit order markets, including markets of countries such as Australia, Canada, France, Germany, Hong Kong, Singapore, the United Kingdom, and the United States, at most five to ten best bid- and offer data are provided either by the exchanges or the brokerage firms. Second, the market is extremely liquid, and hundreds of transactions occur per minute. In fact, there are many low latency traders in the market who trade on a millisecond basis. Due to such extreme liquidity, transacting, submitting, and canceling orders are executed so frequently that the distributional shape of all the limit orders changes very fast and very dramatically, making it impractical to analyze the distributional shapes of limit orders and use them in trading as in Cao, Hansch, and Wang (2008, 2009). Third, order prices are pre-fixed at multiples of 0.05 around the last transacted price. Pre-fixed order prices, together with extreme liquidity, make traders tend to compete with the quantities of limit orders, rather than limit prices, by simply filling in pre-specified limit prices with the desired order quantities. Put differently, the traders cannot make new limit order prices other than pre-fixed prices so that they simply put orders to one of the many pre-fixed order prices set by the exchange. This practice allows us to examine the pure effect of the order book imbalance in isolation.

We analyze order book information and other market data at 5-, 10-, and 20-minute intervals. Our main findings are as follows. First, order imbalance, measured by the buy-order quantities minus the sell-order quantities of all the order book quantities, is significantly positively related to the direction of future returns. That is, excess sell limit orders lower future returns, while excess buy limit orders work in the opposite way. This predictive power remains significant at least up to 20 minutes. Second, trading intensity, measured by the ratio of trading volume to open interests, is significantly negatively related with the direction and size of future returns. Third, the larger the relative trading volume, the larger the future volatility. These findings are obtained by controlling for the effects of lagged return and volatility.

An order book imbalance contains information about potential or unrealized expectations on asset values. Thus, it is likely to affect the direction of future returns. In comparison, trading intensity can be regarded as a measure of the extent of the information digested. The higher the trading intensity, the larger the executed orders, so more information is digested and incorporated into prices. Thus, the trading intensity is likely to affect the magnitude of future price movements. Our findings are consistent with these expectations.

Our findings have practically important implications. First, in a highly liquid market, where the distributional shape of the limit orders changes very quickly, it could easily become impractical for traders to implement trading strategies based on the distributional shape of limit orders. Our findings suggest that using the information of order book quantities could provide an easier, practical way of using order book information in predicting future returns and fluctuations as well. Second, trading intensity is another important variable that can be used in predicting the direction, size, and volatility of future returns.

Our findings are also generally consistent with the predictions of theoretical models by Rosu (2009) in that only the numbers of the buy and sell limit orders, rather than the shape of the limit orders, matters in equilibrium, and higher trading activity leads to lower volatility.

The rest of this paper proceeds as follows. The second section describes the data. The third section describes the variables used in our analysis. The fourth section presents the empirical methodology, and the fifth section presents the results. Concluding remarks are provided in the sixth section.

Data

Our transaction data on the KOSPI200 futures contracts are from□ Blashnet□and span the period from January 1, 2007 to March 29, 2013.[2] □Blashnet□ stacks tick-by-tick market transaction data and provides information on transaction prices, trading volumes, open interests, and remaining limit orders, etc. It is noteworthy that the KOSPI200 futures contract is a very popular investment vehicle for traders in global financial markets, and it is also well known as one of the most actively-traded financial securities in the world. To give a sense of its importance, the average daily trading volume during the sample period was 18,401,610,000,000 in Korean won (KRW), which amounts to 18,401,610,000 in U.S. dollars (USD) using an exchange rate of 1000 KRW per USD.

The Korea Exchange (KRX) lists four futures contracts maturing on the second Thursday of March, June, September, and December. The daily price limits of the futures contracts are ±10% of the closing price of the previous day. Trading hours are 09:00 to 15:15 Korean time (09:00 to 14:50 on the last trading day).

[2] Blashnet is a local data vendor specialized in intra-day trading support. http://www.blashnet.com

Most importantly, the KRX provides information of the entire limit orders of the KOSPI200 futures contracts. This practice is unique in the sense that most major exchanges (at least all major exchanges such as the U.S., Canada, Hong Kong, Singapore, and most European countries) do not reveal the entire limit order book data; they provide only the information of the first five buy and sell limit orders.

Another distinctive characteristic of KOSPI 200 futures contracts is that the prices that trades can submit limit orders (i.e., the tick size) are pre-set at multiples of ±0.05 around the last transacted market price. These pre-set prices are always fulfilled by limit orders very quickly due to the extreme liquidity of the KOSPI200 futures market. Hence, traders trading over a very short time horizon, such as 5- to 10-minute horizons, usually do not compete with limit-order prices by generating new limit order prices. This practice for the KOSPI200 futures contracts enables us to examine the pure effect of the order imbalance in isolation by simply examining the total order quantities from both the sell side and the buy side order book.

More detailed information on the KOSPI200 futures contracts is available from the webpage of the KRX (http://eng.krx.co.kr).

Variables

We consider three time intervals: 5-, 10-, or 20-minute intervals. We do not use the first and last

5-minute return data of each day for empirical analysis at the 5-minute interval. Likewise, we do not use the first and last 10-minute (20- minute) return data for empirical analysis at the 10-minute (20- minute) interval.

The variables used in this study are summarized in Table 1.

*** Table 1 about here. ***

It is noteworthy that in the definition of the trading intensity variable, we use both the trading volume and the open interests summed across all four futures contracts outstanding. In practice, only the two futures contracts with the two nearest maturities are actively traded. However, we use open interests across all four maturities because of the rollover practice. As each expiration day approaches, traders roll over to contracts maturing the next maturity dates. Thus, if we normalize the trading volume using only the open interests of the contract maturing on the nearest maturity date, the trading intensity variable will always be exaggerated near the expiration dates.

*** Figure 1 about here. ***

Figure 1 presents the density plots of the returns measured over three time intervals. The distributional shapes appear to be fairly symmetric and bell-shaped. Indeed, the skewness statistics for the returns measured over 5-, 10-, and 20- minute intervals are close to zero: -0.214, -0.009, and -0.139, respectively. On the other hand, the kurtosis statistics for the returns measured over 5-, 10-, and 20-minute intervals were 36.67, 35.31, and 25.40, respectively. Thus, although bell-shaped, the distributional shapes are very different from a normal distribution. Still, the kurtosis statistic decreases with the length of the time interval. This is expected since the return distribution measured over a wider interval is likely to be closer to the normal distribution due to the effect of the returns summation.

*** Table 2 about here. ***

Table 2 presents the descriptive statistics of the regression variables, separately for each of the three time intervals. The mean returns are essentially zero for all three time intervals, but the minimum and maximum returns are fairly large ranging from 5% to 8%, in absolute-value terms, depending on the time intervals. Interestingly, the minimum and maximum order imbalances also indicate that there are such times, albeit rare, when almost all traders trade in the same direction of the market (i.e., herd). We tested this with 1% winsorization for robustness, and the results were essentially the same.

*** Table 3 about here. ***

Table 3 presents the contemporaneous correlations among the variables. Double asterisks indicate statistical significance at the 5% level based on a Pearson's two-tailed t-test. The correlations are significant for all pairs of variables. A few points are noteworthy from the table. First, the returns and order imbalances are positively correlated. That is, the order imbalance is positively correlated with the contemporaneous return. Of course, this contemporaneous correlation between the returns and order imbalances does not allow us to distinguish between the three possibilities that i) high returns are driven by large buy limit orders waiting to be executed; ii) high returns induce large buy limit orders of momentum traders; and iii) market-wide news could generate a common shock to both returns and buy limit orders. Second, returns and relative volumes are negatively correlated. Thus, on average, larger transactions occur in times of lower returns. Third, returns and volatilities are negatively correlated, suggesting that markets tend to be more volatile in times of lower returns. Fourth, order imbalances are negatively correlated with both trading intensities and volatilities. Thus, on average, the order imbalance is likely to be higher when the relative volumes or volatilities are lower. Fifth and last, the positive correlation between relative volume and volatility was as expected. This is simply because, unless there are many transactions, there cannot be large price movements.

Methodology

To investigate whether open order book information is helpful in predicting future short-term returns, we employ the following predictive regression equation:

$$r_{t+1} = \alpha + \beta_1 r_t + \beta_2 VOL_t + \beta_3 OI_t + \beta_4 TI_t + +\epsilon_t. \tag{1}$$

The two variables *Order Imbalance (OI)* and *Trading Intensity (TI)* are of our main interest regarding the question of whether they include valuable information that can be used to predict future returns. Note that we are controlling for the effects of the return autocorrelations and the return volatility.

To investigate the predictive power of the two variables—*Order Imbalance (OI)* and *Trading Intensity (TI)*—on the magnitude of the future returns and the future return volatilities, we employ the same regression Equation (1) but by replacing r_{t+1} with r_{t+1}^2 and VOL_{t+1}, respectively.

Results

Table 4 reports the results of the predictive regressions, separately for the three time intervals considered. The first column lists the dependent variables, and the first row lists the explanatory variables. The numbers in each cell are the coefficients and t-statistics in parentheses. The coefficients are scaled by 10^{-4}. The return-squared variable is used as a proxy for the size of the return.

*** Table 4 about here. ***

Overall, Table 4 shows that the future returns, squared future returns, and future volatilities are all predictable with current returns, order imbalances, trading intensities, and volatilities for all three time intervals. The only exception is the relationship between future returns and current returns at the 20-minute interval when the order imbalances are controlled.

In detail, the following are noteworthy from the table. First, the coefficients of the current returns, roughly speaking, the coefficients from the lag-one auto-regressive model (i.e., the

AR(1) model), are significantly negative in all cases except for one. They are significantly negative at the 5- and 10-minute intervals irrespective of whether the bid-ask ratios are controlled for or not. However, when the order imbalances are controlled for, the coefficients become less negative. This can be explained by the positive concurrent relation between the returns and the order imbalances. The overall results suggest that the order imbalances are positively associated with the future returns. Thus, when the order imbalances are controlled for, the current returns should produce larger negative future returns in order to compensate for the positive relationship between the order imbalances and the future returns. This pattern is most salient in the results at the 20-minute interval. When the order imbalances are controlled for, the current returns do not predict the future returns over the next 20-minute interval. When the order imbalances are not controlled for, the returns absorb the effects of the order imbalances and generate positively significant predictability.

The pattern of the AR(1) coefficients in Table 4 suggests that futures prices tend to overshoot current temporary shocks. Thus, prices partially rebound. Nevertheless, this rebound does not last over 20 minutes. To re-emphasize, KOSPI200 futures contracts are highly liquid with hundreds of transactions occurring per minute. Hence, aggregating the data by minute eliminates the bid-ask tick bounce and thus the rebound documented has nothing to do with the bid-ask tick bounce.

Second, the higher the returns, the smaller the magnitude of the return change and volatility in the future. Thus, there is asymmetry. Negative returns generate a larger price movement in the future than positive returns. This phenomenon is often interpreted as the leverage effect (Black,

1976) predoicting that volatility increases more after bad news than after good news.

Third, order imbalance, our main variable of interest, predicts future returns. Its coefficients are significantly positive and remain significant up to 20 minutes. Thus, the larger the buy limit orders a relative sell, the larger the future returns. On a per minute basis, the response of the future returns decreases, i.e., from 1.68/5 min to 2.99/10 min and 4.50/20 min. Nevertheless, the coefficients are all significant. Roughly, the adjusted R^2 doubles when the order imbalances are included in the model.

There are two ways of interpreting the positive relationship between the order imbalances and the future returns: (1) the shock to the order imbalances may be correlated with the future valuation shocks to return. For example, the large buy limit orders relative to sell limit orders are driven by positive news. Some of this news may not be incorporated in the current prices yet, so price will increase in accordance; (2) the shock to the order imbalances may be correlated with future liquidity shocks to return. For example, the large buy limit orders relative to the sell limit orders reflect the trading schedule of institutions which in turn results from the requests of internal or external clients. Therefore, the order imbalance is likely to remain in the future, to generate excess demand resulting in increased prices in future.

Fourth, the higher the order imbalance, the lower the future volatility of the futures price. Since we control for past return, we cannot ascribe this finding entirely to the leverage effect such as “high bid-ask ratio generates a high return, which produces lower volatility.” In contrast, order imbalance predicts the size of the return change negatively in 5 minutes, but positively in 10 minutes and 20 minutes. This finding is puzzling and difficult to account for.

Fifth, the larger the trading intensity, the lower the future returns and the size of the future return movement. Large trading intensity indicates sufficient liquidity. This decreases market impact in execution and results in a decrease in price movement. This also reduces the liquidity risk premium in the short term and thus reduces the expected return, which produces low future returns. The relation between trading intensity and future volatility is more complex. In the short term, volatility increases due to the surge in the trading volume that an open interest may not digest. However, in the mid- (10 min) and long term (20 min), volatility decreases as information diffuses across the traders and the market stabilizes with deeper liquidity.

****** Table 5 about here. ******

Table 5 includes two interaction terms in addition: (order imbalance)*(trading intensity) and (volatility)*(trading intensity). First, the order imbalance is inconclusive concerning the size of the future price movement. For the results at the 5-minute interval, the coefficient on the bid-ask ratio is (0.01-1.46* trading intensity)* order imbalance. Thus, when trading intensity <=0.007, the order imbalance and future returns are positively related, while when trading intensity >0.07, they are negatively related. An order imbalance decreases future price movement with (order imbalance)*(trading intensity), while it increases future movement with (volatility)*(trading intensity). However, order imbalance always increases future movement in the long term. This result is in line with Table 4. Second, the interaction terms are very significant, and the predictive power (adjusted R^2) increases with the interval.

Our interpretation of the interaction terms is as follows. We already know the large relative volume, and that the trading intensity decreases the size of the future stock price movement because the large liquidity decreases the market impact. When the order imbalance is positive, such a large trading intensity is likely to be driven by buy limit orders and increases price. Due to the leverage effect, the change in price should be smaller when the price increases are driven by a positive order imbalance than when the price decrease is driven by a negative order imbalance.

This intuition about the volatility is similar. The size of the price movement varies over time, but in a persistent way (Bollerslev, 1986). The fifth column of Table 5 also shows this pattern. Thus, the larger the present volatility, the larger the future volatility. In addition, if the current relative volume is high, this means that the current high volatility happens even under large market liquidity. High volatility under high liquidity should be more persistent and dominant than those under low liquidity because it is more difficult for prices to be volatile in a liquid market.

Conclusion

We examine the intraday effects of the order imbalances on the future short-term behavior of the KOSPI200 futures prices. The KOSPI200 futures market provides a unique setup for the analysis of the information content of the order books in that i) all the order book data are publicly open to any traders on a real-time basis; ii) it is a highly liquid market; and iii) the limit order prices (bid and ask prices) are pre-fixed by multiples of 0.05. The second and third features make it impractical for traders to use strategies based on the distributional shape of all the limit orders because they change very fast and dramatically.

Using a simple proxy for the order imbalance based on the numbers of buy and sell limit orders, we find that the order imbalances predict the direction of the futures prices positively but the variability of futures prices negatively. Furthermore, we find that the extent of trading activity also has a significant predictive power on the short-term behavior of futures prices. Indeed, the ratio of trading volume to open interests is significantly negatively related to both the direction and the variability of the futures prices. The results are obtained by controlling for the effects of the return autocorrelations and volatility.

Our research is based on price data measured at one-minute intervals aggregated over 5- to 20-minute intervals. However, the behavior of high-frequency traders might be captured better at even higher frequencies. With a tick-by-tick data set, a comparison between the low- and mid-latency trading would be an interesting and meaningful subject for future research.

References

Baruch, S. (2005). Who benefits from an open limit-order book?. Journal of Business, 78(4), 1267-1306.

Black, F. (1976). The pricing of commodity contracts, Journal of Financial Economics, 3, 167-179.

Boehmer, E., Saar, G., & Yu, L. (2005). Lifting the veil: An analysis of pre-trade transparency at the NYSE. Journal of Finance, 60(2), 783-815.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307-327.

Cao, C., Hansch, O., & Wang, X. (2008). Order placement strategies in a pure limit order book market. Journal of Financial Research, 31, 113–140.

Cao, C., Hansch, O., & Wang, X. (2009). The information content of an open limit-order book. Journal of Futures Markets, 29(1), 16-41.

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. Journal of Finance, 50(4), 1175-1199.

Kyle, A. S. (1985). Continuous auctions and insider trading. Econometrica, 53(6), 1315-1335.

Madhavan, A., Porter, D., & Weaver, D. (2005). Should securities markets be transparent? Journal of Financial Markets, 8(3), 265-287.

Rosu, I. (2009). A dynamic model of the limit order book. Review of Financial Studies, 22(11), 4601-4641.

Figure 1: Distribution of Returns

This figure presents the density plots of the returns on the KOSPI200 index futures calculated over three time intervals: 5-, 10-, and 20-minute intervals. The sample period spans from January 2007 to March 2013.

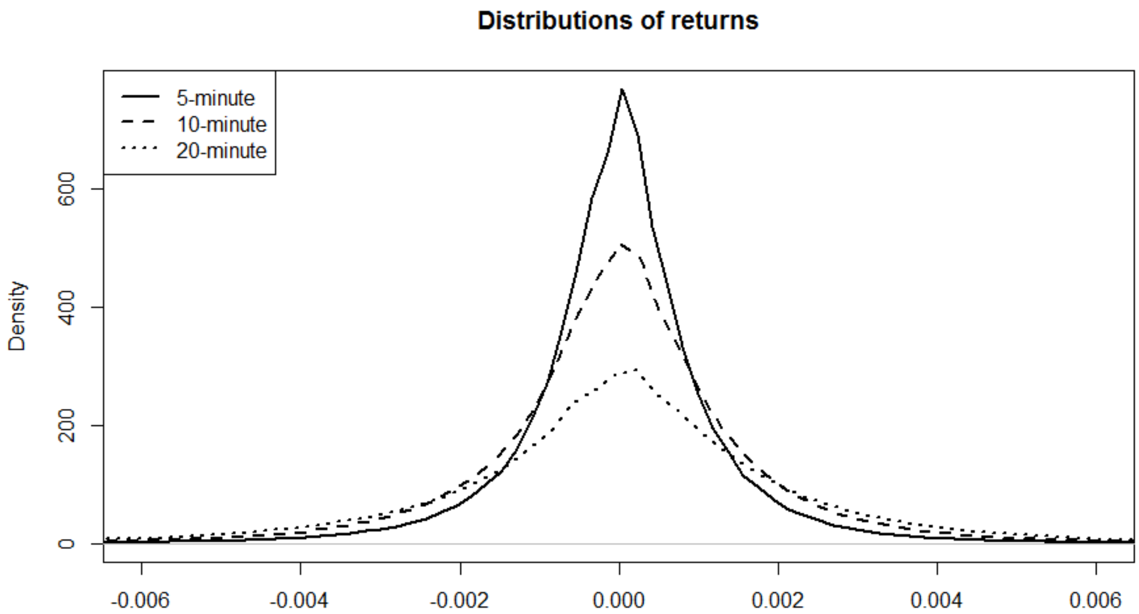

Table 1: Definition of the Variables

This table presents the definitions of the variables used in the paper.

Variable	Definition
Price	The last transacted price nearest to the end of each minute. The KOSPI200 futures contracts are extremely liquid, so that hundreds of transactions occur per minute. Hence, using the prices measured over one-minute intervals eliminates the likely bid-ask bounce effect in consecutive prices.
Return (r)	The simple returns measured at the end of each time interval.
Order Imbalance (OI)	The ratio of the total buy-limit-order quantities minus the total sell-limit-order quantities to all limit order book quantities. The variable is computed at the end of each time interval. This variable can be seen as the bid-ask ratio.
Trading Intensity (TI)	The ratio of trading volume to open interests. The trading volume represents the sum of all of the trading volume across all four futures contracts outstanding within each time interval. The open interests also represent the sum of all the open interests across all four futures contracts outstanding within each time interval. The variable is computed at the end of each time interval. This variable can be seen as the relative trading volume.
Volatility (VOL)	The standard deviation of the returns measured over one-minute intervals within each time interval.

Table 2: Descriptive Statistics

This table presents the descriptive statistics of the data obtained from Blashnet from January of 2007 to March of 2013. We collected the last transaction data at the end of each minute. Panel A, B, and C describe the return, volatility based on transaction prices at each minute, trading intensity and order imbalance over selection of 5-minute, 10-minute, and 20-minute intervals. The sum of all the sell limit orders given in a transacted price is what we call the total remaining sell order book quantities or the remaining quantity of the total sell limit orders. For simplicity, let us call the remaining quantity of the total sell limit orders #sell_orderbook. The sum of all these buy limit orders given a transacted price is what we call the total remaining buy order book quantities or the remaining quantity of the total buy limit orders. For simplicity, let us call the remaining quantity of the total buy limit orders as #buy_orderbook. We define the order imbalance as (#buy_orderbook - #sell_orderbook)/(#buy_orderbook + #sell_orderbook). Thus, this measures the relative stacked size of the bid orders (bid orderbook) over the ask orders (ask orderbook). Trading Intensity is the sum of the total transactions divided by the total open interests of the futures at all expiration dates. Traders should roll over their contracts to the next month contracts. Therefore, using the total open interests for the denominator stabilizes the variables. Volatility is computed at the end of each interval with minute data within the interval. The return is computed over each interval.

Panel A: 5-minute interval

	Mean	Minimum	Q1	Median	Q3	Maximum
Return	-0.0000	-0.0457	-0.0005	0.0000	0.0005	0.0507
Volatility	0.0005	0.0000	0.0003	0.0004	0.0006	0.1730
Trading Intensity	0.0069	0.0000	0.0028	0.0052	0.0090	0.1190
Order Imbalance	0.0087	-0.9868	-0.1347	0.0071	0.1557	0.9998

Panel B: 10-minute interval

	Mean	Minimum	Q1	Median	Q3	Maximum
Return	0.0000	-0.0548	-0.0008	0.0000	0.0008	0.0774
Volatility	0.0067	0.0000	0.0003	0.0004	0.0006	0.0120
Trading Intensity	0.0088	0.0000	0.0038	0.0051	0.0088	0.1190
Order Imbalance	0.0086	-0.9865	-0.1352	0.0070	0.001563	0.9996

Panel C: 20-minute interval

	Mean	Minimum	Q1	Median	Q3	Maximum
Return	-0.0000	-0.0690	-0.0011	0.0000	0.0012	0.0670
Volatility	0.0005	0.0000	0.0003	0.0004	0.0006	0.0093
Trading Intensity	0.0065	0.0000	0.0027	0.0050	0.0085	0.1190
Order Imbalance	0.0086	-0.9865	-0.1363	0.0069	0.1570	0.9996

Table 3: Contemporaneous Correlations

This table describes the contemporaneous correlations. Correlations between the return, order imbalance, trading intensity and volatility are computed and tested with Pearson's two-tailed t-test. Double asterisks indicate statistical significance under 5%. See Table 1 for the definition of each variable.

Panel A: 5-minute

	Return	Order Imbalance	Trading Intensity	Volatility
Return	1	0.223**	-0.033**	-0.024**
Order Imbalance	0.223**	1	-0.069**	-0.084**
Trading Intensity	-0.033**	-0.069**	1	0.336**
Volatility	-0.024**	-0.084**	0.336**	1

Panel B: 10-minute

	Return	Order Imbalance	Trading Intensity	Volatility
Return	1	0.288**	-0.041**	-0.038**
Order Imbalance	0.288**	1	-0.071**	-0.092**
Trading Intensity	-0.041**	-0.071**	1	0.307**
Volatility	-0.038**	-0.092**	0.307**	1

Panel C: 20-minute

	Return	Order Imbalance	Trading Intensity	Volatility
Return	1	0.367**	-0.046**	-0.046**
Order Imbalance	0.367**	1	-0.074**	-0.096*
Trading Intensity	-0.046**	-0.074**	1	0.272**
Volatility	-0.046**	-0.096**	0.272**	1

Table 4: Predictive Regression Analysis

This table presents the results of the predictive regression. The first column lists the dependent variables. The first row lists the independent variables. The numbers in each cell are the coefficients and the t-statistics (parenthesis) of the multivariate regression. The last column shows the adjusted R-squared. The coefficients are scaled with 10^{-4}. The dependent variables are one interval ahead of the independent variables. The square returns are a proxy for the size of the returns.

Panel A: 5-minute

	$return_t$	Order Imbalance	Trading Intensity	$volatility_t$	Adj. R^2 (%)
$return_{t+1}$	-316(-23.02)	1.68(18.30)	-6.45(-1.95)	486(10.02)	0.15
$return_{t+1}$	-261(-19.64)		-8.78(-2.66)	431(8.88)	0.08
$return_{t+1}^2$	-2.25(-20.64)	-0.05(-6.32)	-0.80(-30.42)	91.5(235.84)	9.82
$return_{t+1}^2$	-2.40(-22.59)		-0.79(-30.20)	91.7(236.70)	9.81
$volatility_{t+1}$	-106(-34.39)	-0.46(-22.06)	2.63(3.52)	6103(556.57)	39.53
$volatility_{t+1}$	-120(-40.24)		3.26(4.37)	6118(558.84)	39.48

Panel B: 10-minute

	$return_t$	Order Imbalance	Trading Intensity	$volatility_t$	Adj. R^2 (%)
$return_{t+1}$	-232(-16.62)	2.99(22.93)	-13.86(-2.97)	1254(17.44)	0.16
$return_{t+1}$	-141(-10.53)		-17.98(-3.85)	1140(15.883)	0.07
$return_{t+1}^2$	-5.97(-39.99)	0.04(2.80)	-1.56(-31.11)	192.6(250.24)	11.28
$return_{t+1}^2$	-5.85(-40.88)		-1.56(-31.24)	192.4(250.64)	11.28
$volatility_{t+1}$	-57.02(-32.61)	-0.27(-16.53)	-11.07(-18.91)	7463(828.25)	58.78
$volatility_{t+1}$	-65.25(-38.91)		-10.70(-18.28)	7473(831.16)	58.76

Panel C: 20-minute

	$return_t$	Order Imbalance	Trading Intensity	$volatility_t$	Adj. R^2 (%)
$return_{t+1}$	-5.77(-0.39)	4.50(23.54)	-24.10(-3.56)	2536(23.78)	0.22
$return_{t+1}$	132.7(9.60)		-30.58(-4.52)	2358(22.15)	0.11
$return_{t+1}^2$	-9.66(-53.91)	0.04(16.01)	-2.54(-31.04)	396.2(307.38)	16.87
$return_{t+1}^2$	-8.61(-51.60)		-2.59(-31.71)	394.8(306.94)	16.83
$volatility_{t+1}$	-52.59(-48.21)	-0.07(-4.99)	-6.78(-13.60)	8098(1031)	69.81
$volatility_{t+1}$	-54.57(-53.70)		-6.67(-13.41)	8101(1034)	69.81

Table 5: Predictive Regression with an Interaction Variable

The predicted regressions for the squared return are conducted using a proxy for the future return size. In addition to the independent variables from Table 4, we included the interaction variables between the order imbalance (OI) and trading intensity (TI), and between the volatility and the trading intensity (TI).

Dependent variable: $return_{t+1}^2$

min	$return_t$	OI_t	TI_t	$volatility_t$	$OI_t \times TI_t$	$volatility_t \times TI_t$	Adj. R^2 (%)
5	-1.87(-16.60)	0.01(5.67)	-0.82(-30.88)	91.48(235.6)	-1.46(-13.39)		9.85
	-2.15(-19.68)	-0.00(-6.76)	-1.40(-38.61)	80.16(130.51)		99.94(23.99)	9.92
10	-5.74(-37.31)	0.01(6.41)	-1.57(-31.36)	192.5(250.14)	-1.26(-6.17)		11.28
	-5.70(-38.13)	-0.00(2.01)	-3.47(-47.00)	157.4(124.76)		3300(35.21)	11.48
20	-9.41(-51.18)	0.05(15.52)	-2.57(-31.33)	396.2(307.34)	-2.00(-5.97)		16.88
	-9.27(-51.80)	0.03(15.05)	-6.55(-51.16)	324.2(148.17)		7192(40.67)	17.14